

# KAGGLE – VIDEO GAME SALES

Marko Peedosk   Henrik Innos   Robert Ivask

## Introduction

The primary goal of this project was to analyse trends video game sales and identify features, which make games successful. The secondary goal was to determine whether it is possible to predict sale numbers based on these metrics.

## Data

In this project we used 3 datasets containing different information about video games from the years 1980-2016. These were combined into one to get the highest possible amount of cases and features. Additionally, we scraped data from <https://www.metacritic.com> to fill missing values.

## Methodology

Steps:

1. Merging and cleaning the datasets
2. Feature engineering to determine useful features
3. Exploratory data analysis
4. Creating a predictive model using Linear, Ridge and Lasso regression
5. Hyperparameter tuning
6. Evaluating the model using RMSE

## Results

- The best window to release your game is September through November.
- Family games have the highest average sales, while making up only a small fraction of the market, which could indicate an untapped market.
- Shooter games, despite making up a large part of the market, still have the fifth highest average sales.
- Over time, almost in every genre, average user satisfaction has either decreased or remained the same. There has been little to no increase.
- The industry is top-heavy. Few big releases have very high sales, while majority of the market contends with comparatively low figures.
- The age of physical copies is over. With the rising popularity of digital sales, physical sales have taken a nosedive.
- Due to insufficient features and coverage of only physical sales data, as digital sales gain prominence, our dataset lacked the breadth needed to train a sufficiently accurate predictive model.

## References

<https://www.kaggle.com/datasets/shivamvadalia27/video-games/data>  
<https://www.kaggle.com/datasets/rush4ratio/video-game-sales-with-ratings?rvi=1>  
<https://www.kaggle.com/datasets/jummyegg/rawg-game-dataset?rvi=1>  
<https://www.metacritic.com>  
<https://github.com/rops11/DS-Project-2023>

