Rusa Maja Pedigo

May 2022

Video games sales

Analysis and Prediction

# Overview

## Project Background and Description

This project is intended for the XYZ company who wishes to analyze different market shares of video games makers in US. Company’s goal is to predict sales of their products globally for the next year.

## Data

Project will consist of in-depth analysis of the dataset containing a list of video games with sales greater than 100,000 copies. It was generated by a scrape of [vgchartz.com](http://www.vgchartz.com/).

## Anticipated Data Science Approach:

As the goal of the project is prediction of global sales, in millions, we will conduct the following:

* Initial loading and cleaning of the data
* Exploration of the data and visualization of statistical inferences and trends
* Preprocessing of the data for further model building
* Modelling of the data via Linear regression, Random Forest regressor, Decision Tree Regressor
* Choosing the best model and hypertuning the parameters
* Final predictions and conclusions

## Deliverables

Project outputs will include:

* Project Proposal (this document)
* Jupyter Notebooks with phases of data processing:
  + Data Cleaning
  + Data Exploration and Visualization
  + Preprocessing and Testing
  + Modeling
* Final Report
* Presentation Slide Deck