



---

# Popularity of Steam Platform Games

Eric Graves

Springboard Data Science Capstone Project

# The Data

---

- Scope – games on the Steam platform (application by Valve Corporation that sells/distributes computer games virtually)
- Source – a data extract pulled from SteamSpy, a website by Sergey Galyonkin that aggregates and reports data from Valve's Web API, in 2017
- Size – 10,294 games (rows)
- Features – focus areas included estimated number of players (all-time), genre, release date, and price

# Data Processing

---

- Removed duplicate games and records that were not games
- For certain fields, replaced zeros with nulls (e.g., Metacritic score)
- Replaced all null values with “None” for better processing in Tableau
- One-hot encoded language column into multiple columns per language
- Pivoted genre columns for aggregation of player counts by genre in Tableau
- Binned prices for easier visualization

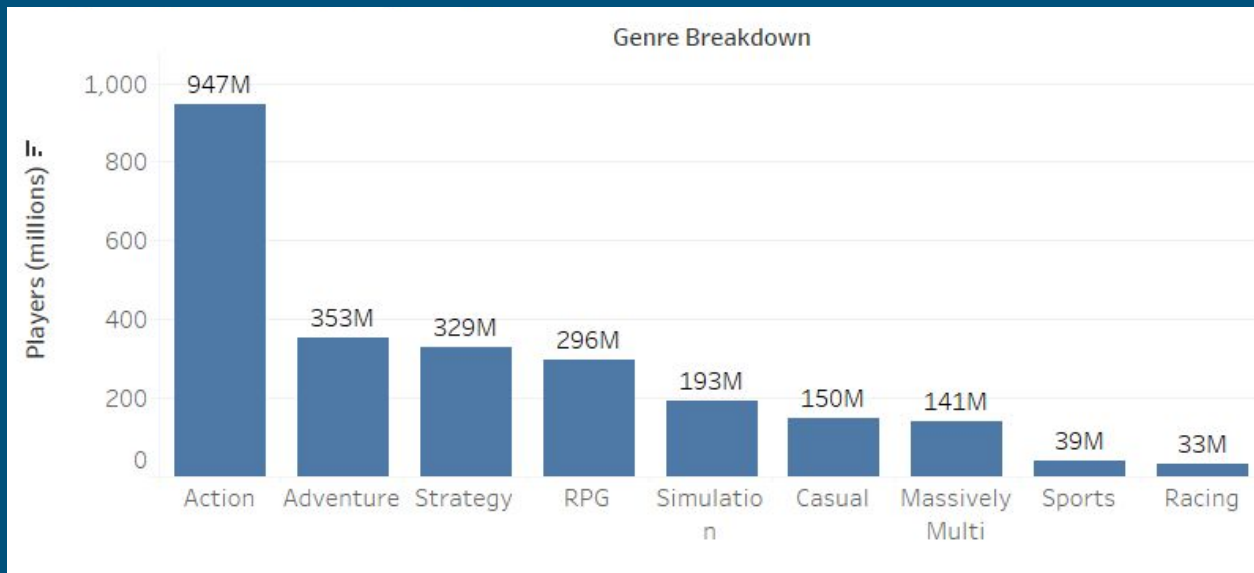
# Snapshot of Top 10 Games

- Most top games on Steam are also made by Valve
- Action is the most popular genre
- Three of four top games are free, but a fuller price analysis is shown later
- The user can drill down to see the top 10 for any combination of genre, price, and release date filters

Game	Players	Release Year	Price (USD)	Action	Adventure	Casual	Massive Multi	Racing	RPG	Simulation	Sports	Strategy
Dota 2	90.7M	2013	\$0	●	○	○	○	○	○	○	○	●
Team Fortress 2	37.9M	2007	\$0	●	○	○	○	○	○	○	○	○
Counter-Strike: Global Offensive	25.2M	2012	\$15	●	○	○	○	○	○	○	○	○
Unturned	21.4M	2014	\$0	●	●	●	○	○	○	○	○	○
Left 4 Dead 2	13.6M	2009	\$20	●	○	○	○	○	○	○	○	○
Garry's Mod	11.9M	2006	\$10	○	○	○	○	○	○	●	○	○
Counter-Strike: Source	11.5M	2004	\$20	●	○	○	○	○	○	○	○	○
The Elder Scrolls V: Skyrim	10.9M	2011	\$20	○	○	○	○	○	●	○	○	○
Sid Meiers Civilization(r) V	9.2M	2010	\$30	○	○	○	○	○	○	○	○	●
Counter-Strike	9.1M	2000	\$10	●	○	○	○	○	○	○	○	○

# Aggregate Trends: Players Across Genres

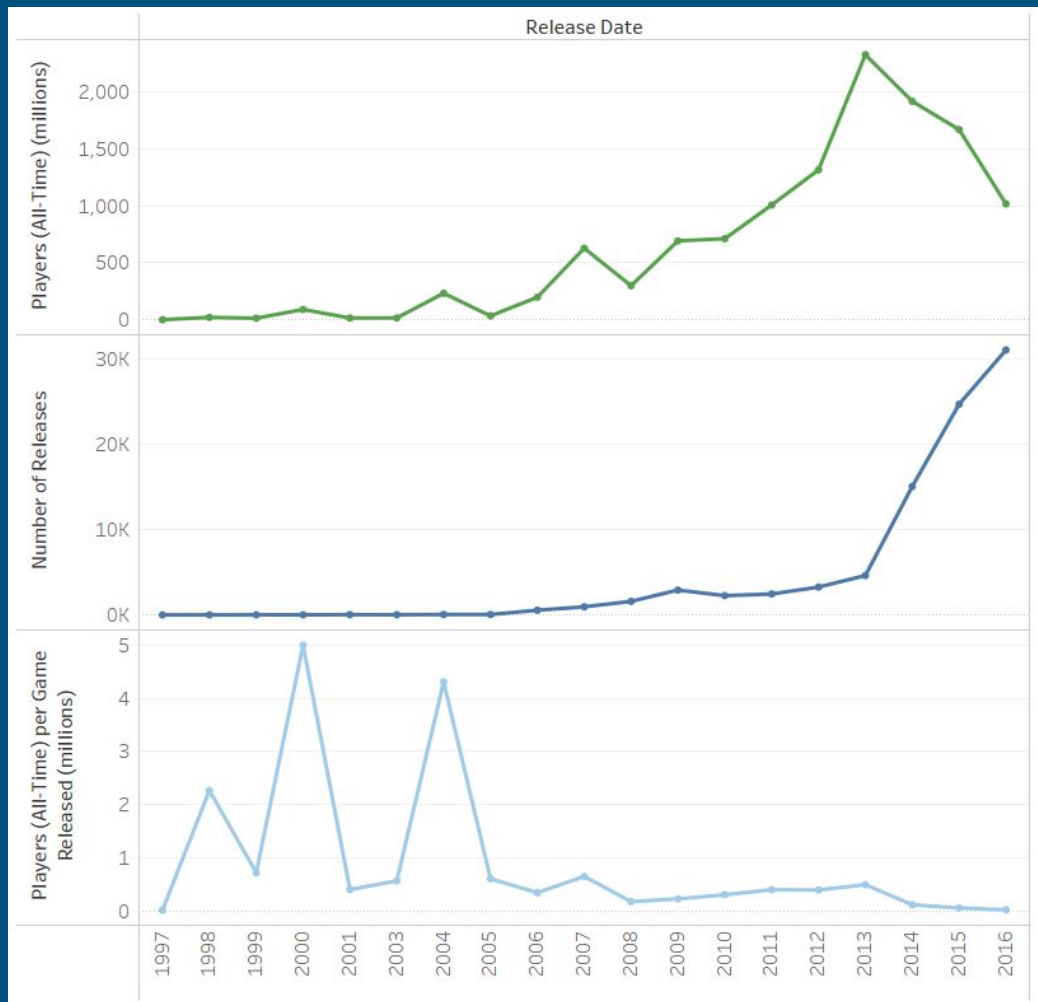
- The action genre leads, followed by adventure, strategy, and RPG
- Importantly, games can have multiple genres on Steam
- Thus, the same game and its estimated player count can be included under multiple genres in the graph



# Aggregate Trends: Release Dates

- The top view shows total players for the games released in each year
  - As a reminder, a game's player count is "all-time" and thus grows over time
- Since total players is influenced by total games released, the second view shows total games released each year as additional context
- The third view combines the first two, showing the average players per game released each year

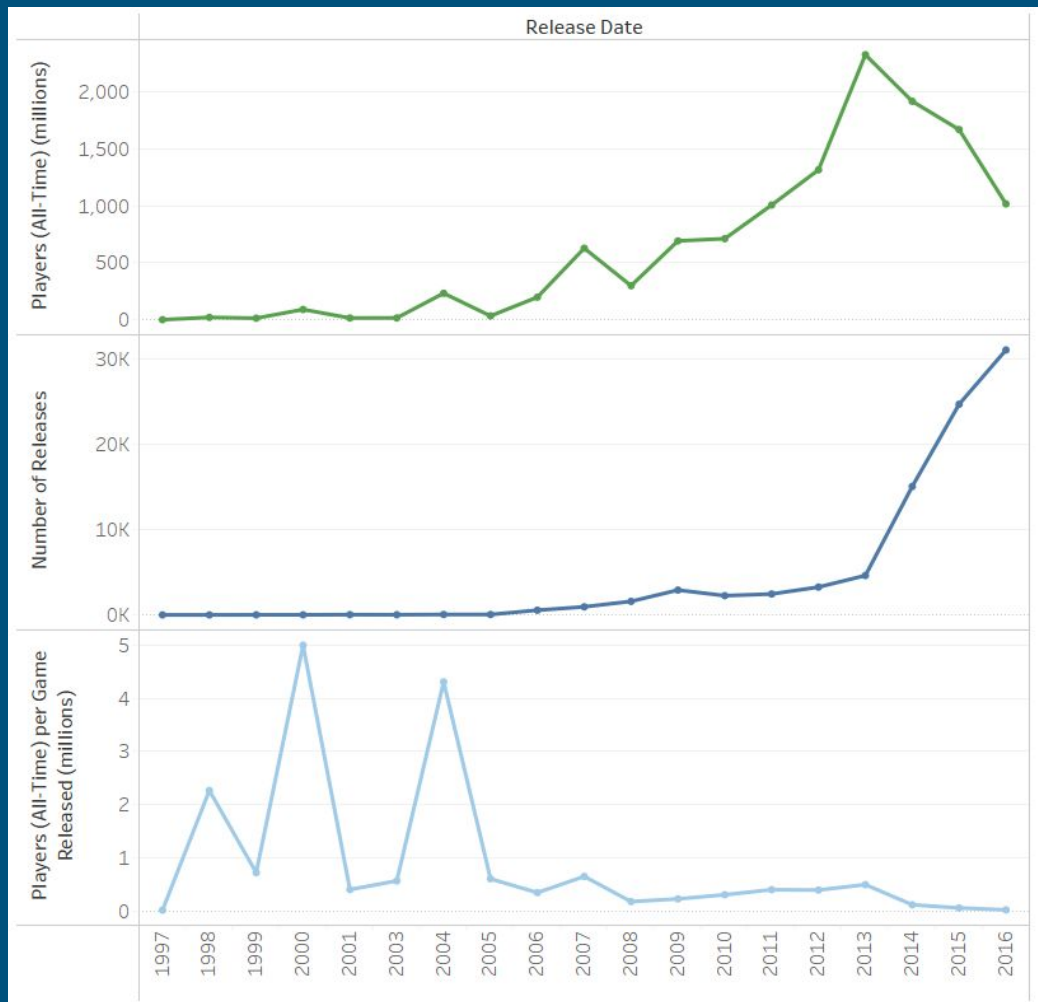
*(continued on following page)*



# Aggregate Trends: Release Dates (cont.)

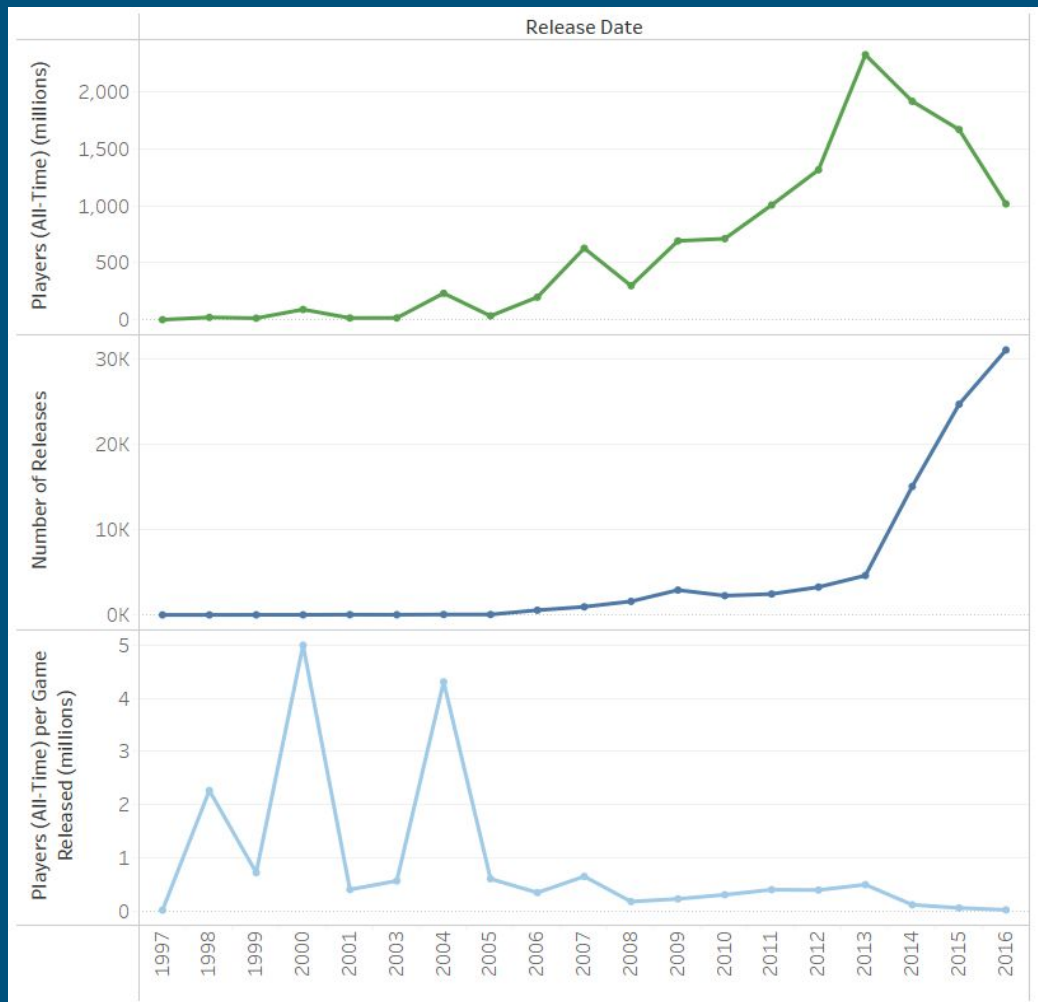
- Total players, and average players per game, appear to decrease in the most recent years in the data set
  - This is likely because newer games take time to grow in popularity and accumulate players
  - We would expect these figures to normalize over time

*(continued on the following page)*



# Aggregate Trends: Release Dates (cont.)

- Earlier years exhibit higher average players per game, which may be primarily driven by two factors:
  - Steam was launched in 2003, so when loading old titles onto Steam, Valve may have focused only on their more popular games
  - Steam didn't start offering third-party titles until 2005





# Aggregate Trends: Pricing

- Games with both an upfront price and in-game purchases available represent the most-played pricing category
  - Truly free games (no upfront cost and no in-game purchases) represent the second-largest category
  - This suggests price is driven by game quality/popularity, rather than popularity being driven by price
- This is further evident in the second figure, which illustrates that while the average free game may be more popular than games with low upfront costs, games with higher upfront costs are the most popular

Is Free	Purchase Avail		Grand Total
	False	True	
False	11M	936M	947M
True	358M	73M	431M
Grand Total	370M	1,009M	1,378M

