

Players engagement with respect to game genres, on Steam public data - Milestone 1

Team "no preference"

Ulysse Widmer, Jack Lau, Luis Busta

EPFL, COM-480 Data Visualization

April 2023

Contents

1	Dataset	1
2	Problematic	2
3	Exploratory Data Analysis	2
4	Related work	2

1 Dataset

We've chosen to work on Steam data. We found a few interesting datasets on Kaggle:

- <https://www.kaggle.com/datasets/tamber/steam-video-games>
- <https://www.kaggle.com/datasets/nikdavis/steam-store-games>
- <https://www.kaggle.com/datasets/jackogozaly/steam-player-data>
- <https://www.kaggle.com/datasets/trolukovich/steam-games-complete-dataset>
- <https://www.kaggle.com/datasets/souyama/steam-dataset>

We've also looked at already existing data visualizations on Steam public data, including the two well-known websites

- <https://steamcharts.com/>

- <https://steamdb.info/>

Finally, we've looked directly at Steam API to see how to automatically gather data in real-time directly from Steam:

- https://developer.valvesoftware.com/wiki/Steam_Web_API

We've agreed that the most interesting would be to use Steam API directly. Therefore there is no data-cleaning necessary.

Pre-processing Depending on what kind of visualizations we want to achieve (see next section) we intend to do some pre-processing to aggregate some statistics to get novel visualization that goes beyond what can be found on *steamcharts* and *steamdb*.

2 Problematic

General Topic We want to work on Steam public data, namely the time series of player count per game, per category, per genre, and also the playtime per game per user of public profiles. Our goal will be to highlight player engagement toward game genres over time. We want to go toward something like this: https://www.youtube.com/watch?v=54oPNwB_zQY and show which genres are most popular over time, according to which metrics, etc.

Visualization Our first idea is to make a big interactive chart with sliders that allow the user to navigate through different metrics and through time.

Target audience Our visualization targets curious gamers on one hand, those who want to see what is or was hot; and game developers who may want to get an insight into the evolution of general trends in the gaming community of Steam.

3 Exploratory Data Analysis

You can find the exploratory data analysis on our Jupyter Notebook `no_preference.ipynb`

4 Related work

Existing work with Steam data Existing visualizations on Steam data are mainly the ones that we can find on the previously mentioned websites *steamcharts* and *steamdb*, and also the visualization present on Steam itself, which are limited to time series of player count per game. Other than visualizations, most of the data present on Kaggle is used for machine learning, for the purpose of recommender systems.

Originality of our visualization The novelty of our visualization will be to focus on game genres instead of individual games, to get a feeling if there's a transversal engagement of certain types of players across a whole genre rather than an individual popular title.

Inspiration

- https://www.youtube.com/watch?v=54oPNwB_zQY