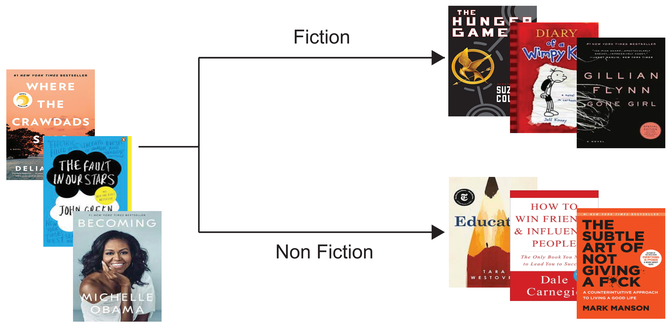
# Amazon's Top 50 Best Selling Books from 2009 - 2019 Analysis

Alex Cheung

## Introduction

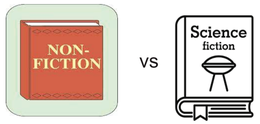
Amazon is one of the largest online marketplaces for purchasing books.
In this data analysis, we observe data about their top 50 bestsellers from 2009 - 2019.

Data has been categorized into fiction and non-fiction using Goodreads.




Research Questions:

1. What effect does genre have on the sale price of non fiction versus fiction books?
2. How does the price of books vary over time?




## Methods

**Find a dataset on Kaggle**

kaggle

**Inspect the data**

**Clean data with R**

R

**EDA: Correlations and distributions**

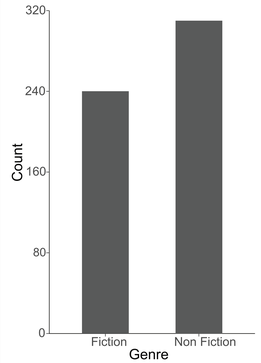
**Find significant relationships**

$p < .05$

**Visualize results**

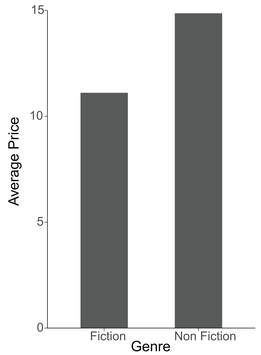
## Main Results

Which genre sold more unique books during 2009 - 2019?



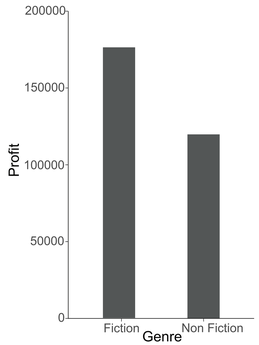

More unique non fiction books were sold compared to fiction books from 2009-2019.

Which genre had higher pricing?




Non fiction books had a higher average sale price in contrast to fiction books

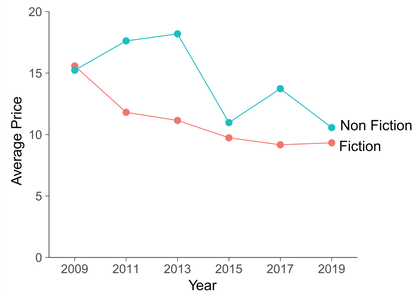
Which genre had higher overall profit?




Fiction books brought in a greater amount of profit.

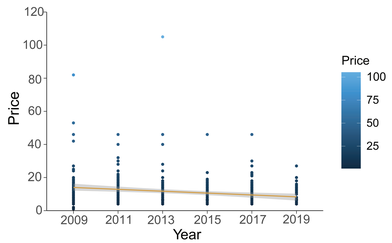
## Additional Results

What do you notice about the price of books for the two genres?




Fiction books seem to decline in price at a slower rate in comparison to non fiction books.

How does the price change over time?




The price of books has steadily fallen over time regardless of genre.

## Conclusion

**Findings**

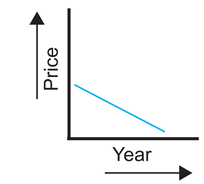
Non fiction books took up a larger count of the marketplace compared to fiction books.

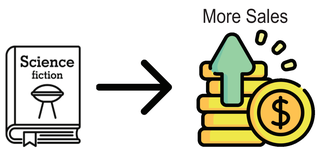
Non fiction books on average were around $2.5 more expensive than fiction books

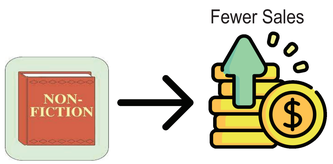
Yet, fiction books had greater profit of around $60k, indicating that they earn more money overall.

There is a negative correlation between year and sale price.

Price of books has steadily declined over the years in both genres.




**Implications and Future Directions**

Future authors and publication companies should focus more on fiction books as they bring in more profit overall compared to non fiction books.

This study can be expanded to other online marketplaces and be a point of comparison in consumers' behavior of purchasing books online.

Future studies can look into books that are categorized as fiction and their sub-genres (e.g. science fiction & historical fiction).

Further research can applied to sub-genres of fiction books and how they differ from one another in sale price, reviews, and ratings.

## Acknowledgements & References

https://www.kaggle.com/datasets/sootersaalu/amazon-top-50-bestselling-books-2009-2019

https://bublish.com/blog/2018/03/think-like-a-reader-thursday-what-does-price-mean-to-book-buyers-in-your-genre/