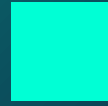
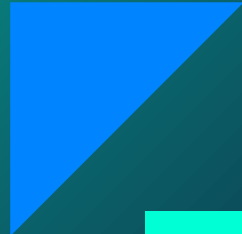
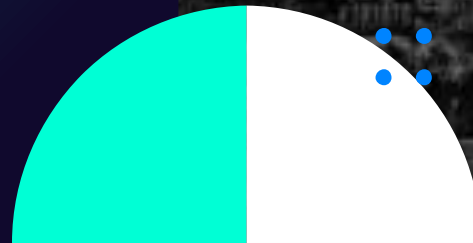


Project 5:



Steam Games Analysis



PROJECT 5: AN ANALYSIS OF GAMES ON STEAM

Overview: This project has two components: 1) Guided data analysis with CareerFoundry tasks including exploratory visual analysis, geographical visualizations, supervised machine learning and unsupervised machine learning; and 2) An independent project that analyzed different variables to determine why over 75% of Steam games never get played.



Data

- [Steam Games Dataset](#)
- [Video Game Developers Regional Dataset](#)



Skills

- Python
- Data cleaning, wrangling, and merging.
- Geographical visualizations.
- Supervised and unsupervised machine learning.

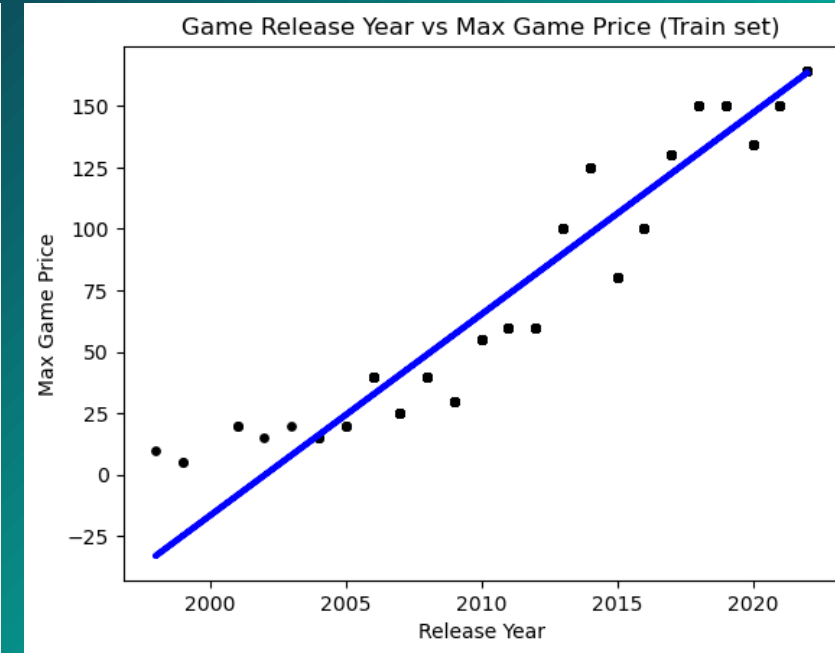
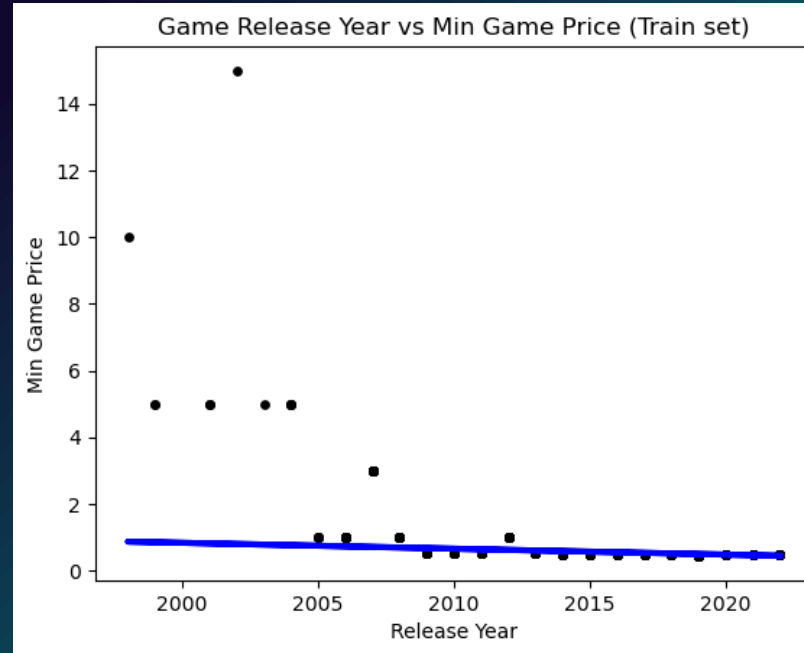
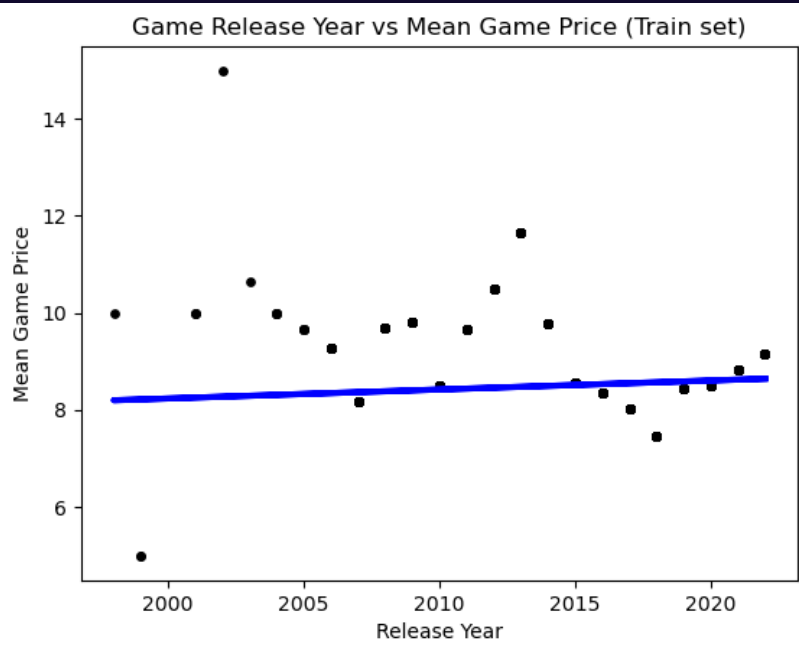


Tools

- Jupyter Notebook (Anaconda 3)
- Libraries: pandas, NumPy, Seaborn, Matplotlib, Folium, Json, Statsmodels, Scipy, Sklearn, and PyLab
- Microsoft Excel
- Tableau

Project 5: Supervised Machine Learning – Regression Analysis

To test the hypothesis that game prices increase with years, linear regression was run between year, and the mean, minimum, and maximum game prices. Regression analysis was able to predict maximum prices through time well. However, this did not work well with average and minimum game prices. This may be due to the high number of free-to-play games that have become increasingly common on the Steam platform, which would throw off averages, and minimum prices but have little effect on maximum game prices.

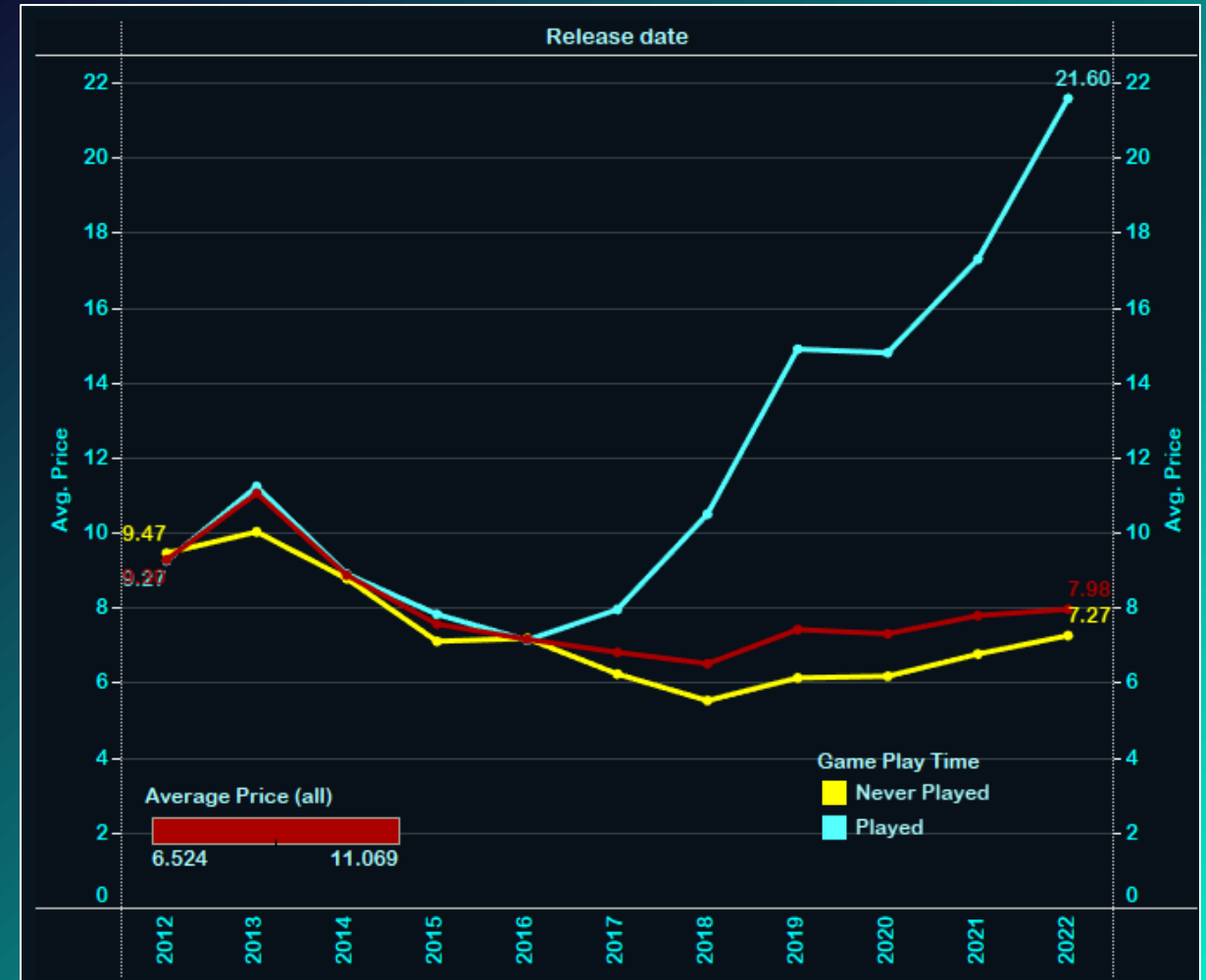
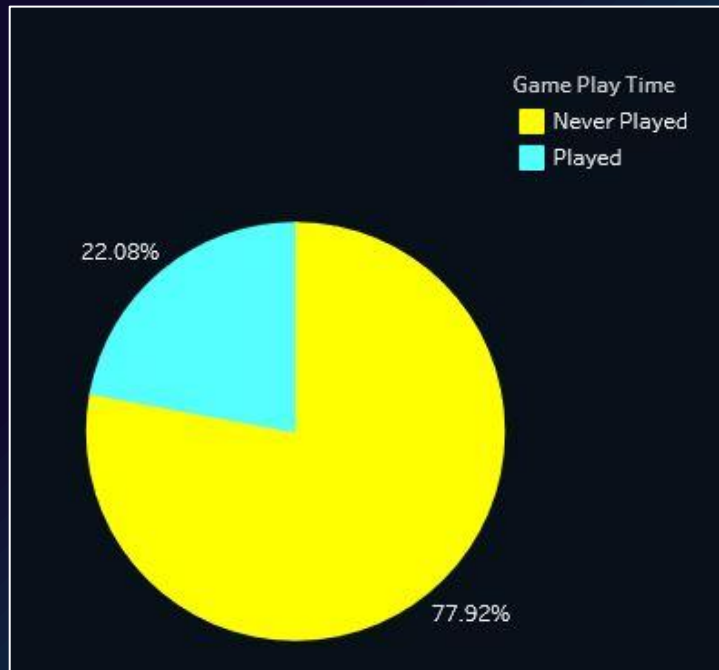


Python scripts and deliverables on Github

Project 5: Analysis of Unplayed Games

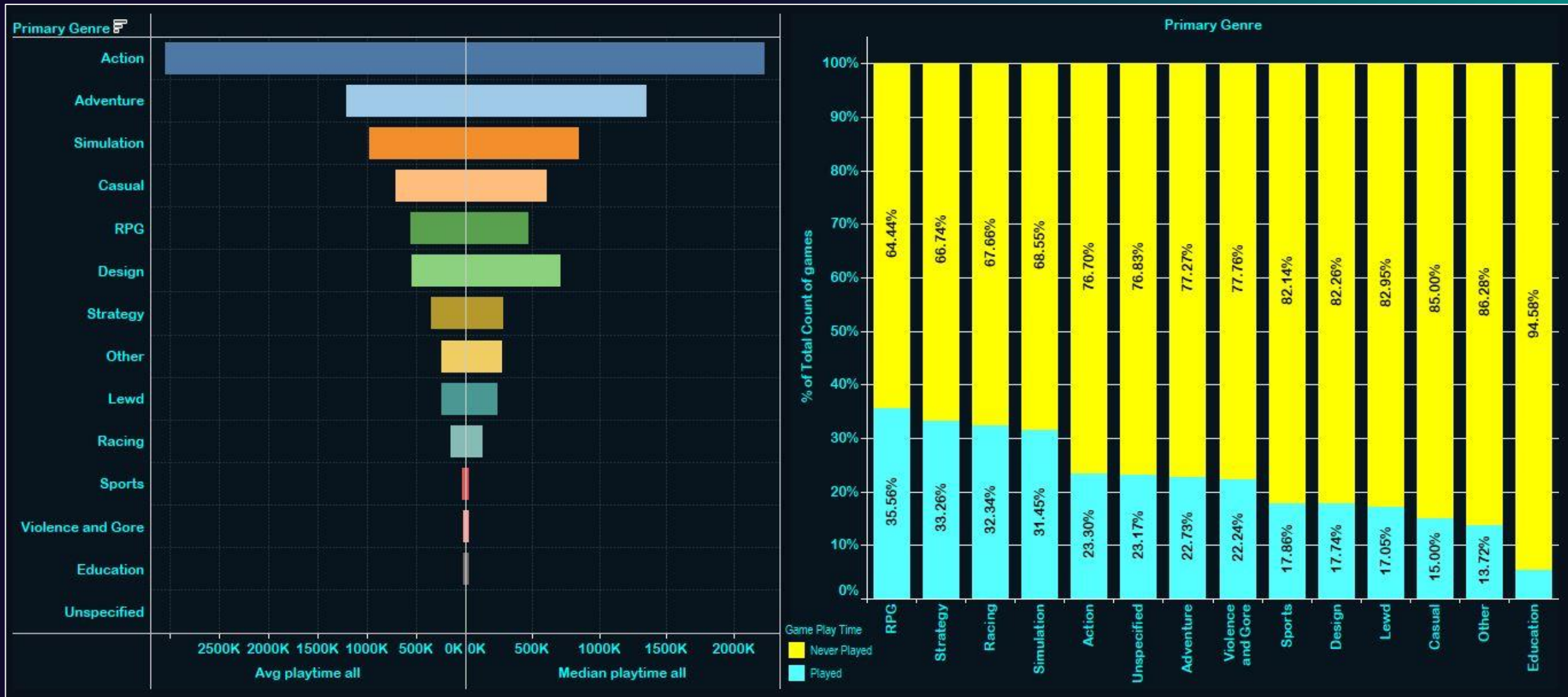
As of the end of 2022, the Steam library contained nearly 70,000 games. However, nearly 78% of those games never get played.

An analysis of average price over the past 10 years has shown that game prices have dropped, but people are more likely to play games that have a higher price point.



Project 5: Unplayed games – Genre

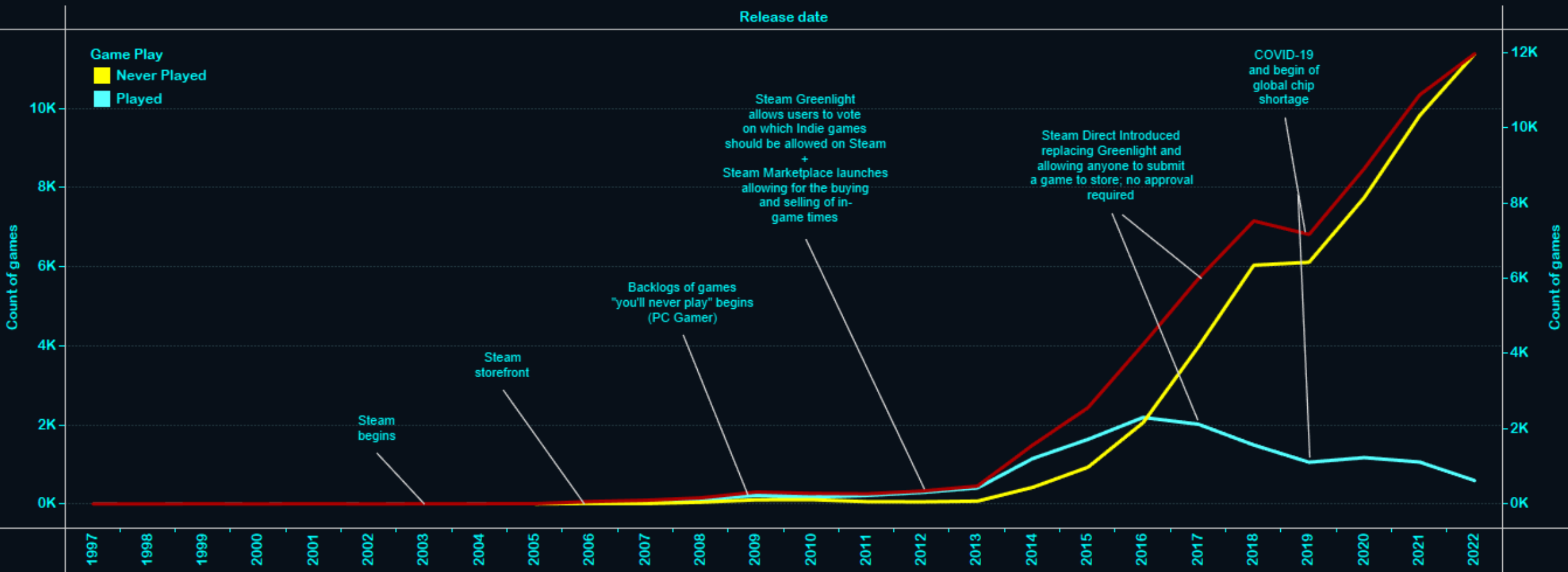
Action games have the highest average and median game play times. However, RPG and strategy games are more likely to be played at least once. Education games are the most likely to never be played.



Project 5: A Look Through Time

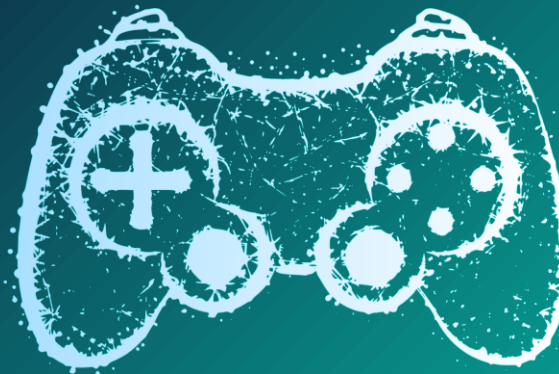
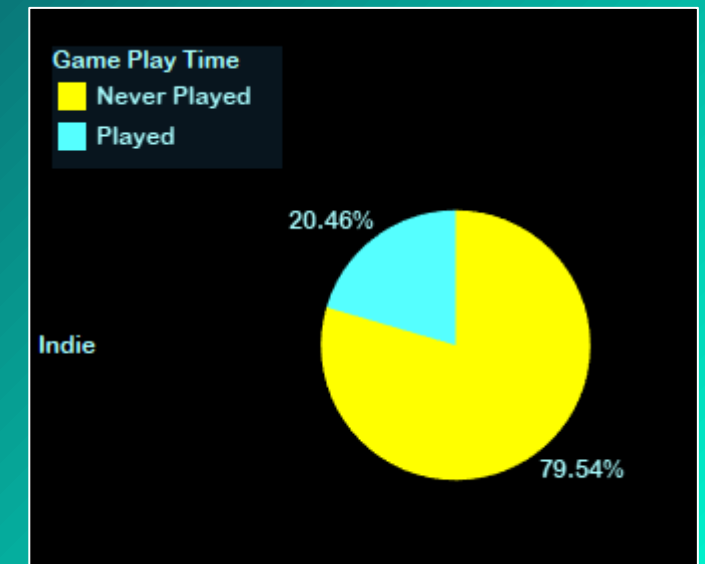
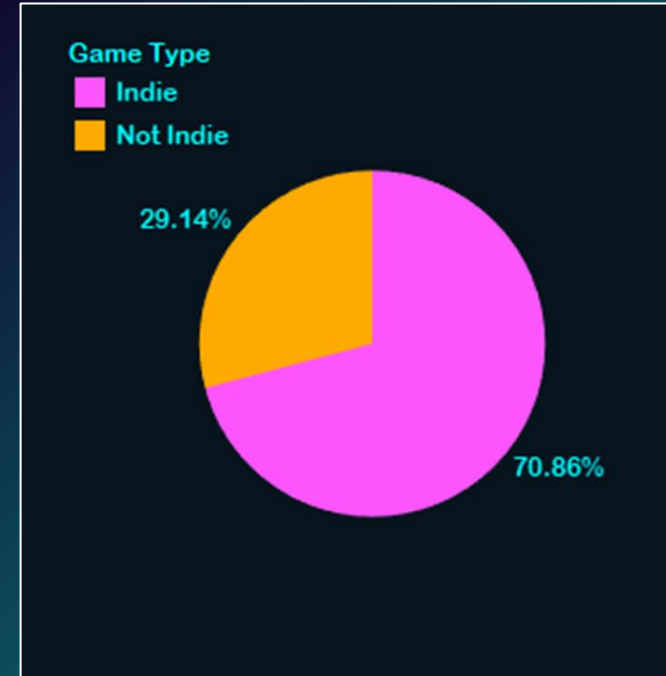
Steam's library of games has grown rapidly since 2013. With the addition of thousands of indie games, free-to-play games, and large discounts, many users have a large backlog of games they may never play.

While the trend here shows a decrease in game play, eMarketer research has shown an overall increase in video game play with most of this increase being in the form of mobile gaming.



Project 5: Conclusions and Recommendations

- Steam has an over abundance of games, particularly indie games, which can create individual backlogs of games that players will never get around too and it can make finding games difficult.
- While there are many free and low priced games, people are willing to pay more for quality.
- People are spending less time gaming on PCs and more time gaming on their mobile devices. As Steam does have a mobile app, they may want to consider adding mobile games to their library while also being more selective of indie games allowed into the library.
- Independent game developers may want to consider an alternative distribution platform, such as itch.io, which is specific to indie games if they want their games to be noticed.



PROJECT 5 TAKEAWAYS

The original dataset I had picked for this project was an archaeological dataset I had worked on creating for several years while I was a doctoral candidate. This dataset, because it drew from archaeological sites with poor preservation and was based mostly on qualitative data, did not work for this project and I had to change gears.

Being a lifelong gamer, I decided to go lookup databases related to gaming and decided to use the one for Steam, a platform I utilize frequently myself. With a large number of variables, it was difficult to decide on a track to follow. I ended up just “playing” with the data in Tableau until I found that most games on Steam are never played. From here, I was able to form questions and narrow down my analysis. While I was able to find changes through time, finding how why these changes occurred required some extra research into Steam’s history.

I was disappointed that my initial dataset did not work well for the project requirements but was very pleased with the results of analyzing the Steam dataset.

