Statistical Analysis of User Experience in Steam Video Games

Mark Sverdlov   Or Yacobovich

July 2024

- Steam is a video game digital distribution service and storefront managed by Valve
- Due to its enormous size and unique business model, it has access to rich data both about video games and user experience.

Background

- Steam is a video game digital distribution service and storefront managed by Valve
- Due to its enormous size and unique business model, it has access to rich data both about video games and user experience.

- We aim to use the rich data to learn general phenomena about the connection between user experience and other features

- We aim to use the rich data to learn general phenomena about the connection between user experience and other features

Research Question

Main Question

What are the relationships between pricing, user experience, game quality, and actual use?

Research Question

Main Question

What are the relationships between pricing, user experience, game quality, and actual use?

- Is pricing fair in general? Do more expensive games tend to be of better quality?
- How do different metrics for user experience and game quality relate? Do they complement each other?
- Can we statistically differentiate between games from different periods?

- We focus on a well-regularized part of the data - only games with Metacritic scores.
- We define and derive 15 statistics that we believe are useful for our research.
- We use various statistical tools and tests to explore and understand the relationships between the statistics.

Statistics

User Experience
- `metacritic_score`
- `user_ratio`
- `median_playtime`

Pricing
- `price`
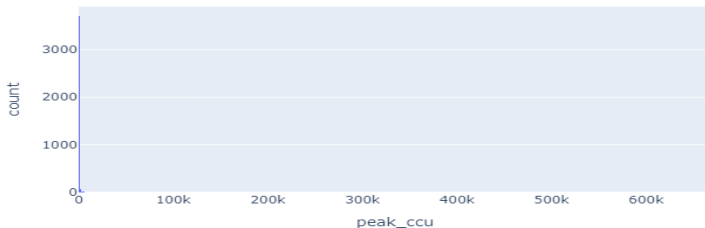
Period
- `release_date`
- `age`

Popularity
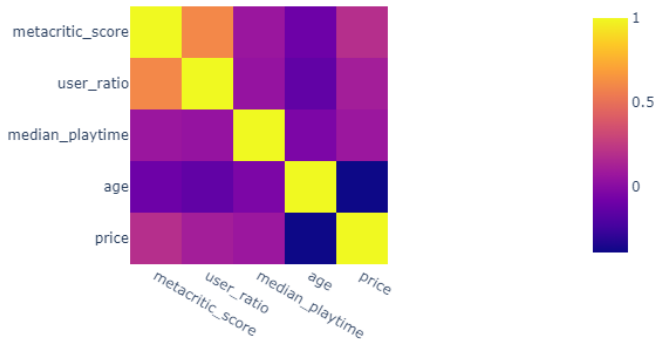- `estimated_owners`
- `peak_ccu`

Genres
- `action`
- `sport`
- `massively multiplayer`
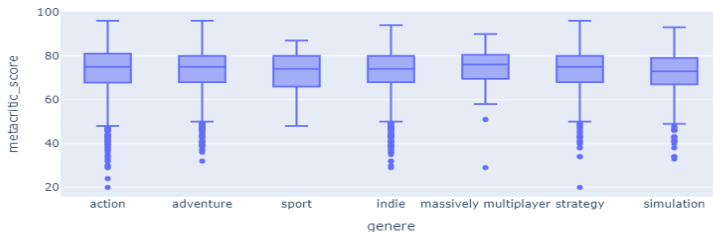- `adventure`
- `indie`
- `simulation`
- `strategy`

estimated_owners	0-20,000	20,000-50,000	50,000-100,000	100,000-200,000	200,000-500,000	500,000-1,000,000
release_date						
2006-2008	0.000149	0.156821	0.399077	0.646302	0.054379	0.011672
2009-2011	0.000003	0.157383	0.188836	0.094387	0.209501	0.205565
2012-2014	0.000000	0.015499	0.670774	0.029459	0.000344	0.401351
2015-2017	0.448981	0.000069	0.428553	0.443582	0.186639	0.097974
2018-2020	0.000001	0.895796	0.765696	0.249248	0.008644	0.997289
2021-2023	0.000000	0.447881	0.509054	0.237833	0.233270	0.214286

Initial Results - Relationships Between Features

	action	adventure	sport	indie	massively multiplayer	strategy	simulation
action	0.500006	p=0.709, diff=0.00	p=0.777, diff=1.00	p=0.982, diff=1.00	p=0.268, diff=-1.00	p=0.578, diff=0.00	p=0.999, diff=2.00
adventure	p=0.291, diff=0.00	0.500007	p=0.723, diff=1.00	p=0.940, diff=1.00	p=0.208, diff=-1.00	p=0.402, diff=0.00	p=0.999, diff=2.00
sport	p=0.223, diff=-1.00	p=0.277, diff=-1.00	0.500423	p=0.462, diff=0.00	p=0.154, diff=-2.00	p=0.237, diff=-1.00	p=0.780, diff=1.00
indie	p=0.018, diff=-1.00	p=0.060, diff=-1.00	p=0.538, diff=0.00	0.500005	p=0.104, diff=-2.00	p=0.058, diff=-1.00	p=0.979, diff=1.00
massively multiplayer	p=0.732, diff=1.00	p=0.792, diff=1.00	p=0.847, diff=2.00	p=0.896, diff=2.00	0.500736	p=0.757, diff=1.00	p=0.975, diff=3.00
strategy	p=0.422, diff=0.00	p=0.598, diff=0.00	p=0.763, diff=1.00	p=0.942, diff=1.00	p=0.243, diff=-1.00	0.500016	p=0.999, diff=2.00
simulation	p=0.001, diff=-2.00	p=0.001, diff=-2.00	p=0.220, diff=-1.00	p=0.021, diff=-1.00	p=0.025, diff=-3.00	p=0.001, diff=-2.00	0.500034

- The shape of data is very skewed: 95% of the games are 'small'.
- Among the 'user experience' variables, "user ratio" and "Metacritic score" are correlated but not overlapping, but "median playtime" seems to be entirely unrelated.
- There seems to be statistically significant bias against certain genres, but the bias differs between different user experience metrics.

Plans for Further Work

- Continue to delve into the data and improve our understanding, utilizing different a-parametric tests.
- Answer questions about the pricing mechanisms and test statistical hypotheses about it.
- Try to model the ever-enigmatic idea of 'quality' by fitting a multi-linear model between the features.