STEAM GAMES

USING GAME FEATURES TO PREDICT REVENUE

What is Steam?
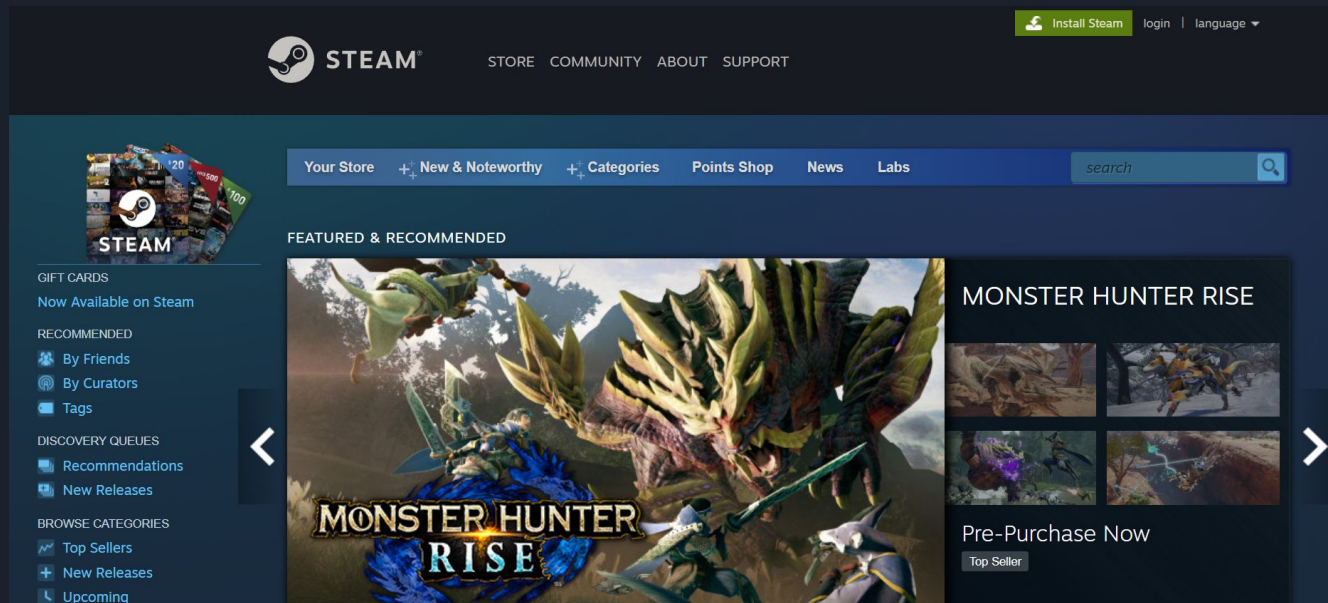
A digital platform for purchasing, playing, and forming communities around PC games

Steam Store

The Steam store uses a user-defined tagging system for more effect search, categorization, and recommendation.

Popular user-defined tags for this product: (?)

- Metroidvania
- Souls-like
- Platformer
- Great Soundtrack
- Difficult
- 2D
- Indie
- Atmospheric
- Exploration
- Singleplayer

Objective

1) Design an effective model to predict game revenue.

2) Identify which game features are the most influential in predicting revenue.

Data

- Sourced from Nik Davis on kaggle

- After cleaning, dataset contained 26,355 games and 35 features.

- Features include info such as a) genre, b) type of gameplay, c) season of release, and d) average playtime.

- **Target Feature:** Revenue (Game Price * Avg Owners)

Key Findings

- Most games do not generate much revenue.

    - The median revenue is about $58,000, while the mean revenue is closer to $1,000,000.

- Summer seems to be the most influential season in generating revenue.

- Indie, RPG, and Co-op games are highly predictive of revenue.

Influential Game Features

- Indie

- RPG (Role-Playing Game)

- Co-op (Cooperative)

- Steam achievements

- Tags per game
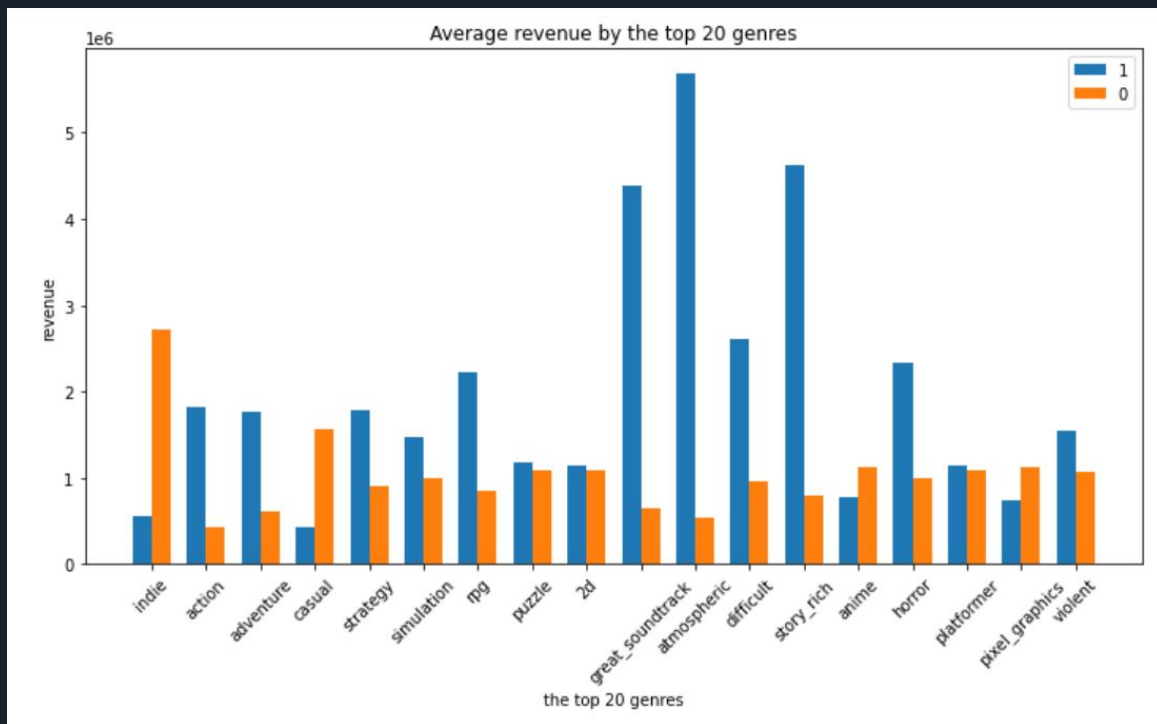
- Released in summer

Influential Model Predictors

- Total ratings

- Median playtime

- Percent positive ratings

- Years since release
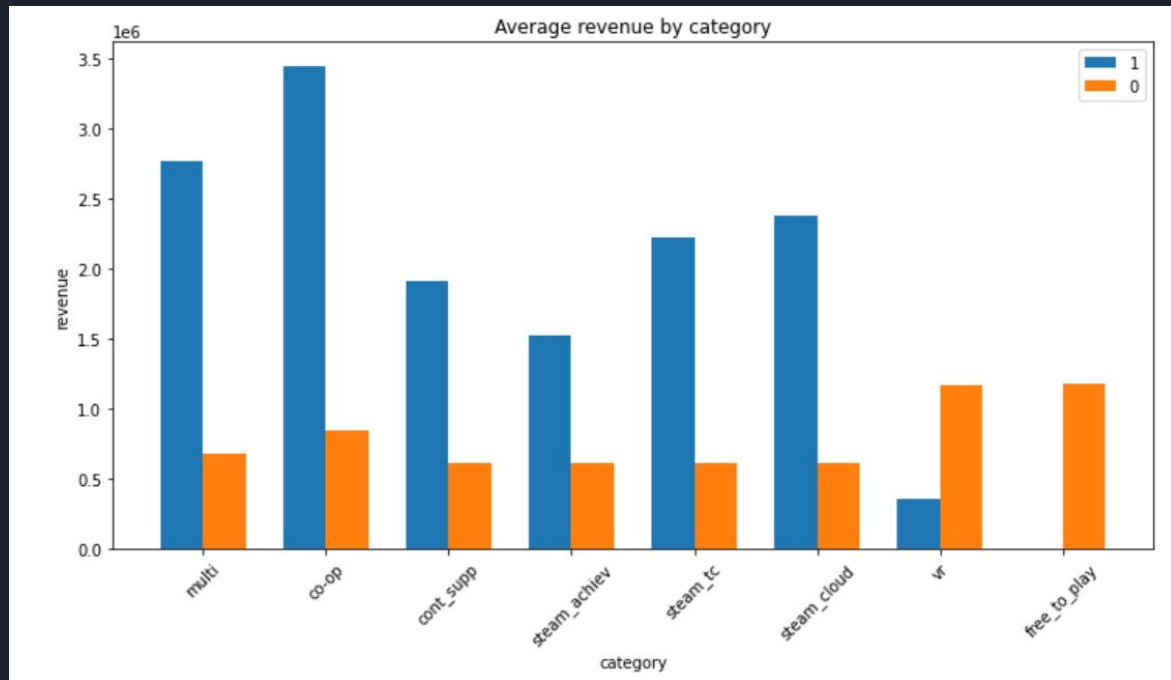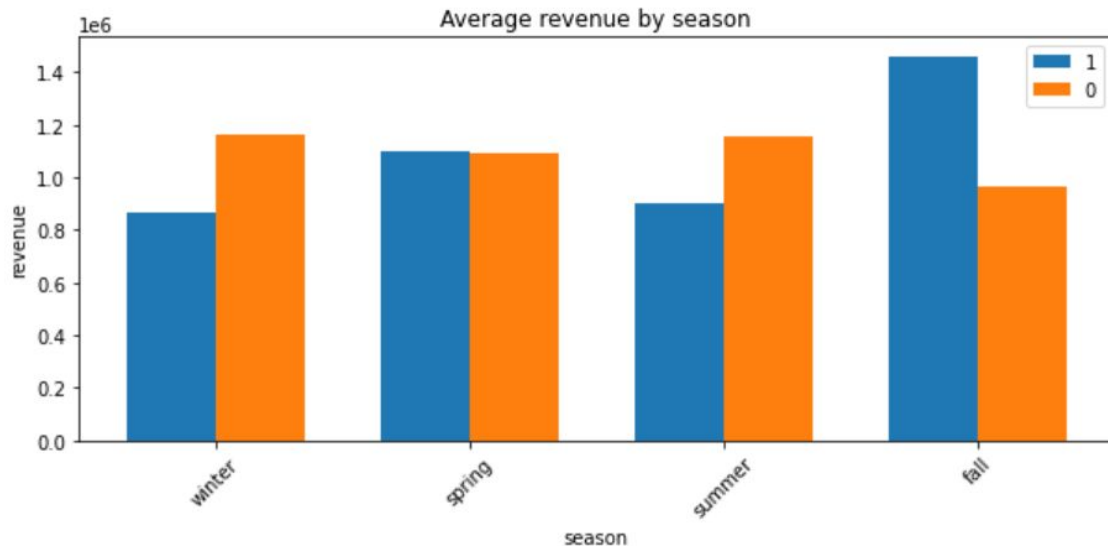
Genres

Average revenue by the top 20 genres

- ✅ Great soundtrack, atmospheric, and story-rich games.
- ❌ Indie and casual games.

Game Categories

Average revenue by category

- ✅ Co-Op and Multiplayer
- ❌ VR

Season of Release

Average revenue by season

- ✅ Fall
- ❌ Winter and Summer

Summary

- Game features do have some level of influence in predicting revenue, though a more extensive analysis may yield a more accurate model.

- Some popular tags such as RPG and co-op are associated with higher revenue.

- Summer and winter are surprisingly associated with lower revenue.

Limitations

- The distribution of our target feature, revenue, was not normally distributed.

  - Many games in the store were small and independently produced, meaning they did not generate much revenue.

  - A few games were highly successful, and generated significantly more revenue than other games.

	price	revenue
count	24,515.00	24,515.00
mean	6.71	1,259,020.88
std	8.01	15,439,153.55
min	0.32	7,200.00
25%	2.09	28,900.00
50%	4.79	69,650.00
75%	7.49	156,750.00
max	421.99	2,024,250,000.00

Future Directions

- The scope of this analysis was limited to the top 20 most popular genres.

  - An exploration of less popular genres might demonstrate a potential impact of niche interests.

- Our model evaluates each game independently.

  - This does not take into consideration franchises with a pre-existing fanbase, or well-known creators.

Thank you!