

## Data Visualization - “How does the Recommended Retail Price (RRP) of books differ across different bestseller lists?”

Overall, I found the research process lengthy but rewarding in terms of the information that I was able to retrieve. I started my topic with the research question “Are certain bestseller lists targeted at certain socioeconomic groups?”. However, upon further review and after considering advice from Dr. Fitzsimmons and my fellow classmates, I decided that my data set couldn't entirely answer this specific research question. As a result, I changed my research question to “How does the Recommended Retail Price (RRP) of books differ across different bestseller lists?”. This chart  I believe, allows me to use my data set to comprehensively answer my research question.

Firstly, I had to set some guidelines to follow which would allow substantial yet meaningful comparison of the Recommended Retail Prices (RRP) between bestseller lists. I decided on using the RRP's for the 'Hardcover Fiction' bestseller lists. This was simply due to the fact that there was far less variation in editions and price-points for hardcover novels as for paperbacks. This made it easier to compare and contrast the similarities and differences in prices between the 4 national bestseller lists. Further to this, I decided to measure the data in the first 26 weeks on 2017, a large enough time period to avoid error due to anomalous results.

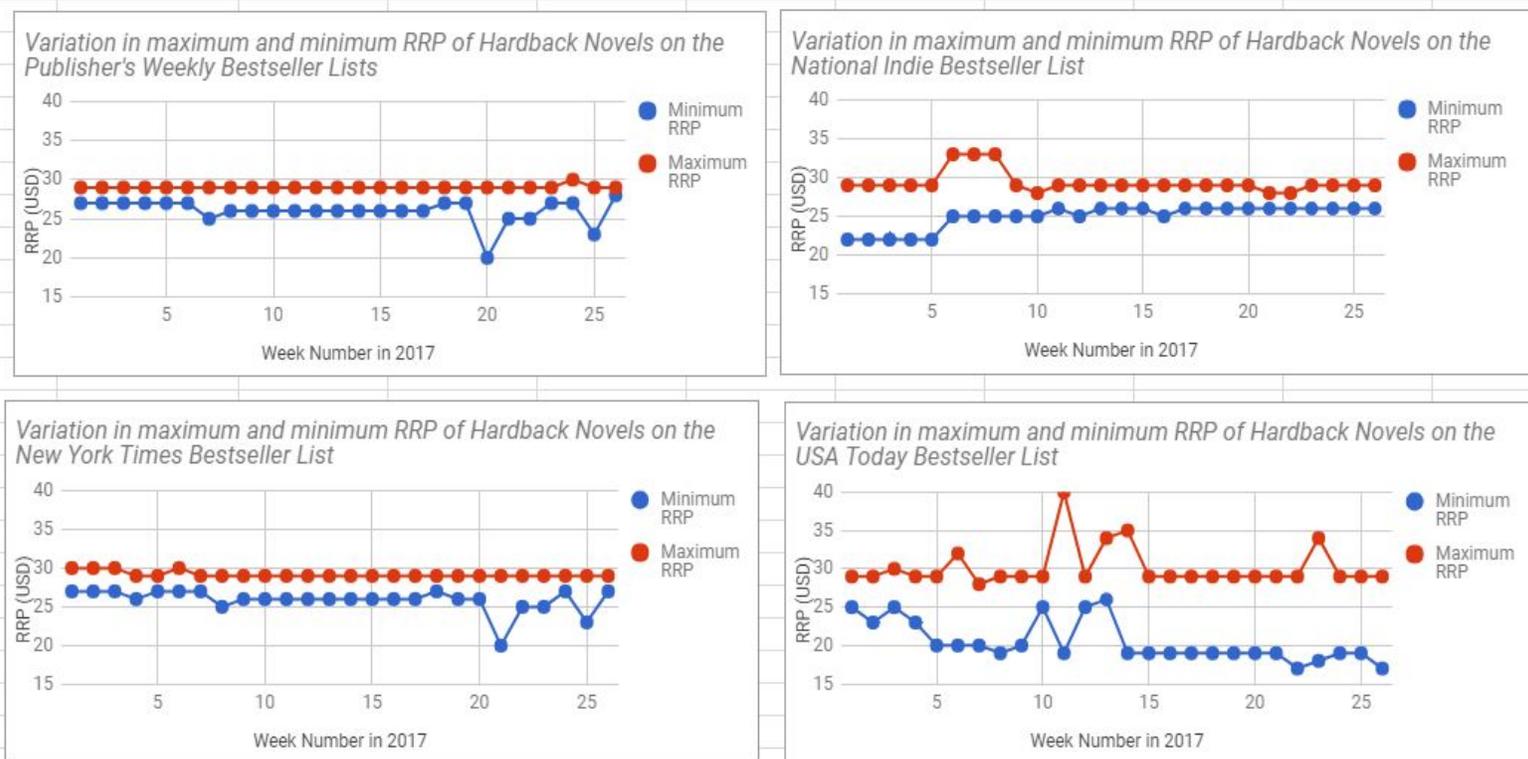
Furthermore, my data set includes the top 10 hardback fiction bestsellers for each week, which I thought would give a usable and comparable average RRP for each weekly bestseller list. I found that the most difficult part of the research was finding a reliable source of RRP data to use for the two bestseller lists which don't publish RRP's themselves. Specifically, the two lists which don't publish the RRP for their bestselling books are *The New York Times* and *USA Today*. This was initially a challenge as these are two of the most influential bestseller lists in the US, and I didn't want to leave them out simply because they didn't publish RRP data. To resolve this problem, I proceeded to find other data sources. First of all, I sourced from the *Publisher's Weekly* and *National Indie* bestseller lists to get RRP data for books that may appear on multiple lists. However, not all  books on either *The New York Times* or *USA Today* appeared on the other two, so I then searched on Amazon and took the 'List Price' for the hardcover version of the novel as the RRP.

I created 3 different data visualizations to visually display the data that I had collected. The process of creating these visualizations, as well as the individual general interpretations of the visualizations, are outlined in each of the figure legends. To summarize my findings, the data visualizations show that, as well as having a lower average RRP value than the other lists (*Figure 2, Figure 3*), the *USA Today* bestseller list had an extremely wide range of price points in its top 10 from week to week (*Figure 1*). This, I believe, was mainly due to the fact that the *USA Today* list featured many more genres of books than other lists, such as romance, literary classics, and children's literature. Also, I noticed the unusual similarity between the *New York Times* and *Publisher's Weekly* lists. Coincidentally, the overall average RRP for the entire 26 weeks was exactly the same for both lists (*Figure 2*). Furthermore, the similarities continue when comparing the average RRP for the individual weeks, with both lines in *Figure 3* almost 'trailing' each other.

To comprehensively answer my research question, the Recommended Retail Price of hardback fiction books does vary between bestseller lists. However, the variation is only seen between some lists. Lists such as *USA Today* seem to have wildly different price points than other popular lists whilst *Publisher's Weekly* and *The New York Times* have average RRP values which almost mirror each other over a sustained period of time.

## Data Visualizations

Four dual-line graphs to show the variation in maximum and minimum RRP of hardback fiction bestsellers in the top 10 of the 4 separate bestseller lists over the same 26 week period



**Figure 1 (Above)** - This figure shows the maximum of minimum price of the top 10 books on each bestseller list over the 26 week period that I measured. Each graph has the same axis scales and so both the variance in area between the lines (range of prices) over time and the volatility of maximum and minimum prices over time can be seen for each bestseller list. It is evident that *USA Today* list has both the widest range of book prices on its bestseller list and also the most frequently changing book prices.

Sources used for this visualization;

*New York Times Bestseller List, Hardcover Fiction,*

<https://www.nytimes.com/books/best-sellers/>

*Publisher's Weekly Bestseller List, Hardcover Fiction,*

<https://www.publishersweekly.com/pw/nielsen/hardcoverfiction.html>

*National Indie Bestseller List, Hardcover Fiction,*

<http://www.bookweb.org/national-indie-bestsellers>

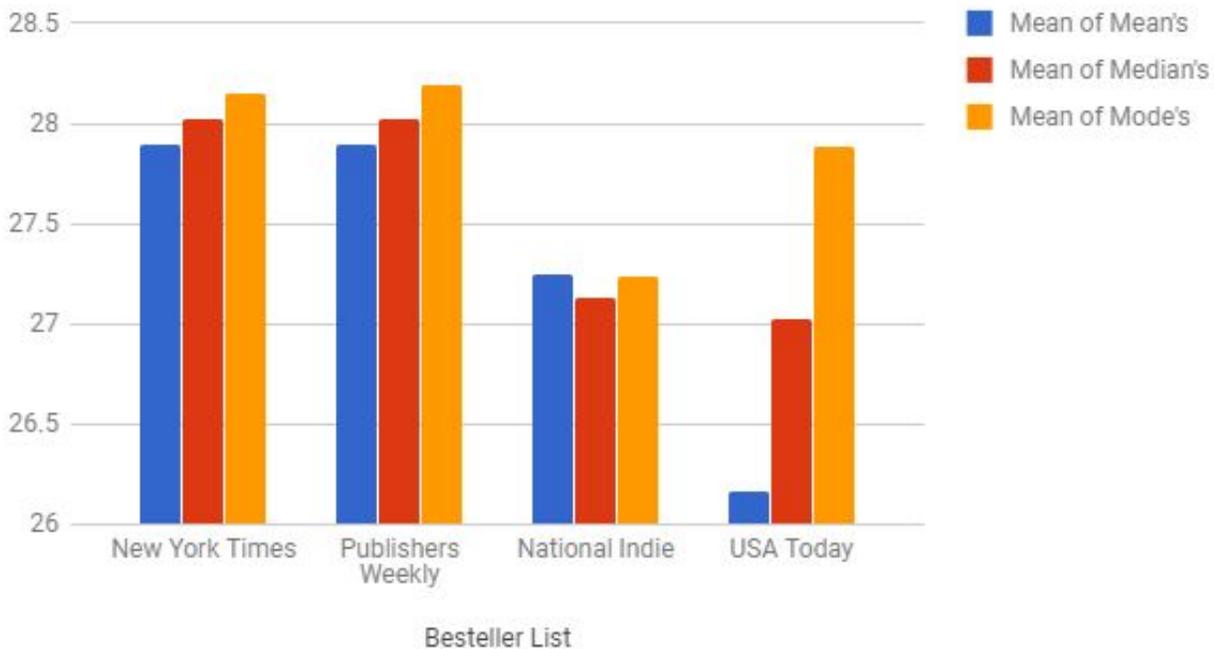
*USA Today Bestseller List, Fiction,*

<https://www.usatoday.com/life/books/best-selling/week/2017/48/?sort=1&classId=1&genrel=0>

*Amazon, Multiple Book Listings,* <https://www.amazon.com/>

A multi-bar graph to show the average mean, median, and mode of the top 10 hardback fiction bestsellers for the 26 week period for each of the 4 bestseller lists

### Mean of Mean's , Mean of Median's and Mean of Mode's



**Figure 2 (Above)** - A graph to show how the common averages of the 4 bestseller lists varied over the 26 weeks. For each bestseller list, a mean, median, and mode were calculated from the prices of the top 10 bestsellers each week. Then, a mean of these values was calculated for the 26 week period. This yielded a value of “Mean of Means”, “Mean of Medians”, and “Mean of Modes” for each bestseller list over the 26 week period. The graphs show that the *Publisher’s Weekly* and *New York Times* had very similar values, whilst *USA Today* varied wildly between its mean, median, and mode of prices in the top 10 books on its list.

Sources used for this visualization;

*New York Times Bestseller List, Hardcover Fiction,*

<https://www.nytimes.com/books/best-sellers/>

*Publisher’s Weekly Bestseller List, Hardcover Fiction,*

<https://www.publishersweekly.com/pw/nielsen/hardcoverfiction.html>

*National Indie Bestseller List, Hardcover Fiction,*

<http://www.bookweb.org/national-indie-bestsellers>

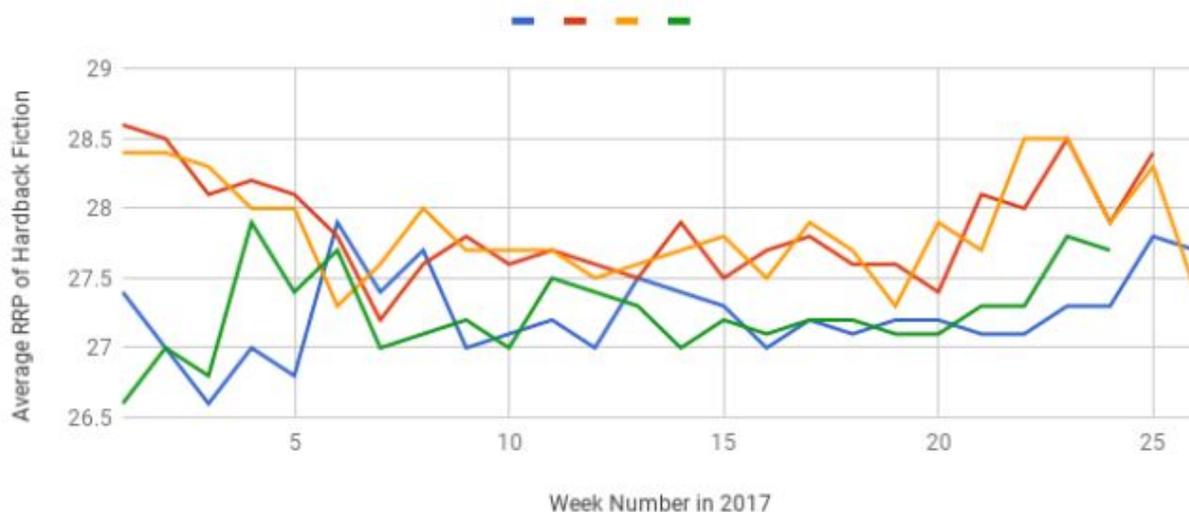
*USA Today Bestseller List, Fiction,*

<https://www.usatoday.com/life/books/best-selling/week/2017/48/?sort=1&classId=1&genreId=0>

*Amazon, Multiple Book Listings,* <https://www.amazon.com/>

A multi-line graph to show how the average (mean) RRP of the top 10 hardcover bestsellers on the 4 bestseller lists changes over the 26 week period

A chart to show how the RRP of Hardcover Fiction Bestsellers vary for different the top 10 books on 4 different Bestseller Lists



Red- New York Times

Yellow - Publishers Weekly

Blue - National Indie

Green - USA Today

**Figure 3 (Above)** - This figure shows how the average (mean) RRP of the books on each bestseller list changes over the 26 weeks measured. For each bestseller list, the total average price for that week's top 10 list was calculated and plotted. This allows us to easily compare the average RRP for the 4 bestseller lists and also see the trend in average price over the 26 weeks. The graph generally shows that *Publisher's Weekly* and *The New York Times* had a higher average RRP of its hardcover bestsellers than *USA Today* and *National Indie*. However, there were some crossover points show in the graph.

Sources used for this visualization;

*New York Times Bestseller List, Hardcover Fiction,*

<https://www.nytimes.com/books/best-sellers/>

*Publisher's Weekly Bestseller List, Hardcover Fiction,*

<https://www.publishersweekly.com/pw/nielsen/hardcoverfiction.html>

*National Indie Bestseller List, Hardcover Fiction,*

<http://www.bookweb.org/national-indie-bestsellers>

*USA Today Bestseller List, Fiction,*

<https://www.usatoday.com/life/books/best-selling/week/2017/48/?sort=1&classId=1&genreId=0>

*Amazon, Multiple Book Listings,* <https://www.amazon.com/>

Link to Google Sheet containing raw data

<https://docs.google.com/spreadsheets/d/1y-O9SsdL5yoI5QplincxKr2KhOuPKvTEy7yEi2dk2F4/edit?usp=sharing>