Cover design by Twisha Patni

Index

Letter from the Editor

Over the past months, The Economics Review at New York University has grown tremendously. We have increased our staff of writers, editors and web designers to roughly 40 people. I have never been so proud of the talent that we have on board, and it is an honor to present our 3rd printed publication.

With the recent growth that we have experienced, we are even closer to our goal, which is to give more NYU students the opportunity to conduct research and get published. Working closely with each individual writer and lifting up the ambitious economics students at NYU is the core of our organization.

This edition of The Economics Review at NYU contains four intriguing research paper covering very different topics; from social media to Russian politics, and from NBA salaries to real estate prices in New York. I am confident that you will find each piece fascinating, and eye opening. We always strive to increase students' understanding of economics while also touching their fields of interest.

Lastly, I am happy to announce that Prabhod Mudlapur will take over as Editor-in-Chief after this semester. Prabhod has been the Managing Editor of our publication the past months, and I know he will do a great job and take The Economics Review at NYU to a new level. It has been a pleasure to see the publication develop from scratch, and I look forward to seeing the future of it.

Happy reading!

Sincerely,

Ewa Staworzynska

THE VIDEO WAR – A RISING FACEBOOK CHALLENGES YOUTUBE'S DOMINANCE

Rishabh Ranawat & Shivaditya Sinha

In this paper we study the competition in the online video market by comparing its two largest companies, namely, Facebook and YouTube. This paper was undertaken with the knowledge that limited work has been conducted in this sphere. To examine both the state of the market and the economic possibilities that lie in the same, we looked at the views obtained by videos posted on either platform. For this, we obtained the data through the API(s) of the respective sites. Using Stata to plot the views and length of the videos, we show that top content creators are treating the two sites as separate. This is an important victory for both sides as it allows them to cultivate their own niche.

We live in the world of videos. They require little effort to watch and provide unlimited entertainment. The numbers seem to back this up - a study by a broadband company showed that 70% of North American downstream traffic during peak evening hours came from video streaming (D'Onfro, 2015). And that was two years ago. Based on the kind of content and number of users, only two companies stand a chance of dominating this market – Facebook and YouTube.

On the 7th of June 2014, Facebook announced that "there [had] been an average of more than 1 billion video views on Facebook every day" (Simo, 2014). Back then, 1 billion views was a big deal. Now, Facebook averages four times that amount. It has taken a little more than two years for Facebook to garner a four-hundred percent increase in views, a feat that perhaps no other company can match (Facebook Media, n.d.). During the same time, its posts have also risen. Between 2014 and 2015, Facebook claimed a 75% increase in the number of videos that each user put up (Peterson, 2015).

It is this massive growth that has lead us to believe that one of the few companies that can challenge YouTube is Facebook. YouTube has dominated the video sphere for a long time. Back in 2014, when Facebook was celebrating 1 billion views, YouTube claimed 4 billion (Simo, 2014). As it no longer releases any such figures, a direct comparison is not really possible.

However, even if Facebook has caught up in that way, YouTube may not be interested in a straightforward battle. A large portion of its views now come from established channels with millions of subscribers. It is no longer the dynamic, young player it was in its early days. It has become more professional and more organised over time. Instead of focusing on making it big with one video, these channels are constantly trying to expand their base.

Methodology

The best way to gauge Facebook's impact is to examine the top channels of YouTube. For this, the top 500 YouTube channels, by subscribers, were extracted from vidstax.com and the YouTube username(s) were pulled from the same site. In order to get the corresponding Facebook pages we used Facebook's Graph Application Program Interface (API) search endpoint and then filtered the verified page(s). Essentially, we use the usernames as unique identifiers to extract data relevant to a particular YouTube channel. In order to get the corresponding Facebook pages we used Facebook's Graph API search endpoint and then filtered the verified page(s). The Graph API is primarily a tool built by Facebook for developers to programmatically interact with their core functionalities.

Once we had the channels and the corresponding verified Facebook pages, we got the channel details from Google's API using the 'playlistItems' end point. Next, we used the videos endpoint to get 'viewCount' in the statistics data. YouTube's API was then accessed by making requests using Python's 'urllib' module. Similar to Facebook's Graph API, Google and YouTube's APIs allow us to extract relevant information related to users, videos and channels. By leveraging these services programmatically we were able to extract the views for the 10 most recent videos for every channel.

A similar technique was used to extract the view count from Facebook. Here, we first got the list of videos by the user ID that we had collected from Facebook's Graph Search API endpoint. Then, we got all the videos and their IDs that were uploaded by the specific user. Finally, we used selenium to scrape the views because Facebook's Graph API only provides the view count for cross-posted videos.

At this point, we had the two datasets. In order to match them, we started off by taking the 10 most recent videos (as returned by the API) from the YouTube for each of the users. Then, we got the max videos returned by Facebook's API and did a title fuzzy match. In such a case, the matching variable, in this case the title of the video, is not matched exactly but instead, is matched with a certain degree of confidence. For instance, if the title of one video was 'New Challenges with Bikes' and the other was 'New Bike Challenge', a simple match would not help us. However, a fuzzy match would allow us to match these videos. It should be noted that we first filtered by channel and then conducted a fuzzy match within the channels. So, a 'Buzzfeed' video would not match with one uploaded by 'Eminem'. If the match ratio was over a certain threshold, we considered them to be the same video. In case there was no Facebook counterpart, the view count for Facebook would be zero (0).

Further, we also looked at the data from Facebook's perspective. The procedure was the same as above but, the Facebook videos were the primary set used to find the corresponding YouTube matches.

It should be noted that using YouTube's top channels for Facebook is problematic but if we had taken the top Facebook pages by likes, there was no guarantee that those pages would be video centric. After all, they could be pages that revolved around text or used podcasts as their primary method of communication. Thus, for lack of a better alternative, we used the verified Facebook pages of the top YouTube channels.

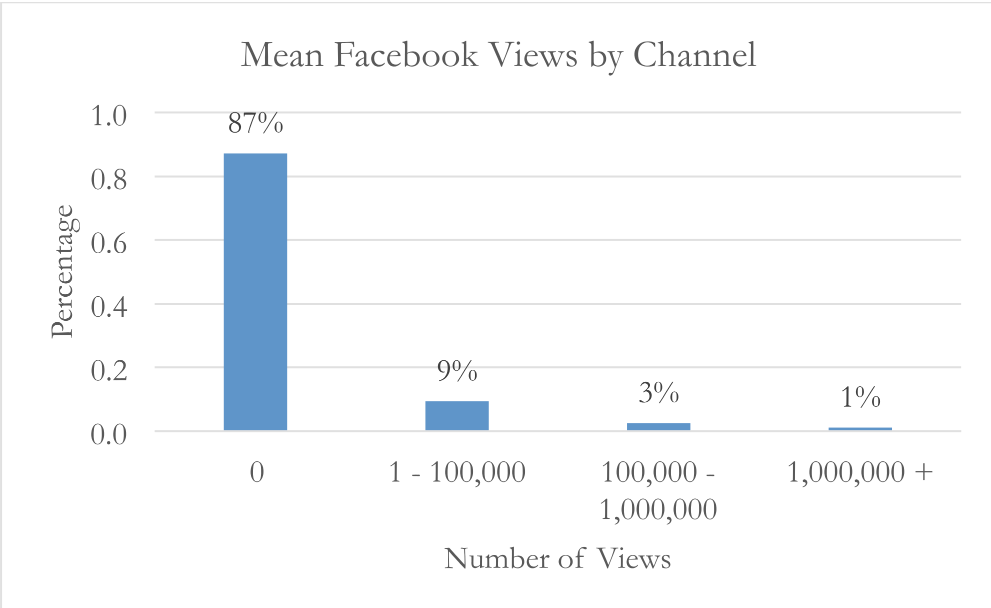
In later sections of this paper, we also explore certain kinds of videos in greater detail for instance, clips from talk shows and music videos. Our reason for picking these areas for further analysis was due to the high number of these videos posted on both platforms and the because of the high level of activity surrounding these topics.

Data Representation and Analysis

Part A: First we try gain some insight into the market space by looking at the broader picture from the perspectives of the two firms.

1. Data from YouTube's Frame of Reference

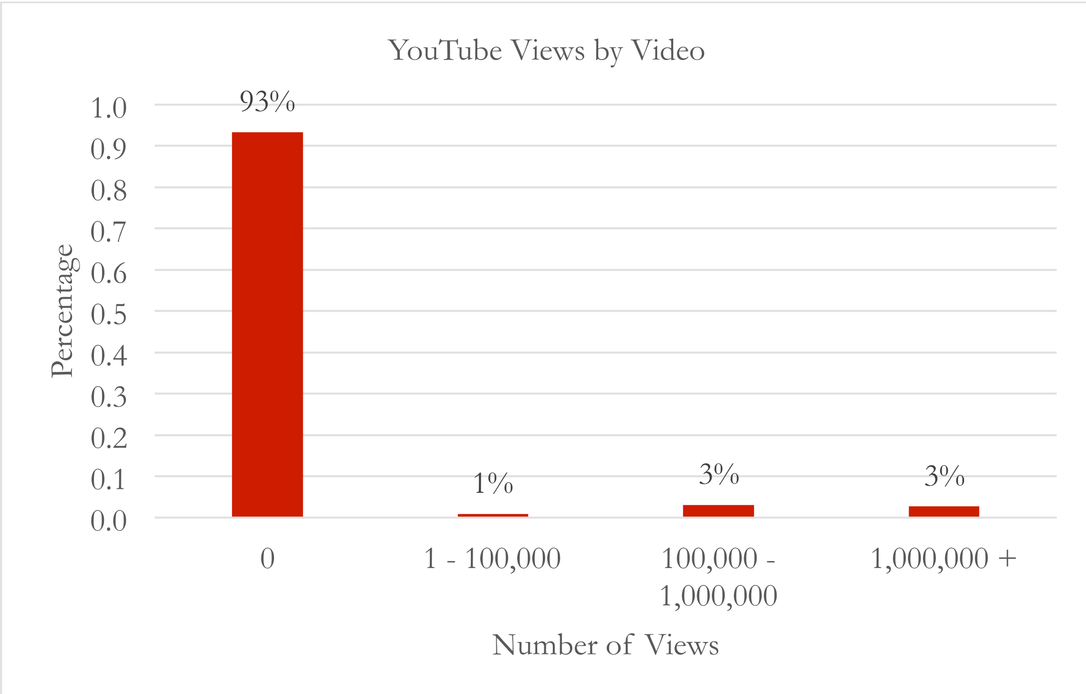
Figure 1: Views for YouTube Videos posted on Facebook (as a Facebook video and not a YouTube link) by the Top 500 YouTube Channels, for all the channels that had a corresponding verified Facebook page



Of the top channels in our sample, more than 87% had a total of zero views on Facebook. That is to say, of the 10 videos per channel, for 87% of the channels, not a single one was on Facebook. If we are to look at the videos instead of the channels, we see that just a little over 3% of the videos matched. That means that more than 96% of the videos had no Facebook counterpart.

2. Data from Facebook's Frame of Reference

Figure 2: Views for Facebook Videos posted on YouTube (as a YouTube video) for the Top 500 YouTube Channels, for all the channels which had a corresponding verified Facebook page



Much like before, there are a low number of matches. A little more than 93% of the videos simply had no YouTube counterpart. And, here, this result is quite surprising considering that it is YouTube's own top channels that are being talked about.

3. Analysis

From the data that was presented above we make a number of crucial observations. Firstly, there is the idea of comfort – the top contributors are comfortable with the revenue that they get from YouTube and don't want to put in the effort required to get a Facebook following. However, this is a little less likely as putting your video up on Facebook will lead to greater visibility, which, given time, could turn into profit. Further, the data shows a low number of matches on both sides. This means that lots of channels are willing put up videos on Facebook, but, they do not put up the same video(s) on YouTube.

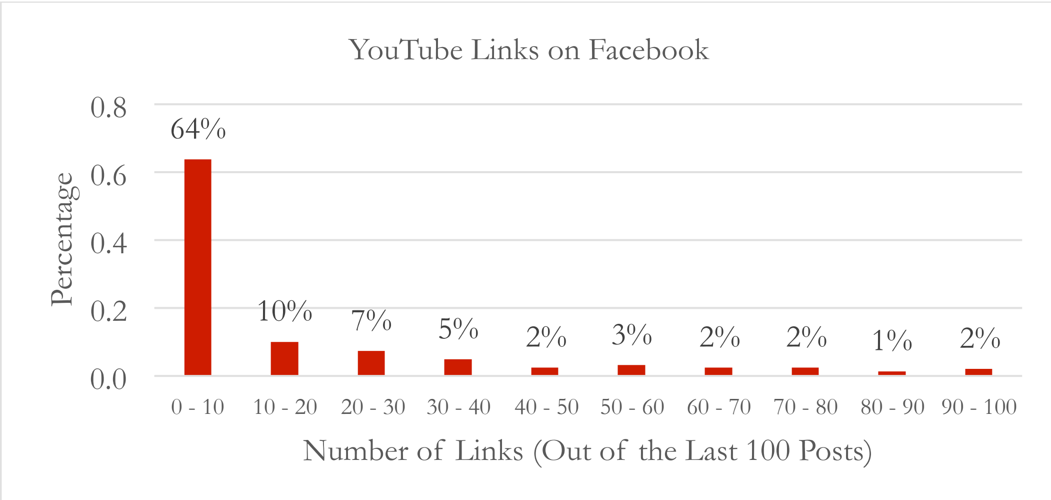
Secondly, it could be argued that since Facebook and YouTube are different platforms, the content is different. This is a more likely solution as here, the content creators themselves view Facebook and YouTube as separate platforms that are to be used in distinct ways. Based on our data, if a video is to be found on Facebook, it is unlikely to be found on YouTube and vice-versa. Only a few channels are posting the same videos on both platforms. Since the number was so small, we manually went through the list in order to identify the similarities between these channels. However, there wasn't much in common. The channels were diverse – music, talk shows, news, comedy, among others – and there seemed to be no particular connection between them. One thing that did stand out was that these channels all had high production value, that is, they had the money required to purchase certain equipment. But, this is not unexpected as these are some of the top channels and thus, are likely to earn greater amounts when compared to other, smaller (in terms of subscribers) channels.

Finally, there is the possibility of a contractual obligation, namely, that YouTube has an agreement with its top contributors, which prevents these channels from posting directly on Facebook. This will be explored further in Part B.

Part B: In this part, we analyze the phenomenon of YouTube video links being posted on Facebook.

1. Data from Posts of verified Facebook pages

Figure 3: Number of YouTube links out of the last 100 posts on the verified Facebook pages of the top 500 channels



The high number of YouTube links being posted on Facebook makes the comparison a little bit harder. As noted in the figure above, assuming that the last 100 posts are representative, close to 10% of these top channels put up 50 YouTube links for every 100 posts, an astonishingly high number. Nearly 44% of these channels posted between 1 and 20 links out of the last 100 posts. This is problematic as these embedded videos are watched on Facebook but, the view goes to YouTube.

2. Analysis

From the data represented in the above section and the analysis from Part A, we can support certain claims regarding the power play between YouTube and Facebook. This data shows us that a contractual obligation that prevents these top channels from uploading videos directly to Facebook is quite probable. Nearly 70% of these top channels post YouTube links on Facebook instead of uploading the video through Facebook, at least some of the time. And, a number of articles have shown that Facebook really prefers it when you upload the video directly (Williams, 2015). The reason for this seems obvious – a link takes you out of Facebook's domain and onto another platform. Thus, it means less time spent on Facebook, which can eventually lead to a loss in revenue. So, the fact that these channels are continuing to post links, despite all the issues that come with such a decision, indicates the likelihood of an incentive, in the form of a contract or, even a greater share of the ad revenue.

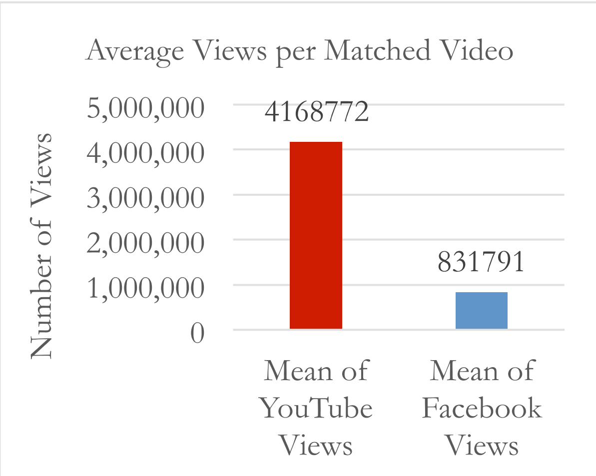
If we look at this from the content creators' point of view, this shows a certain loyalty to YouTube, regardless of whether money is involved. But, by putting up the video on Facebook, the content creator is acknowledging the power of that platform. Further, the creators that are putting up links on Facebook are also the ones who might switch to Facebook Video, in a short time, if they feel that it is beneficial for them.

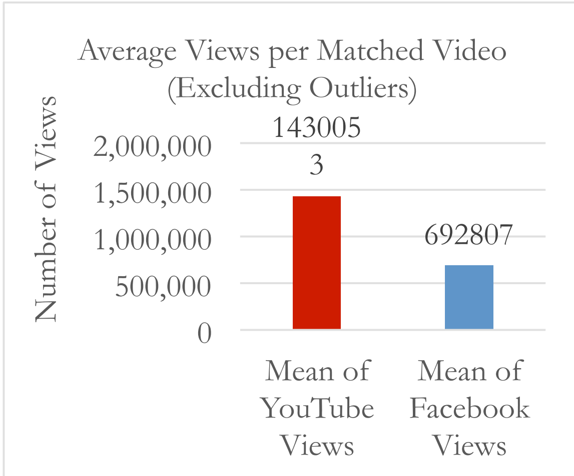
The percentage of views (of a YouTube video) that come from a YouTube link on Facebook is beyond what our data can capture due to the fact that only the channel/verified page has access to these numbers. As a result, we cannot estimate the impact that such links may have on the overall views that we have collected.

Part C: In this part we dive into the specifics of the rivalry based on the content of the videos.

1. Matching the Videos

Figure(s) 4&5: Average Views for the Matched Videos of the Top 500 YouTube Channels with and without outliers



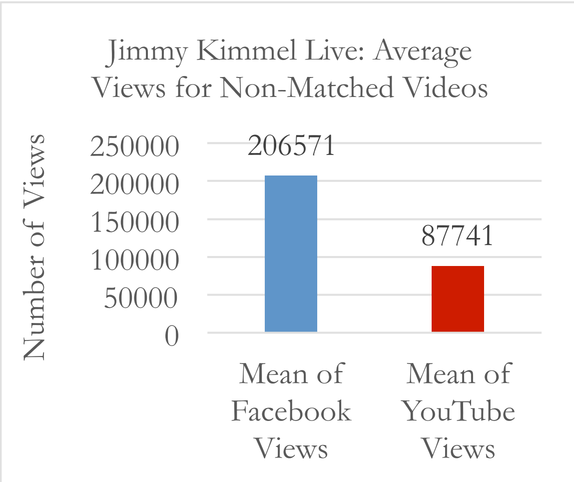


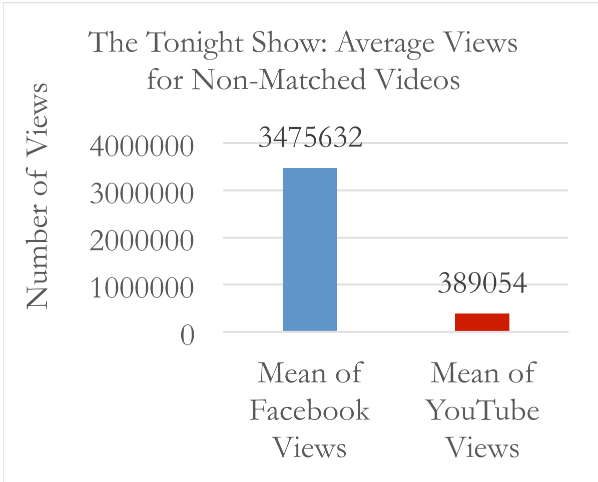
The matched videos provide a glimpse at the most direct competition between the two. Here, Facebook, though still behind, is quite ready for a fight. Though it seems like YouTube is far ahead at first, if we exclude 3 major outliers[1], each with more than 18 million views when the average is just 4 million, we get a completely different picture. Instead of a 5:1 ratio in YouTube's favour (i.e. for every one Facebook view, there are 5 YouTube views), it becomes a 2:1 ratio. While YouTube is ahead, it should be noted that we are talking about the top YouTube channels and so, we would expect them to have a greater number of views on YouTube. The fact that Facebook is competitive makes this result somewhat unexpected.

[1] These are: Skrillex – 'Skrillex & Rick Ross - Purple Lamborghini [Official Video]' – with 194 million views; Steve Kardynal – 'SONGS IN REAL LIFE 4!!' – with 46 million views; Clash of Clans – 'Clash-A-Rama! The

2. Talk Shows

Figure(s) 6&7: Average Views for the Non-Matched Videos of the Top Two Talk Shows



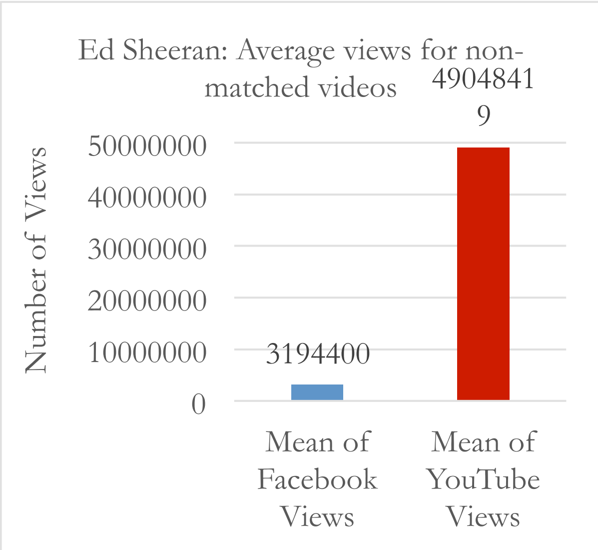


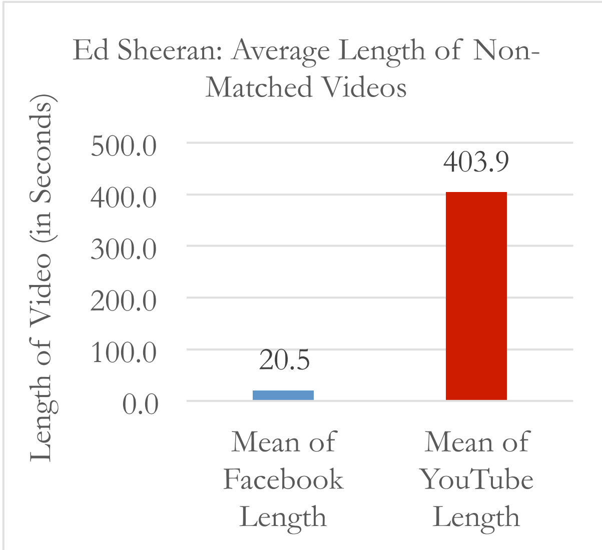
In the case of the top talk shows, like 'The Tonight Show Starring Jimmy Fallon' or 'Jimmy Kimmel Live', the videos posted on Facebook and YouTube differ. These channels are ready to post on both platforms, but, they don't put up the same video twice. In such a case, the videos don't match, but, if we compare just the views on either platform, Facebook moves ahead.

Perhaps that's because such videos are more likely to be seen by someone if they are shared by a friend. It should also be noted that these videos are more topical, that is, they are unlikely to get an audience even a month from now. These videos rely on buzz and thus, play better on Facebook where they are more likely to create conversation.

3. Music

Figure 8&9: Average Views and Length for Non-Matched Videos of 'Ed Sheeran'





On the other hand, the channel, 'Ed Sheeran[2]', (and other music channels), get most of their views from YouTube. Further, they don't even put up their music on Facebook. The average length of a video on Ed Sheeran's Facebook page is just a little above 20 seconds, hardly enough time for a song.

An even better example of this is VEVO, which has 19 channels in the top 100 on YouTube. Here, each artist has a separate VEVO channel. However, they have just one Facebook page and the average length of their videos on this is page is under 40 seconds[3]. Thus, they aren't even putting up music on Facebook whilst, on YouTube, they are willing to create a different channel for each artist!

This domination is probably due to the fact that music videos often build their views over a period of time. And YouTube gives them a chance to re-live what they heard before. On Facebook, the feed changes every minute and people are less likely to search for a video that they saw long ago. So, Facebook's value lies in the now while YouTube continues to extend its hold over those videos with a repeat value.

---

[2] For instance, Ed Sheeran's 'Ed Sheeran - Sing [Official Video]' has more than 160 million views on YouTube, but just 2.5 million views of Facebook.

[3] For this number, we took the last 25 videos on VEVO's verified Facebook page and calculated the average length, which turned out to be 39.55 seconds.

Conclusion

Facebook has certainly made its mark. Even when the same video is up on YouTube, that is when the videos match, Facebook manages to get a significant share of the views. And when it comes to topical, short content, Facebook might even lead the way. After all, such videos are points of discussion, an area that clearly favors Facebook.

But YouTube is hardly out of the ring. It dominates music and continues to increase its professional nature. The top channels on YouTube are no longer the amateurs of years past, they have big budgets, scriptwriters, narrators, video editors, graphic designers and more. And better still, they understand the platform and are willing to show it loyalty. Despite Facebook's appeal, these top channels don't seem ready to make the shift or even spread themselves between the two. If anything, it is these top channels that seem to be stopping Facebook from rising any further.

If we see a rise in the number of top channels putting up content on Facebook, even in the form of links, this could indicate that they are ready to make the switch. However, given the possibility of an incentive, they may not follow this path.

In the current scenario, Facebook and YouTube are not battling in the same space, rather, they can coexist because content creators treat them as separate platforms. Certain types of videos, like music etc., may play better on one platform but, the creators still use both in order to build visibility and recognition. As there is little overlap between the two, at least with regard to some of the top creators, a situation exists where both sides can see the silver lining.

References

Constine, J. (2014, September 7). *Facebook Highlights Its 1-Billion-Video-Views-Per-Day Reach By Adding View Counts*. TeachCrunch. Retrieved from https://techcrunch.com/2014/09/07/facebook-puts-its-video-reach-in-the-spotlight/

D'Onfro, J. (2015, December 7). *More than 70% of internet traffic during peak hours now comes from video and music streaming*. Business Insider. Retrieved from http://www.businessinsider.com/sandvine-bandwidth-data-shows-70-of-internet-traffic-is-video-and-music-streaming-2015-12

Facebook Media. (n.d.). Retrieved from Tell Rich Stories With Facebook Video: https://www.facebook.com/facebookmedia/best-practices/facebook-video

Griffith, E. (2015, January 3). *How Facebook's video-traffic explosion is shaking up the advertising world*. Fortune. Retrieved from http://fortune.com/2015/06/03/facebook-video-traffic/

Peterson, T. (2015, January 7). *Facebook Users Are Posting 75% More Videos Than Last Year*. Advertising Age. Retrieved from http://adage.com/article/digital/facebook-users-posting-75-videos-year/296482/

Python 3.60 [Computer Software]. (2016). Retrieved from https://www.python.org/

Simo, F. (2014, September 7). *The Latest on Facebook Video*. Retrieved from Facebook Newsroom: http://newsroom.fb.com/news/2014/09/the-latest-on-facebook-video/

STATA 13.0 [Computer Software]. (2013). Retrieved from http://www.stata.com/

Vidstatsx. (n.d.). *YouTube Top 500 Most Subscribed Channels List - Top by Subscribers*. Retrieved from https://vidstatsx.com/youtube-top-500-most-subscribed-channels

Williams, O. (2015, July 22). *Facebook throws shade at YouTube when you try to paste a link*. The Next Web. Retrieved from https://thenextweb.com/facebook/2015/07/22/facebook-throws-shade-at-youtube-when-you-try-to-paste-a-link/

RIGGED ELECTIONS IN THE SOVIET UNION AND IN CONTEMPORARY RUSSIA: THE PEOPLE'S PERCEPTION OF THE VOTE

Yulia Lapina

The purpose of this research paper is to compare and contrast the electoral processes in the former Soviet Union and in contemporary Russia. The Soviet Union was known for its rigged elections and corrupt political class, a reputation which modern Russia has sought to rid itself of. The first half of the research paper examines several of the ways in which the political elite rigged elections to their favor, with much of the information presented being either from contemporary history books on the subject or from a direct source who served as a member of the electoral commission in Moscow's precinct. The latter half of the paper looks at the election process in contemporary Russia, further identifying abuses of the election process. The evidence provided is cited from either international monitoring groups such as OSCE or from books seeking to analyse the failure of democratising Russia. By comparing the election abuses and public distrust towards the election process in both the Soviet regime and Putin's regime so directly, it helps draw the conclusion that whilst on paper Russia is not longer an autocratic regime, beneath the surface, Putin's Russia and its predecessor may not be so different.

Controlled and piloted elections were a typical practice across the entire Soviet Union. They were meant to maintain an appearance of democracy and to demonstrate the unity of subjects and rulers (Merl, 2011). Despite this democratic facade, the Soviet regime is commonly categorized as a totalitarian dictatorship, where there was no space for individual freedom and consequently for electoral choice. The phenomenon of rigged elections, however, did not disappear with the collapse of the Soviet Union; it continues to be a widespread practice in contemporary Russia, despite the fact that it is defined as a democratic state in the constitution (St. 1 Konstitutsii Rossiiskoi Federatsii, 1993). I will outline how elections in the Soviet Union and in contemporary Russia are similar and how they differ. In particular, the main issue I am going to address is whether and how people's perception of elections changed in the transition from the Soviet regime to the 'democratic' Russia.

I will start by providing a description of the electoral process in the Soviet Union. I will present the experience of my mother, who served as a member of an electoral commission, as a primary source of what I have previously described. I will also consider the case of the elections that took place in Western Ukraine and Western Belorussia in 1939 as a further example of how people's expression of free will was distorted. I will then compare these violations of the freedom of expression in the Soviet Union with the irregularities that have been observed during elections in contemporary Russia in the last twenty years, focusing in particular on the presidential election of 2004 and on the legislative election of 2007, as the most blatant examples of malfeasance. According to Jessen and Richter (2011, p. 9), "Any acceptable definition of a democratic order includes the following: universal suffrage, a secret ballot, and competing candidates." I will demonstrate how these conditions were systematically violated in the Soviet Union, how it is happening in contemporary Russia as well. Elections in the Soviet Union and contemporary Russia present many similarities. The main distorting characteristic of elections in the Soviet Union, the lack of choice for who is going to rule the country, exists in today's Russia. In the Soviet Union, people were aware of their inability to influence the results of the elections and this made them feel as though they were participating in a process of moral corruption. A similar sense of powerlessness is still common to many people in contemporary Russia and, as it emerges from the interview I have conducted, this fact influences their political behavior.

Elections in the Soviet Union

In 1936, Stalin passed an electoral law dictating that elections were to be universal, equal, direct, and secret (Gross 2002, p.73). However, as Merl (2011, p. 284) states, "Voting from 1937 onwards had nothing to do with choosing candidates or political alternatives." First of all, taking part in the elections was perceived to be mandatory; whoever did not participate in this process was self-excluding from the rest of the political community and was seen by the authorities as an enemy of the people (Merl 2011, p. 281). Furthermore, elections in the Soviet Union were elections without choice. Candidates were already preselected by the authorities, leaving people with no options of who to vote for. There were also problems with the secret ballot; voting regulations were frequently violated, and many voters were compelled to vote openly in order to avoid suspicion of voting against the regime. Finally, there is ample evidence that the results of the elections were manipulated. As a result, people perceived the act of voting as subjugation under a process of moral corruption (Merl, 2011).

My mother served as a member of the electoral commission in a Moscow's precinct in 1978. I interviewed her about her experience, and I will present here some episodes that she shared with me. First of all, she described how elections worked. "The elections were direct: every person was to cast his or her vote directly for the candidate. Every ballot had one name on it; there was no choice between candidates. If you didn't want to vote for that person, you crossed off the name. If there were no marks on the ballot, the vote was positive. If somebody wrote anything on the ballot, such a ballot would not have been considered in the count - it was spoiled." The work my mother was doing was very "simple", as she described it. She had a list of persons whom, as they were coming to vote, she had to give the ballot, and then check off their names from the list. "The lists of voters were prepared long before the election day, and they were checked thoroughly. They included names, phone numbers, addresses, positions, and other personal information." My mother got the job because there was a lack of volunteers, and so post-graduate students of the main universities like her were called to serve in the commissions. "The main work was performed by students like me. Members of the Communist party and the Komsomol were only supervising." I then asked her whether she witnessed some violations of the electoral process, and if she could describe them. "Absolutely yes," was her answer. "After the closing of the polls, if there was a large number of unused ballots, they would simply put these ballots in the box and would sign next to names of people on the lists who did not come to vote, as if they came. The goal was to reach the result of 99.7 percent turnout to the polls, and 99 percent of the ballots should have been positive." Finally, I asked her the most important question for the argument I am discussing: "How did you feel about the elections? Did you feel like taking part in a process or moral corruption, since results were clearly rigged?" Her answer was: "Before working in the commission, I did not realize that there is something wrong with the elections because from the outside everything was very nice; there was music outside the polls, the staff was polite, there was a nice atmosphere, you could buy coffee and treats that you could not find in regular shops. It felt like a party and most people were eager to go to vote. It felt like responding to a call of the nation. After I worked at the commission, however, I understood that my participation did not influence the voting results at all. I felt bad. I felt so bad I have not voted ever since."

Another example indicative of the distortions of the electoral process that I have just described can be found in the elections for the National People's Assemblies of the Western Ukraine and of Western Belorussia that took place in 1939. I am presenting this example in order to illustrate that violations of the voting process were taking part in different periods of the Soviet rule and across the whole territory of the Union. Firstly, the pre-election campaign made it clear that abstention from elections would be injurious. The authorities monitored every adult citizen in Western Ukraine and Western Belorussia, even provisional residents. This led to the consequence that the overwhelming majority of the population performed according to the authority's expectation and went to vote. Those who failed to do so were identified and punished. A process of mass mobilization followed the beginning of the campaign, which took place two weeks preceding the vote (Gross 2002, p. 74-75). As Gross (2002, p. 75) puts it, "Residents of the Western Ukraine and Western Belorussia were ensnared in a web of intimidation, coercion, abuse, and, more familiar, electoral palm-greasing and corruption." Authorities, through the means of terror, made everyone understand that refusal to contribute to the electoral effort would invite severe sanctions. Physical intimidation and verbal threats were used to reinforce the message. Thus, people were drafted to serve on electoral commissions or even as candidates for deputy. Another issue is the obscurity of candidates. They were often unknown to voters, since they were appointed by higher authorities and not selected within the communities. As previously stated, voters had no say in the selection of candidates. They also had no say in the process of voting itself. A massive wave of arrests all over Western Ukraine and Western Belorussia preceded the election day (Gross 2002, p. 92); the population interpreted this event as a preventive measure enacted by the authorities to ensure compliance during elections. Voting stations opened at 4 AM, and there was pressure on residents to go to the polls as early as possible. Superintendents were pounding on apartment doors and waking people up, returning every hour. They were then replaced by militiamen, who began to check on the inhabitants. The ubiquitous presence of military and armed civilians at polling stations contributed to a climate of intimidation. Voters were constantly kept under the guidance of some official or in the commission's field of vision. They received not only instructions on how to vote, but also 'advice', such as to vote as the law says and not according to their wishes. Some received hints that it made no difference if candidates' names were crossed off the ballot or left on. It was suggested that those who were appointed as deputies would become deputies anyway, so there was no point of a negative vote. Tactics to monitor voters' behavior were employed, such as the numbering of ballots: ballots were assigned the number under which a voter's name appeared on the electoral list, or they were consecutively numbered and were then recorded next to voters' names on the list as they were handed out. Whoever crossed a ballot or marked it with obscenities or patriotic slogans was thus identified and arrested. Another violation of the secrecy of the vote was the practice of spying voters who entered the boots. Despite discipline, intimidation and coercion exercised on election day, the organizers still had to falsify the results. Reliable reports from all over Western Ukraine and Western Belorussia confirm that vote tallies were rigged. In some precincts, the Soviet in charge of the electoral commission removed the ballot box, and there was no way to know what was in it except for the official communiqué. In other places, only trusted members of the electoral commissions were allowed to stay for the votes' count. Many witnesses report about the widespread fraud committed at this final stage of the voting procedure. For example, crossed-off ballots were counted as good votes. As a result of all these practices, people felt shame and discomfort during elections. Everybody knew they were taking part in something wrong, but were also aware that not participating would cause trouble (Gross, 2002).

The elections that took place in Western Ukraine and Western Belorussia in 1939 present five typical distortions of Soviet elections that have been described by Merl and my mother: participation was mandatory, there was no choice for candidates, the secrecy of the vote was violated, results were manipulated, and people felt like they were taking part in moral corruption. Elections in contemporary Russia present all of these distortions, too. Beginning with the 1995 parliamentary contest through the 2004 presidential elections and the 2007 parliamentary elections, and including more recent elections too, all elections in Russia were "anything but free and fair" (Myagkov, Ordeshook & Shakin 2009, p. 71). Zimmermann (2014, p. 267), describing the presidential elections of 2000, 2004 and 2008 underlines how each of them "was progressively more authoritarian and less uncertain than its predecessor." I will focus my attention on two elections in particular, which I believe are representative of the overall violations that took place in Russian elections since the collapse of the Soviet Union; the presidential elections of 2004 and the legislative elections of 2007.

Elections in Contemporary Russia

In the presidential elections that were held on 14 March 2004, Incumbent President Vladimir Putin was seeking a second four-year term. The first characteristic of this election that is interesting to notice is the lack of any meaningful competition; indeed, none of the potentially serious candidates participated (Zimmermann, 2014). This lack of "a vibrant political discourse and meaningful pluralism" also acknowledged in the Election Observation Mission Report, written by the Organization for Security and Co-Operation in Europe (OSCE) and its Office for Democratic Institutions and Human Rights (2004, p. 1). According to the Mission, the electoral process overall failed to reflect principles necessary for a healthy democratic election. The main concerns regarded the discriminatory treatment of candidates running against Putin on the state-controlled media, the equal opportunities for such candidates and the secrecy of ballot. Indeed, the practice of open voting, which directly challenges the principle of secret vote, has widely and persistently been observed. Apparently, open voting was actively encouraged by the responsible election commission. Episodes of group voting were also noted. Instances of voters' intimidation were observed, too; some students, for example, were compelled to turn out to vote. Police were present in numerous polling stations inside the room where voting was taking place, and sometimes this presence aroused concern. A policeman was observed taking note of voters' names and passing this information to unidentified persons. Furthermore, private security guards were observed in the voting room at some polling stations, despite the fact that their presence there is strictly prohibited. The processes of vote counting and tabulation were also problematic. It has been observed in many instances that procedures were not strictly adhered to, and the transparency of the process was not safeguarded. For example, before the ballot boxes were opened, unused ballots were not counted and subsequently canceled. In some instances, pages of the voter list were not properly reviewed and certified. Correct procedures for the handling and recording of ballots from early voting and election-day mobile ballot boxes were not always followed. Moreover, instances of result falsification were directly observed; it is clear that some procedural breaches were made with the intent to commit fraud (OSCE, 2004).

In the legislative elections that were held in Russia on 2 December 2007, 450 seats were at stake in the 5th State Duma, the lower house of the Federal Assembly of Russia (Election Guide, n.d.). Also in this election, authorities made their best in order to discourage meaningful competition; indeed, the threshold for representation was raised from 5 to 7 percent, thereby locking out any of the liberal parties that could have opposed United Russia. During the campaign, all explicitly anti-Kremlin opposition was muffled, barred from the ballot, harassed, jailed and cowered into submission (Myagkov et al. 2009, p. 117). These facts are confirmed by the Commission on Security and Cooperation in Europe (2010, pp. 1-2): "Based on credible reports from numerous sources, including the OSCE Parliamentary Assembly, there can be little doubt that Russian authorities used the full range of so-called administrative resources—intimidation, confiscation of campaign literature and, at times, even physical abuse—to overwhelm the already weak and divided opposition." Indeed, the OSCE Parliamentary Assembly reported authorities' clampdowns on opposition rallies and demonstrations, harassment of opposition candidates, detentions, confiscation of election material. Threats against voters were observed, too (OSCEPA, 2007). State workers such as doctors, teachers, and university deans were forced to vote for United Russia; had they done otherwise, they were risking losing their jobs. Students were told they were risking failing exams or being removed from courses if they did not vote for United Russia (Harding & Parfitt, 2007). There were also other episodes of malfeasance that took place during the elections. For example, people coming from small towns were given free tours of Moscow for their vote, applications for ballots were found to be 'written' by people who knew nothing about these applications, by blind people, and by a person that had died one day before the elections, and armed groups were seen at polling stations, taking ballots from the desk, filling them out and putting them in the urns. In addition, some people who presented a photo with the ballot with a vote cast for United Russia received 100 rubles (Myagkov et al. 2009, p. 117-118). The secrecy of the vote failed to be protected in this election, too: voting arrangements, such as the use of electronic boxes and voting booths, did not provide adequate privacy. Moreover, the seals on some ballot boxes were inadequate (OSCEPA, 2007). This election has been described as "unfair" and "non-competitive" by Freedom House (2007), and it has raised concerns regarding the violation of human rights among international organizations such as Amnesty International (2007).

The Perception of the Vote

As we can see, elections in contemporary Russia also present the five electoral distortions outlined earlier. Participation is mandatory for some categories of citizens; state workers and students were forced to go to vote, and were even compelled to vote for the party supported by the government. There is no choice for candidates; all meaningful opposition is ousted in more or less legal ways, with the result that the state-supported party or presidential candidate appears as the only plausible choice. The secrecy of the vote is gravely and frequently violated. Results are knowingly and willingly manipulated. People who take part in elections do perceive the wrongness and immorality of the whole process. As I have previously stated, the most distorting characteristic of elections in the Soviet Union was the lack of choice between candidates. This is not only a present day assessment; it was also frequently perceived, as I previously described, by many Soviet voters (Merl, 2011). The lack of choice is also the main distorting characteristic of elections in contemporary Russia and, similarly to their Soviet predecessors, many voters in today's Russia perceive a sense of powerlessness when it comes to voting. This feeling is expressed by many different scholars and publicly known individuals. Gel'man (2015) claims that from 1996 the nature of elections did not imply any democratic uncertainty; elections outcomes were set up by the ruling group well in advance. Many even refuse to refer to what takes place in Russia as elections. Alexander Yakovlev, a politician considered by many as the grandfather of Russian democracy, when talking about the 2004 presidential elections, expressed the view that it was useless to vote; it was known beforehand who would win (MacKinnon, 2007). He claimed that what took place on 14 March 2004 was not an election, and compared it to what used to happen in the Soviet period. Myagkov, Odershook and Shakin (2009, p. 6) excluded an analysis of the 2008 presidential election from their book "for the simple reason that calling it an election denigrates the meaning of the word." Golosov (in Gel'man 2015, p. 96) referred to the voting procedure not as an election, but as an "election like event" with no electoral relevance. Navalny, talking about the presidential elections of 2012, described them as a procedure and not really an election (Barry & Schwirtz, 2012). Trudolyubov affirms that the word "election" is a "misnomer" for what happens in Russia (2015). This feeling of powerlessness leads to important consequences. When I asked my mother whether she participates in elections, she answered: "I do not go to Russian elections because it does not make any sense. My vote will be spoiled or not taken into consideration. I don't see the point of going to vote." Then I asked her: "What do you think of the Soviet elections compared to the contemporary ones?" This was her answer: "After finding out about the manipulations when I was working in the commission, I felt really bad. But now I think that these manipulations of the votes are very mild compared to the ones that are taking place in Russia now. Maybe the previous elections were not real elections, but at least candidates did something for the people. Now candidates are buying people's votes." It is possible to conclude, therefore, that people's perception of elections, at least the perception that educated people had of it, has not changed much since the Soviet times. They are both perceived as unreal. During the Soviet period, elections were perceived as a process of moral corruption. In contemporary Russia, people feel "they have been conned into playing bit parts in Putin's bad theatre" (Gessen, 2016).

Conclusion

By examining the election processes in both Russia and the Soviet union, it can be concluded that these two systems present more similarities than differences. Both regimes present five of what I called electoral distortions: participation in the election is, at least for some categories of citizens, mandatory; there is no choice in candidates, as they are already preselected by higher authorities; the secrecy of the vote is violated; results are willingly and fraudulently manipulated; and people perceive the futility of the whole process. To overcome this distortions, a transition from a despotic regime to a democratic one would have been needed. Unfortunately, this did not happen after the collapse of the Soviet Union; democratization failed in contemporary Russia.

To conclude, I want to underline that there is a similar perception of the vote in the Soviet Union and in Russia. In both countries, people are conscious of having no real choice between candidates; this makes them feel as though they are taking part in someone else's staged game. As I have stated before, this can lead to severe consequences such as abstention from vote. Indeed, if people are convinced that their vote does not make any difference because the electoral results are already predetermined, it is likely that they would refrain from voting. A question thus arises: would people feel better if they had refused to participate in elections? What is worse, taking part in someone's nasty game or abstaining from vote? I agree with Masha Gessen's (2016) answer to that question: "*Oba khuzhe.*" There is no better choice.

References

Amnesty International. (2007, November 28). *Russian Federation: Systematic repression on eve of elections*. Retrieved from https://www.amnesty.org/en/press-releases/2007/11/russian-federation-systematic-repression-eve-elections-20071128/

Barry, E. & Schwirtz, M. (2012, March 5). After election, Putin faces challenges to legitimacy. *The New York Times*. Retrieved from http://www.nytimes.com/2012/03/06/world/europe/observers-detail-flaws-in-russian-election.html

Commission on Security and Co-Operation in Europe. (2010). *Post analysis of the Russia Duma elections: December 6, 2007, briefing of the Commission on Security and Cooperation in Europe*. Washington: U.S. G.P.O.

Election Guide. (n.d.). *Election for State Duma*. Retrieved from http://www.electionguide.org/elections/id/414/

Freedom House. (2007, December 3). *Russian elections lack legitimacy; Meaningful political competition absent*. Retrieved from https://freedomhouse.org/article/russian-elections-lack-legitimacy-meaningful-political-competition-absent?page=70&release=596

Gel'man, V. (2015). *Authoritarian Russia: Analyzing post-Soviet regime changes*. Pittsburgh: University of Pittsburgh Press.

Gessen, M. (2016, September 21). Russia's election: Every choice was a bad one. *The New Yorker*. Retrieved from http://www.newyorker.com/news/news-desk/russias-election-every-choice-was-a-bad-one

Gross, J. T. (2002). *Revolution from abroad: The Soviet conquest of Poland's Western Ukraine and Western Belorussia*. Princeton, N.J.: Princeton University Press.

Harding, L. & Parfitt, T. (2007, November 29). Fraud, intimidation and bribery as Putin prepares for victory. *The Guardian*. Retrieved from https://www.theguardian.com/world/2007/nov/30/russia.politics

Jessen, R, & Hedwig, R. (2011). Non-competitive elections in 20th century dictatorships: Some questions and general considerations. In R. Jessen & R. Hedwig (Eds.), *Voting for Hitler and Stalin: Elections under 20th century dictatorships* (pp. 9-36). Frankfurt/New York: Campus Verlag.

Konstitutsiya Rossiiskoi Federatsii ot 12 dekabrya 1993 goda. Rossiiskaya Gazeta, 25 December 1993.

MacKinnon, M. (2007). *The new Cold War: Revolutions, rigged elections and pipeline politics in the former Soviet Union*. New York: Carroll & Graf Publishers.

Merl, S. (2011). Elections in the Soviet Union, 1937-1989: A view into a paternalistic world from below. In R. Jessen & R. Hedwig (Eds.), *Voting for Hitler and Stalin: Elections under 20th century dictatorships* (pp. 276-308). Frankfurt/New York: Campus Verlag.

Myagkov, M., Ordeshook, P. C. & Shakin, D. (2009). *The forensics of election fraud: Russia and Ukraine*. New York: Cambridge University Press.

OSCE. (2004). Russian Federation presidential election 14 March 2004. *OSCE/ODIHR Election Observation Mission Report*.

OSCEPA. Russian Duma elections 'not held on a level playing field,' say parliamentary observers. *OSCEPA 2007 parliamentary election press release*. Retrieved from https://www.oscepa.org/documents/all-documents/election-observation/election-observation-statements/russian-federation/press-releases-19/2139-2007-parliamentary-7/file.

Trudolyubov, M. (2015, September 16). Russia's latest fake election. *The New York Times*. Retrieved from https://www.nytimes.com/2015/09/17/opinion/maxim-trudolyubov-russias-latest-fake-election.html?_r=0

Zimmerman, W. (2014). *Ruling Russia: Authoritarianism from the Revolution to Putin*. Princeton, N.J.: Princeton University Press.

NBA STATISTICS AND HOW THEY DICTATE SALARY

David Bernheim, Austin Feit & Ammar Monawar

In this paper, we researched NBA player statistics from the 2014-2015 season and created a regression model to determine which statistics would be useful to include in evaluating the players' salaries. We ran our model six times: once on all 397 players in our data and then five more times after splitting the data into the five NBA positions of Center, Power Forward, Small Forward, Shooting Guard, and Point Guard. The results of the regressions showed that the model was significant for all 397 players but not as useful when broken down by position. Using the results from the regression model, we concluded that our model consistently gives an overestimation for what a player actually earns and therefore cannot be used to attain a player's actual salary.

In the history of sports, salary has always been an important indicator of a player's value to his or her team. As the sports market has grown and revenues for the leagues have skyrocketed in the last number of years, player salaries have tremendously increased, along with the salary caps for individual teams. With this recent development, the value attributed to a player based on his or her salary has changed over time. However, many factors can play into determining what NBA teams are willing to pay their all-star players and benchwarmers. The goal of the research is to analyze the data that can be found for the 2014-2015 NBA season, and use those statistics from that year to determine if there is a correlation between any (or some) of those variables and the salary that each player receives. Below, we will discuss which variables were included or excluded and the rationale behind those decisions. Additionally, through the research, a new variable will be developed and explained that might be a better determinant of salary allocated towards each player (Player's percentage of team points scored).

We will run our model six times, once altogether and five times based on the player's position to determine if the variables we include in our model affect the salaries of players at different positions differently from how they affect all players as a whole. An additional goal of our research is to determine if our model can function to predict player salaries based the statistics we include in the regression.

Literature Review

The most extensive research done on the relation between a player's scoring statistics and their value to the team was conducted last year by Dr. Robert Lyons Junior. In his report, Lyons goes through multiple regressions to see what relationship, if any, there is between these scoring performance variables and the respective player's salary. It delved more deeply than previous studies on the direct relationship between the actual salary contract awarded and the player scoring statistics at the time of his contract being drawn up.

In his regressions, Lyons found that the two most significant indicators of a player's salary were his percentage of shots made and points per game (Lyons). He measured points scored by players in his testing as whole, so we decided to test this further by breaking the points distribution based on the types of points earned (free throw, 3-pointer, etc.)

One detail Lyons noted in his research was that when data for player salary was analyzed for evidence of racial discrimination, he found that there was a premium paid to white players over the course of their careers in the range of 16-20%. However, the finding was not a significant enough factor in the overall model to constitute a form of pay discrimination (Lyons). For this reason, we have decided not to break down players by race, as it does not appear to have significance in the same way scoring statistics did in earlier research.

One major exclusion in the study by Lyons was for rookie players, as the contracts have rookie salary caps, which remain in force for three years. This means exceptional performance in the first few years would not be reflected in the player's salary. However, we chose to include the rookie players, as they constitute a large percentage of the players and their data could help us test for trends of age in relation to the player salary.

Another important note from Lyons' work is how a basketball player is expected to be able to play defensively as well as offensively, thus making point scoring a good metric, since all players must engage in the activity. However, Lyons does not make distinctions between positions, wherein some may be more disposed to scoring than others. To that end, we chose to break our analysis of players by position played as well.

Ultimately, Lyons concluded that the most statistically significant variables in determining player salary were field goal percentage and points per game; however, "rebounds, fouls, and assists were significant contributors as well" (Lyons). Lyons hypothesized that the tendency of these new NBA personnel would begin focusing on giving players with better defensive statistics improved salaries (although not enough to displace scoring as the best metric to potential salary).

Additional research was available to us through the research done by NylonCalculus, a subsection of Sports Illustrated subsidiary Fansided, which writes about various statistics as they relate to basketball. Specifically, they published a series of articles on how players in the NBA are given their valuations.

When analyzing position, NylonCalculus created a formula to simplify how to determine a player's relative worth based off of his position. They compared each player's percentage of salary cap to the mean salary of that position, and it became clear that the players who excelled in those positions had a higher salary (Schimanski). Although this uses the player's valuation to determine his position in comparison to similar players, it still helps point out the relationship between value and relative skill.

On the topic of role, NylonCalculus conducted a regression using the player's percentage of the total salary cap of the team and broke players down based off of their respective roles. They came to find through their study that although star players made more, a strong support role player on a team with a star could stand to gain a larger percentage of the team's total salary cap (Schimanski). However, this information is not as conducive to our own study, as we are seeking to find trends across the board instead of within teams.

Model

Our model incorporates different statistics to determine if there is a correlation between a specific player and the salaries they earn. The proliferation of statistics that are recorded for each player allows us to pick which variables we believe to be the most important in determining salary. It also enables us to limit the amount of variables we use to prevent potential statistical flaws such as multicollinearity. It is easy to predict that one of the most important variables to use is points that a player scores per game. However, picking apart the ways in which points are scored can yield better information on how to determine what goes into a player's salary.

The first three variables that were chosen to be used in the model are points scored from three-point shots per game, points scored from two-point shots per game, and points scored from free throw shots per game. These factors are the only ways in which players can score points, and our prediction is that points scored is the most important factor for how players are paid when they become free agents. It is important to differentiate the ways that players can score because every player has their own specific strengths in how they accumulate points. The priori assumption on the coefficients for these variables is that they will be positive; it is virtually impossible that a study of every NBA player will show that the more points a player scores, the less he will be paid.

The next variable selected is total rebounds per game. Every rebound that a player gets is an extra possession for his team, which contributes directly to the goal of scoring more points than the other team. The more rebounds a player is able to achieve, the higher his value will be. The coefficient for rebounds is expected to be positive. Rebounds are a key statistic and should positively influence a player's salary, especially for the forward/center position players.

The third of the three main statistics that are highlighted by teams and the media, along with points and rebounds, is assists per game. The variable, assists per game (APG), is critical to include in the model because it is a great measure of playmaking ability for NBA players. Assists can be just as important as points, as every assist directly leads to a scored basket for someone else. This variable should also furthermore have a positive coefficient; teams value playmakers that can create scoring opportunities and the statistic should have an impact on how they pay their players, especially for the guard position players.

Another variable that is significant to player salaries is a player's age. In the NBA, a player with a lot of previous success can have a wide range of salary potential. There are intangibles tied to a player with a significant amount of playing experience in the league, such as his impact on team/player morale and veteran experience with in-game management. Data could potentially be skewed, since a player may have been worth a large contract when he signed one, but will not be worth that same amount of money in the latter years of the contract following a drop in skill. Previous success might not manifest at the end of his contract, which usually leaves a higher salary on the remaining years of the initial deal, thus the statistical performance doesn't correlate as it should. Additionally, there is an NBA veteran minimum based on player experience in the league. If an older player were to be put on a roster, there is a certain amount they must be paid, regardless of the team's perception of the player's value. Ultimately, this coefficient too is assumed to be positive.

For this regression, it is also important to note how many games in the season a player has started for the team. The team's perception of value should directly correlate with this statistic. If a player starts more games for the team, it means that they are most likely the best player at that particular position for the team. That will translate into a higher salary paid for that talent relative to the skill of other members of the team. While there is an award titled the "Sixth Man of the Year", which acknowledges the best player in the league that comes off the bench (usually doesn't start the game), it is likely that these types of players will get some starts during the season, due to off days and injuries (NBA & ABA Sixth Man of the Year Award Winners). It won't drastically change our data but it is important to acknowledge that there is value within the league attributed to non-starters. The coefficient should again be positive for this variable, regardless of position held by the player.

The first negative coefficient variable that will be introduced to the model is turnovers per game. A player turning over the ball is giving the opponent an extra possession, which translates into another chance for the opposing team to score more points. A player that turns the ball over more should be paid less. While we predict this variable to have a negative coefficient, it would also not be surprising if turnovers proved to be irrelevant and may therefore randomly have a positive (or insignificant) coefficient. Turnovers are an important part of the game, but owners may look past this statistic if a player can score twenty-five points per game or secure ten rebounds per game. Quality players may have more turnovers than the average player, but this just means that teams are putting the basketball in the hands of their talented players, and the more often they have the ball, the more likely they will turn the ball over to the other team.

The final variable that we decided to include in our model is one that we created on our own using existing data. This statistic is the percentage of a team's total points per game that a player scores. We believe that this statistic is important because of the league-mandated salary cap. Unlike baseball in the MLB, where a team can pay out uncapped salaries to its players, NBA owners are only allowed to spend a maximum amount on its players every season. In the 2014-2015 season, this capped amount was $63,065,000 (NBA Salary Cap History). The salary cap forces teams to properly allocate their funds to fill out a roster, even if they believe that a player may be worth more than they are allowed to pay him. Therefore, the percentage of a team's points that a player scores is significant factor for how much a team decides to pay that player in relation to others on the team.

This statistic may not be as useful in determining how good a player is compared to the rest of the league, but it is important in determining how good he is compared to the rest of his team. It is essential to note that a player with a low percentage in this category who is on a successful team, may in reality be more valuable to the league than a player with a high percentage on a poor team that has trouble scoring.

The model will be run multiple times to figure out the different ways that certain variables affect salary. Running the model on data from every NBA player will give us valuable information on our hypothesis, but in the effort to find more accurate results, we will run the model separately for players of each of the five positions: Center, Power Forward, Small Forward, Shooting Guard, Point Guard. For example, Centers spend most of their time playing in the paint, the area right underneath the basket. Therefore we predict that statistics such as rebounds and blocked shots will have a greater impact on the salaries of players at that position.

There are a number of statistics which were accumulated that we decided to not include in our regressions. One of the main categories that these variables fall under is defense. As mentioned earlier in the paper, there are always going to be intangible aspects to a basketball player that will go into a team's perception of their value. It is much more difficult to quantify a player's defensive value from statistics. Even though we have the data for steals and blocked shots, the data sets typically consist of smaller values, and less players have more of those statistics. Because of that, we believed that they shouldn't be used in the regression model to come.

Another statistic that was found and excluded from the model is personal fouls. The variable for fouls don't seem to accurately quantify the value of the player. Many times, there are in-game situations where a player is either "forced" to commit a foul (based on game strategy) or performs a basketball move and "draws" a foul (based on the referee's decision). This seems to focus more on external factors, rather than a player's intrinsic value. Any player in such a position will be committing this foul, thus the variable is less correlated to player salaries than other statistics that are more heavily reliant on skill set.

Additionally, it was decided that games played should not be a statistic used for our model. We incorporated games started because the variable definitively factors into a player's value. However while using data for total games played, the same "game" will be counted for a player, regardless of whether he plays 3 minutes or the entire game. For example, playing 30 minutes per game for 80 games will be valued the same as playing 3 minutes per game for 80 games.

The same logic applies for shooting percentages for the scope of our model. While shooting percentage may be important, it doesn't accurately measure how many shots/points a player is scoring, since it is only a percentage. (For example, a player who shoots 50% hitting 50 out of his 100 shots will seem to be more valued than a player who hits 60% hitting 3 out of his 5 shots.)

Data

The decision was made to use the 2014-2015 NBA season for our data. The NBA team salary cap has slowly increased over the past few years but the biggest jump was from the ~$63 million 2014-15 season cap to the ~$70 million 2015-16 season, an 11% increase (NBA Salary Cap History). The previous increase was 7.4%. The biggest jump has come from the 2015-16 season to the current 2016-17 season (to ~$94 million) and increase of 33% (NBA Salary Cap History). Because of these jumps, we thought it would be best to try to balance the most recent data with salaries that didn't seem to "overvalue" players just because the salary cap has significantly increased, since that has been the general trend as seen from the Free Agent signings in the summer of 2016.

All of the data for variables, excluding the players' salaries, was either found on Basketballreference.com or derived from a statistic already found on that website, such as Points Scored from 3 pointers made per game. The players' salaries for the 2014-2015 NBA season were pulled from ESPN.com. During the 2014-2015 NBA season, 74 players were active on multiple teams (2014-15 NBA Season Summary). This is due to the player in question being traded from one team to another in the middle of the season, or being released during the season (and potentially being signed by a different team afterwards). Based on this fact, it is difficult to assess the value of the player based on our dependent variable (salary) for the three following reasons. Firstly, we would need to manually input the amount of games "X" player played for team "A" versus how many played for team "B" and then divide the statistics accordingly to properly assess which statistics help each team during the season. Additionally, a large majority of those players played insignificant roles during the season. Their makeup was mostly bench players with small roles that didn't play a significant amount of time in the game. Lastly, only the first team "valued" the "X" player by signing him to a contract, and not the second team, so the salary only dictates how that first team valued the player. Therefore, we decided we would cut those players from the data, leaving 397 players left, which was still enough to get significant results from our regression models.

The following table notes the mean and the standard deviation we calculated for each variable from our data for the 2014-2015 season:

Table 1 – Mean and Standard Deviation

	Mean	Standard Deviation
Age	26.53	4.26
G	54.95	23.29
GS	27.31	28.35
FG	3.19	2.09
FG%	0.44	0.09
3P	0.66	0.70
3P*3	1.99	2.09
3P%	0.30	0.13
2P	2.53	1.87
2P*2	5.05	3.75
2P%	0.47	0.09
FT	1.48	1.34
FT%	0.73	0.14
TRB	3.75	2.52
AST	1.84	1.76
STL	0.66	0.44
BLK	0.42	0.47
TO	1.17	0.77
PF	1.79	0.71
PPG	8.52	5.71
Salary	$4,511,948.27	$4,880,289.73
Team PPG	99.72	4.40
ppg%tp	8.55	5.70

Table 2

Key

Pos	=	Position
Tm	=	Team
G	=	Games
GS	=	Games Started
FG	=	Field Goals per Game
FG%	=	Field Goal Percentage
3P	=	3 Pointers made per-game
3P*3	=	Points Scored From 3-Point shots per-game
3P%	=	Percentage of Successful 3 Point Shots
2P	=	2 Pointers made per-game
2P*2	=	Points Scored From 2-Point shots per-game
2P%	=	Percentage of Successful 2 Point Shots
FT	=	Free Throws Made per-game
FT%	=	Percentage of Successful Free Throws
TRB	=	Total Rebounds per-game
AST	=	Total Assists per-game
STL	=	Steals per-game
BLK	=	Blocks per-game
TO	=	Turnovers per-game
PF	=	Personal Fouls per-game
PPG	=	Points per-game
Salary	=	Salary
Team PPG	=	Team Points per-game
PPG%TP	=	Player Points per-game as a Percentage of Team Points per-game

Figure 1

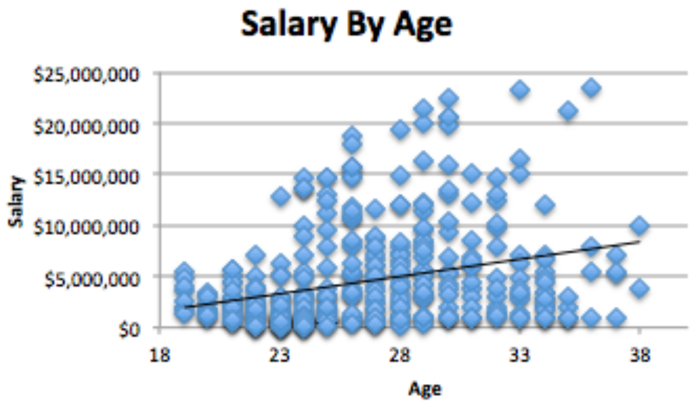


Figure 2

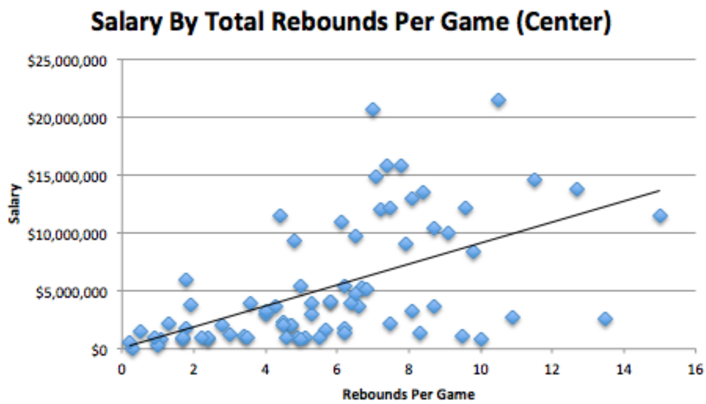


Figure 3

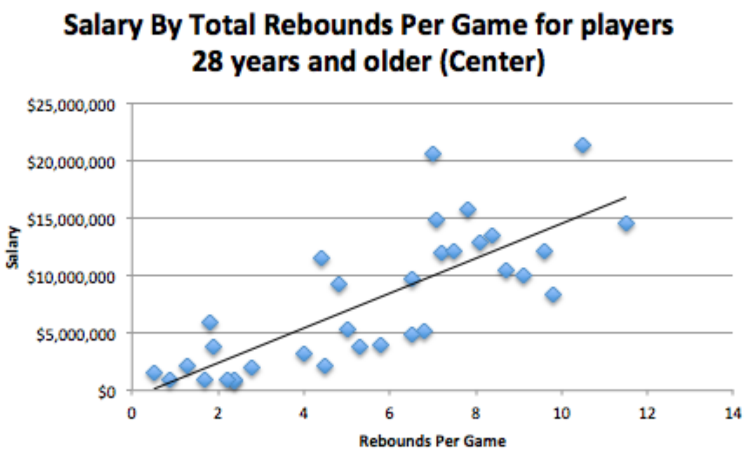
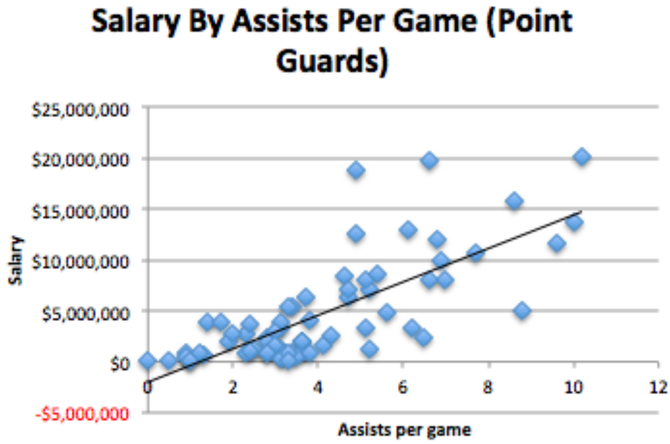


Figure 4



When directly pitting NBA salaries against age, Figure 1 shows that there is a positive relationship between the two variables. Despite the fact that as age increases there are still many players who earn lower salaries, we can see that all of the highest earning players are older. This does not mean that all older players perform better than the younger players, rather it shows that the young players do not even have the opportunity to earn as much because of their rookie contracts.

Figures 2 analyzes a Center's salary against the amount of rebounds per game by each center and figure 3 does the same but for centers 28 years and older. We created the two charts to see if there is a difference in the line of best fit based on what we observed in figure 1 and we can see that the line in the chart for players 28 years and older is steeper. The low salaries of young players pull down the line of best fit because the difference in salaries of young players based on skill is smaller that of older players. These graphs also show that even when disregarding age, rebounds play an important role in how Centers get paid.

Figure 4 emphasizes the correlation between a point guard's salary and his amount of assists per game. The positive relationship is a good sign, as the better players do get paid more, but we can see that the data spreads out for players with more assists. This effect is likely caused by the different ages of the players. The players with a higher amount of assists that are below the line of best fit are most likely to be younger and in their rookie contracts, while those above the line are players who have signed larger contracts for higher salaries because of their skill. There is less deviation in salary based on players with fewer assists because the veterans with fewer assists will still be paid less because of their lack of skill.

Regressions Results

Key for Regression Models

X1 = Games Started
X2 = Age
X3 = points from 2 pointers
X4 = points from 3 pointers
X5 = points from FT
X6 = Total Rebounds
X7 = Assists
X8 = Turnovers
X9 = PPG as a percentage of Total Team points

The Regression function for the model including all 397 players is:

$Y = -8906933 + 3258.957X_1 + 316085.7X_2 + 751906.8X_3 + 476700.1X_4 + 871842.1X_5 + 173790.8X_6 + 162542.4X_7 + 1307761X_8 - 418831.3X_9$

The regression model run on all 397 players in our data proved to be relatively significant. The R-squared value of .536302 shows that there is a strong correlation between the dependent variable, Salary, and all of the independent variables employed. The adjusted R-squared of .525518 explains that the abundance of variables is not unnecessary and is not just there to increase the R-squared. The high F-statistic for the model as a whole also tells us that the model we used is relatively sufficient.

However, looking at the results for each individual variable tell us which of those are statistically significant. The signs of the coefficients for the variables are mostly what we predicted them to be, outside of those for turnovers and percentage of team points per game (PPG%TP), the statistic we engineered. According to the model, every 1% increase in turnovers leads to an additional $1.3 Million in salary. The low p-value of .0223 and the t-Statistic of 2.2940 show that it is significant. One explanation for this is that the amount of turnovers committed by a player is directly related to his production. The reason why players commit so many turnovers is because they have possession of the basketball very often. The teams want their best players to have the ball as often as possible, so a deeper understanding of basketball will show that it makes sense for a player who scores 30 points per game to turn the ball over more than a player who scores 4 points per game and rarely touches the ball. While frequently turning the ball over is not a good quality to have in a player, it indirectly correlates to salary.

It is more difficult to explain the negative coefficient for the statistic we derived on our own. According to the model, a one percent increase in the percentage of points per game scored by a player out of the team's total amount of points per game will decrease the salary by $.42 million. The p-value of .2621 is high and the t-Statistic of -1.1231 has a low absolute value so this statistic is not extremely significant, but it is still alarming that it yields such results. A possible explanation for why the coefficient could be negative is that players who score a high percentage of their team's points tend to be on worse teams whose talent is as well distributed. A team with a good distribution of talent may be more successful and therefore reward higher salaries.

The other variables all have coefficients that correspond to our predictions but the p-Values and t-Statistics show that not all of them are significant. Outside of the created statistic, the other statistics that are not as significant to our model are games started (GS) and Assists (AST). An explanation for the large insignificance (p-value of .7539, t-Statistic of .3138) of Games Started is that bench players in the NBA have become significantly valuable. There is even an award that goes to the best bench player titled “The Sixth Man of the Year”. Therefore there are NBA veterans, such as point guard Rodney Stuckey of the Indiana Pacers who averaged a substantial 12.6 points-per-game, who start games on the bench but play big roles once they check in (NBA & ABA Sixth Man of the Year Award Winners).

The insignificance of assists per game (p-Value of .3974 and t-Statistic of .8471) shows that while assists are an important factor in scoring points, teams don’t value assists too highly when determining salaries. Teams focus more on statistics that directly correlate to wins, which are points scored which is shown by the low p-values for points scored from 2-point shots made (_2P_2) of .0580 and free throws (FT) made of .0508. The t-Statistics of those values of 1.9010 and 1.4255 respectively are not as high as we would have expected but they still exemplify the importance of these statistics. Points scored from three-point shots made (_3P_3) are not as significant with a p-value of .2054 and a t-Statistic of 1.2685 because they are less frequent. Total Rebounds (TRB) falls closer to the middle in terms of significance with a P-value of .3974 and a t-Statistic of 1.4255. While rebounds are important in the NBA, as they represent an extra possession by either prolonging a possession with an offensive rebound or taking away an opposing possession with a defensive possession, they do not represent major significance in our regression of all 397 players.

Interestingly, the most significant statistic according the model is age with a p-Value of

.0000 and a t-Statistic of 7.8127. A 1% increase in a player's age will result in an increase of \$316,086 in salary. This equates to an average salary increase of approximately \$1.2 Million per season based on the average age of our players, 26.53 years old. The explanation for this is not so complex when looking into the history of NBA contracts. Players typically sign low paying deals as rookies and if their careers pan out well, they can earn large contracts when they reach free agency a few years later. Therefore, even if a player in his 3rd year and a player in his 10th year have identical statistics, the 10th year player is likely already on his 3rd or 4th contract and will be earning much more than the 3rd year player on his rookie contract.

When thinking about the potential results of our data, we predicted a positive intercept. Our thoughts were that this would equate to a player that is sitting on the bench, taking up a spot on a roster and putting up zero statistics/points during the entire game. It might even amount to the veteran minimum salary, or the cost of having any player on your roster. However, our results yielded an approximately -\$8.9 million figure. This was shocking at first, since it was both negative and such a large number. But after much consideration, we came to the following conclusion. The large negative number is more representative of a player that is playing on the court that is not only scoring zero statistics but is taking away the opportunity for someone else to contribute to the team's performance. In essence, it is the average value of a player on the court playing at all times to a team. When this player is on the court he is costing his team \$8.9 million.

The regression results for players based on that they play gave us different results. Our model for players at the Center position is:

$$Y = -15183029 + 42395.70X_1 + 580821.4X_2 + 670587.0X_3 + 1513450X_4 + 1442637X_5 - 74780.92X_6 - 186761.8X_7 + 941658.0X_8 - 404060.7X_9$$

The regression gave us an R^2 of .562416 and an adjusted R^2 of .501828, both of which are not relatively different from our first regression results. However, the low F-statistic of 9.282556 tells us that the model does not perform as well as the complete regression. When examining the p-values and the t-Statistics of the model, we see that the variables are not as significant as they were above. Points Scored From 2 Point Shots, Free throws, and Total Rebounds have respective P-values of .5959, .3097, and .8287 and respective t-Statistics of .5329, 1.0238, and -.2173, all of which show major insignificance of the variables. The only variable that remained significant in this model is Age.

While some of the signs of the coefficients are surprising, the results proving the insignificance of the variables make them more understandable. It is still puzzling that Total Rebounds have a negative coefficient. Rebounds are an important statistic for Centers because they spend most of their time close to the basket. Our data shows six out of the top seven rebounding players are Centers, with the exception of the seven-foot tall Power Forward, Pau Gasol. We assumed that rebounds would have a large correlation to salary for Centers but our model is not able to give us a firm answer because of the insignificance.

The regression results for Power Forwards tell a similar story:

$$Y = -5535536 - 5551.210X_1 + 208821.8X_2 + 524480.5X_3 + 611084.6X_4 + 1025524X_5 - 230825.6X_6 - 244000.5X_7 + 2315423X_8 - 163003.8X_9$$

Running the model on Power Forwards is also flawed; as the F-Statistic is a low 9.2408 even though the R-squared is .519251 and the adjusted R-squared is .46306. The individual p-values and t-Statistics again show massive insignificance. Points from Two-Point Shots, Total Rebounds, and Assists have respective P-values of .5777, .4199, and .7258 and t-Statistics of .5591, -.8109, and -.3520. While the significant p-value and t-Statistics of .0172 and 2.4245 for Age are not shocking, the results for the same tests of .0996 and 1.6668 for Turnovers are. It is difficult to explain the relative significance of Turnovers, outside of the explanation used to explain the significance in the original regression of all players, as it is due to the better players having possession of the ball more often and therefore having a higher exposure to turnovers.

The negative coefficients for Games Started, Total Rebounds, and Assists are all startling, but again it is difficult to determine if they mean much because of the low significance. It does not make much sense that a player would be paid less because he started more games, grabbed more rebounds or passed out more assists. Based on our knowledge from watching and studying basketball over the years, we noticed that assists, rebounds, and points scored are the most important statistics on the surface and it is difficult to determine why they have negative coefficients in this regression.

The model run on Small Forwards give a regression equation of:

$Y = -5142444 - 53994.10X_1 + 161563.2X_2 + 633691.0X_3 + 747396.1X_4 + 1849989X_5 + 48274.54X_6 + 1509807X_7 + 1079740X_8 - 456240.9X_9$

The R-squared of .6595 and adjusted R-squared of .6138 are both higher than the regression of all the players and the F-statistic for the model is a little higher than the previous two at 14.41 but it is still not high enough to be considered a good model in this case. Similar to the results of the previous regressions analyzing the players based on their positions, most of the variables in this regression are insignificant with the exception of games started, age, free throws made, and assists. Those variables have p-values of .0131, .0842, .1063, and .0496 and t-Statistics of -2.5491, 1.7527, 1.6369, and 1.9994. The absolute values of these t-Statistics are lower than what we what saw in the original regression and the coefficients also tell an interesting story. Games Started has the lowest p-value, which shows its significance, but at the same time it has a negative coefficient. The value of the coefficient and the significance for assists is surprising as our data shows that small forwards tend to not be at the top of the league in that category. Even though a few of the variables are statistically significant, it is still not accurate to call this regression significant.

The model run on the 88 shooting guards in our data set gave us the regression:

$Y = -7643698 + 16132.37X_1 + 267612.7X_2 - 303020.5X_3 - 305807.2X_4 - 324968.6X_5 - 258901.4X_6 + 700280.4X_7 + 1831638X_8 + 439575.5X_9$

Again, the regression appears relevant on the surface with an R-squared of .6452 and an adjusted R-squared of .604228 but the F-statistic is a low 15.7581 and the p-values and t-Statistics show that most of the variables are insignificant. The most significant variables are age and turnovers with p-values of .0006 and .1537 and t-Statistics of 3.6031 and 1.4407. Based on their name, the prime responsibility of shooting guards is to shoot and score points. Therefore, it is surprising that the three scoring variables in our regression: points scored from two point shots, points scored from three points shots, and free throws made - have negative coefficients and similarly high respective p-values of .6847, .6535, and .6695. These values should be the most significant to the shooting guard but they prove to be insignificant in this regression.

The final regression that we ran was on the 70 Point Guards in the model is as follows:

$Y = -9365643 + 51665.40X_1 + 320476.0X_2 + 1317889X_3 + 910788.6X_4 + 1046434X_5 - 687717.9X_6 + 882108.9X_7 + 871949.7X_8 - 1062540X_9$

The R-squared of .6740 and the adjusted R-squared of .6251 are high but the F-statistic of 13.7814 is lower than what we would like to consider it a good regression. However, the p-values for Games Started, Age, points from two point shots, and assists of .0784, .0034, .1317, and .0198 show that this regression has more significant variables than the regression for players at other positions. The extremely low p-value for assists and its t-Statistic of 2.3934 are important, as they align with our prediction for point guards that assists would be an important factor in their salaries. Our data shows that most of the assist leaders in the NBA are point guards as it is primarily their job to carry the ball down the court and start the play. The significance of points from two point shots for point guards also shows their value as well as scorers in the NBA.

In the six regression models that we ran, the one variable that remained significant throughout was Age. While a player is not guaranteed to have a higher salary because of his age, he is much more likely to earn more in his later years because of the opportunities in free agency that are not there for rookies and other players in their first few years in the league. This shows that if a player wants to earn a large salary, he needs to prove his worth to the league long enough so that he reaches free agency while he is still a relevant player.

To put our model into practical use, we decided to take player statistics from previous years and attempted to forecast what their projected salary should be, based on our regression. We chose Michael Jordan's 1995-1996 season (his "comeback year") to see what his value would be if he were playing nowadays, players that signed during the summer of 2014, and other significant players we thought were important to note (Michael Jordan 1995-96 Splits). Some of these players were: Lebron James (his last year with the Miami Heat before signing as a free agent with the Cleveland Cavaliers), Lance Stephenson (signed with a different team during the Free Agency of 2014), Giannis Antetokuonmpo (the youngest player in the league during the 2013-2014 season), and Joakim Noah (the 2013-2014 Defensive Player of the Year) (2013-14 NBA Player Stats: Per Game). The results of the regression for these players are listed below in Table 3. For the players that were selected, we noticed that each player's projected salary is more than what they ended up making during the 2014-2015 season (excluding Michael Jordan), showing that our model is producing salary estimates that are significantly higher than the actual player salaries for the season to follow.

Table 3
Forecasted Salaries Based on our Model

Name	Season	Projected Salary	Actual 2014-15 Salary	Percent Increase of the Projection
Michael Jordan	1995-1996	$28,919,435	Didn't Play	N/A
Lebron James	2013-2014	$27,096,904	$20,644,400	31%
Lance Stephenson	2013-2014	$13,846,493	$9,000,000	54%
Giannis Antetokuonmpo	2013-2014	$5,150,691	$1,873,200	175%
Joakim Noah	2013-2014	$16,049,266	$12,200,000	32%
Vince Carter	2013-2014	$13,248,094	$3,911,981	239%
Carmelo Anthony	2013-2014	$29,938,018	$22,458,401	33%
Kyle Lowry	2013-2014	$17,129,800	$12,000,000	43%

Summary and Conclusions

While not as accurate as we would have liked, our regression model served its purpose of using NBA players' statistics to predict their salaries. The variables we chose: Games Started, Age, Points Scored from 2-Point Shots per game, Points Scored from 3-Point Shots per game, Free Throws per game, Total Rebounds per game, Assists Per Game, Turnovers per game, and Player Points per-game as a Percentage of Team Points per-game, gave us a model that put us in the ballpark range of the actual NBA salaries when we forecasted. Our model didn't take into account the same variables that were used in the models by other researchers, such as blocked shots and the ethnicity of the player. The outside research also determined that rebounds are much more significant than they are when analyzing the results of our model. Running the model five different times for players of each of the five positions did not give us results with enough significance to consider them useful in forecasting player salaries. Another important discovery was that the Y-intercept was -$8.9 million. As discussed earlier, this translated into the estimated "cost" per player on the court. If this holds true, with five players on the court for the team, the average amount of money invested in players on the court at a time should be $8.9 million * 5 = $44.5 million. With a team salary being $63.065 million, that would mean approximately 70% of the team salary should be actively playing on the court. The major finding in our model is that Age is the most directly related variable to a player's salary. Older players who have similar statistics to younger players are likely to have higher salaries because of the young players still playing under their less lucrative salaries. To finalize, our model is able to give us respectable results, despite minor flaws, but cannot precisely measure player salary based on our chosen independent variables.

Appendix

Table 4 - Regression of All Players

Dependent Variable: SALARY
Method: Least Squares
Date: 11/28/16 Time: 14:44
Sample (adjusted): 1 397
Included observations: 397 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-8906933.	1111575.	-8.012892	0.0000
GS	3258.957	10385.98	0.313784	0.7539
AGE	316085.7	40457.87	7.812713	0.0000
_2P_2	751906.8	395535.7	1.900983	0.0580
_3P_3	476700.1	375809.4	1.268462	0.2054
FT	871842.1	444923.7	1.959532	0.0508
TRB	173790.8	121912.9	1.425532	0.1548
AST	162542.4	191870.0	0.847148	0.3974
TURN_OVERS	1307761.	570068.9	2.294040	0.0223
PPG_TP	-418831.3	372940.7	-1.123051	0.2621

R-squared	0.536302	Mean dependent var	4511948.
Adjusted R-squared	0.525518	S.D. dependent var	4880290.
S.E. of regression	3361672.	Akaike info criterion	32.91864
Sum squared resid	4.37E+15	Schwarz criterion	33.01899
Log likelihood	-6524.350	Hannan-Quinn criter.	32.95839
F-statistic	49.73274	Durbin-Watson stat	2.132024
Prob(F-statistic)	0.000000		

Table 5 - Regression for Centers

Dependent Variable: SALARY
Method: Least Squares
Date: 11/28/16 Time: 15:09
Sample: 1 75
Included observations: 75

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-15183029	3350093.	-4.532122	0.0000
GS	42395.70	34828.75	1.217262	0.2279
AGE	580821.4	116178.4	4.999394	0.0000
_2P_2	670587.0	1258278.	0.532940	0.5959
_3P_3	1513450.	1454222.	1.040728	0.3019
FT	1442637.	1409055.	1.023833	0.3097
TRB	-74780.92	344175.2	-0.217276	0.8287
AST	-186761.8	823394.2	-0.226819	0.8213
TURN_OVERS	941658.0	1829548.	0.514694	0.6085
PPG_TP	-404060.7	1165418.	-0.346709	0.7299

R-squared	0.562416	Mean dependent var	5529723.
Adjusted R-squared	0.501828	S.D. dependent var	5564949.
S.E. of regression	3927814.	Akaike info criterion	33.32863
Sum squared resid	1.00E+15	Schwarz criterion	33.63763
Log likelihood	-1239.824	Hannan-Quinn criter.	33.45201
F-statistic	9.282556	Durbin-Watson stat	2.179348
Prob(F-statistic)	0.000000		

Table 6 - Regression for Power Forwards

Dependent Variable: SALARY
Method: Least Squares
Date: 11/28/16 Time: 15:23
Sample: 1 87
Included observations: 87

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-5535536.	2304267.	-2.402298	0.0187
GS	-5551.210	23105.76	-0.240252	0.8108
AGE	208821.8	85776.91	2.434476	0.0172
_2P_2	524480.5	938157.8	0.559054	0.5777
_3P_3	611084.6	937357.5	0.651923	0.5164
FT	1025524.	1210622.	0.847105	0.3996
TRB	-230825.6	284663.1	-0.810873	0.4199
AST	-244000.5	693086.6	-0.352049	0.7258
TURN_OVERS	2315423.	1389147.	1.666795	0.0996
PPG_TP	-163003.8	946677.8	-0.172185	0.8637

R-squared	0.519251	Mean dependent var	4391007.
Adjusted R-squared	0.463060	S.D. dependent var	4232964.
S.E. of regression	3101754.	Akaike info criterion	32.84062
Sum squared resid	7.41E+14	Schwarz criterion	33.12405
Log likelihood	-1418.567	Hannan-Quinn criter.	32.95475
F-statistic	9.240766	Durbin-Watson stat	2.116376
Prob(F-statistic)	0.000000		

Table 7 - Regression for Point Guards

Dependent Variable: SALARY
Method: Least Squares
Date: 11/28/16 Time: 15:26
Sample: 1 70
Included observations: 70

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-9365643.	2893091.	-3.237245	0.0020
GS	51665.40	28852.58	1.790668	0.0784
AGE	320476.0	104990.6	3.052425	0.0034
_2P_2	1317889.	862373.3	1.528211	0.1317
_3P_3	910788.6	756284.2	1.204294	0.2332
FT	1046434.	867344.4	1.206480	0.2324
TRB	-687717.9	523531.7	-1.313613	0.1940
AST	882108.9	368558.4	2.393403	0.0198
TURN_OVERS	871949.7	1073430.	0.812303	0.4198
PPG_TP	-1062540.	759451.8	-1.399089	0.1669

R-squared	0.673971	Mean dependent var	4606191.
Adjusted R-squared	0.625066	S.D. dependent var	5006171.
S.E. of regression	3065371.	Akaike info criterion	32.84080
Sum squared resid	5.64E+14	Schwarz criterion	33.16201
Log likelihood	-1139.428	Hannan-Quinn criter.	32.96839
F-statistic	13.78138	Durbin-Watson stat	1.969368
Prob(F-statistic)	0.000000		

Table 8 - Regression for Small Forwards

Dependent Variable: SALARY
Method: Least Squares
Date: 11/28/16 Time: 15:34
Sample: 1 77
Included observations: 77

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-5142444.	2648245.	-1.941831	0.0564
GS	-53994.10	21181.58	-2.549106	0.0131
AGE	161563.2	92178.29	1.752725	0.0842
_2P_2	633691.0	890307.1	0.711767	0.4791
_3P_3	747396.1	842382.5	0.887241	0.3781
FT	1849989.	1130178.	1.636900	0.1063
TRB	48274.54	424777.6	0.113647	0.9099
AST	1509807.	755092.8	1.999499	0.0496
TURN_OVERS	1079740.	1650092.	0.654351	0.5151
PPG_TP	-456240.9	832096.7	-0.548303	0.5853

R-squared	0.659503	Mean dependent var	4328880.
Adjusted R-squared	0.613764	S.D. dependent var	5142973.
S.E. of regression	3196249.	Akaike info criterion	32.91348
Sum squared resid	6.84E+14	Schwarz criterion	33.21787
Log likelihood	-1257.169	Hannan-Quinn criter.	33.03523
F-statistic	14.41899	Durbin-Watson stat	2.191900
Prob(F-statistic)	0.000000		

Table 9 - Regression for Shooting Guards

Dependent Variable: SALARY
Method: Least Squares
Date: 11/28/16 Time: 15:38
Sample: 1 88
Included observations: 88

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-7643698.	1888976.	-4.046478	0.0001
GS	16132.37	18555.72	0.869401	0.3873
AGE	267612.7	74272.30	3.603130	0.0006
_2P_2	-303020.5	743557.3	-0.407528	0.6847
_3P_3	-305807.2	678529.2	-0.450691	0.6535
FT	-324968.6	758510.2	-0.428430	0.6695
TRB	-258901.4	393574.7	-0.657820	0.5126
AST	700280.4	583848.3	1.199422	0.2340
TURN_OVERS	1831638.	1271376.	1.440673	0.1537
PPG_TP	439575.5	693810.4	0.633567	0.5282

R-squared	0.645169	Mean dependent var	3849313.
Adjusted R-squared	0.604228	S.D. dependent var	4470487.
S.E. of regression	2812404.	Akaike info criterion	32.64362
Sum squared resid	6.17E+14	Schwarz criterion	32.92514
Log likelihood	-1426.319	Hannan-Quinn criter.	32.75704
F-statistic	15.75814	Durbin-Watson stat	2.083468
Prob(F-statistic)	0.000000		

Table 10 – Data Used to Create the Regression

Player	Pos	Age	Tm	G	GS	FG	FG%	3P	3P*3	3P%	2P	2P*2	2P%	FT	FT%	TRB	AST	STL	BLK	TO	PF	PPG	Salary	Team PPG	ppg%tp
Joakim Noah	C	29	CHI	67	67	2.8	0.445	0	0	0	2.8	5.6	0.447	1.6	0.603	9.6	4.7	0.7	1.1	1.8	3	7.2	$12,200,000	94.2	7.6433
Marc Gasol	C	30	MEM	81	81	6.5	0.494	0	0	0.176	6.5	13	0.5	4.3	0.795	7.8	3.8	0.9	1.6	2.2	2.6	17.4	$15,829,688	98.3	17.7009
DeMarcus Cousins	C	24	SAC	59	59	8.4	0.467	0	0	0.25	8.4	16.8	0.469	7.2	0.782	12.7	3.6	1.5	1.7	4.3	4.1	24.1	$13,701,250	101.3	23.7907
Al Horford	C	28	ATL	76	76	6.8	0.538	0.1	0.3	0.306	6.7	13.4	0.547	1.4	0.759	7.2	3.2	0.9	1.3	1.3	1.6	15.2	$12,000,000	102.5	14.8293
Tim Duncan	C	38	SAS	77	77	5.4	0.512	0	0	0.286	5.4	10.8	0.514	3	0.74	9.1	3	0.8	2	1.7	2.1	13.9	$10,000,000	103.2	13.4690
Andrew Bogut	C	30	GSW	67	65	3	0.563	0	0		3	6	0.563	0.3	0.524	8.1	2.7	0.6	1.7	1.6	2.8	6.3	$12,972,973	110	5.7273
Zaza Pachulia	C	30	MIL	73	45	3.3	0.454	0	0	0	3.3	6.6	0.456	1.7	0.788	6.8	2.4	1.1	0.3	1.8	2.3	8.3	$5,200,000	97.8	8.4867
Chris Bosh	C	30	MIA	44	44	7.8	0.46	1.4	4.2	0.375	6.4	12.8	0.485	4.1	0.772	7	2.2	0.9	0.6	2.2	1.6	21.1	$20,644,400	94.7	22.2809
Gorgui Dieng	C	25	MIN	73	49	3.5	0.506	0	0	0.167	3.5	7	0.51	2.6	0.783	8.3	2	1	1.7	1.7	2.6	9.7	$1,413,480	97.8	9.9182
Nikola Vucevic	C	24	ORL	74	74	8.5	0.523	0	0	0.333	8.5	17	0.524	2.2	0.752	10.9	2	0.7	0.7	2	3	19.3	$2,751,260	95.7	20.1672
Al Jefferson	C	30	CHO	65	61	7.5	0.481	0	0	0.4	7.4	14.8	0.482	1.7	0.655	8.4	1.7	0.7	1.3	1	2.1	16.6	$13,500,000	94.2	17.6221
Jason Smith	C	28	NYK	82	31	3.1	0.434	0.2	0.6	0.357	3	6	0.44	1.5	0.83	4	1.7	0.4	0.5	1.3	2.5	8	$3,278,000	91.9	8.7051
Kelly Olynyk	C	23	BOS	64	13	3.9	0.475	1	3	0.349	3	6	0.538	1.5	0.684	4.7	1.7	1	0.6	1.5	3.3	10.3	$2,075,760	101.4	10.1578
Nerlens Noel	C	20	PHI	75	71	4	0.462	0	0		4	8	0.462	1.9	0.609	8.1	1.7	1.8	1.9	1.9	2.8	9.9	$3,315,120	92	10.7609
Andrea Bargnani	C	29	NYK	29	22	5.7	0.454	0.5	1.5	0.366	5.1	10.2	0.466	3	0.813	4.4	1.6	0.1	0.9	1.4	1.8	14.8	$11,500,000	91.9	16.1045
Cody Zeller	C	22	CHO	62	45	2.8	0.461	0	0	1	2.8	5.6	0.46	2	0.774	5.8	1.6	0.5	0.8	1	2.5	7.6	$4,030,560	94.2	8.0679
Tiago Splitter	C	30	SAS	52	35	3.3	0.558	0	0		3.3	6.6	0.558	1.7	0.75	4.8	1.5	0.7	0.7	1.2	2	8.2	$9,250,000	103.2	7.9457
Jordan Hill	C	27	LAL	70	57	5.1	0.459	0	0	0.273	5.1	10.2	0.462	1.8	0.738	7.9	1.5	0.5	0.7	1.5	2.3	12	$9,000,000	98.5	12.1827
Tyler Zeller	C	25	BOS	82	59	4.1	0.549	0	0		4.1	8.2	0.549	1.9	0.823	5.7	1.4	0.2	0.6	0.9	2.5	10.2	$1,703,760	101.4	10.0592
Anderson Varejao	C	32	CLE	26	26	4.3	0.555	0	0	0	4.3	8.6	0.561	1.3	0.733	6.5	1.3	0.7	0.6	1.3	2.2	9.8	$9,704,545	103.1	9.5053
Player	**Pos**	**Age**	**Tm**	**G**	**GS**	**FG**	**FG%**	**3P**	**3P*3**	**3P%**	**2P**	**2P*2**	**2P%**	**FT**	**FT%**	**TRB**	**AST**	**STL**	**BLK**	**TO**	**PF**	**PPG**	**Salary**	**Team PPG**	**ppg%tp**
Rudy Gobert	C	22	UTA	82	37	3.1	0.604	0	0	0	3.1	6.2	0.607	2.1	0.623	9.5	1.3	0.8	2.3	1.4	2.1	8.4	$1,127,400	95.1	8.8328
Marcin Gortat	C	30	WAS	82	82	5.4	0.566	0	0	0	5.4	10.8	0.569	1.5	0.703	8.7	1.2	0.6	1.3	1.2	2.3	12.2	$10,434,782	98.5	12.3858
Dwight Howard	C	29	HOU	41	41	6.1	0.593	0	0	0.5	6.1	12.2	0.594	3.5	0.528	10.5	1.2	0.7	1.3	2.8	3.3	15.8	$21,436,271	103.9	15.2069
Cole Aldrich	C	26	NYK	61	16	2.4	0.478	0	0		2.4	4.8	0.478	0.8	0.781	5.5	1.2	0.6	1.1	1	2	5.5	$915,243	91.9	5.9848
Lavoy Allen	C	25	IND	63	0	2.2	0.472	0	0		2.2	4.4	0.472	0.5	0.702	5.1	1.2	0.2	0.7	0.6	1.6	5	$915,243	97.3	5.1387
Tyson Chandler	C	32	DAL	75	75	3.9	0.666	0	0		3.9	7.8	0.666	2.5	0.72	11.5	1.1	0.6	1.2	1.4	2.3	10.3	$14,596,888	105.2	9.7909
Roy Hibbert	C	28	IND	76	76	4.2	0.446	0	0	0	4.2	8.4	0.448	2.2	0.824	7.1	1.1	0.2	1.6	1.4	2.8	10.6	$14,898,938	97.3	10.8941
Henry Sims	C	24	PHI	73	32	3.3	0.474	0.1	0.3	0.182	3.2	6.4	0.488	1.5	0.774	4.9	1.1	0.5	0.4	1.4	1.8	8	$915,243	92	8.6957
Ronny Turiaf	C	32	MIN	2	0	0		0	0		0	0		0		0.5	1	0	0	0	0	0	$1,500,000	97.8	0.0000
Samuel Dalembert	C	33	NYK	32	21	1.8	0.438	0	0		1.8	3.6	0.438	0.4	0.7	5.3	0.9	0.4	1.3	1.2	2.2	4	$3,867,282	91.9	4.3526
Chris Kaman	C	32	POR	74	13	3.8	0.515	0	0		3.8	7.6	0.515	1	0.706	6.5	0.9	0.2	0.7	1.5	1.9	8.6	$4,800,000	102.8	8.3658
Nikola Pekovic	C	29	MIN	31	29	4.5	0.424	0	0		4.5	9	0.424	3.5	0.837	7.5	0.9	0.6	0.4	1.4	1.9	12.5	$12,100,000	97.8	12.7812
Omer Asik	C	28	NOP	76	76	2.8	0.517	0	0		2.8	5.6	0.517	1.7	0.582	9.8	0.9	0.4	0.7	1.3	1.9	7.3	$8,374,646	99.4	7.3441
Marreese Speights	C	27	GSW	76	9	4.2	0.492	0.1	0.3	0.278	4.1	8.2	0.498	2	0.843	4.3	0.9	0.3	0.4	1.1	2.5	10.4	$3,657,500	110	9.4545
Robin Lopez	C	26	POR	59	59	4	0.535	0	0	0	4	8	0.537	1.7	0.772	6.7	0.9	0.3	1.4	1.2	2.1	9.6	$5,340,229	102.8	9.3385
Larry Sanders	C	26	MIL	27	26	3.2	0.5	0	0		3.2	6.4	0.5	0.9	0.5	6.1	0.9	1	1.4	1	3.9	7.3	$11,000,000	97.8	7.4642
Mason Plumlee	C	24	BRK	82	45	3.4	0.573	0	0	0	3.4	6.8	0.576	1.9	0.495	6.2	0.9	0.8	0.8	1.3	2.5	8.7	$1,357,080	98	8.8776
John Henson	C	24	MIL	67	11	2.9	0.566	0	0		2.9	5.8	0.566	1.2	0.569	4.7	0.9	0.4	2	1.3	2.3	7	$1,987,320	97.8	7.1575
Steven Adams	C	21	OKC	70	67	3.1	0.544	0	0	0	3.1	6.2	0.547	1.5	0.502	7.5	0.9	0.5	1.2	1.4	3.2	7.7	$2,184,960	104	7.4038
J.J. Hickson	C	26	DEN	73	8	3	0.475	0	0	0	3	6	0.477	1.6	0.577	6.2	0.8	0.5	0.5	1.3	1.9	7.6	$5,381,750	101.5	7.4877
Robert Sacre	C	25	LAL	67	18	1.9	0.412	0	0	0	1.9	3.8	0.413	0.8	0.671	3.5	0.8	0.4	0.6	0.5	1.8	4.6	$915,243	98.5	4.6701
Jusuf Nurkic	C	20	DEN	62	27	2.8	0.446	0	0	0	2.8	5.6	0.449	1.4	0.636	6.2	0.8	0.8	1.1	1.4	3.3	6.9	$1,762,680	101.5	6.7980

Player	Pos	Age	Tm	G	GS	FG	FG%	3P	3P*3	3P%	2P	2P*2	2P%	FT	FT%	TRB	AST	STL	BLK	TO	PF	PPG	Salary	Team PPG	ppg%tp
Chris Andersen	C	36	MIA	60	20	2	0.58	0.1	0.3	0.308	1.9	3.8	0.598	1.3	0.667	5	0.7	0.4	1	0.7	1.5	5.3	$5,375,000	94.7	5.5966
Chuck Hayes	C	31	TOR	29	0	0.8	0.478	0	0		0.8	1.6	0.478	0.2	0.545	1.8	0.7	0.3	0.1	0.3	1.3	1.7	$5,958,750	104	1.6346
DeAndre Jordan	C	26	LAC	82	82	4.6	0.71	0	0	0.25	4.6	9.2	0.713	2.3	0.397	15	0.7	1	2.2	1.3	3	11.5	$11,440,123	106.7	10.7779
Brook Lopez	C	26	BRK	72	44	7	0.513	0	0	0.1	7	14	0.517	3.1	0.814	7.4	0.7	0.6	1.8	1.4	2.9	17.2	$15,719,063	98	17.5510
Alexis Ajinca	C	26	NOP	68	8	2.7	0.55	0	0		2.7	5.4	0.55	1.2	0.818	4.6	0.7	0.3	0.8	1	2.2	6.5	$981,084	99.4	6.5392
Kevin Seraphin	C	25	WAS	79	0	2.9	0.513	0	0	0	2.9	5.8	0.516	0.7	0.707	3.6	0.7	0.1	0.7	1.2	2.5	6.6	$3,898,692	98.5	6.7005
Andre Drummond	C	21	DET	82	82	6	0.514	0	0	0	6	12	0.515	1.7	0.389	13.5	0.7	0.9	1.9	1.5	3.5	13.8	$2,568,360	98.5	14.0102
Elton Brand	C	35	ATL	36	4	1.2	0.442	0	0	0	1.2	2.4	0.447	0.3	0.522	2.8	0.6	0.5	0.7	0.5	1.5	2.7	$2,000,000	102.5	2.6341
Meyers Leonard	C	22	POR	55	7	2.3	0.51	0.9	2.7	0.42	1.4	2.8	0.586	0.5	0.938	4.5	0.6	0.2	0.3	0.7	2.1	5.9	$2,317,920	102.8	5.7393
Ian Mahinmi	C	28	IND	61	6	1.9	0.552	0	0		1.9	3.8	0.552	0.5	0.304	5.8	0.5	0.5	0.8	1	2.8	4.3	$4,000,000	97.3	4.4193
Aron Baynes	C	28	SAS	70	17	2.6	0.566	0	0	0.25	2.6	5.2	0.57	1.3	0.865	4.5	0.5	0.2	0.3	0.9	2.3	6.6	$2,077,000	103.2	6.3953
Kosta Koufos	C	25	MEM	81	3	2.2	0.508	0	0		2.2	4.4	0.508	0.7	0.647	5.3	0.5	0.4	0.8	0.9	1.8	5.2	$3,000,000	98.3	5.2899
Jonas Valanciunas	C	22	TOR	80	80	4.7	0.572	0	0	0	4.7	9.4	0.573	2.7	0.786	8.7	0.5	0.4	1.2	1.4	2.8	12	$3,678,360	104	11.5385
Alex Len	C	21	PHO	69	44	2.6	0.507	0	0	0.333	2.6	5.2	0.509	1.1	0.702	6.6	0.5	0.5	1.5	1.1	3.1	6.3	$3,649,920	102.4	6.1523
Ognjen Kuzmic	C	24	GSW	16	0	0.5	0.667	0	0		0.5	1	0.667	0.3	1	1.1	0.4	0.1	0.1	0.3	0.8	1.3	$816,482	110	1.1818
Ryan Hollins	C	30	SAC	46	9	1.2	0.646	0	0		1.2	2.4	0.646	0.7	0.574	2.2	0.3	0.1	0.4	0.5	1.4	3	$915,243	101.3	2.9615
Bernard James	C	29	DAL	16	2	0.8	0.444	0	0		0.8	1.6	0.444	1.3	0.87	2.4	0.3	0.1	0.9	0.4	1.1	2.8	$915,243	105.2	2.6616
Jerome Jordan	C	28	BRK	44	0	1.1	0.532	0	0		1.1	2.2	0.532	0.9	0.864	2.4	0.3	0.2	0.3	0.5	1.3	3.1	$816,482	98	3.1633
Joel Freeland	C	27	POR	48	8	1.5	0.49	0	0		1.5	3	0.49	0.4	0.84	4	0.3	0.2	0.5	0.5	1.9	3.5	$3,013,512	102.8	3.4047
Jeff Withey	C	24	NOP	37	0	0.9	0.5	0	0		0.9	1.8	0.5	0.9	0.68	1.7	0.3	0.1	0.5	0.3	0.7	2.6	$816,482	99.4	2.6157
Bismack Biyombo	C	22	CHO	64	21	1.6	0.543	0	0		1.6	3.2	0.543	1.6	0.583	6.4	0.3	0.3	1.5	0.8	2.2	4.8	$3,873,398	94.2	5.0955
Player	**Pos**	**Age**	**Tm**	**G**	**GS**	**FG**	**FG%**	**3P**	**3P*3**	**3P%**	**2P**	**2P*2**	**2P%**	**FT**	**FT%**	**TRB**	**AST**	**STL**	**BLK**	**TO**	**PF**	**PPG**	**Salary**	**Team PPG**	**ppg%tp**
Sim Bhullar	C	22	SAC	3	0	0.3	0.5	0	0		0.3	0.6	0.5	0		0.3	0.3	0	0.3	0	0	0.7	$35,000	101.3	0.6910
Greg Stiemsma	C	29	TOR	17	0	0.4	0.75	0	0		0.4	0.8	0.75	0.1	0.5	0.9	0.2	0.1	0	0.4	0.9	0.8	$915,243	104	0.7692
Dewayne Dedmon	C	25	ORL	59	15	1.5	0.562	0	0	0	1.5	3	0.565	0.6	0.531	5	0.2	0.3	0.8	0.9	2.4	3.7	$816,482	95.7	3.8662
Festus Ezeli	C	25	GSW	46	7	1.7	0.547	0	0		1.7	3.4	0.547	1.1	0.628	3.4	0.2	0.2	0.9	0.7	1.7	4.4	$1,112,880	110	4.0000
Alex Kirk	C	23	CLE	5	0	0.2	0.25	0	0		0.2	0.4	0.25	0.4	1	0.2	0.2	0	0	0	0.2	0.8	$507,336	103.1	0.7759
Lucas Nogueira	C	22	TOR	6	0	0.3	0.25	0	0		0.3	0.6	0.25	0.3	0.5	1.8	0.2	0.3	0	0.3	1	1	$1,762,680	104	0.9615
Clint Capela	C	20	HOU	12	0	1.2	0.483	0	0		1.2	2.4	0.483	0.3	0.174	3	0.2	0.1	0.8	0.4	1.2	2.7	$1,189,200	103.9	2.5987
Nazr Mohammed	C	37	CHI	23	0	0.6	0.433	0	0		0.6	1.2	0.433	0	0.333	1.7	0.1	0.2	0.2	0.3	0.7	1.2	$915,243	94.2	1.2739
Brendan Haywood	C	35	CLE	22	1	0.6	0.467	0	0		0.6	1.2	0.467	0.3	0.538	1.3	0.1	0.1	0.5	0.5	0.8	1.6	$21,213,688	103.1	1.5519
Joel Anthony	C	32	DET	49	0	0.7	0.581	0	0		0.7	1.4	0.581	0.3	0.682	1.9	0.1	0.2	1	0.2	1.1	1.8	$3,800,000	98.5	1.8274
Hassan Whiteside	C	25	MIA	48	32	5.1	0.628	0	0		5.1	10.2	0.628	1.6	0.5	10	0.1	0.6	2.6	1.2	2.7	11.8	$790,881	94.7	12.4604
Miroslav Raduljica	C	27	MIN	5	0	0.6	0.375	0	0		0.6	1.2	0.375	0.4	1	1	0	0.2	0	0.4	1.6	1.6	$300,000	97.8	1.6360
Blake Griffin	PF	25	LAC	67	67	8.6	0.502	0.1	0.3	0.4	8.4	16.8	0.504	4.6	0.728	7.6	5.3	0.9	0.5	2.3	2.9	21.9	$14,693,906	106.7	20.5248
Draymond Green	PF	24	GSW	79	79	4.3	0.443	1.4	4.2	0.337	2.9	5.8	0.523	1.7	0.66	8.2	3.7	1.6	1.3	1.7	3.2	11.7	$915,243	110	10.6364
David West	PF	34	IND	66	66	4.9	0.471	0.1	0.3	0.2	4.8	9.6	0.479	1.8	0.739	6.8	3.4	0.7	0.7	1.8	2.4	11.7	$12,000,000	97.3	12.0247
Paul Millsap	PF	29	ATL	73	73	6.1	0.476	1.1	3.3	0.356	5	10	0.513	3.5	0.757	7.8	3.1	1.8	0.9	2.3	2.8	16.7	$9,500,000	102.5	16.2927
Boris Diaw	PF	32	SAS	81	15	3.6	0.46	0.7	2.1	0.32	2.9	5.8	0.512	0.9	0.774	4.3	2.9	0.4	0.3	1.6	1.8	8.7	$8,000,000	103.2	8.4302
Pau Gasol	PF	34	CHI	78	78	7.3	0.494	0.2	0.6	0.462	7.2	14.4	0.495	3.8	0.803	11.8	2.7	0.3	1.9	2	1.9	18.5	$7,128,000	94.2	19.6391
Markieff Morris	PF	25	PHO	82	82	6.2	0.465	0.7	2.1	0.318	5.5	11	0.494	2.2	0.763	6.2	2.3	1.2	0.5	2.1	3	15.3	$2,989,239	102.4	14.9414
Jared Sullinger	PF	22	BOS	58	49	5.4	0.439	0.9	2.7	0.283	4.5	9	0.494	1.7	0.744	7.6	2.3	0.8	0.7	1.3	2.6	13.3	$1,424,520	101.4	13.1164
Zach Randolph	PF	33	MEM	71	71	6.4	0.487	0.1	0.3	0.35	6.3	12.6	0.49	3.2	0.765	10.5	2.2	1	0.2	2.2	2.5	16.1	$16,500,000	98.3	16.3784

Player	Pos	Age	Tm	G	GS	FG	FG%	3P	3P*3	3P%	2P	2P*2	2P%	FT	FT%	TRB	AST	STL	BLK	TO	PF	PPG	Salary	Team PPG	ppg%tp
Kevin Love	PF	26	CLE	75	75	5.5	0.434	1.9	5.7	0.367	3.6	7.2	0.48	3.4	0.804	9.7	2.2	0.7	0.5	1.6	1.9	16.4	$15,719,062	103.1	15.9069
Anthony Davis	PF	21	NOP	68	68	9.4	0.535	0	0	0.083	9.4	18.8	0.54	5.5	0.805	10.2	2.2	1.5	2.9	1.4	2.1	24.4	$5,607,240	99.4	24.5473
Greg Monroe	PF	24	DET	69	57	6.1	0.496	0	0		6.1	12.2	0.496	3.7	0.75	10.2	2.1	1.1	0.5	2.2	2.1	15.9	$5,479,934	98.5	16.1421
Dirk Nowitzki	PF	36	DAL	77	77	6.3	0.459	1.4	4.2	0.38	5	10	0.486	3.3	0.882	5.9	1.9	0.5	0.4	1.1	2.1	17.3	$7,974,482	105.2	16.4449
Josh McRoberts	PF	27	MIA	17	4	1.6	0.528	0.5	1.5	0.421	1.2	2.4	0.588	0.5	0.615	2.6	1.9	0.7	0.2	1.3	2.4	4.2	$5,305,000	94.7	4.4351
Patrick Patterson	PF	25	TOR	81	4	3	0.449	1.3	3.9	0.371	1.7	3.4	0.536	0.8	0.788	5.3	1.9	0.7	0.5	0.7	1.8	8	$5,831,326	104	7.6923
Nene Hilario	PF	32	WAS	67	58	4.6	0.511	0	0	0.2	4.6	9.2	0.513	1.8	0.604	5.1	1.8	1	0.3	1.9	2.7	11	$13,000,000	98.5	11.1675
Donatas Motiejunas	PF	24	HOU	71	62	5	0.504	0.7	2.1	0.368	4.3	8.6	0.536	1.4	0.602	5.9	1.8	0.8	0.5	1.7	2.9	12	$1,483,920	103.9	11.5496
Ryan Kelly	PF	23	LAL	52	34	2	0.337	0.9	2.7	0.336	1.1	2.2	0.337	1.5	0.832	2.8	1.8	0.6	0.5	0.7	2.3	6.4	$1,650,000	98.5	6.4975
David Lee	PF	31	GSW	49	4	3.3	0.511	0	0	0	3.3	6.6	0.514	1.4	0.654	5.2	1.7	0.6	0.5	1	1.7	7.9	$15,012,000	110	7.1818
LaMarcus Aldridge	PF	29	POR	71	71	9.3	0.466	0.5	1.5	0.352	8.8	17.6	0.475	4.3	0.845	10.2	1.7	0.7	1	1.7	1.8	23.4	$16,256,000	102.8	22.7626
Luc Mbah a Moute	PF	28	PHI	67	61	3.7	0.395	0.9	2.7	0.307	2.8	5.6	0.435	1.4	0.589	4.9	1.6	1.2	0.3	1.5	1.6	9.9	$4,382,576	92	10.7609
Amir Johnson	PF	27	TOR	75	72	4	0.574	0.3	0.9	0.413	3.7	7.4	0.59	1.1	0.612	6.1	1.6	0.6	0.8	1.5	3	9.3	$7,000,000	104	8.9423
Marcus Morris	PF	25	PHO	81	35	4.1	0.434	1.4	4.2	0.358	2.7	5.4	0.487	0.9	0.628	4.8	1.6	0.8	0.2	0.9	2.3	10.4	$2,943,221	102.4	10.1563
Derrick Favors	PF	23	UTA	74	74	6.5	0.525	0	0	0.167	6.5	13	0.527	3	0.669	8.2	1.5	0.8	1.7	1.6	2.8	16	$12,833,333	95.1	16.8244
Nick Collison	PF	34	OKC	66	2	1.6	0.419	0.2	0.6	0.267	1.4	2.8	0.466	0.7	0.692	3.8	1.4	0.5	0.4	0.7	2.4	4.1	$2,242,003	104	3.9423
James Johnson	PF	27	TOR	70	17	3.4	0.589	0.2	0.6	0.216	3.2	6.4	0.643	1	0.657	3.7	1.4	0.8	1	1.1	2.2	7.9	$2,500,000	104	7.5962
Luis Scola	PF	34	IND	81	16	3.7	0.467	0.1	0.3	0.25	3.6	7.2	0.474	2	0.699	6.5	1.3	0.6	0.2	1.2	2.4	9.4	$4,868,499	97.3	9.6608
Carlos Boozer	PF	33	LAL	71	26	5.2	0.499	0	0		5.2	10.4	0.499	1.3	0.627	6.8	1.3	0.6	0.2	1.3	2.6	11.8	$3,251,000	98.5	11.9797
Channing Frye	PF	31	ORL	75	51	2.5	0.392	1.8	5.4	0.393	0.7	1.4	0.39	0.4	0.886	3.9	1.3	0.6	0.5	1	2.5	7.3	$8,579,088	95.7	7.6280
Brandon Bass	PF	29	BOS	82	43	4.2	0.504	0.1	0.3	0.281	4.1	8.2	0.515	2.1	0.79	4.9	1.3	0.5	0.4	1	1.7	10.6	$6,900,000	101.4	10.4536
Marvin Williams	PF	28	CHO	78	37	2.7	0.424	1.2	3.6	0.358	1.5	3	0.5	0.8	0.713	4.9	1.3	0.9	0.5	0.8	1.9	7.4	$7,000,000	94.2	7.8556
Player	**Pos**	**Age**	**Tm**	**G**	**GS**	**FG**	**FG%**	**3P**	**3P*3**	**3P%**	**2P**	**2P*2**	**2P%**	**FT**	**FT%**	**TRB**	**AST**	**STL**	**BLK**	**TO**	**PF**	**PPG**	**Salary**	**Team PPG**	**ppg%tp**
Mirza Teletovic	PF	29	BRK	40	4	3.1	0.382	1.6	4.8	0.321	1.5	3	0.476	0.8	0.717	4.9	1.2	0.4	0.4	1.3	1.9	8.5	$3,368,100	98	8.6735
Spencer Hawes	PF	26	LAC	73	15	2.2	0.393	0.8	2.4	0.313	1.5	3	0.452	0.6	0.647	3.5	1.2	0.3	0.7	0.8	2.4	5.8	$5,305,000	106.7	5.4358
Kenneth Faried	PF	25	DEN	75	71	5	0.507	0	0	0.125	5	10	0.511	2.7	0.691	8.9	1.2	0.8	0.8	1.6	2.8	12.6	$2,249,768	101.5	12.4138
Ed Davis	PF	25	LAL	79	24	3.6	0.601	0	0		3.6	7.2	0.601	1.2	0.487	7.6	1.2	0.6	1.2	0.7	2.6	8.3	$981,084	98.5	8.4264
Kyle O'Quinn	PF	24	ORL	51	17	2.3	0.492	0.2	0.6	0.279	2.1	4.2	0.538	0.9	0.772	3.9	1.2	0.6	0.8	1.1	2.2	5.8	$915,243	95.7	6.0606
Nikola Mirotic	PF	23	CHI	82	3	3.1	0.405	1.2	3.6	0.316	1.9	3.8	0.495	2.8	0.803	4.9	1.2	0.7	0.7	1.1	2.1	10.2	$5,305,000	94.2	10.8280
Taj Gibson	PF	29	CHI	62	17	4.1	0.502	0	0		4.1	8.2	0.502	2.1	0.717	6.4	1.1	0.6	1.2	1.2	2.6	10.3	$8,000,000	94.2	10.9342
Trevor Booker	PF	27	UTA	79	5	2.9	0.487	0.4	1.2	0.345	2.5	5	0.518	1.1	0.581	5	1.1	0.5	0.5	1.1	1.8	7.2	$5,000,000	95.1	7.5710
Mike Scott	PF	26	ATL	68	0	3	0.444	1	3	0.344	2	4	0.517	0.9	0.792	2.9	1.1	0.4	0	0.6	1.2	7.8	$3,333,333	102.5	7.6098
Terrence Jones	PF	23	HOU	33	24	4.8	0.528	0.4	1.2	0.351	4.4	8.8	0.553	1.7	0.606	6.7	1.1	0.5	1.8	1.1	2.4	11.7	$1,618,680	103.9	11.2608
Drew Gooden	PF	33	WAS	51	7	2.2	0.399	0.5	1.5	0.39	1.7	3.4	0.401	0.7	0.773	4.4	1	0.4	0.2	0.5	1.8	5.4	$915,243	98.5	5.4822
Jason Thompson	PF	28	SAC	81	63	2.5	0.47	0	0	0	2.5	5	0.471	1.1	0.622	6.5	1	0.4	0.7	1	3	6.1	$6,037,500	101.3	6.0217
Ersan Ilyasova	PF	27	MIL	58	36	4.5	0.472	1.3	3.9	0.389	3.3	6.6	0.515	1.2	0.645	4.8	1	0.6	0.3	0.8	2.6	11.5	$7,900,000	97.8	11.7587
Darrell Arthur	PF	26	DEN	58	4	2.7	0.404	0.4	1.2	0.236	2.2	4.4	0.471	0.8	0.78	2.9	1	0.8	0.4	0.8	2.5	6.6	$3,457,149	101.5	6.5025
Quincy Acy	PF	24	NYK	68	22	2.2	0.459	0.3	0.9	0.3	2	4	0.494	1.1	0.784	4.4	1	0.4	0.3	0.9	2.2	5.9	$915,243	91.9	6.4200
Kris Humphries	PF	29	WAS	64	17	3.3	0.473	0	0	0	3.3	6.6	0.481	1.4	0.744	6.5	0.9	0.5	0.4	0.7	2	8	$4,250,000	98.5	8.1218
Jeff Adrien	PF	28	MIN	17	0	1.1	0.432	0	0		1.1	2.2	0.432	1.3	0.579	4.5	0.9	0.2	0.5	0.5	1.8	3.5	$742,962	97.8	3.5787
Ryan Anderson	PF	26	NOP	61	5	4.8	0.399	2	6	0.34	2.8	5.6	0.457	2.1	0.854	4.8	0.9	0.5	0.3	1	1.9	13.7	$8,491,500	99.4	13.7827
Serge Ibaka	PF	25	OKC	64	64	5.8	0.476	1.2	3.6	0.376	4.6	9.2	0.511	1.4	0.836	7.8	0.9	0.5	2.4	1.5	3	14.3	$12,350,000	104	13.7500
Pero Antic	PF	32	ATL	63	3	1.7	0.365	0.8	2.4	0.301	0.9	1.8	0.455	1.4	0.715	3	0.8	0.3	0.2	0.8	2.1	5.7	$1,250,000	102.5	5.5610

Player	Pos	Age	Tm	G	GS	FG	FG%	3P	3P*3	3P%	2P	2P*2	2P%	FT	FT%	TRB	AST	STL	BLK	TO	PF	PPG	Salary	Team PPG	ppg%tp
Dante Cunningham	PF	27	NOP	66	27	2.3	0.457	0	0	0.1	2.3	4.6	0.468	0.4	0.617	3.9	0.8	0.7	0.6	0.5	1.5	5.2	$716,043	99.4	5.2314
Anthony Bennett	PF	21	MIN	57	3	2.2	0.421	0.1	0.3	0.304	2.1	4.2	0.431	0.7	0.641	3.8	0.8	0.5	0.3	0.6	1.5	5.2	$5,563,920	97.8	5.3170
Reggie Evans	PF	34	SAC	47	7	1.2	0.423	0	0	0	1.2	2.4	0.43	1.3	0.619	6.4	0.7	0.5	0.1	1	1.9	3.7	$1,768,653	101.3	3.6525
Udonis Haslem	PF	34	MIA	62	25	1.7	0.448	0	0	0.2	1.7	3.4	0.459	0.7	0.703	4.2	0.7	0.3	0.2	0.7	1.8	4.2	$2,732,000	94.7	4.4351
Matt Bonner	PF	34	SAS	72	19	1.3	0.409	0.6	1.8	0.365	0.7	1.4	0.462	0.4	0.811	1.6	0.7	0.1	0.2	0.2	0.9	3.7	$915,243	103.2	3.5853
Jon Leuer	PF	25	MEM	63	6	1.9	0.443	0.1	0.3	0.241	1.8	3.6	0.467	0.6	0.627	3.3	0.7	0.3	0.1	0.6	1.4	4.5	$967,500	98.3	4.5778
Furkan Aldemir	PF	23	PHI	41	9	1	0.513	0	0	0	1	2	0.548	0.3	0.481	4.3	0.7	0.4	0.4	0.4	2.3	2.3	$2,963,232	92	2.5000
Derrick Williams	PF	23	SAC	74	6	2.9	0.447	0.7	2.1	0.314	2.3	4.6	0.511	1.8	0.684	2.7	0.7	0.5	0.1	0.8	0.9	8.3	$6,331,404	101.3	8.1935
Aaron Gordon	PF	19	ORL	47	8	2	0.447	0.3	0.9	0.271	1.7	3.4	0.5	0.9	0.721	3.6	0.7	0.4	0.5	0.8	1.8	5.2	$3,992,040	95.7	5.4336
Andrew Nicholson	PF	25	ORL	40	3	2.1	0.437	0.3	0.9	0.317	1.8	3.6	0.47	0.4	0.6	2.1	0.6	0.2	0.3	0.6	1.3	4.9	$1,545,840	95.7	5.1202
Mike Muscala	PF	23	ATL	40	8	2.1	0.55	0.2	0.6	0.409	1.9	3.8	0.574	0.6	0.88	3	0.6	0.4	0.5	0.5	1.3	4.9	$816,482	102.5	4.7805
Glen Davis	PF	29	LAC	74	0	1.6	0.459	0	0	0	1.6	3.2	0.466	0.8	0.632	2.3	0.5	0.6	0.3	0.5	1.8	4	$7,515,243	106.7	3.7488
Tristan Thompson	PF	23	CLE	82	15	3.3	0.547	0	0		3.3	6.6	0.547	1.9	0.641	8	0.5	0.4	0.7	1	2.3	8.5	$5,138,430	103.1	8.2444
Johnny O'Bryant	PF	21	MIL	34	15	1.3	0.367	0	0		1.3	2.6	0.367	0.4	0.444	1.9	0.5	0.1	0.1	0.7	1.3	2.9	$600,000	97.8	2.9652
Joey Dorsey	PF	31	HOU	69	17	1.2	0.552	0	0	0	1.2	2.4	0.556	0.3	0.289	4	0.4	0.6	0.4	0.6	2.3	2.7	$948,163	103.9	2.5987
Carl Landry	PF	31	SAC	70	15	2.7	0.515	0	0		2.7	5.4	0.515	1.8	0.82	3.8	0.4	0.2	0.2	0.8	1.9	7.2	$6,500,000	101.3	7.1076
Malcolm Thomas	PF	26	PHI	17	0	1.1	0.45	0	0	0	1.1	2.2	0.545	0.5	0.692	3.3	0.4	0.2	0.1	0.8	1	2.6	$948,163	92	2.8261
Mitch McGary	PF	22	OKC	32	2	2.8	0.533	0	0	0	2.8	5.6	0.54	0.8	0.625	5.2	0.4	0.5	0.5	1	2.3	6.3	$1,400,040	104	6.0577
Jason Maxiell	PF	31	CHO	61	0	1.3	0.422	0	0		1.3	2.6	0.422	0.7	0.577	3.3	0.3	0.3	0.7	0.5	1.6	3.3	$915,243	94.2	3.5032
Charlie Villanueva	PF	30	DAL	64	1	2.3	0.414	1.3	3.9	0.376	1	2	0.475	0.3	0.571	2.3	0.3	0.2	0.3	0.4	1	6.3	$915,243	105.2	5.9886
Tyler Hansbrough	PF	29	TOR	74	8	1.2	0.521	0	0	0.143	1.1	2.2	0.538	1.3	0.698	3.6	0.3	0.4	0.2	0.3	1.9	3.6	$3,326,235	104	3.4615
Jeff Ayres	PF	27	SAS	51	0	1.1	0.579	0	0		1.1	2.2	0.579	0.5	0.75	2.3	0.3	0.2	0.2	0.5	1.2	2.7	$1,828,750	103.2	2.6163
Player	**Pos**	**Age**	**Tm**	**G**	**GS**	**FG**	**FG%**	**3P**	**3P*3**	**3P%**	**2P**	**2P*2**	**2P%**	**FT**	**FT%**	**TRB**	**AST**	**STL**	**BLK**	**TO**	**PF**	**PPG**	**Salary**	**Team PPG**	**ppg%tp**
Cory Jefferson	PF	24	BRK	50	1	1.5	0.449	0	0	0.133	1.5	3	0.48	0.6	0.574	2.9	0.3	0.2	0.4	0.4	1.2	3.7	$507,336	98	3.7755
Shayne Whittington	PF	23	IND	20	0	1	0.452	0.1	0.3	0.167	0.9	1.8	0.5	0.9	0.783	1.5	0.3	0.3	0.1	0.3	0.9	2.9	$507,336	97.3	2.9805
Ekpe Udoh	PF	27	LAC	33	0	0.3	0.458	0	0		0.3	0.6	0.458	0.2	0.778	0.8	0.2	0.2	0.2	0.2	0.6	0.9	$915,243	106.7	0.8435
Drew Gordon	PF	24	PHI	9	0	0.9	0.421	0	0	0	0.9	1.8	0.5	0.1	0.5	2	0.2	0.1	0	0.9	0.8	1.9	$40,000	92	2.0652
Greg Smith	PF	24	DAL	42	2	0.7	0.612	0	0		0.7	1.4	0.612	0.5	0.513	1.9	0.2	0.2	0.3	0.3	1.3	1.9	$948,163	105.2	1.8061
Jarnell Stokes	PF	21	MEM	19	2	1.1	0.568	0	0		1.1	2.2	0.568	0.8	0.536	1.8	0.2	0.3	0.3	0.4	1.3	3	$725,000	98.3	3.0519
Noah Vonleh	PF	19	CHO	25	0	1.2	0.395	0.2	0.6	0.385	1	2	0.397	0.7	0.692	3.4	0.2	0.2	0.4	0.4	0.8	3.3	$2,524,200	94.2	3.5032
DeJuan Blair	PF	25	WAS	29	0	0.9	0.456	0	0		0.9	1.8	0.456	0.1	0.667	1.9	0.1	0.2	0	0.4	1.4	1.9	$2,000,000	98.5	1.9289
Cameron Bairstow	PF	24	CHI	18	1	0.2	0.214	0	0		0.2	0.4	0.214	0.2	0.8	0.4	0.1	0.1	0.1	0.2	0.4	0.6	$507,336	94.2	0.6369
Jack Cooley	PF	23	UTA	16	0	0.6	0.409	0	0		0.6	1.2	0.409	0.6	0.429	1.6	0.1	0.4	0.2	0.2	1.4	1.7	$65,000	95.1	1.7876
James Michael McAdoo	PF	22	GSW	15	0	1.6	0.545	0	0		1.6	3.2	0.545	0.9	0.56	2.5	0.1	0.3	0.6	0.4	1.4	4.1	$35,000	110	3.7273
Jerrelle Benimon	PF	23	UTA	2	0	0		0	0		0	0		0		1.5	0	0	0	0.5	0	0	$35,000	95.1	0.0000
Eric Moreland	PF	23	SAC	3	0	0.3	1	0	0		0.3	0.6	1	0		0.3	0	0	0	0	0.3	0.7	$507,336	101.3	0.6910
Julius Randle	PF	20	LAL	1	0	1	0.333	0	0		1	2	0.333	0	0	0	0	0	0	1	1	2	$2,997,360	98.5	2.0305
Chris Paul	PG	29	LAC	82	82	6.9	0.485	1.7	5.1	0.398	5.2	10.4	0.523	3.5	0.9	4.6	10.2	1.9	0.2	2.3	2.5	19.1	$20,068,563	106.7	17.9007
John Wall	PG	24	WAS	79	79	6.6	0.445	0.8	2.4	0.3	5.7	11.4	0.478	3.6	0.785	4.6	10	1.7	0.6	3.8	2.3	17.6	$13,701,250	98.5	17.8680
Ty Lawson	PG	27	DEN	75	75	5.4	0.436	0.9	2.7	0.341	4.5	9	0.463	3.5	0.73	3.1	9.6	1.2	0.1	2.5	1.7	15.2	$11,595,506	101.5	14.9754
Ricky Rubio	PG	24	MIN	22	22	3.5	0.356	0.6	1.8	0.255	3	6	0.387	2.6	0.803	5.7	8.8	1.7	0	2.9	2.7	10.3	$5,070,686	97.8	10.5317
Russell Westbrook	PG	26	OKC	67	67	9.4	0.426	1.3	3.9	0.299	8.1	16.2	0.457	8.1	0.835	7.3	8.6	2.1	0.2	4.4	2.7	28.1	$15,719,063	104	27.0192
Stephen Curry	PG	26	GSW	80	80	8.2	0.487	3.6	10.8	0.443	4.6	9.2	0.528	3.9	0.914	4.3	7.7	2	0.2	3.1	2	23.8	$10,629,213	110	21.6364

Player	Pos	Age	Tm	G	GS	FG	FG%	3P	3P*3	3P%	2P	2P*2	2P%	FT	FT%	TRB	AST	STL	BLK	TO	PF	PPG	Salary	Team PPG	ppg%tp
Jeff Teague	PG	26	ATL	73	72	5.6	0.46	1	3	0.343	4.6	9.2	0.496	3.8	0.862	2.5	7	1.7	0.4	2.8	1.9	15.9	$8,000,000	102.5	15.5122
Jrue Holiday	PG	24	NOP	40	37	6	0.446	1.3	3.9	0.378	4.7	9.4	0.469	1.6	0.855	3.4	6.9	1.6	0.6	2.3	2.8	14.8	$9,904,495	99.4	14.8893
Kyle Lowry	PG	28	TOR	70	70	6.1	0.412	1.9	5.7	0.338	4.3	8.6	0.457	3.6	0.808	4.7	6.8	1.6	0.2	2.5	3	17.8	$12,000,000	104	17.1154
Deron Williams	PG	30	BRK	68	55	4.4	0.387	1.3	3.9	0.367	3.1	6.2	0.395	3	0.834	3.5	6.6	0.9	0.3	2.3	2.3	13	$19,754,465	98	13.2653
Brandon Jennings	PG	25	DET	41	41	5.3	0.401	1.9	5.7	0.36	3.5	7	0.428	2.9	0.839	2.5	6.6	1.1	0.1	2.2	1.6	15.4	$8,000,000	98.5	15.6345
Elfrid Payton	PG	20	ORL	82	63	3.7	0.425	0.1	0.3	0.262	3.5	7	0.435	1.4	0.551	4.3	6.5	1.7	0.2	2.5	2.4	8.9	$2,397,840	95.7	9.2999
Damian Lillard	PG	24	POR	82	82	7.2	0.434	2.4	7.2	0.343	4.8	9.6	0.5	4.2	0.864	4.6	6.2	1.2	0.3	2.7	2	21	$3,340,920	102.8	20.4280
Eric Bledsoe	PG	25	PHO	81	81	5.8	0.447	1.1	3.3	0.324	4.7	9.4	0.491	4.4	0.8	5.2	6.1	1.6	0.6	3.4	2.3	17	$13,000,000	102.4	16.6016
Darren Collison	PG	27	SAC	45	45	5.8	0.473	1.3	3.9	0.373	4.4	8.8	0.514	3.2	0.788	3.2	5.6	1.5	0.3	2.5	2.1	16.1	$4,797,664	101.3	15.8934
Mike Conley	PG	27	MEM	70	70	5.6	0.446	1.5	4.5	0.386	4.1	8.2	0.473	3.1	0.859	3	5.4	1.3	0.2	2.2	2	15.8	$8,694,216	98.3	16.0732
Kyrie Irving	PG	22	CLE	75	75	7.7	0.468	2.1	6.3	0.415	5.6	11.2	0.491	4.2	0.863	3.2	5.2	1.5	0.3	2.5	1.9	21.7	$7,070,730	103.1	21.0475
Tony Wroten	PG	21	PHI	30	15	5.8	0.403	1.2	3.6	0.261	4.6	9.2	0.473	4	0.667	2.9	5.2	1.6	0.3	3.8	2.4	16.9	$1,210,080	92	18.3696
George Hill	PG	28	IND	43	36	5.9	0.477	1.6	4.8	0.358	4.3	8.6	0.544	2.6	0.79	4.2	5.1	1	0.3	1.6	2.6	16.1	$8,000,000	97.3	16.5468
Kemba Walker	PG	24	CHO	62	58	6.1	0.385	1.4	4.2	0.304	4.7	9.4	0.418	3.8	0.827	3.5	5.1	1.4	0.5	1.6	1.5	17.3	$3,272,091	94.2	18.3652
Tony Parker	PG	32	SAS	68	68	5.9	0.486	0.6	1.8	0.427	5.4	10.8	0.493	1.9	0.783	1.9	4.9	0.6	0	2.1	1.6	14.4	$12,500,000	103.2	13.9535
Derrick Rose	PG	26	CHI	51	51	6.6	0.405	1.5	4.5	0.28	5.1	10.2	0.465	3	0.813	3.2	4.9	0.7	0.3	3.2	1.2	17.7	$18,862,876	94.2	18.7898
Jose Calderon	PG	33	NYK	42	42	3.5	0.415	1.4	4.2	0.415	2.1	4.2	0.415	0.7	0.906	3	4.7	0.7	0	1.8	1.8	9.1	$7,097,191	91.9	9.9021
Jarrett Jack	PG	31	BRK	80	27	4.5	0.439	0.5	1.5	0.267	4	8	0.477	2.5	0.881	3.1	4.7	0.9	0.2	2.4	1.8	12	$6,300,000	98	12.2449
Jeremy Lin	PG	26	LAL	74	30	3.7	0.424	0.9	2.7	0.369	2.9	5.8	0.444	2.9	0.795	2.6	4.6	1.1	0.4	2.2	2.6	11.2	$8,374,646	98.5	11.3706
Trey Burke	PG	22	UTA	76	43	4.9	0.368	1.6	4.8	0.318	3.2	6.4	0.4	1.4	0.752	2.7	4.3	0.9	0.2	1.6	1.6	12.8	$2,548,560	95.1	13.4595
Dennis Schroder	PG	21	ATL	77	10	3.7	0.427	0.7	2.1	0.351	3	6	0.449	1.9	0.827	2.1	4.1	0.6	0.1	1.9	1.6	10	$1,690,680	102.5	9.7561
Player	**Pos**	**Age**	**Tm**	**G**	**GS**	**FG**	**FG%**	**3P**	**3P*3**	**3P%**	**2P**	**2P*2**	**2P%**	**FT**	**FT%**	**TRB**	**AST**	**STL**	**BLK**	**TO**	**PF**	**PPG**	**Salary**	**Team PPG**	**ppg%tp**
Ronnie Price	PG	31	LAL	43	20	1.8	0.345	0.7	2.1	0.284	1.1	2.2	0.404	0.8	0.8	1.6	3.8	1.6	0.1	1.2	2.7	5.1	$915,243	98.5	5.1777
Mario Chalmers	PG	28	MIA	80	37	3.3	0.403	0.9	2.7	0.294	2.4	4.8	0.467	2.6	0.774	2.6	3.8	1.5	0.1	2.2	3.1	10.2	$4,000,000	94.7	10.7709
Greivis Vasquez	PG	28	TOR	82	29	3.6	0.408	1.6	4.8	0.379	2	4	0.436	0.6	0.758	2.6	3.7	0.6	0.1	1.5	2.2	9.5	$6,400,000	104	9.1346
Steve Blake	PG	34	POR	81	0	1.5	0.373	1	3	0.352	0.6	1.2	0.417	0.4	0.707	1.7	3.6	0.5	0.1	1.3	1.5	4.3	$2,077,000	102.8	4.1829
C.J. Watson	PG	30	IND	57	21	3.2	0.434	1.2	3.6	0.4	1.9	3.8	0.459	2.4	0.826	2.9	3.6	1	0.2	1.8	1.9	10	$2,077,000	97.3	10.2775
Donald Sloan	PG	27	IND	53	21	2.7	0.408	0.8	2.4	0.313	2	4	0.464	1.1	0.779	2.7	3.6	0.4	0	1.2	0.9	7.4	$948,163	97.3	7.6053
Zach LaVine	PG	19	MIN	77	40	3.7	0.422	0.7	2.1	0.341	3	6	0.449	1.9	0.842	2.8	3.6	0.7	0.1	2.5	2.1	10.1	$2,055,840	97.8	10.3272
Jordan Clarkson	PG	22	LAL	59	38	4.5	0.448	0.6	1.8	0.314	3.9	7.8	0.482	2.2	0.829	3.2	3.5	0.9	0.2	1.6	1.8	11.9	$507,336	98.5	12.0812
J.J. Barea	PG	30	DAL	77	10	2.9	0.42	0.7	2.1	0.323	2.2	4.4	0.463	1	0.809	1.7	3.4	0.4	0	0.9	1.4	7.5	$5,429,359	105.2	7.1293
Patrick Beverley	PG	26	HOU	56	55	3.6	0.383	2.1	6.3	0.356	1.6	3.2	0.426	0.8	0.75	4.2	3.4	1.1	0.4	1.5	3.3	10.1	$915,243	103.9	9.7209
Shaun Livingston	PG	29	GSW	78	2	2.5	0.5	0	0	0	2.5	5	0.503	0.8	0.714	2.3	3.3	0.6	0.3	1.3	1.4	5.9	$5,305,000	110	5.3636
Gal Mekel	PG	26	NOP	4	0	0.8	0.15	0	0	0	0.8	1.6	0.176	0		0.3	3.3	0.5	0	0.5	0.3	1.5	$816,482	99.4	1.5091
Langston Galloway	PG	23	NYK	45	41	4.5	0.399	1.4	4.2	0.352	3.2	6.4	0.424	1.4	0.808	4.2	3.3	1.2	0.3	1.4	2.9	11.8	$30,000	91.9	12.8400
Aaron Brooks	PG	30	CHI	82	21	4.2	0.421	1.5	4.5	0.387	2.7	5.4	0.442	1.8	0.833	2	3.2	0.7	0.2	1.9	2.3	11.6	$915,243	94.2	12.3142
Devin Harris	PG	31	DAL	76	3	2.9	0.418	1.3	3.9	0.357	1.6	3.2	0.481	1.7	0.815	1.8	3.1	1	0.2	1.1	1.9	8.8	$3,878,896	105.2	8.3650
Rodney Stuckey	PG	28	IND	71	36	4.7	0.44	0.8	2.4	0.39	3.9	7.8	0.452	2.5	0.819	3.5	3.1	0.8	0.1	1.7	1.8	12.6	$1,227,985	97.3	12.9496
Lorenzo Brown	PG	24	MIN	29	7	1.8	0.426	0.2	0.6	0.214	1.6	3.2	0.489	0.4	0.632	2.4	3.1	1	0.2	1	1.7	4.2	$283,367	97.8	4.2945
Kendall Marshall	PG	23	MIL	28	3	1.6	0.455	0.6	1.8	0.391	1	2	0.509	0.3	0.889	1	3.1	0.8	0	1.3	0.9	4.2	$915,243	97.8	4.2945
Spencer Dinwiddie	PG	21	DET	34	1	1.5	0.302	0.4	1.2	0.185	1.1	2.2	0.375	0.9	0.912	1.4	3.1	0.6	0.2	1	1.6	4.3	$700,000	98.5	4.3655
Marcus Smart	PG	20	BOS	67	38	2.6	0.367	1.4	4.2	0.335	1.3	2.6	0.41	1.2	0.646	3.3	3.1	1.5	0.3	1.3	2.6	7.8	$3,283,320	101.4	7.6923
Jerryd Bayless	PG	26	MIL	77	4	2.9	0.426	0.5	1.5	0.308	2.4	4.8	0.462	1.6	0.883	2.7	3	0.8	0.2	1.7	2.1	7.8	$3,000,000	97.8	7.9755

Player	Pos	Age	Tm	G	GS	FG	FG%	3P	3P*3	3P%	2P	2P*2	2P%	FT	FT%	TRB	AST	STL	BLK	TO	PF	PPG	Salary	Team PPG	ppg%tp
Shane Larkin	PG	22	NYK	76	22	2.5	0.433	0.5	1.5	0.302	2	4	0.481	0.8	0.782	2.3	3	1.2	0.1	1.1	2	6.2	$1,606,080	91.9	6.7465
Beno Udrih	PG	32	MEM	79	12	3.1	0.487	0.3	0.9	0.268	2.8	5.6	0.539	1.1	0.853	1.8	2.8	0.6	0.1	1.1	1.1	7.7	$2,077,000	98.3	7.8332
Sebastian Telfair	PG	29	OKC	16	1	2.9	0.368	1.1	3.3	0.3	1.8	3.6	0.431	1.5	0.706	1.9	2.8	0.6	0	1.1	1.8	8.4	$915,243	104	8.0769
Shelvin Mack	PG	24	ATL	55	0	2.1	0.401	0.7	2.1	0.315	1.4	2.8	0.467	0.5	0.806	1.4	2.8	0.5	0	0.9	0.6	5.4	$2,433,333	102.5	5.2683
Ray McCallum	PG	23	SAC	68	30	3	0.438	0.5	1.5	0.306	2.5	5	0.479	0.8	0.679	2.6	2.8	0.7	0.2	1.3	1.4	7.4	$816,482	101.3	7.3050
Shabazz Napier	PG	23	MIA	51	10	1.7	0.382	0.8	2.4	0.364	0.9	1.8	0.4	0.9	0.786	2.2	2.5	0.8	0.1	1.6	1.5	5.1	$1,238,640	94.7	5.3854
Cory Joseph	PG	23	SAS	79	14	2.6	0.504	0.2	0.6	0.364	2.4	4.8	0.52	1.3	0.734	2.4	2.4	0.6	0.2	0.8	1.3	6.8	$1,023,261	103.2	6.5891
Dante Exum	PG	19	UTA	82	41	1.8	0.349	1	3	0.314	0.8	1.6	0.408	0.2	0.625	1.6	2.4	0.5	0.2	1.4	1.8	4.8	$3,615,000	95.1	5.0473
Brian Roberts	PG	29	CHO	72	10	2.4	0.389	0.9	2.7	0.321	1.6	3.2	0.44	0.9	0.892	1.5	2.3	0.5	0.1	0.8	1	6.7	$2,732,000	94.2	7.1125
Phil Pressey	PG	23	BOS	50	0	1.3	0.368	0.3	0.9	0.246	1	2	0.44	0.7	0.673	1.6	2.3	0.6	0.1	0.8	1.1	3.5	$816,482	101.4	3.4517
Luke Ridnour	PG	33	ORL	47	0	1.6	0.426	0.4	1.2	0.317	1.2	2.4	0.487	0.4	0.857	1.4	2	0.4	0.1	0.8	1.4	4	$2,750,000	95.7	4.1797
Jordan Farmar	PG	28	LAC	36	0	1.7	0.386	1	3	0.361	0.7	1.4	0.426	0.3	0.909	1.2	1.9	0.6	0.1	0.9	1.4	4.6	$2,077,000	106.7	4.3112
Patrick Mills	PG	26	SAS	51	0	2.5	0.381	1.2	3.6	0.341	1.3	2.6	0.429	0.6	0.825	1.5	1.7	0.5	0	0.7	1.1	6.9	$3,842,105	103.2	6.6860
Raymond Felton	PG	30	DAL	29	3	1.5	0.406	0.3	0.9	0.294	1.1	2.2	0.458	0.4	0.8	0.9	1.4	0.4	0.1	0.6	0.6	3.7	$3,793,693	105.2	3.5171
Darius Morris	PG	24	BRK	38	0	0.9	0.34	0.2	0.6	0.212	0.8	1.6	0.397	0.1	0.444	0.7	1.3	0.2	0	0.6	0.5	2.2	$702,756	98	2.2449
Jimmer Fredette	PG	25	NOP	50	0	1.3	0.38	0.2	0.6	0.188	1.1	2.2	0.458	0.9	0.956	0.8	1.2	0.3	0	0.7	0.9	3.6	$915,243	99.4	3.6217
Will Cherry	PG	23	CLE	8	0	0.6	0.263	0.3	0.9	0.222	0.4	0.8	0.3	0.4	0.5	0.6	1	0.8	0.1	0.5	1.3	1.9	$125,000	103.1	1.8429
Bryce Cotton	PG	22	UTA	15	0	1.9	0.42	0.5	1.5	0.35	1.5	3	0.449	1	0.833	1.2	1	0.3	0	0.8	0.4	5.3	$50,000	95.1	5.5731
Jannero Pargo	PG	35	CHO	9	0	1.7	0.429	1	3	0.409	0.7	1.4	0.462	0.2	1	0.3	0.9	0	0	0.3	0.9	4.6	$915,243	94.2	4.8832
Erick Green	PG	23	DEN	43	1	1.4	0.377	0.3	0.9	0.298	1.1	2.2	0.409	0.2	0.833	0.7	0.9	0.3	0	0.3	0.6	3.4	$507,336	101.5	3.3498
Seth Curry	PG	24	PHO	2	0	0	0	0	0	0	0	0	0	0		1	0.5	0	0	0	1	0	$150,000	102.4	0.0000
Player	**Pos**	**Age**	**Tm**	**G**	**GS**	**FG**	**FG%**	**3P**	**3P*3**	**3P%**	**2P**	**2P*2**	**2P%**	**FT**	**FT%**	**TRB**	**AST**	**STL**	**BLK**	**TO**	**PF**	**PPG**	**Salary**	**Team PPG**	**ppg%tp**
Kalin Lucas	PG	25	MEM	1	0	0	0	0	0		0	0	0	0		0	0	1	0	0	1	0	$50,000	98.3	0.0000
LeBron James	SF	30	CLE	69	69	9	0.488	1.7	5.1	0.354	7.3	14.6	0.536	5.4	0.71	6	7.4	1.6	0.7	3.9	2	25.3	$20,644,400	103.1	24.5393
Nicolas Batum	SF	26	POR	71	71	3.4	0.4	1.4	4.2	0.324	2	4	0.481	1.2	0.857	5.9	4.8	1.1	0.6	1.9	1.5	9.4	$11,765,500	102.8	9.1440
Kevin Durant	SF	26	OKC	27	27	8.8	0.51	2.4	7.2	0.403	6.4	12.8	0.565	5.4	0.854	6.6	4.1	0.9	0.9	2.7	1.5	25.4	$17,888,932	104	24.4231
Gordon Hayward	SF	24	UTA	76	76	6.4	0.445	1.6	4.8	0.364	4.8	9.6	0.481	4.9	0.812	4.9	4.1	1.4	0.4	2.7	1.7	19.3	$14,746,000	95.1	20.2944
Rudy Gay	SF	28	SAC	68	67	7.5	0.455	1.2	3.6	0.359	6.3	12.6	0.479	5	0.858	5.9	3.7	1	0.6	2.7	2.3	21.1	$19,317,326	101.3	20.8292
Carmelo Anthony	SF	30	NYK	40	40	9	0.444	1.5	4.5	0.341	7.4	14.8	0.474	4.7	0.797	6.6	3.1	1	0.4	2.2	2.2	24.2	$22,458,401	91.9	26.3330
Andre Iguodala	SF	31	GSW	77	0	3	0.466	1	3	0.349	2	4	0.553	0.9	0.596	3.3	3	1.2	0.3	1.1	1.3	7.8	$12,289,544	110	7.0909
Trevor Ariza	SF	29	HOU	82	82	4.5	0.402	2.4	7.2	0.35	2.1	4.2	0.485	1.5	0.853	5.6	2.5	1.9	0.2	1.7	2.3	12.8	$8,579,089	103.9	12.3195
Kawhi Leonard	SF	23	SAS	64	64	6.2	0.479	1	3	0.349	5.1	10.2	0.519	3.2	0.802	7.2	2.5	2.3	0.8	1.5	2	16.5	$2,894,059	103.2	15.9884
Chandler Parsons	SF	26	DAL	66	66	5.8	0.462	2	6	0.38	3.8	7.6	0.521	2.1	0.72	4.9	2.4	1	0.3	1.5	2.1	15.7	$14,700,000	105.2	14.9240
Joe Ingles	SF	27	UTA	79	32	1.9	0.415	0.9	2.7	0.356	0.9	1.8	0.493	0.4	0.75	2.2	2.3	0.9	0.1	1.2	1.6	5	$507,336	95.1	5.2576
Khris Middleton	SF	23	MIL	79	58	5.1	0.467	1.4	4.2	0.407	3.8	7.6	0.494	1.7	0.859	4.4	2.3	1.5	0.1	1.4	2.3	13.4	$915,243	97.8	13.7014
Solomon Hill	SF	23	IND	82	78	3.1	0.396	0.8	2.4	0.327	2.2	4.4	0.43	1.9	0.824	3.8	2.2	0.8	0.2	1.4	2.2	8.9	$1,302,840	97.3	9.1470
Andrew Wiggins	SF	19	MIN	82	82	6.1	0.437	0.5	1.5	0.31	5.6	11.2	0.453	4.3	0.76	4.6	2.1	1	0.6	2.2	2.3	16.9	$5,510,640	97.8	17.2802
Paul Pierce	SF	37	WAS	73	73	4	0.447	1.6	4.8	0.389	2.4	4.8	0.496	2.2	0.781	4	2	0.6	0.3	1.3	2.2	11.9	$5,305,000	98.5	12.0812
Kostas Papanikolaou	SF	24	HOU	43	1	1.6	0.35	0.7	2.1	0.292	0.9	1.8	0.418	0.3	0.722	2.7	2	0.7	0.3	1.3	1.7	4.2	$4,797,664	103.9	4.0423
Luol Deng	SF	29	MIA	72	72	5.1	0.469	1.1	3.3	0.355	4	8	0.513	2.7	0.761	5.2	1.9	0.9	0.3	1.5	1.5	14	$9,714,461	94.7	14.7835
Mike Dunleavy	SF	34	CHI	63	63	3.3	0.435	1.7	5.1	0.407	1.6	3.2	0.468	1.1	0.805	3.9	1.8	0.6	0.3	1	2	9.4	$3,326,235	94.2	9.9788
Tobias Harris	SF	22	ORL	68	63	6.5	0.466	1.3	3.9	0.364	5.2	10.4	0.5	2.8	0.788	6.3	1.8	1	0.5	1.7	2	17.1	$2,380,594	95.7	17.8683
DeMarre Carroll	SF	28	ATL	70	69	4.5	0.487	1.7	5.1	0.395	2.8	5.6	0.567	1.8	0.702	5.3	1.7	1.3	0.2	1.1	2.2	12.6	$2,442,455	102.5	12.2927

Player	Pos	Age	Tm	G	GS	FG	FG%	3P	3P*3	3P%	2P	2P*2	2P%	FT	FT%	TRB	AST	STL	BLK	TO	PF	PPG	Salary	Team PPG	ppg%tp
Wilson Chandler	SF	27	DEN	78	75	5.4	0.429	1.8	5.4	0.342	3.6	7.2	0.49	1.4	0.775	6.1	1.7	0.7	0.4	1.4	3	13.9	$6,757,913	101.5	13.6946
Jabari Parker	SF	19	MIL	25	25	5.2	0.49	0.2	0.6	0.25	5	10	0.506	1.8	0.697	5.5	1.7	1.2	0.2	1.9	1.7	12.3	$4,930,560	97.8	12.5767
P.J. Tucker	SF	29	PHO	78	63	3.3	0.438	1.1	3.3	0.345	2.2	4.4	0.506	1.3	0.727	6.4	1.6	1.4	0.3	1.2	2.3	9.1	$5,700,000	102.4	8.8867
Wesley Johnson	SF	27	LAL	76	59	3.8	0.414	1.2	3.6	0.351	2.6	5.2	0.451	1.2	0.804	4.2	1.6	0.8	0.6	1.1	2.1	9.9	$915,243	98.5	10.0508
Matt Barnes	SF	34	LAC	76	74	3.6	0.444	1.8	5.4	0.362	1.9	3.8	0.569	1	0.779	4	1.5	0.9	0.7	1.1	3.2	10.1	$3,396,250	106.7	9.4658
Omri Casspi	SF	26	SAC	67	19	3.1	0.489	0.5	1.5	0.402	2.6	5.2	0.512	2.1	0.733	3.9	1.5	0.5	0.1	1.3	1.6	8.9	$915,243	101.3	8.7858
Robert Covington	SF	24	PHI	70	49	4.3	0.396	2.4	7.2	0.374	1.9	3.8	0.426	2.5	0.82	4.5	1.5	1.4	0.4	1.8	2.7	13.5	$1,125,000	92	14.6739
Thabo Sefolosha	SF	30	ATL	52	7	2	0.418	0.5	1.5	0.321	1.5	3	0.462	0.9	0.776	4.3	1.4	1	0.4	0.7	1.3	5.3	$4,150,000	102.5	5.1707
Danilo Gallinari	SF	26	DEN	59	27	3.9	0.401	1.8	5.4	0.355	2.1	4.2	0.453	2.9	0.895	3.7	1.4	0.8	0.3	1	1.6	12.4	$10,854,850	101.5	12.2167
Harrison Barnes	SF	22	GSW	82	82	3.9	0.482	1.1	3.3	0.405	2.8	5.6	0.519	1.3	0.72	5.5	1.4	0.7	0.2	0.9	1.8	10.1	$3,049,920	110	9.1818
Michael Kidd-Gilchrist	SF	21	CHO	55	52	4.1	0.465	0	0		4.1	8.2	0.465	2.7	0.701	7.6	1.4	0.5	0.7	1.1	2.1	10.9	$5,016,960	94.2	11.5711
Michael Beasley	SF	26	MIA	24	1	3.8	0.434	0.3	0.9	0.235	3.5	7	0.472	0.8	0.769	3.7	1.3	0.6	0.5	1.5	2.8	8.8	$777,778	94.7	9.2925
Jerami Grant	SF	20	PHI	65	11	1.9	0.352	0.8	2.4	0.314	1.2	2.4	0.383	1.8	0.591	3	1.2	0.6	1	1.3	2.2	6.3	$884,879	92	6.8478
Francisco Garcia	SF	33	HOU	14	0	1.2	0.27	0.7	2.1	0.222	0.5	1	0.389	0.1	0.25	1.2	1.1	0.6	0.4	0.7	1.4	3.2	$915,243	103.9	3.0799
Caron Butler	SF	34	DET	78	21	2.1	0.407	1.1	3.3	0.379	1	2	0.441	0.7	0.902	2.5	1	0.6	0.1	0.6	1.5	5.9	$4,500,000	98.5	5.9898
Chris Copeland	SF	30	IND	50	12	2.2	0.361	1	3	0.311	1.2	2.4	0.42	0.7	0.733	2.2	1	0.2	0.2	1.2	1.6	6.2	$3,135,000	97.3	6.3720
Chase Budinger	SF	26	MIN	67	4	2.5	0.433	0.8	2.4	0.364	1.7	3.4	0.474	1	0.827	3	1	0.7	0.1	0.7	1.1	6.8	$5,000,000	97.8	6.9530
Paul George	SF	24	IND	6	0	3	0.367	1.5	4.5	0.409	1.5	3	0.333	1.3	0.727	3.7	1	0.8	0.2	2	1.8	8.8	$13,701,250	97.3	9.0442
Terrence Ross	SF	23	TOR	82	61	3.8	0.41	1.8	5.4	0.372	2	4	0.452	0.5	0.786	2.8	1	0.6	0.3	0.8	1.7	9.8	$2,793,960	104	9.4231
JaKarr Sampson	SF	21	PHI	74	32	2	0.422	0.4	1.2	0.244	1.6	3.2	0.525	0.9	0.67	2.2	1	0.5	0.4	1	1.8	5.2	$507,336	92	5.6522
Shawn Marion	SF	36	CLE	57	24	2.1	0.446	0.2	0.6	0.261	1.9	3.8	0.484	0.5	0.765	3.5	0.9	0.5	0.5	0.6	1	4.8	$915,243	103.1	4.6557
Player	**Pos**	**Age**	**Tm**	**G**	**GS**	**FG**	**FG%**	**3P**	**3P*3**	**3P%**	**2P**	**2P*2**	**2P%**	**FT**	**FT%**	**TRB**	**AST**	**STL**	**BLK**	**TO**	**PF**	**PPG**	**Salary**	**Team PPG**	**ppg%tp**
Mike Miller	SF	34	CLE	52	15	0.7	0.325	0.6	1.8	0.327	0.1	0.2	0.313	0.1	0.75	1.8	0.9	0.3	0.1	0.4	1.4	2.1	$2,732,000	103.1	2.0369
Dorell Wright	SF	29	POR	48	2	1.5	0.379	0.9	2.7	0.38	0.6	1.2	0.378	0.7	0.81	2.3	0.9	0.4	0.2	0.4	1.1	4.6	$3,135,000	102.8	4.4747
Bojan Bogdanovic	SF	25	BRK	78	28	3.3	0.453	1.2	3.6	0.355	2.2	4.4	0.531	1.1	0.821	2.7	0.9	0.4	0.1	1	1.3	9	$3,278,000	98	9.1837
Cleanthony Early	SF	23	NYK	39	7	1.9	0.355	0.6	1.8	0.262	1.4	2.8	0.415	0.9	0.75	2.5	0.9	0.6	0.3	1	1.5	5.4	$507,336	91.9	5.8760
Tony Snell	SF	23	CHI	72	22	2.2	0.429	1	3	0.371	1.2	2.4	0.494	0.6	0.8	2.4	0.9	0.4	0.2	0.7	1.2	6	$1,472,400	94.2	6.3694
Otto Porter	SF	21	WAS	74	13	2.4	0.45	0.5	1.5	0.337	1.9	3.8	0.491	0.8	0.734	3	0.9	0.6	0.4	0.7	1.3	6	$4,470,480	98.5	6.0914
Rasual Butler	SF	35	WAS	75	1	2.9	0.422	1.2	3.6	0.387	1.7	3.4	0.45	0.7	0.791	2.6	0.8	0.4	0.3	0.6	1.3	7.7	$915,243	98.5	7.8173
Richard Jefferson	SF	34	DAL	74	18	1.9	0.444	0.9	2.7	0.426	1.1	2.2	0.462	1.1	0.684	2.5	0.8	0.4	0.1	0.7	1.6	5.8	$915,243	105.2	5.5133
Damjan Rudez	SF	28	IND	68	2	1.8	0.452	1	3	0.406	0.7	1.4	0.538	0.2	0.696	0.7	0.8	0.2	0.1	0.7	1.2	4.8	$1,100,000	97.3	4.9332
Jeffery Taylor	SF	25	CHO	29	13	1.6	0.395	0.4	1.2	0.306	1.2	2.4	0.436	0.9	0.634	1.8	0.8	0.4	0	0.7	1.2	4.4	$915,243	94.2	4.6709
Al-Farouq Aminu	SF	24	DAL	74	3	2	0.412	0.5	1.5	0.274	1.5	3	0.485	1.1	0.712	4.6	0.8	0.9	0.8	0.7	1.9	5.6	$981,084	105.2	5.3232
James Ennis	SF	24	MIA	62	3	1.6	0.409	0.5	1.5	0.326	1.1	2.2	0.461	1.3	0.84	2.8	0.8	0.4	0.3	0.6	1.4	5	$507,336	94.7	5.2798
Travis Wear	SF	24	NYK	51	1	1.6	0.402	0.2	0.6	0.367	1.4	2.8	0.408	0.4	0.769	2.1	0.8	0.3	0.2	0.7	1.1	3.9	$507,336	91.9	4.2437
Kyle Anderson	SF	21	SAS	33	8	0.9	0.348	0.1	0.3	0.273	0.8	1.6	0.359	0.3	0.643	2.2	0.8	0.5	0.2	0.3	0.8	2.2	$1,093,680	103.2	2.1318
Hedo Turkoglu	SF	35	LAC	62	2	1.3	0.441	1	3	0.432	0.4	0.8	0.469	0.1	0.545	1.6	0.6	0.3	0.1	0.5	1.1	3.7	$915,243	106.7	3.4677
John Salmons	SF	35	NOP	21	0	0.8	0.333	0.4	1.2	0.308	0.4	0.8	0.364	0.1	0.5	1	0.6	0.4	0.2	0.2	1.2	2	$3,000,000	99.4	2.0121
Danny Granger	SF	31	MIA	30	6	2.2	0.401	1	3	0.357	1.2	2.4	0.449	0.9	0.757	2.7	0.6	0.4	0.2	0.8	2.2	6.3	$2,077,000	94.7	6.6526
Landry Fields	SF	26	TOR	26	9	0.8	0.488	0	0	0.5	0.7	1.4	0.487	0.2	0.833	1	0.6	0.4	0	0.3	0.3	1.8	$6,250,000	104	1.7308
Robbie Hummel	SF	25	MIN	45	4	1.8	0.459	0.4	1.2	0.314	1.4	2.8	0.521	0.5	0.828	3	0.6	0.4	0.2	0.4	1.9	4.4	$880,000	97.8	4.4990
Maurice Harkless	SF	21	ORL	45	4	1.4	0.399	0.2	0.6	0.179	1.2	2.4	0.52	0.5	0.537	2.4	0.6	0.7	0.2	0.6	1.5	3.5	$2,771,340	95.7	3.6573
T.J. Warren	SF	21	PHO	40	1	2.8	0.528	0.1	0.3	0.238	2.7	5.4	0.56	0.4	0.737	2.1	0.6	0.5	0.2	0.7	1.3	6.1	$1,953,120	102.4	5.9570

Player	Pos	Age	Tm	G	GS	FG	FG%	3P	3P*3	3P%	2P	2P*2	2P%	FT	FT%	TRB	AST	STL	BLK	TO	PF	PPG	Salary	Team PPG	ppg%tp
Cartier Martin	SF	30	DET	23	0	0.7	0.283	0.3	0.9	0.182	0.4	0.8	0.45	0		0.9	0.5	0.1	0	0.2	0.9	1.6	$1,145,685	98.5	1.6244
Martell Webster	SF	28	WAS	32	0	0.8	0.264	0.3	0.9	0.233	0.4	0.8	0.292	1.5	0.75	1.4	0.5	0.2	0	0.7	0.8	3.3	$5,381,750	98.5	3.3503
Reggie Williams	SF	28	SAS	20	0	0.8	0.385	0.2	0.6	0.158	0.6	1.2	0.6	0.2	1	0.9	0.5	0.1	0	0.1	0.6	1.9	$125,104	103.2	1.8411
Jordan Hamilton	SF	24	LAC	14	2	1	0.4	0.7	2.1	0.476	0.3	0.6	0.286	0		1.1	0.5	0.1	0	0.3	1	2.7	$25,000	106.7	2.5305
James Jones	SF	34	CLE	57	2	1.3	0.368	1.1	3.3	0.36	0.2	0.4	0.423	0.7	0.848	1.1	0.4	0.2	0.1	0.2	1.2	4.4	$915,243	103.1	4.2677
Luke Babbitt	SF	25	NOP	63	19	1.5	0.479	0.9	2.7	0.513	0.5	1	0.429	0.2	0.684	1.8	0.4	0.3	0.2	0.4	1	4.1	$981,084	99.4	4.1247
Darius Miller	SF	24	NOP	5	1	0.2	0.143	0	0	0	0.2	0.4	0.167	0		0.2	0.4	0.2	0	0.2	1.6	0.4	$400,000	99.4	0.4024
Perry Jones	SF	23	OKC	43	13	1.7	0.397	0.3	0.9	0.233	1.4	2.8	0.476	0.6	0.649	1.8	0.4	0.4	0.2	0.6	1.2	4.3	$1,129,200	104	4.1346
Gerald Wallace	SF	32	BOS	32	0	0.4	0.412	0	0	0.333	0.4	0.8	0.419	0.2	0.4	1.8	0.3	0.5	0.1	0.6	0.6	1.1	$10,105,855	101.4	1.0848
Jeremy Evans	SF	27	UTA	38	0	0.8	0.552	0.1	0.3	0.4	0.8	1.6	0.566	0.6	0.828	1.9	0.3	0.3	0.3	0.1	0.8	2.4	$1,794,871	95.1	2.5237
Xavier Henry	SF	23	LAL	9	0	0.3	0.231	0	0		0.3	0.6	0.231	1.6	0.583	0.4	0.3	0.3	0	0.3	0.7	2.2	$1,082,000	98.5	2.2335
Doug McDermott	SF	23	CHI	36	0	1.2	0.402	0.4	1.2	0.317	0.8	1.6	0.455	0.3	0.667	1.2	0.2	0.1	0	0.5	0.8	3	$2,277,960	94.2	3.1847
Andrei Kirilenko	SF	33	BRK	7	0	0	0	0	0		0	0	0	0.4	0.75	1.1	0.1	0.1	0	0.1	0.1	0.4	$3,326,235	98	0.4082
Victor Claver	SF	26	POR	10	0	0.9	0.45	0.6	1.8	0.545	0.3	0.6	0.333	0		2	0.1	0.1	0.1	0.4	1.2	2.4	$1,370,000	102.8	2.3346
Bruno Caboclo	SF	19	TOR	8	0	0.5	0.333	0.3	0.9	0.667	0.3	0.6	0.222	0		0.3	0	0	0.1	0.5	0.4	1.3	$1,458,360	104	1.2500
James Harden	SG	25	HOU	81	81	8	0.44	2.6	7.8	0.375	5.4	10.8	0.48	8.8	0.868	5.7	7	1.9	0.7	4	2.6	27.4	$14,693,906	103.9	26.3715
Tyreke Evans	SG	25	NOP	79	76	6.6	0.447	0.9	2.7	0.304	5.7	11.4	0.482	2.6	0.694	5.3	6.6	1.3	0.5	3.1	2.5	16.6	$11,265,416	99.4	16.7002
Kobe Bryant	SG	36	LAL	35	35	7.6	0.373	1.5	4.5	0.293	6.1	12.2	0.401	5.6	0.813	5.7	5.6	1.3	0.2	3.7	1.9	22.3	$23,500,000	98.5	22.6396
Evan Turner	SG	26	BOS	82	57	3.9	0.429	0.4	1.2	0.277	3.5	7	0.458	1.3	0.752	5.1	5.5	1	0.2	2.4	2.2	9.5	$3,278,000	101.4	9.3688
Dwyane Wade	SG	33	MIA	62	62	8.2	0.47	0.5	1.5	0.284	7.7	15.4	0.489	4.6	0.768	3.5	4.8	1.2	0.3	3.4	1.7	21.5	$15,000,000	94.7	22.7033
Manu Ginobili	SG	37	SAS	70	0	3.6	0.426	1.3	3.9	0.345	2.3	4.6	0.489	2.1	0.721	3	4.2	1	0.3	2.2	2	10.5	$7,000,000	103.2	10.1744
Player	**Pos**	**Age**	**Tm**	**G**	**GS**	**FG**	**FG%**	**3P**	**3P*3**	**3P%**	**2P**	**2P*2**	**2P%**	**FT**	**FT%**	**TRB**	**AST**	**STL**	**BLK**	**TO**	**PF**	**PPG**	**Salary**	**Team PPG**	**ppg%tp**
Monta Ellis	SG	29	DAL	80	80	7.5	0.445	1	3	0.285	6.5	13	0.487	2.9	0.752	2.4	4.1	1.9	0.3	2.5	2.5	18.9	$8,360,000	105.2	17.9658
Victor Oladipo	SG	22	ORL	72	71	6.6	0.436	1.2	3.6	0.339	5.4	10.8	0.464	3.6	0.819	4.2	4.1	1.7	0.3	2.8	2.6	17.9	$4,978,200	95.7	18.7043
Lance Stephenson	SG	24	CHO	61	25	3.4	0.376	0.3	0.9	0.171	3.1	6.2	0.425	1.1	0.627	4.5	3.9	0.6	0.1	2.1	2.2	8.2	$9,000,000	94.2	8.7049
Eric Gordon	SG	26	NOP	61	60	4.7	0.411	2.3	6.9	0.448	2.4	4.8	0.38	1.8	0.805	2.6	3.8	0.8	0.2	2	2.4	13.4	$14,898,938	99.4	13.4809
Joe Johnson	SG	33	BRK	80	80	5.6	0.435	1.5	4.5	0.359	4.1	8.2	0.472	1.8	0.801	4.8	3.7	0.7	0.2	1.7	1.5	14.4	$23,180,790	98	14.6939
DeMar DeRozan	SG	25	TOR	60	60	6.8	0.413	0.4	1.2	0.284	6.4	12.8	0.426	6	0.832	4.6	3.5	1.2	0.2	2.3	2	20.1	$9,500,000	104	19.3269
Jimmy Butler	SG	25	CHI	65	65	6.5	0.462	1.1	3.3	0.378	5.4	10.8	0.484	5.9	0.834	5.8	3.3	1.8	0.6	1.4	1.7	20	$2,008,748	94.2	21.2314
Bradley Beal	SG	21	WAS	63	59	5.8	0.427	1.7	5.1	0.409	4.1	8.2	0.434	2.1	0.783	3.8	3.1	1.2	0.3	2	2.2	15.3	$4,505,280	98.5	15.5330
Matthew Dellavedova	SG	24	CLE	67	13	1.7	0.362	1	3	0.407	0.6	1.2	0.307	0.4	0.763	1.9	3	0.4	0	0.9	2.3	4.8	$816,482	103.1	4.6557
Alec Burks	SG	23	UTA	27	27	4.5	0.403	1	3	0.382	3.5	7	0.409	3.9	0.822	4.2	3	0.6	0.2	1.9	2.4	13.9	$3,034,356	95.1	14.6162
Klay Thompson	SG	24	GSW	77	77	7.8	0.463	3.1	9.3	0.439	4.7	9.4	0.481	2.9	0.879	3.2	2.9	1.1	0.8	1.9	1.6	21.7	$3,075,880	110	19.7273
O.J. Mayo	SG	27	MIL	71	15	4.2	0.422	1.4	4.2	0.357	2.8	5.6	0.465	1.6	0.827	2.6	2.8	0.8	0.3	1.8	2.3	11.4	$8,000,000	97.8	11.6564
Kyle Korver	SG	33	ATL	75	75	3.9	0.487	2.9	8.7	0.492	0.9	1.8	0.47	1.4	0.898	4.1	2.6	0.7	0.6	1.4	1.9	12.1	$6,253,521	102.5	11.8049
Will Bynum	SG	32	WAS	7	0	1.4	0.323	0	0	0	1.4	2.8	0.455	0.3	0.5	0.9	2.6	0.1	0.1	1.1	1.3	3.1	$2,916,000	98.5	3.1472
Gerald Henderson	SG	27	CHO	80	72	4.6	0.437	0.6	1.8	0.331	4.1	8.2	0.457	2.3	0.848	3.4	2.6	0.6	0.3	1.4	1.7	12.1	$6,000,000	94.2	12.8450
Giannis Antetokounmp	SG	20	MIL	81	71	4.7	0.491	0.1	0.3	0.159	4.6	9.2	0.511	3.2	0.741	6.7	2.6	0.9	1	2.1	3.1	12.7	$1,873,200	97.8	12.9857
Jamal Crawford	SG	34	LAC	64	4	5.2	0.396	1.9	5.7	0.327	3.3	6.6	0.448	3.5	0.901	1.9	2.5	0.9	0.2	1.4	1.7	15.8	$5,450,000	106.7	14.8079
Nick Calathes	SG	25	MEM	58	0	1.8	0.421	0.2	0.6	0.256	1.6	3.2	0.456	0.4	0.533	1.8	2.5	1.1	0.1	1.1	1.7	4.2	$816,482	98.3	4.2726
Randy Foye	SG	31	DEN	50	21	3	0.368	1.9	5.7	0.357	1.1	2.2	0.388	0.9	0.818	1.7	2.4	0.7	0.2	1.2	1.9	8.7	$3,000,000	101.5	8.5714
Kevin Martin	SG	31	MIN	39	36	6.8	0.427	1.9	5.7	0.393	4.9	9.8	0.442	4.4	0.881	3.6	2.3	0.8	0	1.9	1.9	20	$6,792,500	97.8	20.4499
Wesley Matthews	SG	28	POR	60	60	5.6	0.448	2.9	8.7	0.389	2.7	5.4	0.534	1.8	0.752	3.7	2.3	1.3	0.2	1.4	2.2	15.9	$7,245,640	102.8	15.4669

Player	Pos	Age	Tm	G	GS	FG	FG%	3P	3P*3	3P%	2P	2P*2	2P%	FT	FT%	TRB	AST	STL	BLK	TO	PF	PPG	Salary	Team PPG	ppg%tp
Kirk Hinrich	SG	34	CHI	66	22	2.2	0.373	0.9	2.7	0.345	1.3	2.6	0.394	0.5	0.7	1.8	2.2	0.7	0.2	1.1	2.2	5.7	$2,732,000	94.2	6.0510
Lou Williams	SG	28	TOR	80	0	4.7	0.404	1.9	5.7	0.34	2.8	5.6	0.464	4.3	0.861	1.9	2.1	1.1	0.1	1.3	1.3	15.5	$5,450,000	104	14.9038
Evan Fournier	SG	22	ORL	58	32	4.4	0.44	1.5	4.5	0.378	2.9	5.8	0.48	1.8	0.728	2.6	2.1	0.7	0	1.4	2	12	$1,483,920	95.7	12.5392
Jason Richardson	SG	34	PHI	19	15	3.3	0.348	1.6	4.8	0.323	1.6	3.2	0.378	0.9	0.773	3.5	2	0.7	0.2	0.8	1.5	9.1	$6,601,125	92	9.8913
Courtney Lee	SG	29	MEM	77	74	3.7	0.448	1.2	3.6	0.402	2.6	5.2	0.473	1.4	0.86	2.3	2	1	0.2	1	1.7	10.1	$5,450,000	98.3	10.2747
Danny Green	SG	27	SAS	81	80	4	0.436	2.4	7.2	0.418	1.6	3.2	0.466	1.4	0.874	4.2	2	1.2	1.1	1.1	2	11.7	$4,025,000	103.2	11.3372
Jason Terry	SG	37	HOU	77	18	2.4	0.422	1.6	4.8	0.39	0.8	1.6	0.504	0.5	0.813	1.6	1.9	0.9	0.2	1	1.5	7	$5,450,000	103.9	6.7372
J.J. Redick	SG	30	LAC	78	78	5.7	0.477	2.6	7.8	0.437	3.2	6.4	0.515	2.3	0.901	2.1	1.8	0.5	0.1	1.2	1.7	16.4	$6,792,500	106.7	15.3702
Jared Dudley	SG	29	MIL	72	22	2.8	0.468	1	3	0.385	1.7	3.4	0.537	0.7	0.716	3.1	1.8	1	0.2	0.9	1.6	7.2	$4,250,000	97.8	7.3620
Avery Bradley	SG	24	BOS	77	77	5.6	0.429	1.6	4.8	0.352	4	8	0.47	1	0.79	3.1	1.8	1.1	0.2	1.4	2.3	13.9	$7,191,011	101.4	13.7081
Tim Hardaway	SG	22	NYK	70	30	4	0.389	1.7	5.1	0.342	2.3	4.6	0.435	1.8	0.801	2.2	1.8	0.3	0.2	1.2	1.7	11.5	$1,250,640	91.9	12.5136
Rodney Hood	SG	22	UTA	50	21	3.1	0.414	1.2	3.6	0.365	1.9	3.8	0.456	1.2	0.763	2.3	1.7	0.6	0.2	0.9	2.4	8.7	$1,290,360	95.1	9.1483
Ben McLemore	SG	21	SAC	82	82	4.4	0.437	1.7	5.1	0.358	2.7	5.4	0.507	1.6	0.813	2.9	1.7	0.9	0.2	1.7	2.7	12.1	$3,026,280	101.3	11.9447
Wayne Ellington	SG	27	LAL	65	36	3.9	0.412	1.4	4.2	0.37	2.5	5	0.439	0.8	0.813	3.2	1.6	0.5	0	0.8	1.2	10	$1,839,033	98.5	10.1523
Leandro Barbosa	SG	32	GSW	66	1	2.8	0.474	0.7	2.1	0.384	2.1	4.2	0.511	0.9	0.784	1.4	1.5	0.6	0.1	0.7	1.3	7.1	$915,243	110	6.4545
Marco Belinelli	SG	28	SAS	62	9	3.2	0.423	1.4	4.2	0.374	1.8	3.6	0.471	1.4	0.848	2.5	1.5	0.5	0	1	1.4	9.2	$2,873,750	103.2	8.9147
Tony Allen	SG	33	MEM	63	41	3.6	0.495	0.2	0.6	0.345	3.4	6.8	0.505	1.3	0.627	4.4	1.4	2	0.5	1.4	2.6	8.6	$4,831,461	98.3	8.7487
Sergey Karasev	SG	21	BRK	33	16	1.6	0.403	0.5	1.5	0.296	1.2	2.4	0.475	0.9	0.763	2	1.4	0.7	0	0.7	1.4	4.6	$1,533,840	98	4.6939
Willie Green	SG	33	ORL	52	2	2.3	0.386	0.8	2.4	0.347	1.5	3	0.411	0.5	0.824	1.5	1.3	0.5	0.1	0.9	1.6	5.9	$1,448,490	95.7	6.1651
Jodie Meeks	SG	27	DET	60	0	3.7	0.416	1.2	3.6	0.349	2.5	5	0.46	2.4	0.906	1.7	1.3	1	0.1	1	1.3	11.1	$6,000,000	98.5	11.2690
Tyler Johnson	SG	22	MIA	32	2	2.2	0.419	0.6	1.8	0.375	1.6	3.2	0.437	1	0.681	2.5	1.3	1	0.3	0.9	1.3	5.9	$70,000	94.7	6.2302
Kentavious Caldwell-Po	SG	21	DET	82	82	4.8	0.401	1.9	5.7	0.345	2.9	5.8	0.449	1.3	0.696	3.1	1.3	1.1	0.2	1.1	2	12.7	$2,772,480	98.5	12.8934
Player	**Pos**	**Age**	**Tm**	**G**	**GS**	**FG**	**FG%**	**3P**	**3P*3**	**3P%**	**2P**	**2P*2**	**2P%**	**FT**	**FT%**	**TRB**	**AST**	**STL**	**BLK**	**TO**	**PF**	**PPG**	**Salary**	**Team PPG**	**ppg%tp**
Vince Carter	SG	38	MEM	66	1	2	0.333	1	3	0.297	1	2	0.379	0.7	0.789	2	1.2	0.7	0.2	0.7	1.5	5.8	$3,911,981	98.3	5.9003
Gerald Green	SG	29	PHO	74	4	4.4	0.416	1.9	5.7	0.354	2.5	5	0.478	1.3	0.825	2.5	1.2	0.6	0.2	1.4	2	11.9	$3,500,000	102.4	11.6211
Elijah Millsap	SG	27	UTA	47	5	1.7	0.34	0.6	1.8	0.311	1.1	2.2	0.358	1.2	0.674	3.2	1.2	1.2	0.3	1.4	2.6	5.3	$241,731	95.1	5.5731
Hollis Thompson	SG	23	PHI	71	23	3.2	0.413	1.6	4.8	0.401	1.5	3	0.426	0.9	0.708	2.8	1.2	0.8	0.4	0.9	2	8.8	$816,482	92	9.5652
Shabazz Muhammad	SG	22	MIN	38	13	5.1	0.489	0.5	1.5	0.392	4.6	9.2	0.503	2.7	0.717	4.1	1.2	0.5	0.2	0.9	1.3	13.5	$1,971,960	97.8	13.8037
Alan Anderson	SG	32	BRK	74	19	2.6	0.443	1	3	0.348	1.6	3.2	0.53	1.1	0.812	2.8	1.1	0.8	0.1	0.8	2	7.4	$1,276,061	98	7.5510
Garrett Temple	SG	28	WAS	52	18	1.3	0.4	0.6	1.8	0.375	0.7	1.4	0.427	0.7	0.729	1.7	1.1	0.8	0.2	0.7	1.5	3.9	$981,084	98.5	3.9594
C.J. Miles	SG	27	IND	70	40	4.7	0.398	2.2	6.6	0.345	2.5	5	0.459	1.9	0.807	3.1	1.1	0.9	0.4	1	1.9	13.5	$2,421,000	97.3	13.8746
Devyn Marble	SG	22	ORL	16	7	0.9	0.318	0.3	0.9	0.182	0.6	1.2	0.455	0.3	0.313	1.9	1.1	0.6	0.1	0.6	0.9	2.3	$884,879	95.7	2.4033
Archie Goodwin	SG	20	PHO	41	2	1.9	0.393	0.3	0.9	0.293	1.6	3.2	0.419	1.5	0.735	1.8	1.1	0.4	0.2	1.2	1.3	5.6	$1,112,280	102.4	5.4688
Nick Young	SG	29	LAL	42	0	4.1	0.366	2	6	0.369	2.2	4.4	0.363	3.1	0.892	2.3	1	0.5	0.3	1	2	13.4	$4,994,420	98.5	13.6041
Kent Bazemore	SG	25	ATL	75	10	1.9	0.426	0.6	1.8	0.364	1.2	2.4	0.467	0.8	0.6	3	1	0.7	0.4	1	1.7	5.2	$2,000,000	102.5	5.0732
Sean Kilpatrick	SG	25	MIN	4	0	1.8	0.35	1	3	0.308	0.8	1.6	0.429	1	1	1.5	1	0.8	0	0.5	0.8	5.5	$35,000	97.8	5.6237
Andre Roberson	SG	23	OKC	67	65	1.4	0.458	0.3	0.9	0.247	1.1	2.2	0.612	0.3	0.479	3.8	1	0.8	0.4	0.7	2.1	3.4	$1,160,880	104	3.2692
C.J. McCollum	SG	23	POR	62	3	2.6	0.436	0.9	2.7	0.396	1.7	3.4	0.46	0.8	0.699	1.5	1	0.7	0.1	0.8	1.3	6.8	$200,000	102.8	6.6148
Jamaal Franklin	SG	23	DEN	3	0	0.3	0.5	0.3	0.9	0.5	0	0		0		0.7	1	0	0.3	1	1.3	1	$163,296	101.5	0.9852
Ben Gordon	SG	31	ORL	56	0	2.3	0.437	0.6	1.8	0.361	1.7	3.4	0.475	1	0.836	1.1	0.9	0.3	0	1	1.1	6.2	$4,500,000	95.7	6.4786
Jeremy Lamb	SG	22	OKC	47	8	2.2	0.416	0.8	2.4	0.342	1.4	2.8	0.475	1	0.891	2.3	0.9	0.4	0.1	0.6	1.1	6.3	$2,202,000	104	6.0577
Nik Stauskas	SG	21	SAC	73	1	1.5	0.365	0.7	2.1	0.322	0.8	1.6	0.408	0.8	0.859	1.2	0.9	0.3	0.2	0.5	1.3	4.4	$2,745,840	101.3	4.3435
Anthony Morrow	SG	29	OKC	74	0	3.9	0.463	1.9	5.7	0.434	1.9	3.8	0.497	1.1	0.888	2.6	0.8	0.7	0.1	0.5	1.8	10.7	$3,200,000	104	10.2885
Player	**Pos**	**Age**	**Tm**	**G**	**GS**	**FG**	**FG%**	**3P**	**3P*3**	**3P%**	**2P**	**2P*2**	**2P%**	**FT**	**FT%**	**TRB**	**AST**	**STL**	**BLK**	**TO**	**PF**	**PPG**	**Salary**	**Team PPG**	**ppg%tp**
Justin Holiday	SG	25	GSW	59	4	1.5	0.387	0.6	1.8	0.321	0.9	1.8	0.444	0.6	0.822	1.2	0.8	0.7	0.2	0.5	0.9	4.3	$816,482	110	3.9091
Markel Brown	SG	23	BRK	47	29	1.6	0.362	0.4	1.2	0.266	1.3	2.6	0.404	1	0.825	2.3	0.8	0.7	0.3	0.6	1.2	4.6	$507,336	98	4.6939
Allen Crabbe	SG	22	POR	51	9	1.2	0.412	0.6	1.8	0.353	0.6	1.2	0.485	0.2	0.75	1.4	0.8	0.4	0.3	0.3	1.5	3.3	$862,000	102.8	3.2101
Shannon Brown	SG	29	MIA	5	2	1.4	0.368	0.6	1.8	0.429	0.8	1.6	0.333	0.6	0.75	0.2	0.6	0.8	0	0.8	1.6	4	$250,000	94.7	4.2239
E'Twaun Moore	SG	25	CHI	56	0	1.1	0.446	0.2	0.6	0.342	0.9	1.8	0.485	0.2	0.6	0.8	0.6	0.4	0.1	0.3	0.8	2.7	$948,163	94.2	2.8662
John Jenkins	SG	23	ATL	24	3	2	0.495	0.9	2.7	0.404	1.2	2.4	0.596	0.7	0.842	1.6	0.5	0.4	0	0.3	0.6	5.6	$1,312,920	102.5	5.4634
Joe Harris	SG	23	CLE	51	1	0.9	0.4	0.6	1.8	0.369	0.3	0.6	0.472	0.2	0.6	0.8	0.5	0.1	0	0.5	1.2	2.7	$884,879	103.1	2.6188
Jared Cunningham	SG	23	LAC	19	0	0.6	0.364	0.2	0.6	0.308	0.4	0.8	0.4	0.4	0.538	0.5	0.5	0.2	0	0.4	0.6	1.8	$915,243	106.7	1.6870
P.J. Hairston	SG	22	CHO	45	2	1.9	0.323	1.1	3.3	0.301	0.8	1.6	0.358	0.7	0.861	2	0.5	0.5	0.3	0.5	1.4	5.6	$1,149,720	94.2	5.9448
Gary Harris	SG	20	DEN	55	6	1.2	0.304	0.4	1.2	0.204	0.8	1.6	0.395	0.6	0.745	1.2	0.5	0.7	0.1	0.7	1.3	3.4	$1,519,200	101.5	3.3498
Jordan Adams	SG	20	MEM	30	0	1.2	0.407	0.3	0.9	0.4	0.8	1.6	0.41	0.5	0.609	0.9	0.5	0.5	0.2	0.5	0.8	3.1	$1,344,120	98.3	3.1536
Brandon Rush	SG	29	GSW	33	0	0.3	0.204	0.1	0.3	0.111	0.2	0.4	0.296	0.2	0.455	1.2	0.4	0.2	0.4	0.3	0.8	0.9	$1,145,685	110	0.8182
Glen Rice	SG	24	WAS	5	0	0.6	0.2	0.2	0.6	0.143	0.4	0.8	0.25	0.8	0.667	0.8	0.4	0	0	0.4	2.2	2.2	$816,482	98.5	2.2335
C.J. Wilcox	SG	24	LAC	21	0	0.8	0.421	0.3	0.9	0.368	0.4	0.8	0.474	0.1	1	0.3	0.4	0.1	0	0.5	0.3	2	$1,109,760	106.7	1.8744
Nick Johnson	SG	22	HOU	28	0	0.9	0.347	0.2	0.6	0.238	0.8	1.6	0.389	0.6	0.68	1.4	0.4	0.3	0.1	0.7	1.2	2.6	$507,336	103.9	2.5024
James Young	SG	19	BOS	31	0	1.2	0.353	0.5	1.5	0.258	0.6	1.2	0.528	0.5	0.552	1.4	0.4	0.3	0.1	0.2	0.7	3.4	$1,674,480	101.4	3.3531
Andre Dawkins	SG	23	MIA	4	0	0.3	0.167	0.3	0.9	0.167	0	0		0		0.5	0.3	0	0	0.3	0.8	0.8	$507,336	94.7	0.8448
Patrick Christopher	SG	26	UTA	4	1	0.8	0.25	0	0	0	0.8	1.6	0.375	0		1.5	0	0.3	0	0.3	1.5	1.5	$379,010	95.1	1.5773
Malcolm Lee	SG	24	PHI	1	0	0	0	0	0		0	0	0	0		0	0	0	0	0	0	0	$50,000	92	0.0000

References

1995-96 NBA Player Stats: Per Game. (n.d.). Retrieved November 16, 2016, from http://www.basketball-reference.com/leagues/NBA_1996_per_game.html

2013-14 NBA Player Stats: Per Game. (n.d.). Retrieved November 16, 2016, from http://www.basketball-reference.com/leagues/NBA_2014_per_game.html

2013-14 NBA Season Summary. (n.d.). Retrieved November 30, 2016, from http://www.basketball-reference.com/leagues/NBA_2014.html

2014-15 NBA Player Stats: Per Game. (n.d.). Retrieved November 16, 2017, from http://www.basketball-reference.com/leagues/NBA_2015_per_game.html

2014-15 NBA Season Summary. (n.d.). Retrieved November 16, 2016, from http://www.basketball-reference.com/leagues/NBA_2015.html

Honkasalo, M. (2015, June 23). The NBA's Inequality Problem: Defense Earns 60 Cents on the Dollar. Retrieved April 30, 2016, from http://nyloncalculus.com/2015/06/23/in-the-nba-defense-doesnt-get-you-paid/

Lyons, R., Jr., Jackson, E. N., Jr., & Livingston, A. (2016, October 12). Determinants of NBA Player Salaries. Retrieved November 16, 2016, from http://thesportjournal.org/article/determinants-of-nba-player-salaries/

Michael Jordan 1995-96 Splits. (n.d.). Retrieved November 15, 2016, from http://www.basketball-reference.com/players/j/jordami01/splits/1996

NBA & ABA Sixth Man of the Year Award Winners. (n.d.). Retrieved November 16, 2016, from http://www.basketball-reference.com/awards/smoy.html

NBA Player Salaries - 2014-2015. (n.d.). Retrieved November 30, 2016, from http://www.espn.com/nba/salaries/_/year/2015

NBA Salary Cap History. (n.d.). Retrieved November 16, 2016, from http://www.basketball-reference.com/contracts/salary-cap-history.html

Schimanski, M. (2016, September 01). NBA dollars & sense — Part I: Player monetary value. Retrieved April 15, 2016, from http://nyloncalculus.com/2016/09/01/nba-dollars-sense-player-monetary-value/

Schimanski, M. (2016, September 05). NBA Dollars & Sense — Part II: Player value by role. Retrieved November 15, 2016, from http://nyloncalculus.com/2016/09/05/nba-dollars-sense-player-value-by-role/

Schimanski, M. (2016, September 23). NBA Dollars & Sense – Part III: Player value by position. Retrieved November 15, 2016, from http://nyloncalculus.com/2016/09/23/nba-dollars-sense-player-value-by-position/

DETERMINANTS OF REAL ESTATE PRICES HEDONIC VALUATION AND TIME SERIES ANALYSIS ACROSS THE NEW YORK METROPOLITAN AREA

Jonathan Panzures & Theodore Plotkin

This papers attempting to find what factors accurately determine residential real estate prices in New York City, across the boroughs of Manhattan, Brooklyn and Queens. Our data, observations, and regressions will span twenty-four neighborhoods within the New York Metropolitan Area. Some factors are structural, while others are environmental. The goal is to build a hedonic valuation model with these variables and be able to define their exact marginal effects. From these results, we hope to gain a sense of the costs or benefits these variables pose on society and real estate. Finally, we plan to use this data to price a sample of properties in each of the respective neighborhoods in our study. Furthermore, we will also be examining median price trends across two pairs of New York City neighborhoods; Park Slope and Flatbush in the borough of Brooklyn, and the Lower East Side and Chelsea in Manhattan. Our goals are to see if there were any statistically significant relationships between changes in the monthly median values amongst neighborhoods, and if so, derive the degree of change.

Introduction

Cross Sectional Study

The price of a residential real estate asset is the resultant of the summation of a bundle of components and their respective marginal values. The components form vectors of physical characteristics and environmental information that describe the property and its surroundings; these vectors are sets of specific qualities that measure an asset's given features. These components each have a certain marginal value that alter their scale. The linear combinations of all the component vectors and their corresponding marginal values is the number seen on the price tag. This reality is the fundamental idea behind hedonic real estate valuation; finding the minimum error function that relates a given set of features to its corresponding price.

$$V_i = \boxed{\quad ? \quad}$$

If the value of residential property *i* is the area of this square, how is this square divided up? What proportion of its area is dependent on the properties' physical characteristics, and what proportion is dependent on the features of the surrounding environment? Which factors belong inside this square to create a parsimonious model of the true value and which factors are excessive, contributing to collinearity? How much of it is devoted to the unmeasurable variability in agent preferences, or geographic locality, and captured by the error term, ε of our model. These are, very generally, the questions we will attempt to answer with our hedonic valuation model.

Hedonic Valuation

We will examine residential real estate assets in New York City across 3 boroughs; Queens, Brooklyn and Manhattan. Our locational scope will be 24 randomly selected neighborhoods, defined by zip codes $j \in J$. The physical characteristics we will look at will be the square footage of each property as well as number of bathrooms. We will also incorporate the number of violent crime occurrences in each neighborhood in the year 2015, the number of loud noise complaints in the year 2015, and the 2016 inflation adjusted median household income in each neighborhood in the year 2014. We will utilize OLS to relate these five independent variables to the relevant dependency factor, the list price of each residential property i.

Our goal is to build and refine a model that derives the marginal value of each of these five attributes and assesses the statistical significance of each attribute in explaining the price. From these results we will get a sense of the degree of the societal cost and benefit of each of these attributes and then use the marginal values of the statistically significant factors to value a property i for each zip code j . We will compare our model price to the list price in order to define the degree of error in each valuation, and this process will provide us with a sense of the model's accuracy. We will also consider how these degrees of error change depending on model specification.

ANOVA of Neighborhoods

We will then derive the standalone marginal values for each of the 24 neighborhoods used in the hedonic valuation study and defined by zip codes $j \in J$. We will examine which neighborhoods form submarkets of the general area of study J. We will employ indicator variables for each location, using the list prices of the relevant properties i in each locational bin j as the dependency factor. We will also employ two indicator variables that account for property type; house or apartment. The locational indicators that are statistically significant imply these neighborhoods are individual submarkets of study area J.

We will compare our results with the 2015 violent crime, 2015 noise complaint, and 2016 inflation adjusted 2014 median income data to see if there are correlations between the marginal values we derive and the relevant environmental factors that defines each neighborhood in the hedonic valuation model.

Time Series Study

We will look at the median residential property values of a few neighborhoods over time to formulate an understanding of general neighborhood price trends and their relationships over the past decade.

Neighborhood Groups

We will randomly select 4 neighborhoods from the 24 seen in the cross-sectional study; 2 from Brooklyn and 2 from Manhattan. Our goal is to create an OLS model that derives the marginal effect of monthly median value changes in one Brooklyn neighborhood on the other, and then run the same model for the Manhattan group. We will calculate the serial correlation for the Brooklyn group series and serial correlation for the Manhattan group series. We will then visualize the median value trends for each group using scatterplots, and examine the explanatory power and statistical significance of each model.

Individual Neighborhood Series

We will next select one neighborhood from the Manhattan group and one neighborhood for the Brooklyn group and examine if the Great Recession had a statistically significant effect on the price trends for these neighborhoods using the Chow Test. If there is a break, we will fit models for the sub-periods, pre and post recession, to see what the degree of the change in the monthly price trend was before and after the crisis for each of the 2 neighborhoods. Finally, we will examine monthly seasonality in these series, model their autocorrelation with lags, examine heteroscedasticity and observe residual patterns after employing alternative fitting techniques.

Literature Review

One of the well-known hedonic residential pricing studies was *Hedonic Housing Prices and The Demand For Clean Air* by Harrison and Rubinfeld in 1976. This early cross sectional study examined the effect of a variety of factors including property age, number of rooms, neighborhood crime and demographic distribution as well as air pollution on the median value of owner-occupied homes in the Boston Metropolitan Area using data from the 1970 U.S. Census. They found many variables we will study to be statistically significant including crime frequency and property size, and we will compare our priori expectations and results to theirs. While their research goal was to ultimately inform policy makers by deriving the willingness to pay for air pollution improvements, the goal for our hedonic valuation model will be to use the pricing models we derive to accurately predict prices for the various properties in our 24 neighborhood study.

Nelson's 1982 study *Highway Noise and Property Values* attempted to define the statistical significance of noise pollution on property values and find the marginal value of silence as measured by decibels of traffic noise for various localities. Though Nelson's findings were meant to inform public officials concerned with the environmental effects of highway noise, we believe noise will be a relevant factor in our hedonic valuation model and have decided to include the number of noise complaints in each of the 24 neighborhoods in our study.

A later study of interest was *Spatial Statistics and Real Estate* by Pace, Barry and Sirmans in 1998. This study discussed methods for addressing the effects of spatial autocorrelation seen in real estate data, both as a means of improving property valuation accuracy and identifying submarkets. One technique they described was utilizing GLS with a spatially weighted variance-covariance matrix in order to account for geographic price variance. We found the technique to be beyond our capabilities at this time, but were nevertheless intrigued by their referencing the use of indicator variables as a means of identifying the value of different localities. Though Pace, Barry and Sirmans recommended against this technique in favor of those of greater complexity, we decided it was ultimately a necessary preliminary step in gaining a sense of the degree of locational price variance in our study, and thus this research inspired our ANOVA neighborhood model.

The Models

Cross Sectional Models

<u>Hedonic Valuation Model</u>

The initial hedonic pricing model for a specific property *i* in each neighborhood $j \in J$ is given by:

$$V_{i,j} = \beta_1 Baths_{1i} + \beta_2 Sqft_{2i} + \beta_3 Noise_{3i,j} + \beta_4 Crime_{4i,j} + \beta_5 MedInc_{5i,j} + \epsilon_{i,j}$$

This model utilizes OLS to derive the degree to which each of the 5 characteristics contribute to the overall price of each property *i.* Our priori expectation is that baths and square feet, will be positive. The more relative living space and bathrooms a respective property has, the higher its inherent value should be, especially given the density of an urban metropolis like New York City. Similarly, we suspected median household income will also be a positive priori. Areas with high neighborhood incomes will likely have higher levels of welfare than low income areas. This reality may be a proxy for the overall quality level of properties in each neighborhood $j \in J$. Additionally, we hypothesize that noise complaint frequency and violent crime occurrence will both be negatively correlated with a property's price; we expect that with each additional incident of violence or additional noise complaint, a locality's respective residencies should lose some degree of value.

If collinearity is discovered in the data set between the bathroom vector and the square feet vector, we will then run the following hedonic valuation model for a specific property *i* in each neighborhood $j \in J$ given by:

$$V_{i,j} = \beta_1 Sqft_{1i} + \beta_2 Noise_{2i,j} + \beta_3 Crime_{3i,j} + \beta_4 MedInc_{4i,j} + \epsilon_{i,j}$$

Barring baths, our priori expectations and their relevant reasoning for this OLS model are the same as those for the original model; we expect square feet and median income to be positive and noise and crime to be negative externalities detracting from the price.

If heteroscedasticity is discovered in the data set, we will log-transform the data and run the following log-transformed hedonic valuation model for a specific property *i* in each neighborhood $j \in J$ given by:

$$V_{i,j} = \beta_1 \, ln\, Sqft_{1i} + \beta_2 \, ln\, Noise_{2i,j} + \beta_3 \, ln\, Crime_{3i,j} + \beta_4 \, ln\, MedInc_{4i,j} + \epsilon_{i,j}$$

This model utilizes OLS to derive the partial elasticities of each attribute. While the interpretation of the coefficients will change to percentage-change form, our priori expectations will remain unchanged.

Ultimately, we decided not to include an intercept in any form of our hedonic valuation model. This decision was made because, collectively, we believe an intercept will capture the inherent value of the twenty-four geographic locations in our study. As an alternative, we ran the 24 neighborhood ANOVA model, where we generated indicator variables for all 24 neighborhoods in order to capture their respective locational effects. We believe this proactive measure, in combination with our price model, will give us a more detailed and analytical picture of the real estate market presented in our study.

<u>ANOVA Neighborhood Model</u>

The ANOVA neighborhood model for the value of each neighborhood $j \in J$ is given by:

$$V_j = \sum_{j=1}^{24} \eta_j \Phi_j + \eta_{25} Apt + \eta_{26} House + \epsilon_j$$

With indicator variables defined as follows:

$$\Phi_j \in [0,1]$$
$$Apt \in [0,1]$$
$$House \in [0,1]$$

This ANOVA neighborhood model will calculate the inherent mean value of living in a specific neighborhood within the New York Metropolitan Area using locational indicator variables $\eta_1, \eta_2, \ldots, \eta_{24}$, while concurrently accounting for the overall mean market value of living in either an apartment or a house using type indicator variables. The list price of each property *i* will remain the dependent variable. There will be eight neighborhood indicator variables for Manhattan properties, eleven for

Brooklyn properties, and five for properties in Queens. Our randomly selected neighborhoods throughout these three boroughs were as follows:

	Location	Zip Code
Manhattan	Central Harlem	10026
	Chelsea	10001
	Financial District	10005
	Gramercy Park	10010
	Lower East Side	10002
	Soho	10014
	Upper East Side	10075
	Upper West Side	10024

	Location	Zip Code
Brooklyn	Bay Ridge	11209
	Brooklyn Heights	11201
	Brownsville	11212
	Bushwick	11221
	Canarsie	11236
	Coney Island	11224
	Crown Heights	11205
	DUMBO	11201
	Flatbush	11226
	Park Slope	11215
	Williamsburg	11211

	Location	Zip Code
Queens	Astoria	11101
	Bayside	11361
	Corona	11368
	Forest Hills	11375
	South Jamaica	11412

We have looked at the noise pollution, crime frequency and median income data for each neighborhood and our priori expectations in regards to the sign for the mean value of each neighborhood $j \in J$ are as follows:

Expected	
Priori Positive	**Priori Negative**
Apartment	Central Harlem
House	Bay Ridge
Chelsea	Brownsville
FIDI	Bushwick
Gramercy Park	Canarsie
LES	Coney Island
Soho	Crown Heights
UES	Flatbush
UWS	Astoria
Brooklyn Heights	Bayside
DUMBO	Corona
Park Slope	South Jamaica
Williamsburg	
Forest Hills	

We assumed neighborhoods with notably high noise and crime rates, and significantly low median household incomes, would intrinsically have negative values. Conversely, we anticipated quieter, safer, and comparatively wealthier localities would be positive. In order to test the validity of our priori we regressed the 24 locational indicators and then correlated the resulting coefficients with their respective crime, noise, and median income data.

Time Series Models

Neighborhood Groups Model

The Brooklyn group model is given by:

$$MVParkSlope_t = \beta_0 + \beta_1 MVFlatbush_t + \epsilon_t$$

This model utilizes OLS to derive the effect of a $1 change in the median residential value of Flatbush, Brooklyn on the median residential value of Park Slope, Brooklyn. The model uses the 2016 inflation-adjusted monthly median values for Flatbush and Park Slope over the years 2003-2016. We specified the model as such because we wanted to see if median residential price changes in a higher crime, less valuable, neighborhood like Flatbush have significant effects on the median residential price changes in a notably safer and more valuable neighborhood like Park Slope. We expect the sign of β_1 to be positive because we believe the improvements in a lower welfare neighborhood that are implied by its increasing median value will act as a positive externality to other neighborhoods within the same borough.

The Manhattan group model is given by:

$$MVLES_t = \beta_0 + \beta_1 MVChelsea_t + \epsilon_t$$

This model also utilizes OLS to derive the effect of a $1 change in the median residential value of Chelsea, Manhattan on the median residential value of Lower East Side, Manhattan. The model uses the 2015 inflation-adjusted monthly median values for Lower East Side and Chelsea over the years 2004-2015. We had to shorten the time range of this study in comparison to our Brooklyn model due to the issue of data availability. We specified the model as such because we wanted to see if changes in the median residential value of a neighborhood in western downtown Manhattan had significant effects on the median residential value changes in a neighborhood in downtown eastern Manhattan. Downtown Manhattan is a highly concentrated real estate market and we expect potential serial dependencies exists between its neighborhoods. Our priori for β_1 is thus positive; the implied quality improvements captured by price increases in a western locality will act as a positive externality for a neighboring eastern locality.

Individual Neighborhood Series

Park Slope and Chelsea median value models are given by:

$$MVChelsea_t = \beta_0 + \beta_1 T_t + \epsilon_t$$
$$MVParkSlope_t = \beta_0 + \beta_1 T_t + \epsilon_t$$

These models utilize OLS to derive the degree of monthly change in the median residential value for Park Slope, Brooklyn and the monthly change in the median value for Chelsea, Manhattan. MV Chelsea_t represents the vector of 2015 inflation-adjusted monthly median value data for Chelsea, Manhattan for the years 2004-2015 and T_t represents the vector of months in the Chelsea series. Similarly, MV ParkSlope_t represents the vector of 2016 inflation-adjusted monthly median value data for Park Slope, Brooklyn for the years 2003-2016 and T_t again represents the vector of months in the Park Slope series. We have chosen the month of June in the year 2008 as the breakpoint for our Chow Test for both series because this is when we believe the effects of the Great Recession began to manifest themselves in the New York City real estate market. If we find that there is a statistically significant structural break, we will fit sub-period models for each neighborhood using the same form given above. We expect the sign for all β_1 to be positive, despite the Great Recession, because both of these areas have historically seen strong growth and been known for exceptional locational and architectural neighborhood attributes that tend to attract higher net-wealth individuals.

The autocorrelation in each series are given by the following lag-models:

$$MVChelsea_t = \beta_0 + \beta_1 MVChelsea_{t-1} + \epsilon_t$$
$$MVParkSlope_t = \beta_0 + \beta_1 MVParkSlope_{t-1} + \epsilon_t$$

These 1 month lag-models will find OLS estimates of how a \$1 change in the median residential value of Chelsea and Park Slope in the previous month effect the median residential value of Chelsea and Park Slope for the following month, for all the months $t \in T$, in each respective neighborhood's series. These lag-models were chosen as a means of detecting and measuring the degree of serial autocorrelation in each series. We again expect the signs for both β_1 to be positive because real estate price trends tend to exhibit serial momentum over time; increases in one period tend to contribute to increases in the next.

The tests for monthly seasonality in each series are given by the following monthly dummy-variable ANOVA models:

$$MVParkSlope_t = \beta_0 + \beta_1 Dec_{1t} + \beta_2 Feb_{2t} + \beta_3 March_{3t} + \beta_4 April_{4t} + \beta_5 May_{5t} + \beta_6 June_{6t} + \beta_7 July_{7t} + \beta_8 Aug_{8t} + \beta_9 Sept_{9t} + \beta_{10} Oct_{10t} + \beta_{11} Nov_{11t} + \epsilon_t$$

$$MVChelsea_t = \beta_0 + \beta_1 Dec_{1t} + \beta_2 Feb_{2t} + \beta_3 March_{3t} + \beta_4 April_{4t} + \beta_5 May_{5t} + \beta_6 June_{6t} + \beta_7 July_{7t} + \beta_8 Aug_{8t} + \beta_9 Sept_{9t} + \beta_{10} Oct_{10t} + \beta_{11} Nov_{11t} + \epsilon_t$$

These ANOVA models will compute OLS estimates of the average of the median value for each month in the overall series, as represented by $\beta_0, \beta_1, \ldots, \beta_{11}$. Information regarding the month of January will be stored in the intercept term, β_0. We expect to see positive signs for all months because New York City properties tend to be in-demand throughout the year, however we expect to see higher means during the spring and summer months because moving is easier and heating costs are lower during these seasons.

The Data

We utilized a series of Python based web-scrapers for our data collection process. These tools scraped a selection of websites for our cross-sectional and time-series data and directed this information to a Digital Ocean Server database. We then setup a program in RStudio that could interact and pull information from the database and deliver it to dataframes to be randomly sampled and analyzed.

Cross Sectional Data

Hedonic Valuation Data

Our hedonic valuation study utilized property data, scraped and collected from *streeteasy.com*. We were able to pull 13,654 observations worth of currently listed properties across all five boroughs. In this dataset, we organized our variables by price, baths, square feet, and location; all encoded to specific addresses and zip codes. The location of each property was also limited to our afore mentioned twenty-four randomly selected subject sites, that resulted in a total 3290 individual property observations. This data was then utilized as the information for our regressions, and respective valuation models.

	Definitions	Sources	Number of Observations	Mean	Standard Deviation
Dependent					
$\mathbf{V}_i$	List price of residential property *i*	*www.streeteasy.com*	3920	2740229	4098716
Structural					
$\mathbf{Baths}_i$	Number of bathrooms in property *i*	*www.streeteasy.com*	3920	2.1	1.5
$\mathbf{Sqft}_i$	Property size in square feet for property *i*	*www.streeteasy.com*	3920	1979	1767
Neighborhood					
$\mathbf{N}_{i,j}$	Average number of noise complaints (Residential, street/sidewalk, vehicle) for property i in zip code j (for year 2015)	*www.nyc.gov*	24	513	*413*
$\mathbf{Cr}_{i,j}$	Average number of violent crimes (robbery, assault, murder, rape, burglary) for property i in zip code j (for year 2015)	*www.nyc.gov*	24	92	52
$\mathbf{MedInc}_{i,j}$	Average household median income for property i in zip code j (year 2014 inflation adjusted for year 2015)	*www.census.gov*	24	90362	44115

Noise Data

For our Noise Pollution statistic, we collected data from the New York 311 complaint database. With our scraper, we were able to amount a data set of roughly 1,954,745 observations, all encoded by their respective zip-codes, across the whole of the New York Metropolitan area; with yearly resolution dating from 2010-2015. We then organized this data into three categories:

Noise – Residential:

All complaints regarding loud music, parties, banging, and pounding from neighboring houses, or apartments.

Noise - Street/Sidewalk:

Loud talking, yelling, or any non-vehicle related noise complaints from the street.

Noise – Vehicle:

Complaints regarding honking or loud music omitted from a car or truck

From this data, we then restricted our observations across all twenty-four previously selected neighborhood zip-codes for the most recent calendar year and took the average of the sums. This data was then used for our cross-sectional models as a means of measuring a respective locality's frequency of noise, based on the average of the sums of complaints for year 2015.

Crime Data

In regard to our crime data, we pulled information directly from all of New York City's NYPD precinct databases. In total, we were able to collect 1,123,465 observations of past violent crime related incidents, dating from 1900 to 2015. For our cross-sectional price models, we organized this data into five distinct variables, only utilizing information from years-end 2015. These major felonies were:

Robbery
Felony Assault
Murder
Rape
Burglary

With these organized subsets of crime data, we then took the sum of the 2015 occurrences, for each of our twenty-four selected zip codes, and took their averages; the results served as a means of evaluating the relative occurrence rate of violent crimes amongst our target locations, and New York City as a whole.

Median Income Data

For the final component of our hedonic valuation dataset, we collected demographic data from the census database. With our scraper, we chose to pull levels of median household income in the past 12 months, across all New York based zip codes; the census backed code for this indicator is B19013_001. In total, we had 531 observations, and yearly resolution, spanning from 2012 to 2014. As a result, we adjusted the 2014 median household income levels, across our 24 randomly selected zip codes, and put them in 2015 inflation adjusted values; this transformation was done so as to keep all cross-sectional data consistent in date and year. We used the median household incomes as a means of numerically comparing the differentials in wealth across our selected neighborhoods, within our hedonic valuation model.

Scattergram and Correlation Matrix

In analyzing our data, we found it beneficial to construct a correlation matrix, to see the explanatory power of our dependent and independent variables, and gauge linear associations with one another (See Appendix A, Fig. 1b). Interestingly, it seems that price is highly correlated with baths (.62) and square feet (.76), respectively. Essentially, as a property accumulates more bathrooms, and square footage, as there is a moderately high linear association with price, the subject site's value should increase. Correlation remains low to moderate amongst all variables considered, and unanimously maintains some degree of positive linear relationships; in our opinion, this data is a beneficial indicator, as low correlations suggest a lesser degree, of collinearity amongst the vectors in the dataset which may allow us to clearly siphon out the individual effects of these factors in explaining property prices. A notable relationship we encountered, however, was square feet on baths, which amounted to an outstanding correlation of .82. Similarly, noise and crime also seem to present a moderately high correlation, roughly .71. As a result, our priori for our data's correlations, would be that a highly linear relation amongst variables, may be a first-stage indicator of multicollinearity.

We also constructed a scattergram of these linear relationships (See Appendix A, Fig. 1a). This process was done in an effort to visually represent all dependent and independent correlations.

The ANOVA Neighborhood Data

For the ANOVA model we took all of our respective locational data from the hedonic valuation model; utilizing our scraped information from *streeteasy.com*. For the twenty-four neighborhoods, just like our Hedonic Valuation Model, we had a total of 3920 observations. In this ANOVA model, however, each unique zip coded neighborhood acted as an independent indicator variable; we also used an aggregate of all apartments and houses, across these twenty-four localities, as indicator variables too. On the following pages we have defined our variables in detail.

	Definitions	Sources	Number of Observations	Mean	Standard Deviation
Dependent					
V_i	List price of residential property *i*	*www.streeteasy.com*	3920	2740229	4098716
Type Indicators					
Apt	1 if residential property *i* is an apartment (0 otherwise)	*www.streeteasy.com*	3138	3941134	6000979
House	1 if residential property *i* is a house (0 otherwise)	“	782	2440960	3401847
Neighborhood Indicators					
Central Harlem	1 if residential property *i* is located in Central Harlem (0 otherwise)	“	114	1517780	1184645
Chelsea	1 if residential property *i* is located in Chelsea (0 otherwise)	“	231	3268294	3183928
FIDI	1 if residential property *i* is located in FIDI (0 otherwise)	“	307	2574983	3228796
Gramercy Park	1 if residential property *i* is located in Gramercy Park (0 otherwise)	“	187	2748631	4418438
LES	1 if residential property *i* is located in LES (0 otherwise)	“	128	2286161	2039845
SoHo	1 if residential property *i* is located in SoHo (0 otherwise)	“	201	5365792	4413923
UES	1 if residential property *i* is located in Upper East Side (0 otherwise)	“	499	4923080	6959236
UWS	1 if residential property *i* is located in Upper West Side (0 otherwise)	“	545	3810298	4803632
Bay Ridge	1 if residential property *i* is located in Bay Ridge (0 otherwise)	“	83	1206807	1565327
Brooklyn Heights	1 if residential property *i* is located in Brooklyn Heights (0 otherwise)	“	129	4350398	6011254

	Definitions	**Sources**	**Number of Observations**	**Mean**	**Standard Deviation**
Neighborhood Indicators (continued)					
Brownsville	1 if residential property *i* is located in Brownsville (0 otherwise)	*www.streeteasy.com*	67	744250	193035
Bushwick	1 if residential property *i* is located in Bushwick (0 otherwise)	"	126	1138997	424533
Canarsie	1 if residential property *i* is located in Canarsie (0 otherwise)	"	78	499788	178391
Coney Island	1 if residential property *i* is located in Coney Island (0 otherwise)	"	43	532209	206460
Crown Heights	1 if residential property *i* is located in Crown Heights (0 otherwise)	"	111	1249135	745082
DUMBO	1 if residential property *i* is located in DUMBO (0 otherwise)	"	93	3586272	1809328
Flatbush	1 if residential property *i* is located in Flatbush (0 otherwise)	"	60	698425	478429
Park Slope	1 if residential property *i* is located in Park Slope (0 otherwise)	"	215	2081293	1764776
Williamsburg	1 if residential property *i* is located in Williamsburg (0 otherwise)	"	327	1681435	1411753
Astoria	1 if residential property *i* is located in Astoria (0 otherwise)	"	79	1238081	765708
Bayside	1 if residential property *i* is located in Bayside (0 otherwise)	"	29	913776	462357
Corona	1 if residential property *i* is located in Corona (0 otherwise)	"	19	948305	731320
Forest Hills	1 if residential property *i* is located in Forest Hills (0 otherwise)	"	219	671540	732768
South Jamaica	1 if residential property *i* is located in South Jamaica (0 otherwise)	"	30	535533	269972

	Definitions	Sources	Number of Observations	Mean	Standard Deviation
Manhattan Group Series					
MV LES$_t$	Monthly 2015 inflation adjusted median residential value for Lower East Side, Manhattan for years 2004-2015 (starting in March 04' and ending in August 15')	*www.Zillow.com*	138	702196	121922
MV Chelsea$_t$	Monthly 2015 inflation adjusted median residential value for Chelsea, Manhattan for years 2004-2015 (starting in March 04' and ending in August 15')	*www.Zillow.com*	138	1581415	339868
Brooklyn Group Series					
MV Park Slope$_t$	Monthly 2016 inflation adjusted median residential value for Park Slope, Brooklyn for years 2003-2016 (starting in Dec 03' and ending in Sept 16')	*www.Zillow.com*	153	694231	167482
MV Flatbush$_t$	Monthly 2016 inflation adjusted median residential value for Flatbush, Brooklyn for years 2003-2016 (starting in Dec 03' and ending in Sept 16')	*www.Zillow.com*	153	475466	83809

Time Series Data

The above monthly median residential value time series data are sourced from *Zillow.com* and will be used for the entirety of the time series models we will run in our study. The 4 neighborhoods, Chelsea, Lower East Side, Park Slope and Flatbush were randomly selected from the 24 neighborhoods seen in our cross-sectional study. In order to keep each series coherent, we ensured each borough group started and ended on the same month and year and that consequently, each borough group had the same number of observations. The Manhattan group series runs almost 2 years shorter than the Brooklyn group series due to the issue of data availability, however this does not harm the validity of our results. Additionally, we decided to adjust the Brooklyn series for 2016 inflation and the Manhattan series for 2015 inflation, in order to get a clearer indicator of the real degree of median value change over the time-span in each data set.

Before running the models, we created plots of each series individually; the plots for each series can be seen in Appendix B, Fig 1-4. We also calculated the monthly serial correlations for each individual series before running the regressions. We found a monthly serial correlation of 0.69 for Flatbush, and 0.93 for Park Slope for the years 2003-2016. We also calculated monthly serial correlations of 0.85 for the Lower East Side and 0.78 for Chelsea for the years 2004-2015.

In order to setup the data in the lag models for Park Slope and Chelsea, we took their respective median value series vectors, duplicated them, and then removed the first month median value data entry from each duplicated vector so that they each begin on month 2 (January 03' for Park Slope and April 04' for the Lower East Side). In order to test for seasonality in the Park Slope and Chelsea series we treated January as our base month intercept and encoded an indicator variable for each of the remaining 11 months, treating them as differential intercepts to January. When we perform the Chow Test on the Park Slope and Chelsea median value data, we will set our breakpoint at June of 2008. This breakpoint corresponds to month number 55 in the Park Slope series and month number 52 in the Chelsea series. If we find a significant break point in each series, we will subset both series at their relevant break month, and fit sub-models for each of the corresponding sub-periods.

The Results

Cross Sectional Results

<u>Hedonic Valuation Results</u>

After regressing number of baths, square feet, noise, crime, and median income on price, we attained all of our expected prior signs (see Appendix A, Fig. 2). We saw that with an increase in bathrooms, square footage, and median income, there was a beneficial, and positive, affect on a property's valuation. Conversely, with increased incidents of noise and crime in a neighborhood, as we expected, a greater discount was imposed on a respective property's median value; noise and crime proved to have a negative relationship with price.

Overall, our individual variables for square feet, noise, crime, and median income were quite significant. Square feet proved to be most significant with a t-value of 31, while noise and crime had a t-values of approximately -5, and median income's was 10; all of these variables, were significant at the 1% level. Nevertheless, baths remained highly insignificant, with a t-value of 1.2, not passing even at the 10% level of significance. We felt this occurrence was due to some degree of multicollinearity between baths and square footage. As previously noted, these two variables are highly correlated with a correlation coefficient of roughly .82. This reality was our primary indicator as to why baths proved to be insignificant in our initial hedonic valuation model.

Nevertheless, this initial hedonic valuation model had an observed F statistic of 1540, on 5 and 2577 degrees of freedom. With this data at hand, when calculated, our model's critical F-statistic was 4.336, at the 5% significance level. Since F-observed was much far greater than the critical value, this result shows that all of the coefficients are simultaneously not zero. Moreover, regardless of the high degree of multicollinearity between baths and square feet, our model had a relatively high adjusted R-squared of .7487; suggesting that variability in baths, square feet, noise, and crime explain 74.87% of the variability in property price.

Adjusting for Multicollinearity

In order to account for our valuation model's multicollinearity, the best course of action was to drop baths from our hedonic valuation model. Upon its removal, the new hedonic valuation model's coefficients' signs remained the same; still meeting our priori expectations (see Appendix A, Fig. 3). An interesting occurrence, however, was that the t-value for square feet increased dramatically, to a value of 60, while noise and crime's t-values remained relatively constant, when juxtaposed to our previous models'. Furthermore, median household income's t-value increased marginally to 10.972. All in all, after adjusting for our hedonic valuation model's multicollinearity, we were left with seemingly more significant variables.

Moreover, if we look at this model's observed F-statistic, we see it increased to a value of 2020, on 4 and 2607 degrees of freedom. The calculated F critical value now is 2.375; as a result, we see that without the presence of baths in our regression equation, at the 5% significance level, our overall model gained more significance, and no longer faces a pronounced case of collinearity. Moreover, if we look at the R-squared, it increased to .7557. Thus when we solely regressed square feet, noise, crime, and median income against price, we saw a higher degree of explanatory power within our model.

Accounting for Heteroscedastic Data

With a seemingly more robust model, we thought it was important to test for heteroscedasticity in our hedonic valuation model; as skewed values for our coefficients' standard errors could potentially invalidate our t-stats, and hypothesis tests. In order to account for this problem, we implemented a studentized Breusch-Pagan test for heteroscedasticity (see Appendix A, Fig. 4). Since we have a generated p-value of 2.2e-16, which is significantly less than .05, we confidently rejected our null hypothesis that the data was homoscedastic and saw that there was significant evidence to suggest that our model was heteroscedastic in nature. Furthermore, when we created a scatterplot of our residual data, we could see a clear heteroscedastic relationship between our dependent and independent variables (see Appendix A, Fig 5).

In order to combat this problem, we initially took the log values of our linear model's dependent and independent variables. We transformed our coefficients from numerical values to percent change interpretation. Immediately, we noticed that the log model was able to explain over 99% of the variability in price with the variability of square feet, noise, crime, and median income (see Appendix A, Fig.6 & Fig. 7). Furthermore, the log-model had an observed F-statistic of 5.189e5 on 4 and 2607 degrees of freedom, and an F-critical value of 2.375, at the 5% significance level; since our F-observed was larger than our F-critical, all the coefficients in our model were indisputably not zero.

All of our coefficients, except for noise, had significance at the 1% level. Moreover, the log of square feet and median income both had exceptionally large t-values, roughly 68 and 66 respectively, while crime had a significant t-value of -7.4. Although noise had a relatively strong t-value of 3.1 and significance at the 5% level, its sign no longer met our prior expectation. As anticipated, square footage and median income maintained a positive relationship with price, while crime still held negatively constant. Noise, however, flipped and became positive; suggesting that an increase in noise complaints correlates to some degree of added value in terms of a property's estimated price. This result was not agreeable amongst us, and we immediately sought for a more robust solution.

Robust Standard Errors

Although taking the log values of our dependent and independent variables seemed to adjust for our data's heteroscedasticity, our priori expectation for noise was not met; we did want to merely accept that its relationship with price became positive, as high noise in a neighborhood should be a negative externality on property values. As a consequence, we decided to calculate the robust standard errors of our model, in order to combat the issue of heteroscedastic data. Most interesting was that this method allowed for us to maintain all prior slope coefficient values and priori expectations, while giving us a more accurate read of the standard errors and t-stats (see Appendix A, Fig. 8). Upon looking at the newly adjusted t-values, we saw that square feet decreased most drastically to a t-value of 21.5, juxtaposed to a value 60 in our heteroscedastic linear model. Comparatively, noise and crime experienced marginal changes in their t-stats, and median income fell to a t-value of 9. These results indicate that, upon adjusting for heteroscedasticity, each slope coefficient was not as significant as we initially believed. Nevertheless, all coefficients still maintained significance above the 1% level.

Another perceived benefit to this procedure is that our F-observed remained unchanged at its previous value of 2020 on 4 and 2607 degrees of freedom, which is larger than the F-critical value of 2.375, at a significance level of 5%; as a result, we can see that our overall model, with calculated robust standard errors, maintains its overall significance. Furthermore, we see that square feet, noise, crime and median income still explain over 77.5% of the variability in property price within the hedonic valuation study.

An Important Note:

With these models, we did run into a notable problem. Since we had an exceptionally large data set with 3920 total cross-sectional observations, amongst our twenty-four randomly selected zip codes, we were unable to account for and omit all N/As for square footage, as this information is not always available in the database. This ultimately, affected our degrees of freedom in some of our regressions, as RStudio automatically omitted them from the results upon performing regression. Regardless, it is important that all models still maintained a high degree of statistical significance.

Coefficients and Price Forecasting – A Field Test (Hedonic Valuation Function vs Log Model):

After accounting for heteroscedasticity by calculating the robust standard errors and preforming log transformations, were able to use our two separate models to forecast individual residential real estate properties' prices throughout all twenty-four of our randomly selected New York City neighborhoods.

Post regression, our hedonic valuation model's slope coefficients proved to display interesting relationships between our independent variables and property price. In regard to square feet, we saw that for every one square foot increase of a subject site's size, the price of said residential property should increase in value by approximately $1790. It is also important to note that for every increase in average noise complaint frequency, a property will experience a discount of -$985. Similarly, for crime, we now understand that for every one-unit increase in a neighborhood's average number of violent offenses, a property will decrease in value by $6888. This finding was of great value to us, as it confirmed our expectation that crime proves to be more detrimental to a property's overall valuation than noise. Finally, we saw that every $1 increase in a neighborhood's median household income led to a $9.73 increase in property value.

As for our log model it is important to note that the coefficients are now to be interpreted in terms of percent change. As a result, we see that for every 1% increase in square feet, we have a .99% increase in a property's price. Similarly, for every 1% increase in median income, property value increases by .67%. As for crime, we see that every for 1% increase in average criminal offense results in a decrease of -.16% in a property's valuation. The noise coefficient, however, violated our priori expectation. We believed that noise would impose a negative value on price, yet in this model, we saw that for every 1% increase in average noise complaints, a property experienced a .05% increase in value. Perhaps this is because taking logs of the data intensified some degree of multicollinearity between noise and crime, thus lowering noise's significance level and relevance in this model.

After gaining a better understanding of our two models, we used them in harmony to price a randomly selected property in each of our twenty-four neighborhoods. This process was done by inputting a subject property's square footage, along with its respective neighborhood's average noise, crime, and median income levels into the hedonic valuation model. We then compiled the data and compared the models' valuation results (See Appendix A, Fig 9a and Fig 9b).

After calculating for each neighborhoods' properties' estimated prices, we then took all forecasted values from the hedonic regression model, and compared them to the those of the log model's; this process was done by taking the differences between the two sets of generated values (See Appendix A, Fig 10).

We were thus able to see some interesting results between the two valuation models. The hedonic valuation model seemed to be the most well rounded estimator, maintaining accuracy when pricing properties throughout Manhattan; while still providing fairly tight price estimations amongst almost all other subject sites in each of the three respective boroughs. Conversely, our log function seemed to price exceptionally well in Queens, and some select neighborhoods throughout Brooklyn, while losing a great degree of accuracy for the Manhattan properties.

ANOVA Neighborhood Model Results

Our ANOVA model calculated the specific values for each of our respective twenty-four neighborhood indicator variables (See Appendix A, Fig. 11). The model's observed F-statistic was 136.2, which is far greater than the critical value of 1.498 at the 5% significance level; thus, we can presume that all indicator variables are not zero. Moreover, upon examining the adjusted R-squared, we see that variability in location by neighborhood explains approximately 47% of the variability in property values across the twenty-four neighborhoods in our study. As this is a cross-section model, we consider this relationship to be satisfactory. Most of the generated dummies gave us a general sense of locational price variability that is fairly consistent with our priori expectations; neighborhoods with higher incidents of noise complaints and crime, coupled with lower median income, tended to have negative valuations. To substantiate these findings, we correlated all of the results for our relative neighborhood indicator coefficients with their respective noise, crime, and median income data included in the hedonic valuation model. (See Appendix A, Fig. 12). We found that value has a negative association with noise (-.29) and crime (-.55), while also maintaining a positive relationship with median household income (.68). These observations confirm statistically that our findings agree with our priori expectations.

We view these indicator variables as the first step in accounting for geographic variation in this respective real estate model. Although we do not view the ANOVA's valuations as the optimal means of performing geospatial statistical analysis, we believe that the 13 out of 24 neighborhoods with significance levels below the 1% level may indicate that there are unique real estate markets within these localities. Although we consider these neighborhoods to be submarkets of the overall area of study, we also acknowledge that in order to fully justify our claims, further research will be necessary.

Time Series Results

Neighborhood Group Results

Brooklyn Group

We first ran the Brooklyn group model, whereby we derived the $1 effect of a change in the median residential value of Flatbush, Brooklyn on the median residential value of Park Slope, Brooklyn over the years 2003-2016 (See Appendix B, Fig 5). We found the Flatbush coefficient to be of the correct, positive sign, indicating that residential price changes in the higher crime, less valuable neighborhood of Flatbush do in fact have positive price effects on the median residential price changes in the more valuable neighborhood Park Slope. It thus appears that a $1 increase in the median price of Flatbush, Brooklyn leads to a $1.62 increase in the median price of Park Slope, Brooklyn. This coefficient is highly statistically significant, with a t-statistic of 17.38; far greater than the critical value of 1.976 for the 5% significance level. The F-statistic for overall model is 302.4 which is far greater than the critical value of 3.9 for the 5% significance level, indicating that all coefficients in the model are not zero. The serial correlation between the median residential value of Flatbush and Park Slope is 0.81. Furthermore, we see that variability in the median residential value trends of Flatbush, Brooklyn explains 66% of the variability in the median residential value of Park Slope. The other 44% may be explained by other neighborhoods not included in the study and other factors related to Park Slope exclusively, such as gentrification and demographic changes within the neighborhood. A detailed plot of the two price trends over the years 2003-2016 is presented in the appendix (See Appendix B, Fig 6). As we can see, for the entirety of the study Flatbush is a less valuable neighborhood than Park Slope. Though both neighborhoods reached their low point at nearly the same time, it appears that after the recessionary downturn, the Park Slope market recovered at a greater rate than the Flatbush market.

Manhattan Group

We next ran the Manhattan model, whereby we derived the $1 effect of a change in the median residential value of Chelsea, Manhattan on the median residential value of Lower East Side, Manhattan over the years 2004-2015 (See Appendix B, Fig 7). We found the Chelsea coefficient to be of the correct, positive sign, indicating that median residential price changes of the western Downtown Manhattan neighborhood do in fact have positive price effects on the median residential price in the eastern Downtown Manhattan neighborhood. It thus appears that a $1 increase in the median price of Chelsea, Manhattan yields a $0.33 increase in the median price of Lower East Side, Manhattan. This coefficient is also highly statically significant with a t-statistic of 33.35; far greater than the critical value of 1.978 for the 5% significance level. The F-statistic for the overall model is 1112, which is far greater than the critical value of 3.9 for the 5% significance level, indicating that all coefficients in the model are not zero. The serial correlation between the median residential value trends of Chelsea and Lower East Side is 0.94. Furthermore, we see that variability in the median residential value of Chelsea, Manhattan explains 89% of the variability in the median residential value of the Lower East Side. The other 11% of the variability in the median residential value of the LES may be related to median price variation in other adjacent eastern neighborhoods not included in the model, as well as neighborhood specific demographic effects. A detailed plot of the two price trends over the years 2004-2015 is presented in the appendix (See Appendix B, Fig 8). As we can see, for the entirety of the study, the Lower East Side is a less valuable area than Chelsea, featuring far less median value growth. Though the Lower East Side seemed to not be as affected by the downturn of the Great Recession as Chelsea, the Chelsea series appeared to recover quickly and then grow at a greater rate than in the pre-recessionary period.

<u>Individual Group Results</u>

Park Slope Median Value Series

We first ran the model whereby we derived the degree of monthly change in the median residential value for Park Slope, Brooklyn over the years 2003-2016 (See Appendix B, Fig 9). We found the Park Slope median value coefficient to be of the correct, positive sign, indicating that the monthly median residential price has been increasing on average throughout the series. It thus appeared that every month in the study period, the median residential price of Park Slope increased by $3530. This coefficient is also highly statistically significant, with a t-statistic of 32.12; far greater than the critical value of 1.976 for the 5% significance level. The F-statistic for the overall model is 1031, which is far greater than the critical value of 3.9 for the 5% significance level, indicating that the slope coefficient of the model is not zero. Furthermore, we see that monthly variability explains 87% of the variability in the median residential value of Park Slope, Brooklyn. We believe part of 13% that is unexplained is due to the linear fit this model uses on the notably non-linear, yet tight trend. Though we will examine this issue greater detail later, a plot of the actual trend and fitted line over the 2003-2016 study is presented in the appendix (See Appendix B, Fig 10).

We next preformed the Chow Test on the median residential value series to see if the Great Recession had a statistically significant structural break in the Park Slope series. As noted before, our breakpoint was June of 2008. We first performed the preliminary test to confirm that the error variances in the two sub-periods were the same. We found an F-statistic of 33.96 which is far greater than the critical value of 1.5 for the 5% significance level; this indicated we have consistent subpopulation error variance and could thus employ the Chow Test. Our computed F-statistic for the actual Chow Test was 63.492, which is far greater than the critical value of 3.05 for the 5% level. We thus confirm that the Great Recession did in fact cause a statistically significant break in the Park Slope series for the month of June in the year 2008 (See ANOVA, Appendix B, Fig 11).

Since we concluded the break exists, we could now fit our sub-period models for the Park Slope median value series; one for 2003-2008 and one for 2008-2016. We found the 2003-2008 sub-period model to have the correct, positive sign, indicating that the monthly median residential price had been increasing on average before the Great Recession (See Appendix B, Fig 12). It thus appeared that every month in the study period leading up to the Great Recession, the median residential price of Park Slope increased by $4322.82 on average. This coefficient is also highly statistically significant, with a t-statistic of 52.83 that is far greater than the critical value of 2.0 for the 5% significance level. The F-statistic for the overall model is 2791; far greater than the critical value of 4.0 for the 5% significance level. Furthermore, we see that monthly variability explains nearly 98% of the variability in the median residential value of pre-recession Park Slope, Brooklyn. Our 2008-2016 sub-period model for Park Slope also appears to have the correct, positive sign, indicating that the monthly median residential price has been increasing on average after the Great Recession (See Appendix B, Fig 13). It appeared that every month in the study period after the Great Recession, the median residential price of Park Slope increased by $4957. This result is interesting when compared with the pre-recession model, because it indicates that median prices in Park Slope have been rising faster in the post-recession period than in the pre-recession period. Perhaps this is because those with capital saw speculative opportunity in the neighborhood during the downturn period. The slope coefficient is also statistically significant for the post-recessionary period, with a t-statistic of 25.81 which is far greater than the critical value of 1.98 for the 5% significance level. The F-statistic for the overall model is 666.1; far greater than the critical value of 3.9 for the 5% significance level. Furthermore, we see that monthly variability in the post-recessionary period explains nearly 87% of the variability in the median residential value of Park Slope, Brooklyn. A piecewise plot of the Park Slope trends for the sub-periods is seen in the appendix (See Appendix B, Fig 14). This plot shows us that the linear fit is more accurate for the pre-recession sub period than the post recession sub period, and this may help explain the difference in the coefficient of variation between the two sub period models.

Returning to the entire 2003-2016 series, we performed the Breusch-Pagan test to examine heteroscedasticity in the data and found a p-value of 2.725e-08%, which is far less than the 5% threshold thus indicating that heteroscedasticity is present in the Park Slope median value series (See Appendix B, Fig 15). We decided to run the same linear model on the series in MATLAB and examined the fit and the residual plot (See Appendix B, Fig 16b). We observed a very clear pattern in the residuals. The linear model underfits the series greatly for July 2005 through November 2008 (corresponding to months 20 to 60 in the series) and underfits it again for November 2011 through September 2016 (corresponding to months 120 through 153 in the series). The linear model overfits the series greatly from November 2008 through November 2011 with the greatest error around March 2012 (month 100 in the series). The wide 95% confidence bound on the fit (as illustrated by the dotted blue lines) further alludes to the incorrect fit of the linear model for this trend (See Appendix B, Fig 16a). We next fit a smoothing spline to the series using a smoothing parameter of 0.9 and re-examined the residual plot and coefficient of variation (See Appendix B, Fig 17a). Upon examining the residual plot for the spline, we saw almost no pattern in the residuals, with much smaller errors (See Appendix B, Fig 17b). The coefficient of variation for this fit increases from .87 in the linear model to nearly .99, indicating that when the series is correctly fit, monthly variability explains nearly 99% of the variability in the median residential value of Park Slope, Brooklyn (See Appendix B, Fig 18).

We next ran our 1 month lag-model on the 2003-2016 median value Park Slope series to examine the degree of autocorrelation in the data (See Appendix B, Fig 19). We found there to be a very strong degree of autocorrelation in the series, with a $1 increase in the previous month's median value leading to a 1$ increase in the following month's median value. This coefficient was highly statistically significant, with a t-value of 231, far greater than the critical value of 1.97 for the 5% significance level. The F-statistic for the lag-model was 5365 which is far greater than the critical value of 3.9 for the 5% significance level. Furthermore, the median price variability in the previous month explains nearly 99% of the variation in the median price of the following month for the span of the series. Thus, we can conclude that there is a strong degree of serial momentum in the data and that increases in one month clearly contribute to increases in the next. For a visual representation of the autocorrelation in the series, refer to the lag-plot for the Park Slope median price data (Appendix B, Fig 20)

We then ran our ANOVA model on the Park Slope data to test for monthly seasonality in the series. We found none of the monthly indicator coefficients to be statistically significant, and the F-statistic for the model was a mere 0.08, far less than the critical value of 1.86, indicating that all of the indicator coefficients are zero. Furthermore, it appeared that variability in months explains only 0.007% of the variability in median price over the span of the study. Thus we conclude that there is no monthly seasonality in the series, which indicates that there is a relatively uniform level of demand for properties in Park Slope throughout the year (See Appendix B, Fig 21).

Chelsea Median Value Series

We next ran the model whereby we derived the degree of monthly change in the median residential value for Chelsea, Manhattan over the years 2004-2015 (See Appendix B, Fig 22). We found the Chelsea median value coefficient to be of the correct, positive sign, indicating that the monthly median residential price has been increasing on average throughout the series. It thus appeared that every month in the study period, the median residential price of Chelsea increased by $6664. This coefficient is also highly statistically significant, with a t-statistic of 14.73; far greater than the critical value of 1.97 for the 5% significance level. The F-statistic for the overall model is 216.9, which is far greater than the critical value of 3.91 for the 5% significance level, indicating that the slope coefficient of the model is not zero. Furthermore, we see that monthly variability explains 61% of the variability in the median residential value of Chelsea, Manhattan. We believe part of 39% that is unexplained is due to the linear fit this model uses on the notably non-linear, yet tight trend. Though we will examine the issue of the fit greater detail later, a plot of the actual trend and fitted line over the 2004-2015 study is presented in the appendix (See Appendix B, Fig 23). Other portions of the unexplained variability may be related to temporal changes in neighborhood relevant demographic variables and median price variability of other adjacent Manhattan neighborhoods not included in the model.

We next preformed the Chow Test on the median residential value series to see if the Great Recession had a statistically significant structural break in the Chelsea series. As noted before, our breakpoint was June of 2008. We first performed the preliminary test to confirm that the error variances in the two sub-periods were the same. We found an F-statistic of 8.99 which is greater than the critical value of 1.54 for the 5% significance level; this indicated we have consistent subpopulation error variance and could thus employ the Chow Test. Our computed F-statistic for the actual Chow Test was 37.58, which is far greater than the critical value of 3.06 for the 5% level. We thus confirm that the Great Recession did in fact cause a statistically significant break in the Chelsea series for the month of June in the year 2008 (See ANOVA, Appendix B, Fig 24).

Since we concluded the break exists, we could now fit our sub-period models for the Chelsea median value series; one for 2004-2008 and one for 2008-2015. We found the 2004-2008 sub-period model to have the correct, positive sign, indicating that the monthly median residential price had been increasing on average before the Great Recession (See Appendix B, Fig 25). It thus appeared that every month in the study period leading up to the Great Recession, the median residential price of Chelsea increased by $17466.5 on average. This coefficient is also highly statistically significant, with a t-statistic of 26.50 that is far greater than the critical value of 2.0 for the 5% significance level. The F-statistic for the overall model is 702.2; far greater than the critical value of 4.0 for the 5% significance level. Furthermore, we see that monthly variability explains nearly 93% of the variability in the median residential value of pre-recession Chelsea, Manhattan. Our 2008-2015 sub-period model for Chelsea also appears to have the correct, positive sign, indicating that the monthly median residential price has been increasing on average after the Great Recession (See Appendix B, Fig 26). It appeared that every month in the study period after the Great Recession, the median residential price of Chelsea increased by $9434.7. This result is interesting when compared with the pre-recession model, because it indicates that median prices in Chelsea have been rising slower in the post-recession period than in the pre-recession period. This may in part be due to the fact that the Chelsea series remained stagnant for 3 years from December of 2009 to June of 2012 (Corresponding to months 70 through 100 in the series) only to begin growing at a greater rate after June of 2012 (See Appendix B, Fig 23). The slope coefficient is also statistically significant for the post-recessionary period, with a t-statistic of 10.63 which is far greater than the critical value of 1.98 for the 5% significance level. The F-statistic for the overall model is 113.1; far greater than the critical value of 3.95 for the 5% significance level. Furthermore, we see that monthly variability in the post-recessionary period explains 57% of the variability in the median residential value of Chelsea, Manhattan. A piecewise plot of the Chelsea trends for the sub-periods is seen in the appendix (See Appendix B, Fig 27). This plot shows us that the linear fit is far more accurate for the pre-recession sub period than the post-recession sub period, and this may help explain the difference in the coefficient of variation between the two sub period models. In order to account for the lack of accuracy of the linear fit for the post-recessionary period, we decided to run a new fit on the 2008-2015 sub-period in MATLAB (See Appendix B, Fig 28). We found a second degree polynomial function captured the trend more accurately as the median price series seemed to remain stagnant from Dec 2009 to June of 2012 and then increase more rapidly after June of 2012. The new polynomial fit showed that monthly variability in the post-recessionary period explains nearly 99% of the variability in the median residential value of Chelsea, Manhattan. (See Appendix B, Fig 29).

Returning to the entire 2004-2015 series, we performed the Breusch-Pagan test to examine heteroscedasticity in the data and found a p-value of 1.502e-05%, which is far less than the 5% threshold thus indicating that heteroscedasticity is present in the Chelsea median value series (See Appendix B, Fig 30). We decided to run the same linear model on the series in MATLAB and examined the fit and the residual plot (See Appendix B, Fig 31b). We observed a very clear pattern in the residuals, similar to that seen in the Park Slope series. The linear model underfits the series greatly for Oct 2005 through Feb 2009 (corresponding to months 20 to 60 in the series) and underfits it again for Feb 2014 through Aug 2015 (corresponding to months 120 through 138 in the series). The linear model overfits the series greatly from Feb 2009 through Feb 2014 with the greatest error around June 2012 (month 100 in the series). The wide 95% confidence bound on the fit (as illustrated by the dotted blue lines) further alludes to the incorrect fit of the linear model for this trend (See Appendix B, Fig 31a). We next fit a smoothing spline to the series using a smoothing parameter of 0.9 and re-examined the residual plot and coefficient of variation (See Appendix B, Fig 32a). Upon examining the residual plot for the spline, we saw far less of a pattern in the residuals compared to the nearly sinusoidal pattern in the linear fit. Smaller errors were generally observed. (See Appendix B, Fig 32b). The coefficient of variation for this fit increases from .61 in the linear model to nearly .99, indicating that when the series is correctly fit, monthly variability explains nearly 99% of the variability in the median residential value of Chelsea, Manhattan (See Appendix B, Fig 33).

We next ran our 1 month lag-model on the 2004-2015 median value Chelsea series to examine the degree of autocorrelation in the data (See Appendix B, Fig 34). We found there to be a very strong degree of autocorrelation in the series, with a $1 increase in the previous month's median value leading to a 1$ increase in the following month's median value. This coefficient was highly statistically significant, with a t-value of 175, far greater than the critical value of 1.97 for the 5% significance level. The F-statistic for the lag-model was 3064, which is far greater than the critical value of 3.9 for the 5% significance level. Furthermore, we see that median price variability in the previous month explains nearly 99% of the variation in the median price of the following month for the span of the series. Thus, we can conclude that there is a strong degree of serial momentum in the data and that increases in one month clearly contribute to increases in the next. For a visual representation of the autocorrelation in the series, refer to the lag-plot for the Chelsea median price data (Appendix B, Fig 35)

We then ran our ANOVA model on the Chelsea data to test for monthly seasonality in the series. We found none of the monthly indicator coefficients to be statistically significant, and the F-statistic for the model was a mere 0.3, far less than the critical value of 1.87, indicating that all of the indicator coefficients are zero. Furthermore, it appeared that variability in months explains only 0.03% of the variability in median price over the span of the study. Thus we conclude that there are no months that are exceptionally above or below average for median value in the series, indicating that there is a relatively uniform level of demand for properties in Chelsea throughout the year with respect to the span of this study (See Appendix B, Figure 36).

Acknowledgments and Limitations

Variable Choice

Due to the time constraints our study faced, we were not able to parse through all the possible factors that could influence the price of a property. However, we attempted to include variables that we believed were the most general contributors to value; such as square footage, noise frequency, crime level, and neighborhood median household income. Ultimately, we acknowledge that other variables, such doorman, elevator, rooftop, in-house laundry, gym and general proximity to cultural and transportation hubs, could also be significant contributors to price. In the future, our continued research will include more of these various factors.

Causation vs. Correlation

In utilizing OLS regression, we have shown household median income levels can explain the variability in residential property prices; our study was based on pure correlation between factors. In order to prove the direction of causality, one would have to conduct a real human social-experiment. This procedure is a great challenge, and can result in inconclusive results. However, this endeavor could ultimately prove to be interesting research topic as it stands; do wealthier neighborhood demographics cause higher property prices, or is the opposite true?

Noise Pollution

In regard to formal noise complaints, we acknowledge that in our 311 dataset there is a potential for bias. There is a potential that various neighborhoods have differing thresholds for noise level tolerance. As a result, they could also differ in their propensities to file formal 311 complaints. This variability in preferences may have ultimately presented itself in our dataset, potentially skewing our estimates.

Adjustments for Seasonality

For our adjustments, we chose to lag our time series data in monthly terms, in order to keep it consistent with its originally recorded resolution. Adjusting by transforming the series into quarterly terms, for example, could have potentially reduced the quality of our estimates; this process could have smoothed out potentially insightful price trends.

Summary and Conclusions

Cross Sectional Conclusions

Our OLS hedonic valuation model underwent multiple transformations. In terms of our cross-sectional data, we found that multicollinearity between bathrooms and square footage of properties proved to be the first challenge in developing an accurate hedonic valuation model. To account for this issue, we removed the variable for baths from our initial model, as it proved to be insignificant, thus leaving solely square feet, noise, crime, and median household income to be regressed on the property price data. We then had to test and account for potential heteroscedasticity in the dataset. In order to accomplish this feat, we opted to implement a Breusch-Pagan Test, and discovered that our data indeed needed to be adjusted for; we approached this procedure using various methods. We initially transformed our data into log terms, and then subsequently went back to our prior linear model and calculated for its robust standard errors. Through running our own "field test", we used both the log and adjusted hedonic valuation model to simultaneously estimate and price a randomly sampled series of twenty-four individual properties, across all of our respective zip codes in the study. The results were intriguing, and ultimately led us to deem the adjusted hedonic valuation model to be the superior means of price estimation; it had a seemingly strong overall ability to value properties across an eclectic range of neighborhoods, whereas the log model simply underperformed.

In comparing the remaining variables in our model with those of Harrison and Rubinfeld's *Hedonic Housing Prices and Demand for Clean Air*, it is evident that our findings for the significance and signs of square feet, crime, and demographic data confirm the results of Harrison and Rubinfeld's early 1976 study. Moreover, if we observe Jon P. Nelson's 1982 *Highway Noise and Property Values,* it becomes apparent that he found higher neighborhood noise levels to be a negative externality that had detrimental effects on the prices of real estate assets. These past studies help serve as a means of confirming our beliefs and expectations that our use of square feet, noise, crime, and median household income, as variables, should sufficiently explain variation in residential real estate prices, as they have proven to be crucial elements of past hedonic pricing research.

As a means of accounting for the geographic nature of property price variability we developed a separate ANOVA model that employed indicator variables for all 24 neighborhoods in our hedonic valuation study as well as 2 indicator variables for property types. The results of the model correlated correctly with our expectations regarding the noise, crime and medium income levels for each of the 24 neighborhoods, conforming the validity of the model. The model also identified 13 neighborhoods as potential submarkets of the overall area of study as they were found to be significant in explaining property prices at the 1% significance level. Though we realize that this model is fairly rudimentary as per the prior research of Pace, Barry and Sirmans in their 1998 study *Spatial Statistics and Real Estate*, we believe it was a valid first step in the highly complex task of accounting for the spatial variance in real estate prices.

Time Series Summary

We first ran an OLS model that derived the marginal effect of monthly median value changes in a Brooklyn neighborhood group on the other, and then ran the same model for a Manhattan group.

For the Brooklyn group, we found that over the years 2003-2016, median residential price changes in a higher crime, less valuable neighborhood like Flatbush do indeed have significant effects on the median residential prices changes in a notably more valuable neighborhood like Park Slope; temporal improvements in a lower welfare neighborhood may act a positive externality for a higher welfare neighborhoods within the same borough.

For the Manhattan group, we found that serial dependencies exist between changes in the median residential value of Chelsea in western Downtown Manhattan and changes in such values for the Lower East Side in eastern Downtown Manhattan, with variability in the median residential value series of Chelsea, Manhattan explaining 89% of the variability in the median residential value series for the Lower East Side for the years 2004-2015. The highly concentrated nature of the Downtown Manhattan real estate market may partially explain effects, although further research will be necessary.

We next selected Chelsea from the Manhattan group series and Park Slope from the Brooklyn group series and performed Chow Tests on the datasets, setting our breakpoint at June of 2008. We found for both series that the Great Recession had a statistically significant effect on the trends in the series. We then fit models for the sub-periods. For Park Slope we saw that median prices for residential properties grew faster in the post-crisis period than they did in the pre-crisis period. For Chelsea we found the opposite, but in both cases our results in the post-crisis periods were constrained by the linear fitting techniques. We turned to MATLAB, fitting smoothing splines and polynomial functions to each respective series, to which we saw great improvements in our coefficient of variation. Then we employed lag-models for both series and noticed that there was a great degree of serial momentum in these two neighborhood's median value trends; changes in the previous month's median price trend explained 99% of the changes in the following month's. Finally, we tested each data for seasonality using monthly indicator variables and found no monthly seasonality in either of the series.

Appendices

Appendix A — Cross Sectional Results

Figure 1a: Hedonic pricing model data scattergram

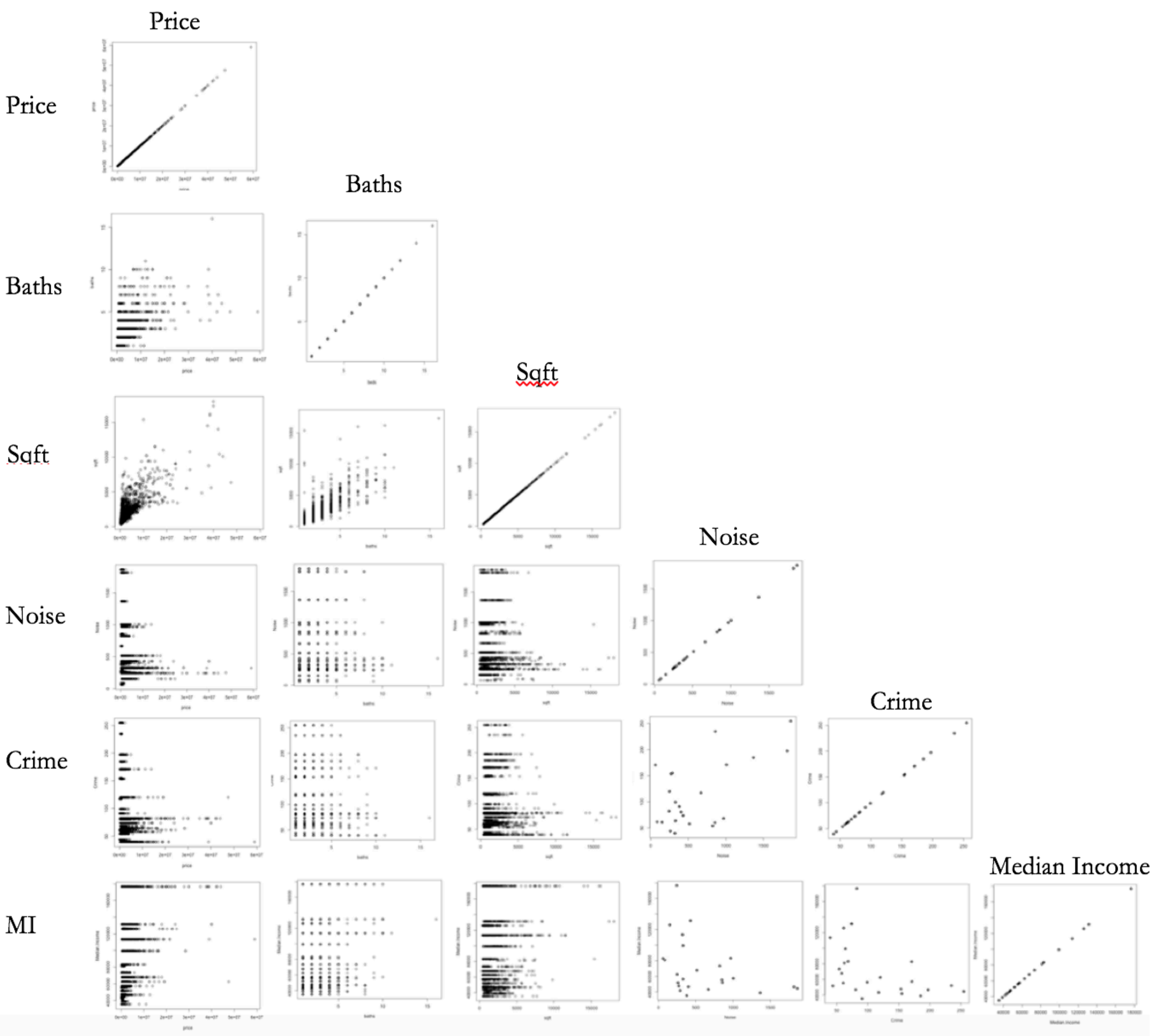


Figure 1b: Hedonic pricing model data correlation matrix

	Price	Bath	Sqft	Noise	Crime	MI
Price	1					
Bath	0.62	1				
Sqft	0.76	0.82	1			
Noise	0.16	0.04	0.04	1		
Crime	0.18	0.01	0.10	0.71	1	
MI	0.30	0.10	0.18	0.48	0.45	1

Figure 2: Linear hedonic model results (with baths included)

```
Call:
lm(formula = price ~ 0 + baths + sqft + Noise + Crime + Median.
    data = d)

Residuals:
      Min        1Q    Median        3Q       Max
-16040960  -1083772   -198952    716849  36694060

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
baths         72216.578  59123.433   1.221    0.222
sqft           1726.419     54.358  31.760  < 2e-16 ***
Noise          -929.951    180.382  -5.155 2.72e-07 ***
Crime         -6986.720   1259.092  -5.549 3.17e-08 ***
Median.Income     9.416      0.938  10.038  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2663000 on 2577 degrees of freedom
  (1338 observations deleted due to missingness)
Multiple R-squared:  0.7492,    Adjusted R-squared:  0.7487
F-statistic:  1540 on 5 and 2577 DF,  p-value: < 2.2e-16
```

Figure 3: Linear hedonic model results (without baths)

```
Call:
lm(formula = price ~ 0 + sqft + Noise + Crime + Median.Income,
    data = d)

Residuals:
      Min        1Q    Median        3Q       Max
-16965149  -1089652   -179780    737884  36622956

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
sqft          1790.8406    29.8241  60.047  < 2e-16 ***
Noise         -958.8093   179.0676  -5.354 9.33e-08 ***
Crime        -6888.7712  1254.0212  -5.493 4.33e-08 ***
Median.Income    9.7348     0.8872  10.972  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2677000 on 2607 degrees of freedom
  (1309 observations deleted due to missingness)
Multiple R-squared:  0.7561,    Adjusted R-squared:  0.7557
F-statistic:  2020 on 4 and 2607 DF,  p-value: < 2.2e-16
```

Figure 4: Breusch-Pagan results for linear hedonic pricing model

```
        studentized Breusch-Pagan test

data:  modela
BP = 335.48, df = 4, p-value < 2.2e-16
```

Figure 5: Heteroscedasticity patterns in linear hedonic pricing model

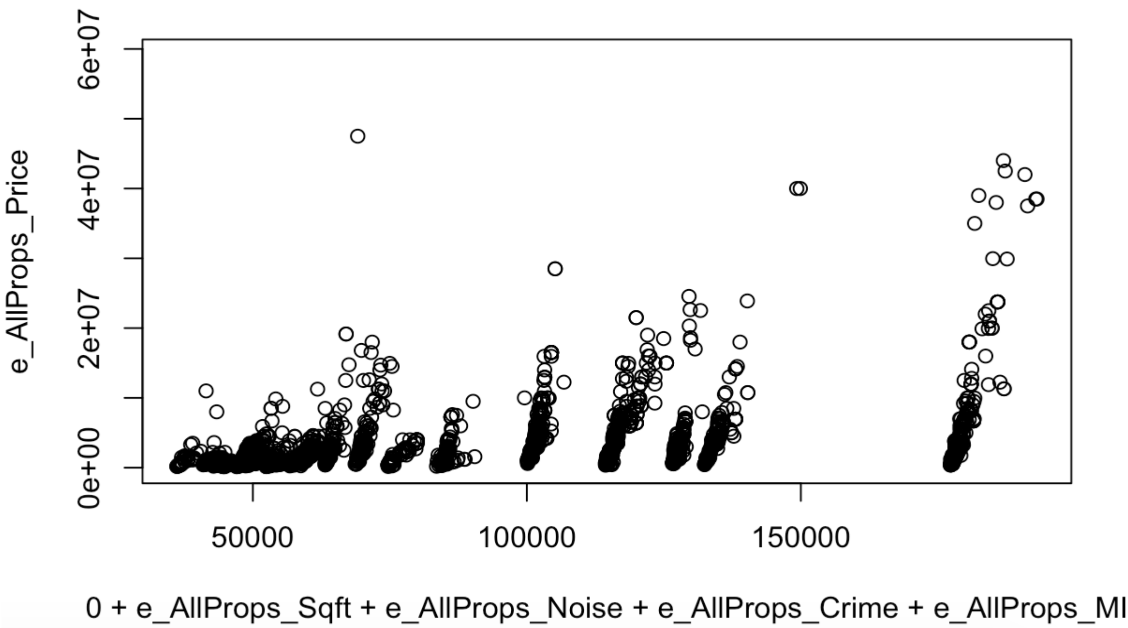


Figure 6: Log-transformed hedonic pricing model results

```
Call:
lm(formula = log(price) ~ 0 + log(sqft) + log(Noise) + log(Crime) +
    log(Median.Income))

Residuals:
    Min      1Q  Median      3Q     Max 
-2.0289 -0.3267  0.0701  0.3605  2.0667 

Coefficients:
                   Estimate Std. Error t value Pr(>|t|)    
log(sqft)           0.99552    0.01449  68.697  < 2e-16 ***
log(Noise)          0.05220    0.01651   3.162  0.00159 ** 
log(Crime)         -0.16561    0.02228  -7.433 1.43e-13 ***
log(Median.Income)  0.67024    0.01015  66.035  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.5125 on 2607 degrees of freedom
  (1309 observations deleted due to missingness)
Multiple R-squared:  0.9987,	Adjusted R-squared:  0.9987 
F-statistic: 5.189e+05 on 4 and 2607 DF,  p-value: < 2.2e-16
```

Figure 7: Log-transformed hedonic pricing model scatterplot

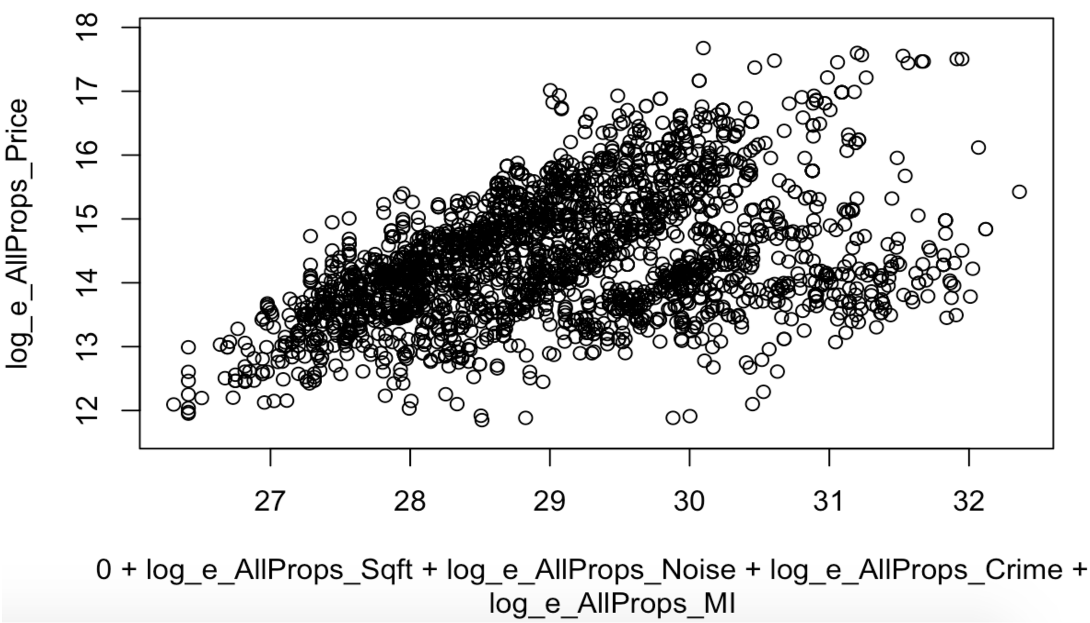


Figure 8: Robust standard error adjusted t-stats for linear hedonic pricing model

```
t test of coefficients:

                Estimate Std. Error t value  Pr(>|t|)
sqft           1790.8406    83.0846 21.5544 < 2.2e-16 ***
Noise          -958.8093   194.1404 -4.9387 8.360e-07 ***
Crime         -6888.7712  1185.4729 -5.8110 6.966e-09 ***
Median.Income     9.7348     1.0778  9.0324 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figure 9a: 24 property price estimations using linear hedonic model

	Hedonic Valuation Model (adjusted for Heteroscedasticity)				
	Neighborhood	**Listed Price**	**Hedonic Valuation**	**Difference**	**% Change**
Manhattan	Central Harlem	1421000	1578181	-157181	-0.11
	Chelsea	2395000	2633875	-238875	-0.10
	FIDI	5285000	4963032	321968	0.06
	Gramercy Park	2650000	3323276	-673276	-0.25
	LES	2315000	1920239	394761	0.17
	Soho	3000000	3025082	-25082	-0.01
	UES	5500000	5397276	102724	0.02
	UWS	4995000	4822758	172242	0.03
Brooklyn	Bay Ridge	260000	428368	-168368	-0.65
	Brooklyn Heights	1999000	2838326	-839326	-0.42
	Brownsville	1100000	3791923	-2691923	-2.45
	Bushwick	995000	2629222	-1634222	-1.64
	Canarsie	4995000	4883270	111730	0.02
	Coney Island	579000	764032	-185032	-0.32
	Crown Heights	3449000	3240533	208467	0.06
	DUMBO	5995000	5821424	173576	0.03
	Flatbush	829000	359708	469292	0.57
	Park Slope	780000	894668	-114668	-0.15
	Williamsburg	1825000	753269	1071731	0.59
Queens	Astoria	999000	1582368	-583368	-0.58
	Bayside	1798000	4760715	-2962715	-1.65
	Corona	249000	168508	80492	0.32
	Forest Hils	1679000	3582413	-1903413	-1.13
	South Jamaica	479000	1808457	-1329457	-2.78

Figure 9b: The same 24 property price estimations using log-transformed hedonic model

	Log Model				
	Neighborhood	**Listed Price**	**Log Valuation**	**Difference**	**% Change**
Manhattan	Central Harlem	1421000	1468191	-47191	-0.03
	Chelsea	2395000	1899795	495205	0.21
	FIDI	5285000	4009593	1275407	0.24
	Gramercy Park	2650000	2021518	628482	0.24
	LES	2315000	1919308	395692	0.17
	Soho	3000000	2313047	686953	0.23
	UES	5500000	5103868	396132	0.07
	UWS	4995000	4174947	820053	0.16
Brooklyn	Bay Ridge	260000	421419	-161419	-0.62
	Brooklyn Heights	1999000	2438062	-439062	-0.22
	Brownsville	1100000	2582285	-1482285	-1.35
	Bushwick	995000	1912069	-917069	-0.92
	Canarsie	4995000	2601929	2393071	0.48
	Coney Island	579000	673203	-94203	-0.16
	Crown Heights	3449000	2638078	810922	0.24
	DUMBO	5995000	5352840	642160	0.11
	Flatbush	829000	1449294	-620294	-0.75
	Park Slope	780000	668892	111108	0.14
	Williamsburg	1825000	463746	1361254	0.75
Queens	Astoria	999000	1063970	-64970	-0.07
	Bayside	1798000	3003777	-1205777	-0.67
	Corona	249000	511921	-262921	-1.06
	Forest Hils	1679000	2023662	-344662	-0.21
	South Jamaica	479000	1268291	-789291	-1.65

Figure 10: Property price estimation difference between log model and linear model

		Difference Between Hedonic and Logged Models			
	Neighborhood	**Hedonic**	**Log**	**Difference**	**% Change**
Manhattan	Central Harlem	1578181	1468191	109990	0.07
	Chelsea	2633875	1899795	734080	0.28
	FIDI	4963032	4009593	953439	0.19
	Gramercy Park	3323276	2021518	1301758	0.39
	LES	1920239	1919308	931	0.00
	Soho	3025082	2313047	712035	0.24
	UES	5397276	5103868	293408	0.05
	UWS	4822758	4174947	647811	0.13
Brooklyn	Bay Ridge	428368	421419	6949	0.02
	Brooklyn Heights	2838326	2438062	400264	0.14
	Brownsville	3791923	2582285	1209638	0.32
	Bushwick	2629222	1912069	717153	0.27
	Canarsie	4883270	2601929	2281341	0.47
	Coney Island	764032	673203	90829	0.12
	Crown Heights	3240533	2638078	602455	0.19
	DUMBO	5821424	5352840	468584	0.08
	Flatbush	359708	1449294	-1089586	-3.03
	Park Slope	894668	668892	225776	0.25
	Williamsburg	753269	463746	289523	0.38
Queens	Astoria	1582368	1063970	518398	0.33
	Bayside	4760715	3003777	1756938	0.37
	Corona	168508	511921	-343413	-2.04
	Forest Hils	3582413	2023662	1558751	0.44
	South Jamaica	1808457	1268291	540166	0.30

Figure 11: Results from 24 neighborhood ANOVA model

```
Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
Apartment           448608     363980   1.233 0.217835
House              4507345     394739  11.419  < 2e-16 ***
Central.Harlem     -283740     495932  -0.572 0.567263
Chelsea            2661554     433276   6.143 8.92e-10 ***
Financial.District 2126375     419248   5.072 4.12e-07 ***
Gramercy.Park      2256614     448253   5.034 5.01e-07 ***
Lower.East.Side    1741677     482936   3.606 0.000314 ***
Soho               4896992     442939  11.056  < 2e-16 ***
Upper.East.Side    4124721     397232  10.384  < 2e-16 ***
Upper.West.Side    2914856     394467   7.389 1.80e-13 ***
Bay.Ridge         -1246720     539295  -2.312 0.020843 *
Brooklyn.Heights   2926434     481566   6.077 1.34e-09 ***
Brownsville       -3792865     817521  -4.639 3.61e-06 ***
Bushwick          -2401982     496465  -4.838 1.36e-06 ***
Canarsie          -3018890     555736  -5.432 5.91e-08 ***
Coney.Island      -1049069     656430  -1.598 0.110092
Crown.Heights     -1320254     502572  -2.627 0.008648 **
DUMBO              2650616     623182   4.253 2.16e-05 ***
Flatbush          -1847197     592075  -3.120 0.001823 **
Park.Slope          386747     439113   0.881 0.378510
Williamsburg        674286     413878   1.629 0.103353
Astoria           -1573842     548561  -2.869 0.004139 **
Bayside           -2473918     764971  -3.234 0.001231 **
Corona            -2063716     902535  -2.287 0.022274 *
Forest.Hills       -444257     436668  -1.017 0.309036
South.Jamaica     -3971812     763492  -5.202 2.07e-07 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3580000 on 3894 degrees of freedom
Multiple R-squared:  0.4763,    Adjusted R-squared:  0.4728
F-statistic: 136.2 on 26 and 3894 DF,  p-value: < 2.2e-16
```

Figure 12: Results of 24 neighborhood ANOVA model with relevant neighborhood attribute data

Neighborhood Indicator Variable	Value	Noise	Crime	Median Income
Central Harlem	-283740	821.6	54	73374.2
Chelsea	2661554	514.6	57.8	67808.7
FIDI	2126375	151.6	61.2	126000.6
Gramercy Park	2256614	248.6	120	62438.2
LES	1741677	968.3	68.2	83155.3
Soho	4896992	327.3	64	99269.8
UES	4124721	242.3	82.2	176557.4
UWS	2914856	321.6	39.6	113601.6
Bay Ridge	-1246720	378	91	35252.1
Brooklyn Heights	2926434	428.6	73.6	131395.6
Brownsville	-3792865	857.6	234.8	52012.8
Bushwick	-2401982	1366	185	38840.9
Canarsie	-3018890	666.6	117.2	43144.4
Coney Island	-1049069	287.6	155	41731.3
Crown Heights	-1320254	1817.3	197.2	46550.5
DUMBO	2650616	428.6	73.6	131395.6
Flatbush	-1847197	1863	254.6	44457.4
Park Slope	386747	402.6	80.4	46800.6
Williamsburg	674286	1004.6	171	57229.1
Astoria	-1573842	327.3	98.6	56381.7
Bayside	-2473918	82.6	61.8	81182.8
Corona	-2063716	852.6	60.4	55356.6
Forest Hils	-444257	260	44	51668.7
South Jamaica	-3971812	60	170.6	82820.5

Appendix B — Time Series Results

Figure 1: Park Slope Median Value

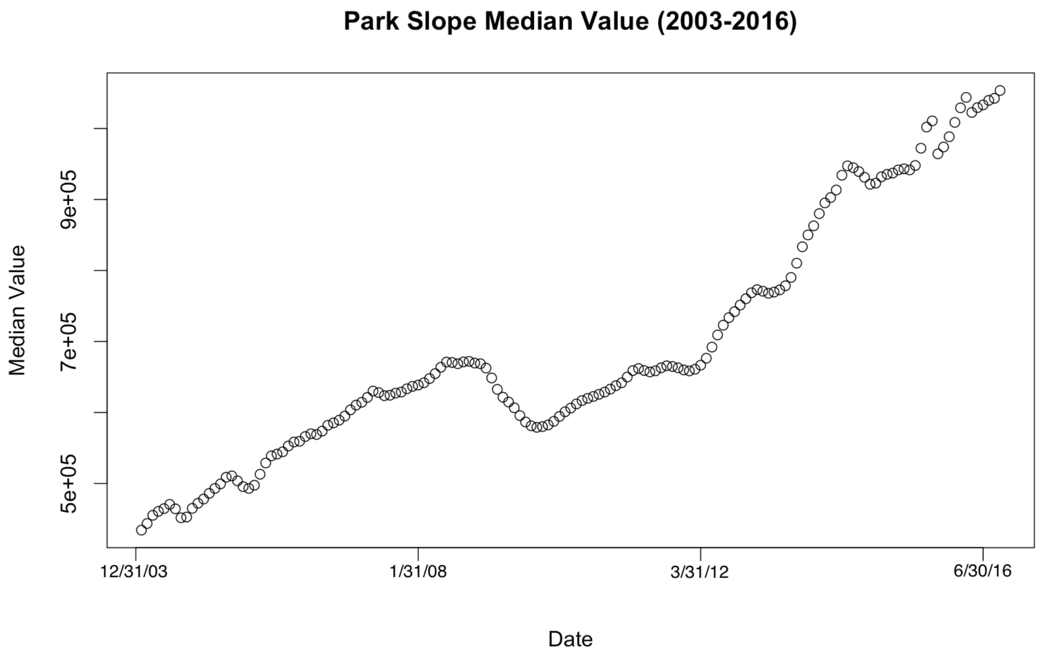


Figure 2: Flatbush Median Value

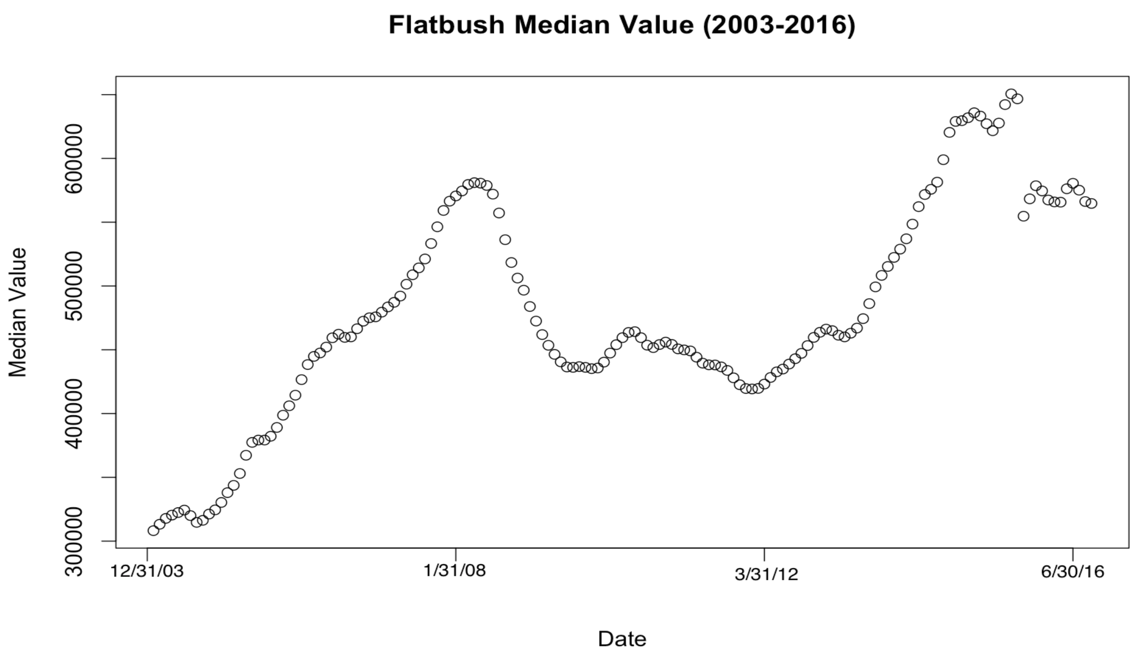


Figure 3: Lower East Side Median Value

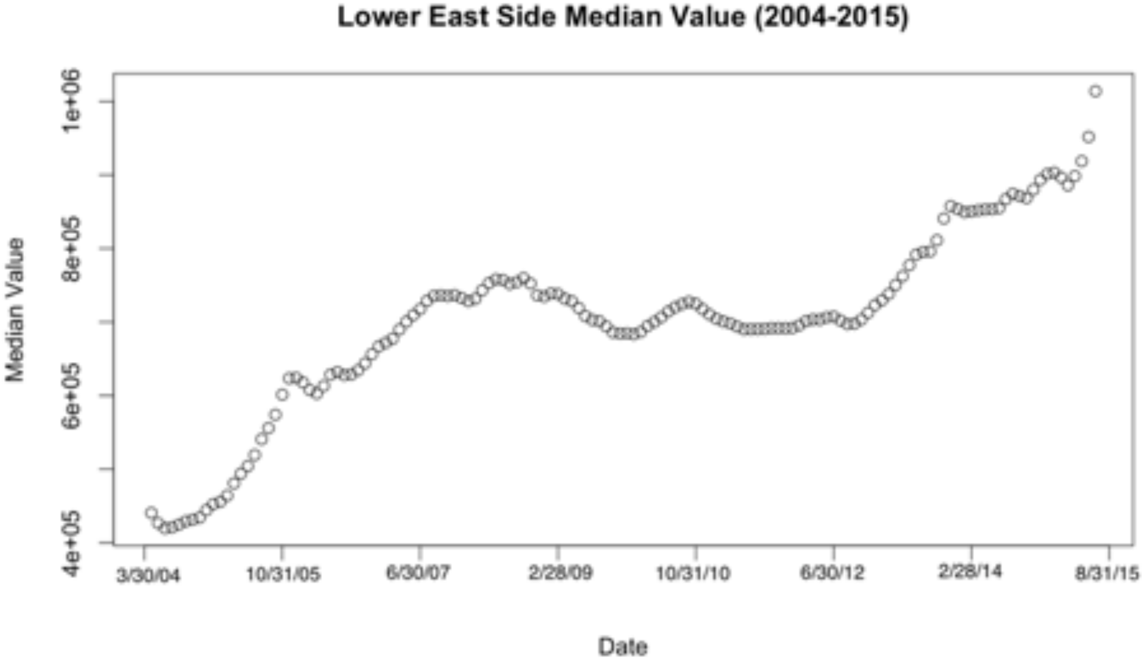


Figure 4: Chelsea Median Value

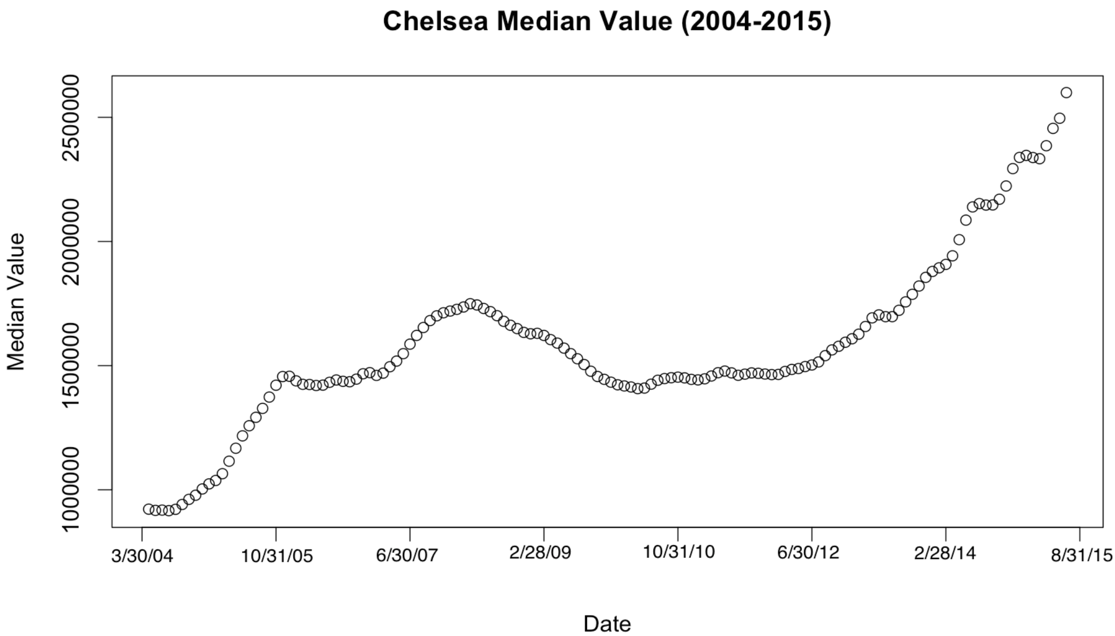


Figure 5: Park Slope on Flatbush median value regression results

```
Call:
lm(formula = Park ~ Flat)

Residuals:
    Min      1Q  Median      3Q     Max
-211306  -55113   -3832   72735  218711

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -8.680e+04  4.560e+04  -1.904   0.0588 .
Flat         1.632e+00  9.385e-02  17.389   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 96980 on 151 degrees of freedom
Multiple R-squared:  0.6669,    Adjusted R-squared:  0.6647
F-statistic: 302.4 on 1 and 151 DF,  p-value: < 2.2e-16

> model1 <- lm(Flat ~ Park)
> summary(model1)
```

Figure 6: Flatbush vs. Park Slope

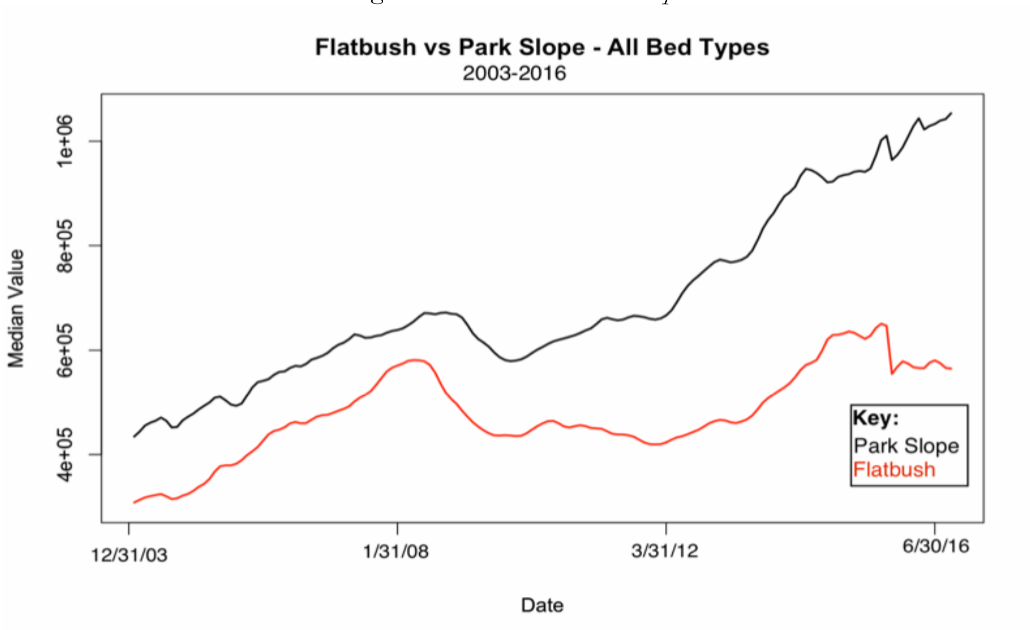


Figure 7: Lower East Side on Chelsea median value regression results

```
Call:
lm(formula = `Lower East Side` ~ Chelsea)

Residuals:
   Min     1Q Median     3Q    Max
-79419 -35391  10602  28913  69399

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 1.667e+05  1.642e+04   10.15   <2e-16 ***
Chelsea     3.386e-01  1.015e-02   33.35   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 40390 on 136 degrees of freedom
Multiple R-squared:  0.8911,    Adjusted R-squared:  0.8903
F-statistic:  1112 on 1 and 136 DF,  p-value: < 2.2e-16
```

Figure 8: Chelsea vs. Lower East Side

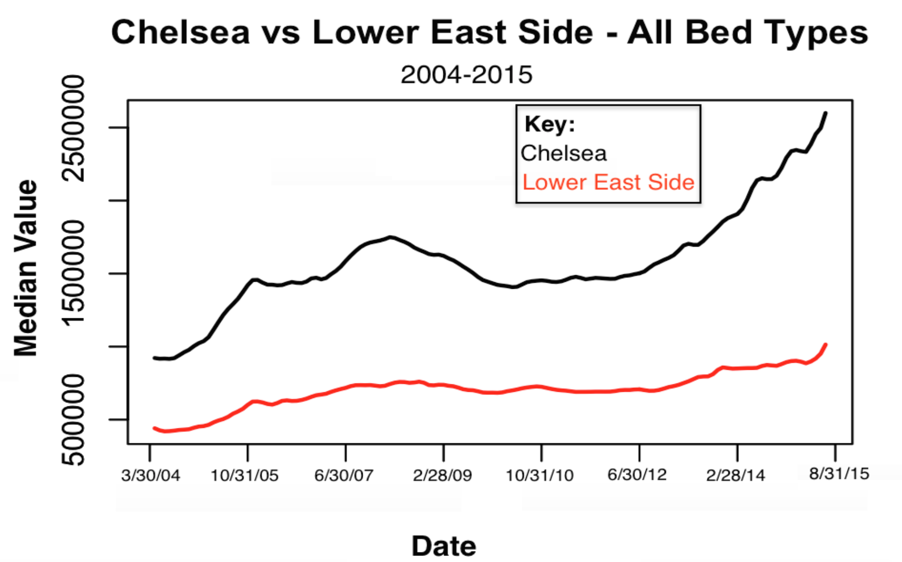


Figure 9: Park Slope median value regression results

```
> model1 <- lm(Park~Date)
> summary(model1)

Call:
lm(formula = Park ~ Date)

Residuals:
    Min      1Q  Median      3Q     Max 
-110994  -54604   24266   42502  102462 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 422413.9     9757.0   43.29   <2e-16 ***
Date          3530.1      109.9   32.12   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 60050 on 151 degrees of freedom
Multiple R-squared:  0.8723,	Adjusted R-squared:  0.8715 
F-statistic:  1031 on 1 and 151 DF,  p-value: < 2.2e-16
```

Figure 10: Park Slope

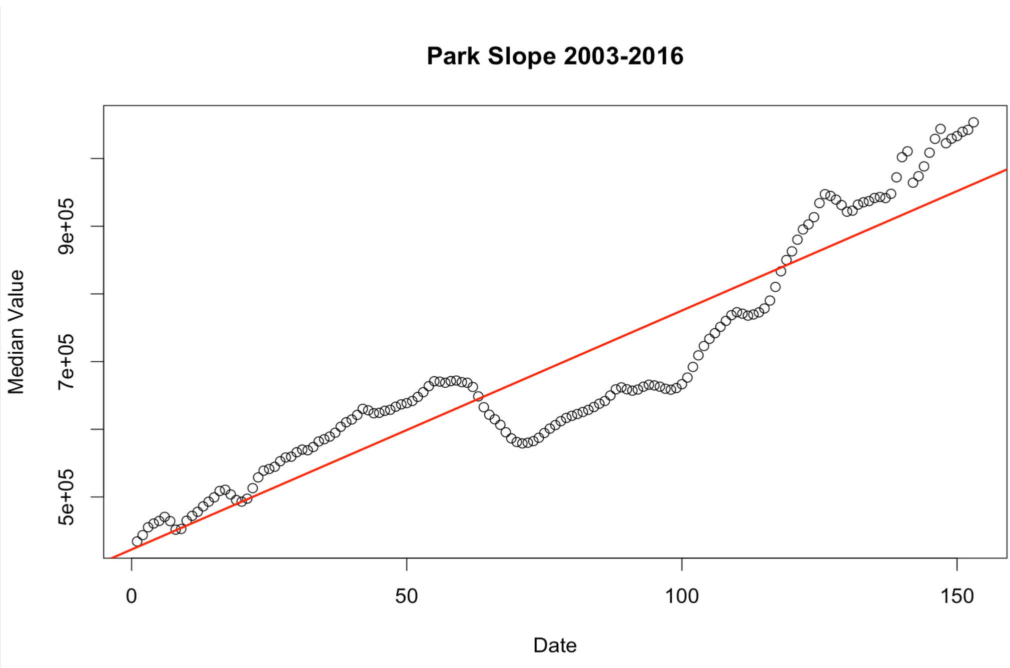


Figure 11: ANOVA for Chow Test — Park Slope Series

```
> anova(modelc)
Analysis of Variance Table

Response: Park_total
            Df     Sum Sq    Mean Sq F value    Pr(>F)
Date_total   1 3.7192e+12 3.7192e+12  1031.4 < 2.2e-16 ***
Residuals  151 5.4448e+11 3.6058e+09
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> model1 <- lm(Park_Rss1~Date_Rss1)
> anova(model1)
Analysis of Variance Table

Response: Park_Rss1
          Df     Sum Sq    Mean Sq F value    Pr(>F)
Date_Rss1  1 2.4512e+11 2.4512e+11    2791 < 2.2e-16 ***
Residuals 52 4.5670e+09 8.7827e+07
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
>
> model2 <- lm(Park_Rss2~Date_Rss2)
> anova(model2)
Analysis of Variance Table

Response: Park_Rss2
          Df     Sum Sq    Mean Sq F value    Pr(>F)
Date_Rss2  1 1.9874e+12 1.9874e+12  666.14 < 2.2e-16 ***
Residuals 97 2.8939e+11 2.9834e+09
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figure 12: Park Slope series (2003-2008) regression results

```
Call:
lm(formula = park1 ~ date1)

Residuals:
   Min     1Q Median     3Q    Max
-24432  -5336   1829   5256  17565

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 431076.03    2586.48  166.66   <2e-16 ***
date1         4322.82      81.83   52.83   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 9372 on 52 degrees of freedom
Multiple R-squared:  0.9817,    Adjusted R-squared:  0.9814
F-statistic:  2791 on 1 and 52 DF,  p-value: < 2.2e-16
```

Figure 13: Park Slope series (2008-2016) regression results

```
Call:
lm(formula = park2 ~ date2)

Residuals:
   Min     1Q Median     3Q    Max
-87239 -40478  -8939  33197 141010

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 257303.4    20718.5   12.42   <2e-16 ***
date2         4957.9      192.1   25.81   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 54620 on 97 degrees of freedom
Multiple R-squared:  0.8729,    Adjusted R-squared:  0.8716
F-statistic: 666.1 on 1 and 97 DF,  p-value: < 2.2e-16
```

Figure 14: Park Slope (Pre and Post Crisis)

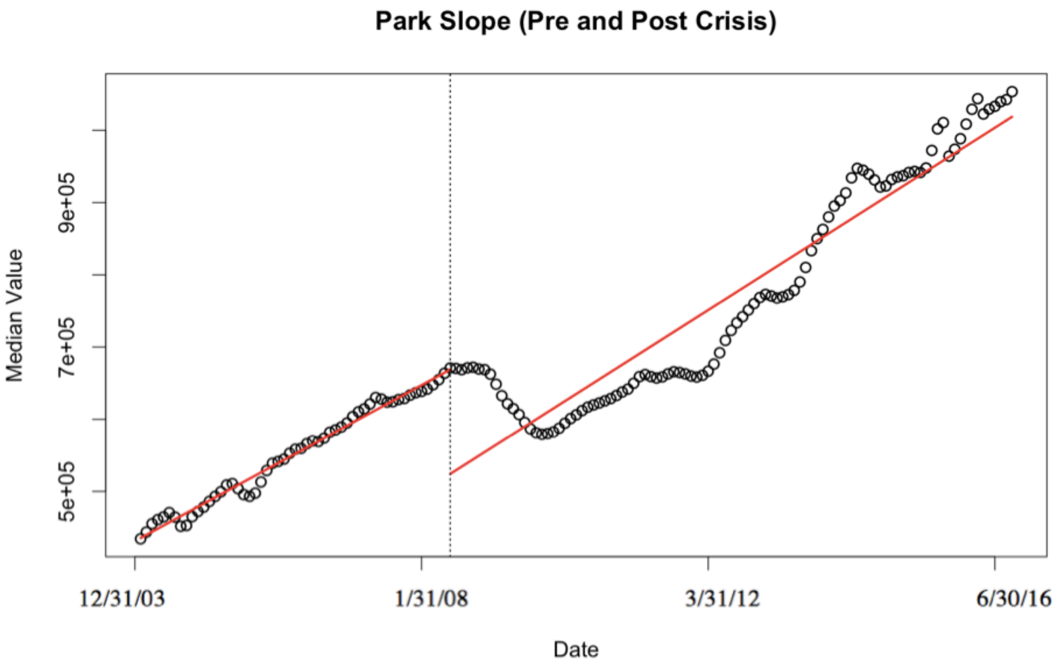


Figure 15: Heteroscedasticity test for Park Slope series

```
> lmtest::bptest(model1)

        studentized Breusch-Pagan test

data:  model1
BP = 30.894, df = 1, p-value = 2.725e-08
```

Figure 16a (top): Park Slope series with 95% confidence bounds
Figure 16b (bottom): Residual plot for Park Slope linear fit

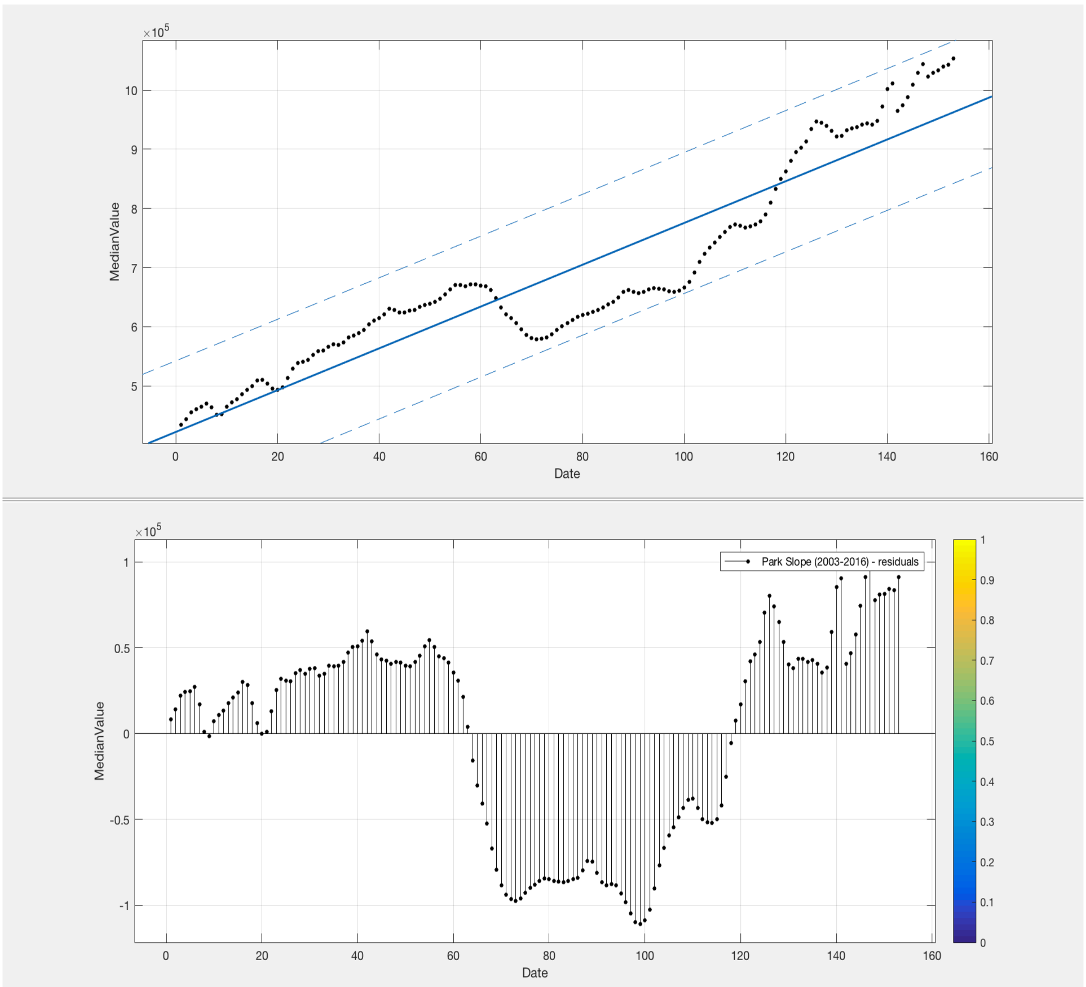


Figure 17a (top): Park Slope series with smoothing spline fit
Figure 17b (bottom): New residual plot for Park Slope with smoothing spline fit

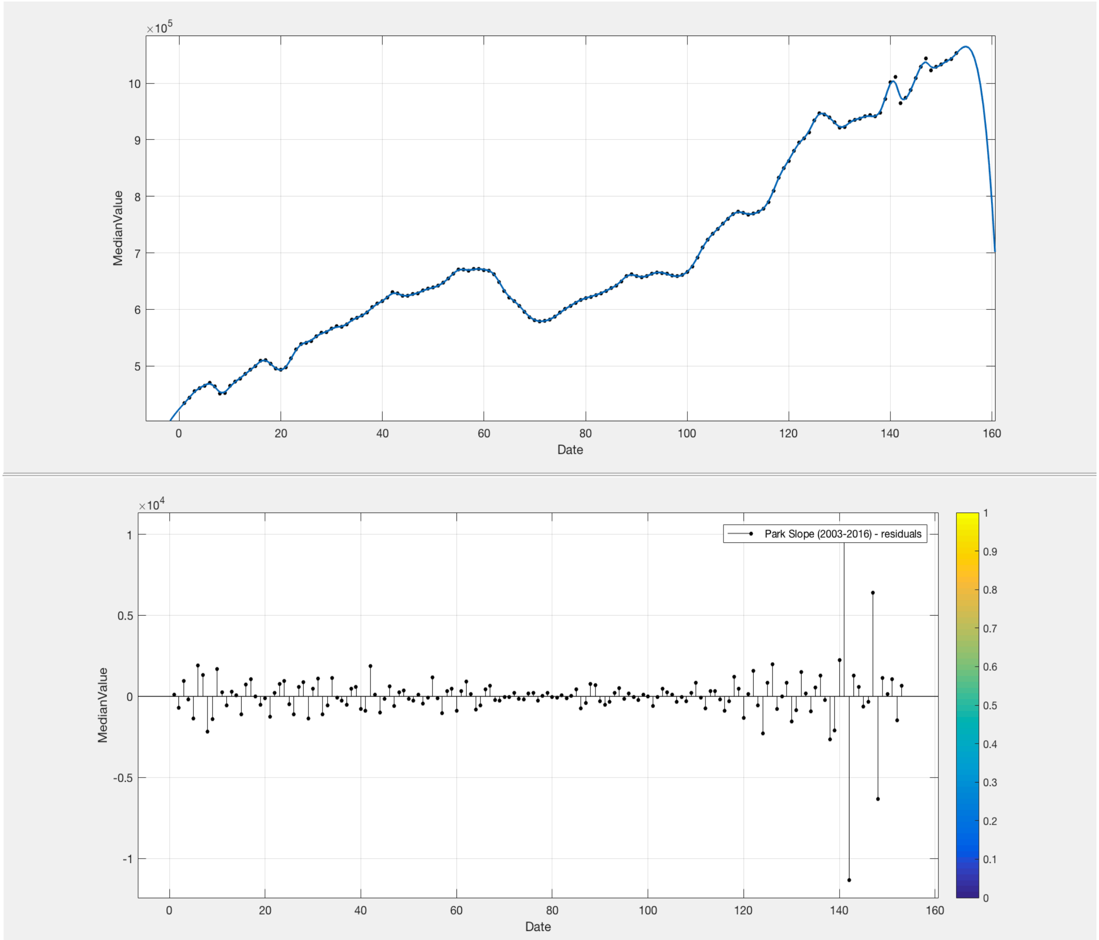


Figure 18: R-Squared for Park Slope smoothing spline fit

Results

Smoothing spline:
f(x) = piecewise polynomial computed from p
Smoothing parameter:
p = 0.9

Goodness of fit:
SSE: 4.195e+08
R-square: 0.9999
Adjusted R-square: 0.9998
RMSE: 2624

Figure 19: Park Slope 1 month lag-model regression results

```
Call:
lm(formula = Value ~ Value_Lagged)

Residuals:
   Min     1Q Median     3Q    Max
-52404  -3995    490   4417  24029

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)  -1.160e+02  3.089e+03  -0.038     0.97
Value_Lagged  1.006e+00  4.343e-03 231.623   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 8831 on 150 degrees of freedom
  (1 observation deleted due to missingness)
Multiple R-squared:  0.9972,    Adjusted R-squared:  0.9972
F-statistic: 5.365e+04 on 1 and 150 DF,  p-value: < 2.2e-16
```

Figure 20: Median Value Lag Plot – Park Slope

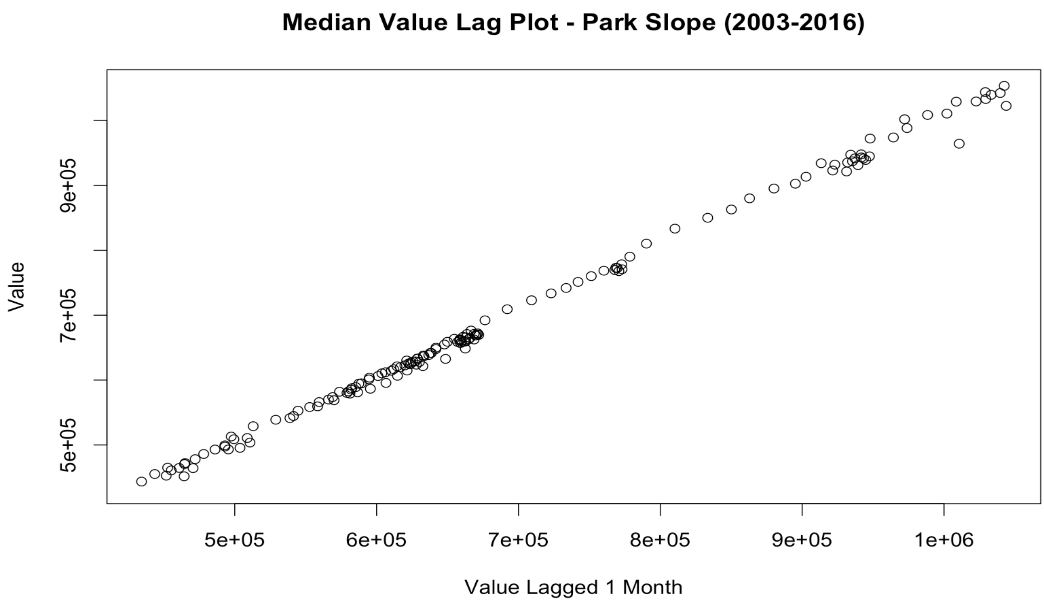


Figure 21: Park Slope monthly seasonality results

```
Call:
lm(formula = Park ~ Dec + Feb + March + April + May + June + 
    July + Aug + Sept + Oct + Nov)

Residuals:
    Min      1Q  Median      3Q     Max 
-228550 -104929  -31258   76237  329450 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)   656600      44886  14.628   <2e-16 ***
Dec            43467      63478   0.685    0.495    
Feb             4125      63478   0.065    0.948    
March           7892      63478   0.124    0.901    
April          11917      63478   0.188    0.851    
May            16750      63478   0.264    0.792    
June           19567      63478   0.308    0.758    
July           21450      63478   0.338    0.736    
Aug            24650      63478   0.388    0.698    
Sept           26092      63478   0.411    0.682    
Oct            31567      63478   0.497    0.620    
Nov            37450      63478   0.590    0.556    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 155500 on 132 degrees of freedom
Multiple R-squared:  0.007177,	Adjusted R-squared:  -0.07556 
F-statistic: 0.08675 on 11 and 132 DF,  p-value: 1
```

Figure 22: Chelsea median value regression results

```
Call:
lm(formula = Chelsea ~ Date)

Residuals:
    Min      1Q  Median      3Q     Max 
-286861 -203363  -13941  168853  561566 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept) 1118233.4    36248.3   30.85   <2e-16 ***
Date           6664.5      452.5   14.73   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 211800 on 136 degrees of freedom
Multiple R-squared:  0.6146,	Adjusted R-squared:  0.6118 
F-statistic: 216.9 on 1 and 136 DF,  p-value: < 2.2e-16
```

Figure 23: Median Value Chelsea

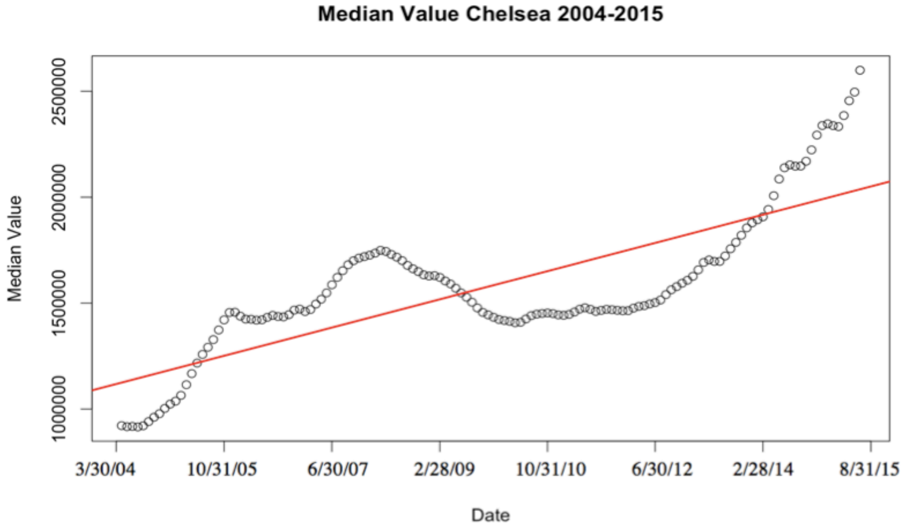


Figure 24: ANOVA for Chow Test — Chelsea Series

```
Analysis of Variance Table

Response: ChelseaRSS3
           Df     Sum Sq    Mean Sq F value    Pr(>F)
DateRSS3    1 9.7268e+12 9.7268e+12  216.92 < 2.2e-16 ***
Residuals 136 6.0982e+12 4.4840e+10
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
> ModelRSS2 <- lm(ChelseaRSS2 ~ DateRSS2)
> ModelRSS1 <- lm(ChelseaRSS1 ~ DateRSS1)
> anova(ModelRSS2)
Analysis of Variance Table

Response: ChelseaRSS2
          Df     Sum Sq    Mean Sq F value    Pr(>F)
DateRSS2   1 4.8839e+12 4.8839e+12  113.07 < 2.2e-16 ***
Residuals 85 3.6716e+12 4.3195e+10
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
> anova(ModelRSS1)
Analysis of Variance Table

Response: ChelseaRSS1
          Df     Sum Sq    Mean Sq F value    Pr(>F)
DateRSS1   1 3.3711e+12 3.3711e+12  702.15 < 2.2e-16 ***
Residuals 49 2.3525e+11 4.8011e+09
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
> 4.3195e+10/4.8011e+09
[1] 8.996897
```

Figure 25: Chelsea median value regression results

```
Call:
lm(formula = Chelsea1 ~ Date1)

Residuals:
   Min     1Q Median     3Q    Max 
-80626 -59705 -16784  32279 174611 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 914493.8    19694.0   46.44   <2e-16 ***
Date1        17466.5      659.2   26.50   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 69290 on 49 degrees of freedom
Multiple R-squared:  0.9348,    Adjusted R-squared:  0.9334 
F-statistic: 702.2 on 1 and 49 DF,  p-value: < 2.2e-16
```

Figure 26: Chelsea sub-period results (2004-2008)

```
Call:
lm(formula = Chelsea2 ~ Date2)

Residuals:
    Min      1Q  Median      3Q     Max 
-250831 -164663  -83785  164777  487653 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept) 1291032.9    44951.2   28.72   <2e-16 ***
Date2          9434.7      887.3   10.63   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 207800 on 85 degrees of freedom
Multiple R-squared:  0.5709,    Adjusted R-squared:  0.5658 
F-statistic: 113.1 on 1 and 85 DF,  p-value: < 2.2e-16
```

Figure 27: Chelsea (Pre and Post Crisis)

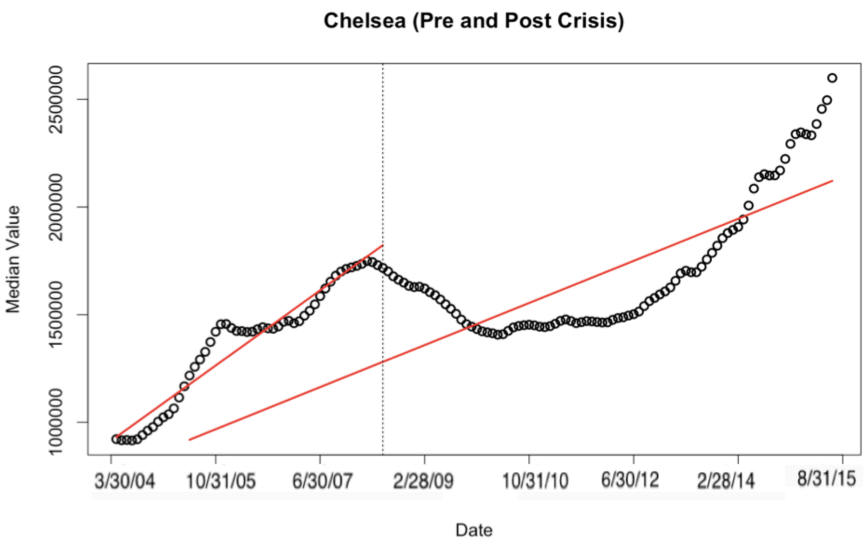


Figure 28: Chelsea Post Crisis (degree 2 polynomial fit)

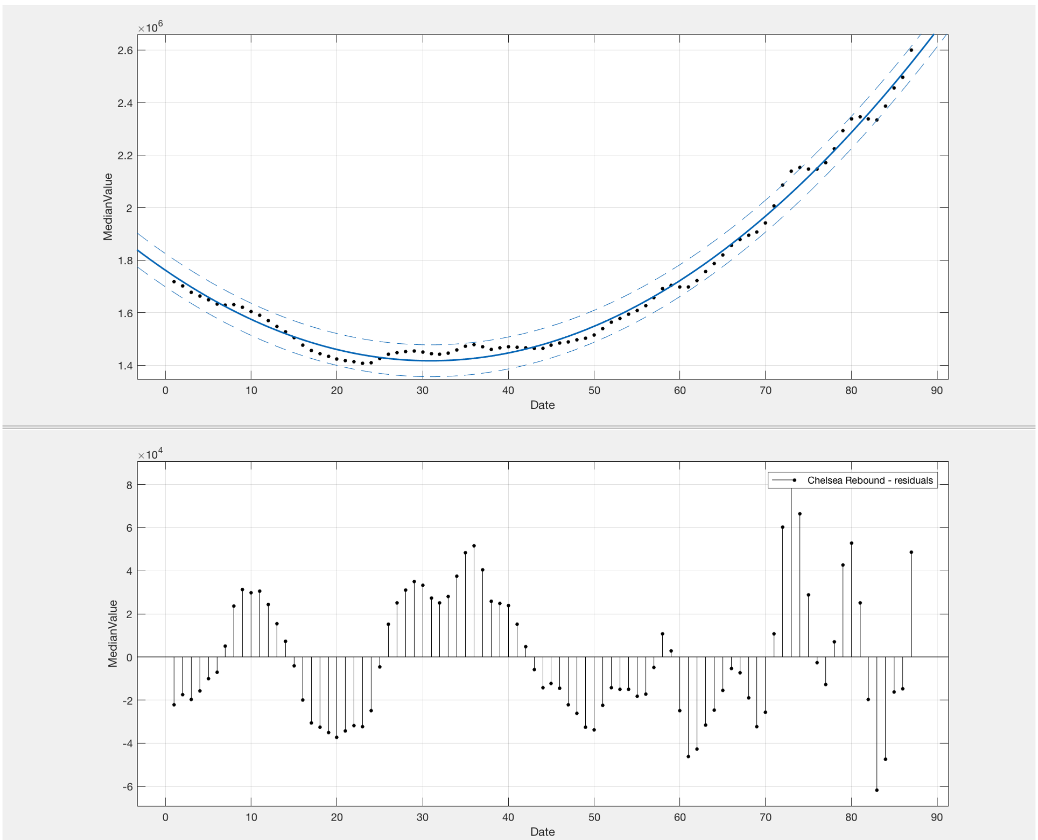


Figure 29: Chelsea (degree 2 polynomial fit) results

```
Results

Linear model Poly2:
    f(x) = p1*x^2 + p2*x + p3
Coefficients (with 95% confidence bounds):
      p1 =       360.4  (349, 371.8)
      p2 =  -2.228e+04  (-2.332e+04, -2.125e+04)
      p3 =   1.762e+06  (1.742e+06, 1.781e+06)

Goodness of fit:
  SSE: 7.641e+10
  R-square: 0.9911
  Adjusted R-square: 0.9909
  RMSE: 3.016e+04
```

Figure 30: Chelsea series heteroscedasticity test

```
        studentized Breusch-Pagan test

data:  modela
BP = 18.735, df = 1, p-value = 1.502e-05
```

Figure 31a (top): Chelsea linear fit (with 95% confidence bounds)
Figure 31b (bottom): Residual plot for Chelsea linear fit

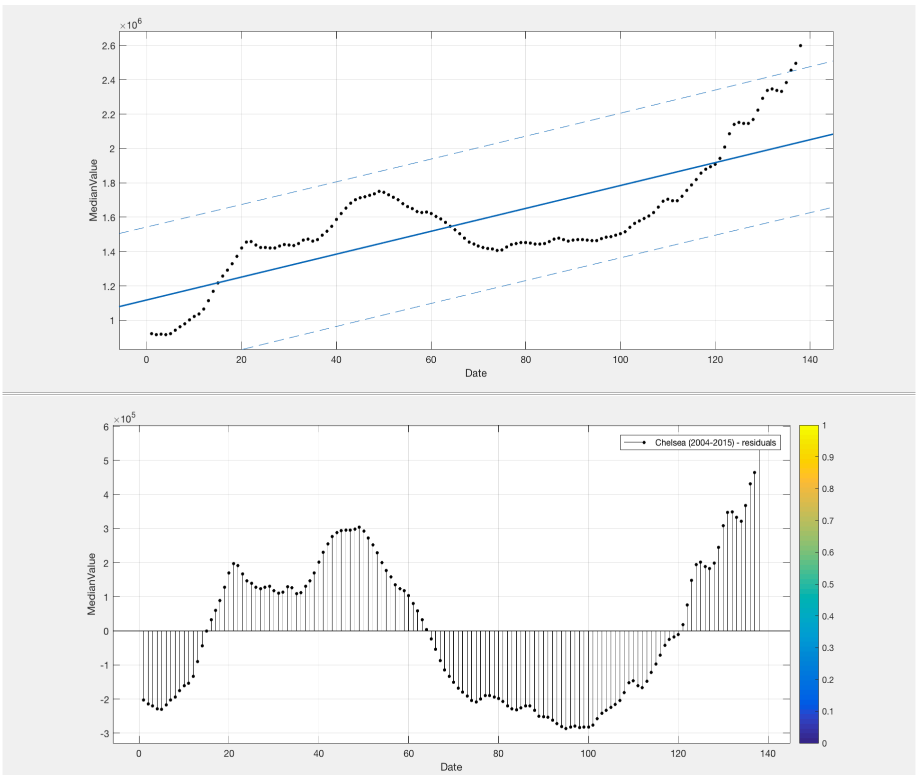


Figure 32a (top): Chelsea series with smoothing spline fit
Figure 32b (bottom): New residual plot for Chelsea with smoothing spline fit

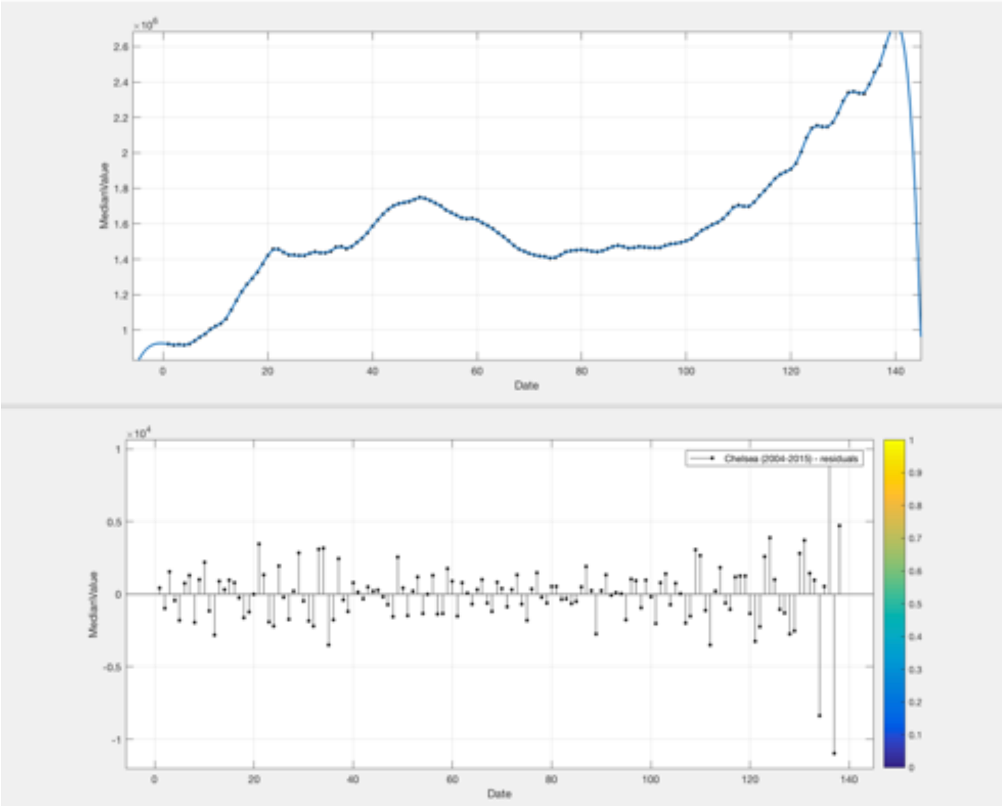


Figure 33: R-Squared for Chelsea smoothing spline fit

Results

Smoothing spline:
f(x) = piecewise polynomial computed from p
Smoothing parameter:
p = 0.9

Goodness of fit:
SSE: 6.276e+08
R-square: 1
Adjusted R-square: 0.9999
RMSE: 3382

Figure 34: Chelsea 1 month lag-model regression results

```
Call:
lm(formula = Value_C ~ Value_Lagged_C)

Residuals:
   Min      1Q Median      3Q     Max
-37987 -15265  -2706  15412  74408

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)    -1.617e+04  9.352e+03  -1.729   0.0862 .
Value_Lagged_C  1.018e+00  5.816e-03 175.032   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 22360 on 135 degrees of freedom
  (1 observation deleted due to missingness)
Multiple R-squared:  0.9956,	Adjusted R-squared:  0.9956
F-statistic: 3.064e+04 on 1 and 135 DF,  p-value: < 2.2e-16
```

Figure 35: Median Value Lag Plot – Chelsea

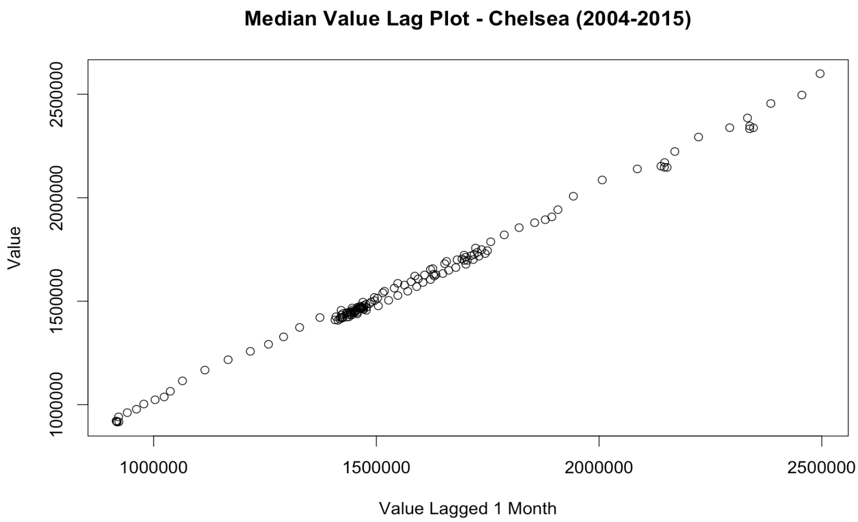


Figure 36: Chelsea monthly seasonality results

```
Call:
lm(formula = Value_10001 ~ February + March + April + May + June +
    July + August + September + October + November + December)

Residuals:
    Min      1Q  Median      3Q     Max
-476920 -146795  -53580  113978  653420

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  1514320      74889  20.221   <2e-16 ***
February        7980     105909   0.075    0.940
March          23620     105909   0.223    0.824
April          37500     105909   0.354    0.724
May            51040     105909   0.482    0.631
June           63420     105909   0.599    0.551
July           73270     105909   0.692    0.491
August         81390     105909   0.768    0.444
September      90540     105909   0.855    0.395
October       101990     105909   0.963    0.338
November      116040     105909   1.096    0.276
December      125560     105909   1.186    0.238
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 236800 on 108 degrees of freedom
Multiple R-squared:  0.03005,	Adjusted R-squared:  -0.06874
F-statistic: 0.3042 on 11 and 108 DF,  p-value: 0.9837
```

References

Barry, R., Pace, R. K., and Sirmans, C.F. (1998). Spatial Statistics and Real Estate. *Journal of Real Estate Finance and Economics Volume 17, Number 1, p. 5-13*. Retrieved from http://www.spatial-statistics.com/pace_manuscripts/spatial_statistics_real_estate/html/web_intro_index.html

Harrison, D., and Rubinfeld, D. (1978). Hedonic Housing Prices and the Demand for Clean Air. *Journal of Environmental Economics and Management Issue 5, p. 81-102*. Retrieved from https://www.law.berkeley.edu/files/Hedonic.PDF

Nelson, J. (1982). Highway Noise and Property Values. *Journal of Transport Economics and Policy Volume 15, Number 2, p.* 117-138. Retrieved from http://www.bath.ac.uk/e-journals/jtep/pdf/Volume_XV1_No_2_117-138.pdf

The Economics Review Team

Ewa Staworzynska	Editor-in-Chief
Nico Van de Bovenkamp	Managing Editor, Online Content
Prabhod Mudlapur	Managing Editor, Printed Publication
Chistopher Chao	Club Treasurer
Angelica Colino	Director of Marketing
Meghna Rangan	Director of Marketing
Victoria Sykora-Lovaas	Web Designer
Hannah Lee	Marketing Associate

Editorial Staff

Arjun Goyal
Surya Ierokomos
Daniel Wang
Yasmine Deswandhy
Meghna Chadha
Dhanya Madugalle
Nikita Mokhov
Samyak Jain
Naman Patel
Aditya Khosla
Shivaditya Sinha
Tim Makhauri
Ahmed El Wakeel
Hazel Choi

Writing Staff

Andrés Rodríguez Brauer
Tristan Baylor
Oleg Biletsky
David Behrens
Zoe Hall
Taber Brown
David Riccione
Eduard Batash
Jeremy Ron Teboul
Damon Aitkin
Adrian Pietrzak