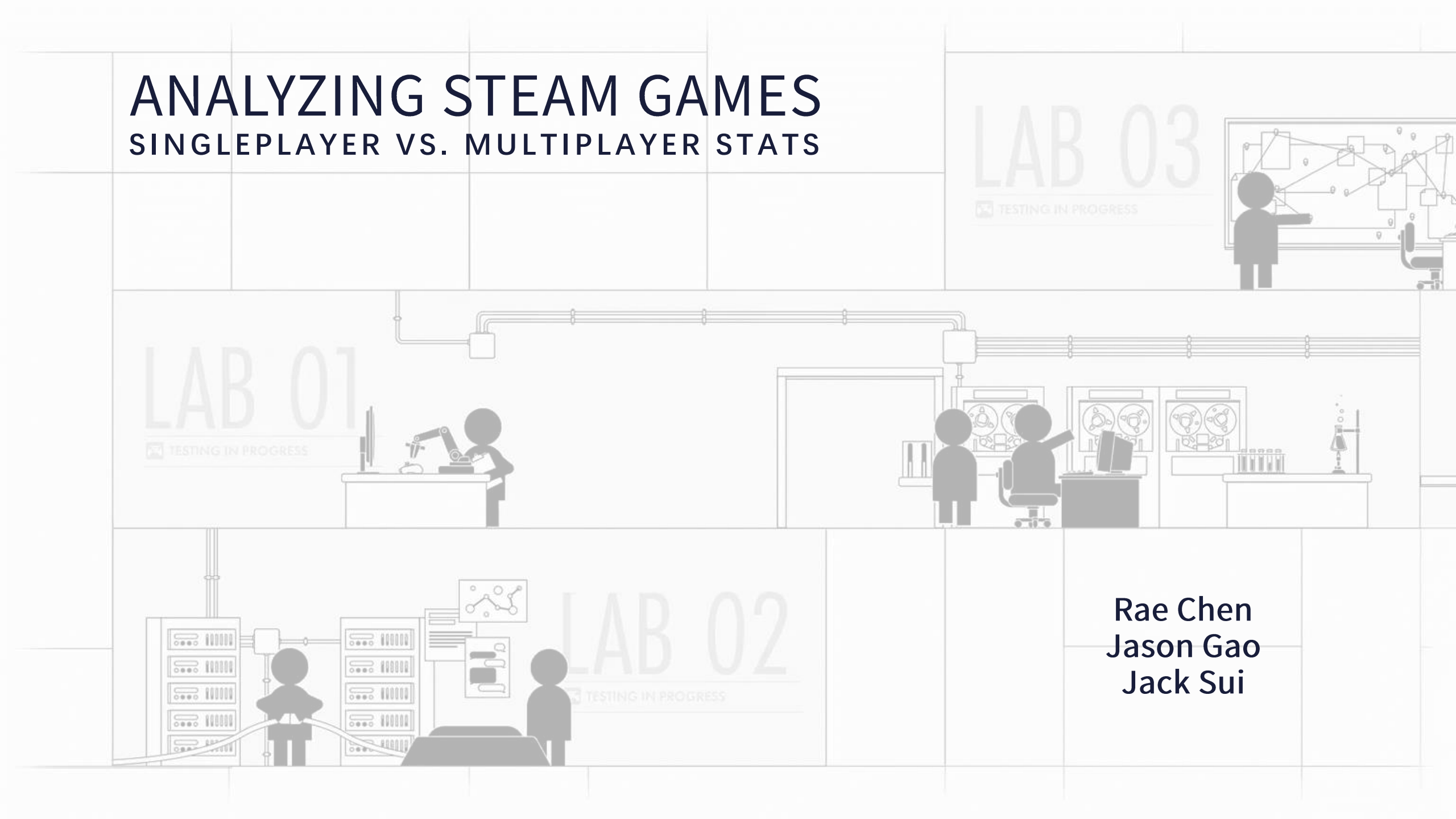
# ANALYZING STEAM GAMES

## SINGLEPLAYER VS. MULTIPLAYER STATS




Rae Chen
Jason Gao
Jack Sui

# The Ultimate Gaming Platform


**109 MILLION GAMERS**
**716 MILLION GAMES**
**1.1 MILLION YEARS OF PLAYTIME**

---

A dataset collected and analyzed for the 2016 ACM Internet Measurement Conference article by Mark O'Neill, Justin Wu, Elham Vaziripour, and Daniel Zappala

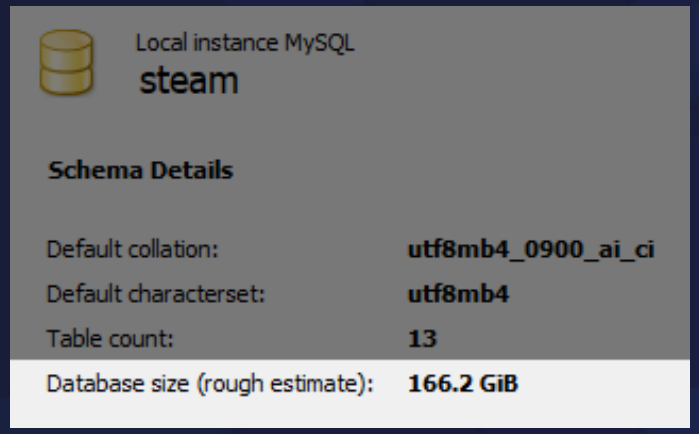
see the paper

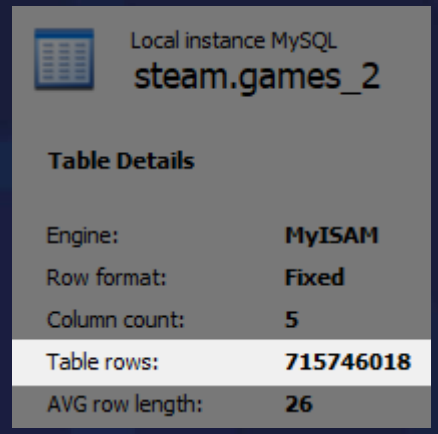
- Player behavior vs. game types?
- Predict game ratings?

# Charts to Plot, Models to Train

- Predicting the Rating for a game out of all total players who own this game with a given price, release date, and whether an "app" contains multiplayer content.
- Whether an "app" contains multiplayer content or not affects Play Time.
- Whether an "app" contains multiplayer content affects Achievement Completion
- How might engagement with games change over time?

# A Wild Dataset Appears!



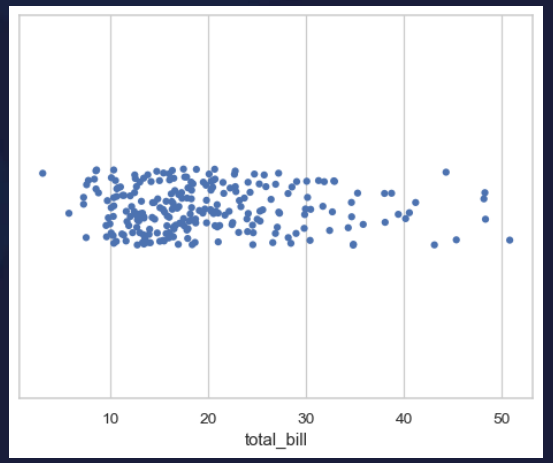

# Plan of Action

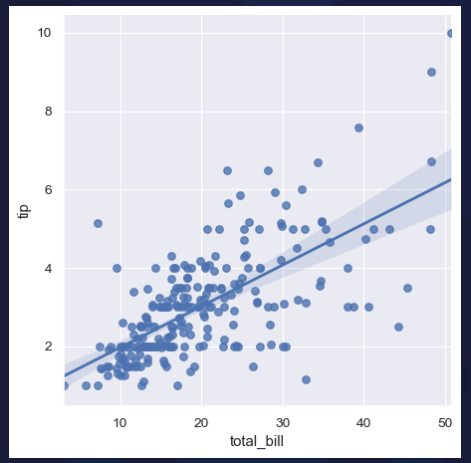
Machine learning with scikit learn

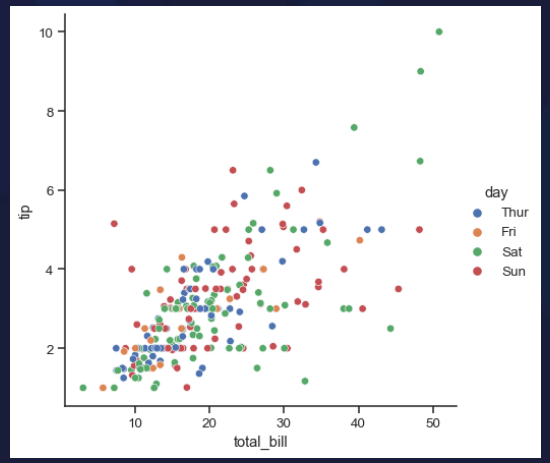
`sklearn.tree.DecisionTreeRegressor`

+

Data visualization with seaborn



# 1. ML Model

The model is not successful and can't really predict rating accurately.

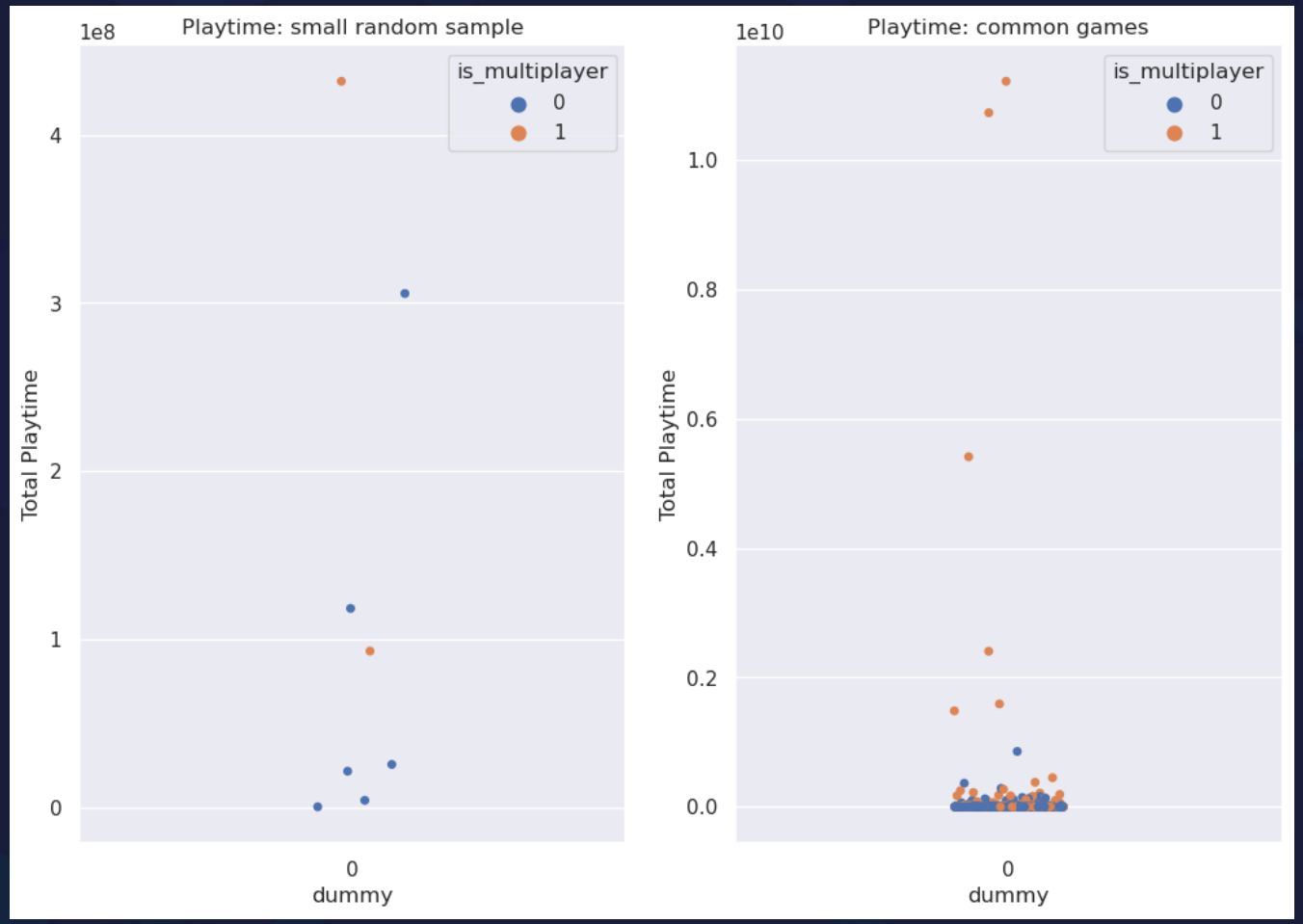
We found only a slight correlation between our selected features and the rating of a game.

The mean-squared error is high.

```
[user@sahara ~]$ python ML.py
ML Regression Model:
Train error: 83.36663701472813
Test error: 154.20593601958555
Test High Prise: [81.]
Test Low Prise: [73.]
[user@sahara ~]$ python ML.py
ML Regression Model:
Train error: 87.90965977507479
Test error: 150.6908414821261
Test High Prise: [81.]
Test Low Prise: [76.]
[user@sahara ~]$ python ML.py
ML Regression Model:
Train error: 89.34138299110272
Test error: 144.29982536605183
Test High Prise: [87.]
Test Low Prise: [75.]
[user@sahara ~]$ python ML.py
ML Regression Model:
Train error: 84.73370923725018
Test error: 158.92435446474977
Test High Prise: [68.]
Test Low Prise: [74.]
[user@sahara ~]$ python ML.py
ML Regression Model:
Train error: 89.34061378102355
Test error: 146.7895591805063
Test High Prise: [81.]
Test Low Prise: [72.]
[user@sahara ~]$
```

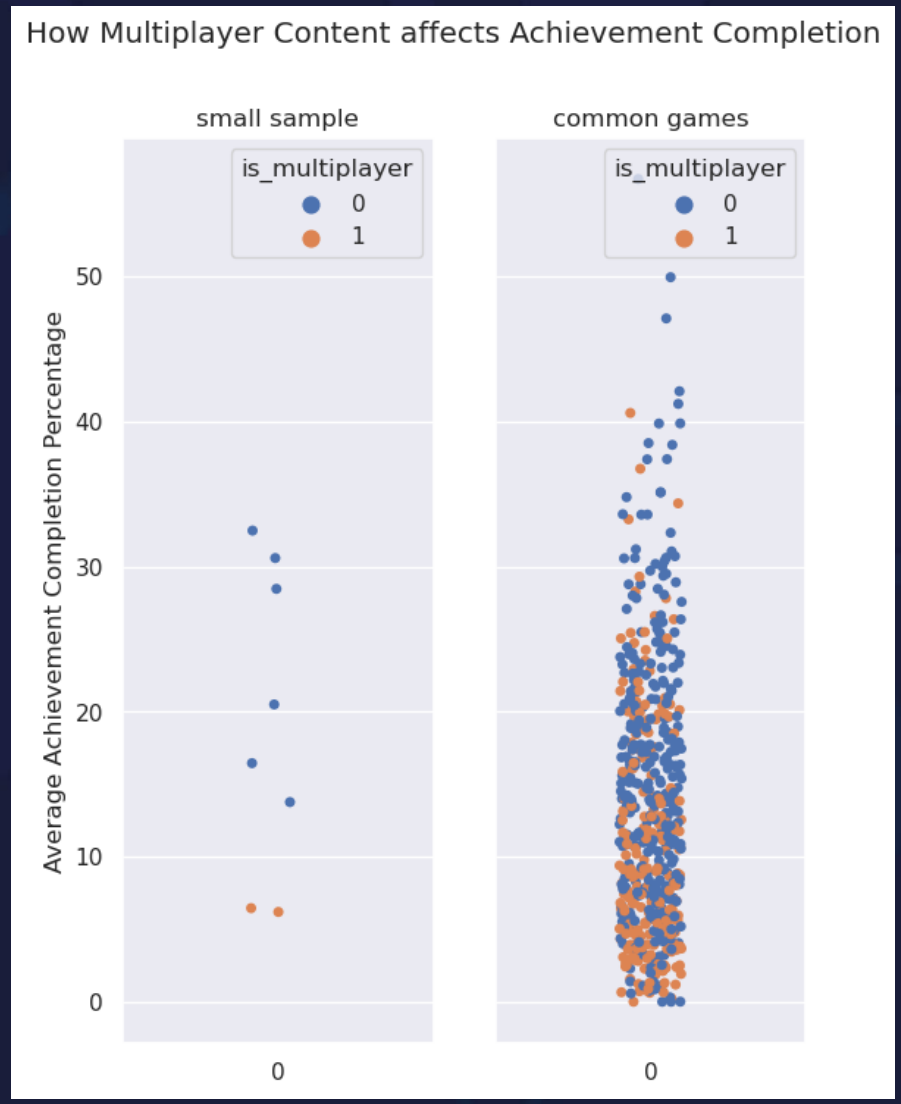
# 2. Playtime Multiplayer vs Single-player

We found that multiplayer content does, in general, have higher playtime.



# 3. Completion Multiplayer vs. Single-player

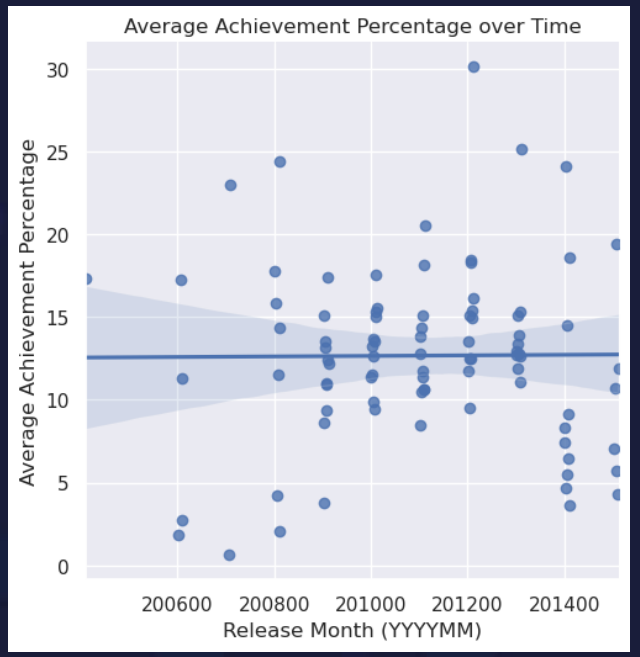
Single-player games have higher achievement completion rates compared to multiplayer ones.

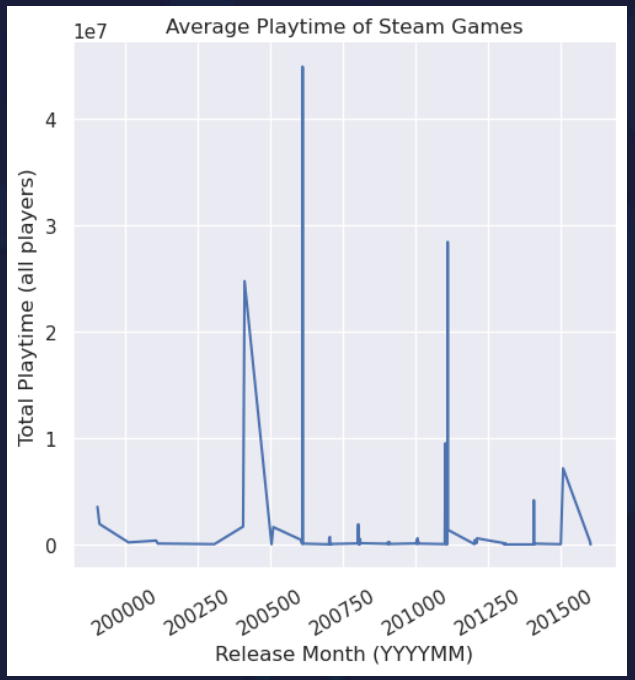


# 4. Completion and Playtime by Release Date

No trend in achievement completion over time

No trend in playtime over time



# Further Exploration

- Improve accuracy of ML model by using different features, or even more data
- Eliminate possible outliers (games that are too popular) to see how less popular games trend
- Gather newer data