

University of Amsterdam

# Big hit analysis

Predicting the genre of the next big hit video game  
for Xbox, Playstation, Nintendo and PC

Floris Kappeyne, Hester Verdenius, Xavier van Dommelen

December 20, 2018



## Introduction

## Analysis

The data set

## Analysis

Multi Linear regression

Noise tests

## Conclusion



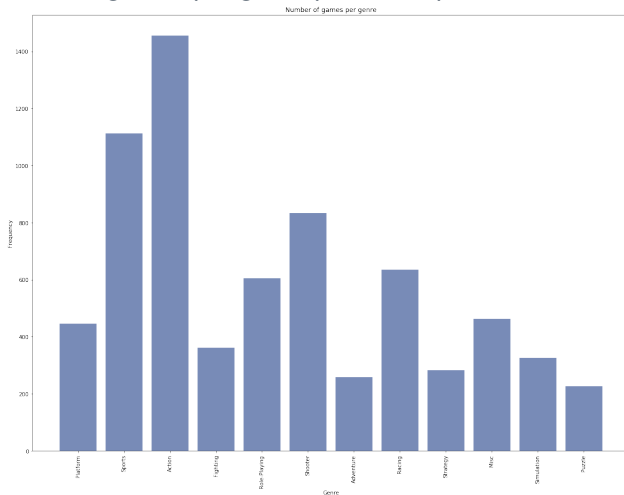
- ▶ Big hit
- ▶ Problem
- ▶ Interest
- ▶ The data



It is possible to predict the genre of the next big hit with significant accuracy based on statistical analysis for each of the following platforms: Xbox, Playstation, Nintendo and PC.

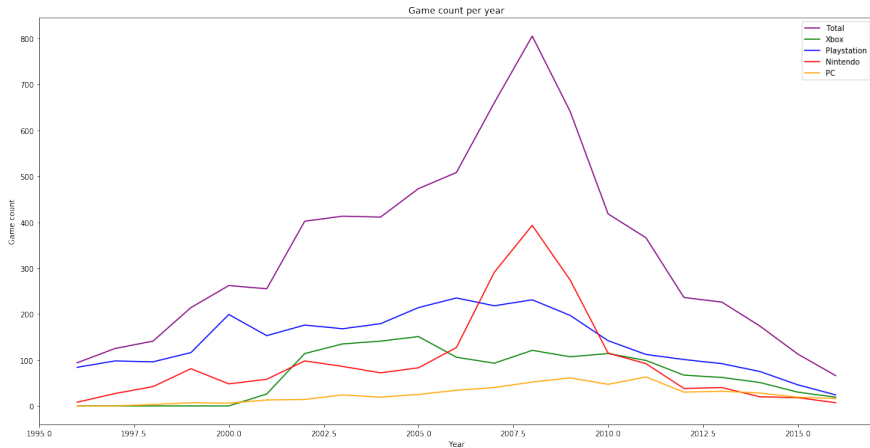


## Number games per genre plot for all platforms.

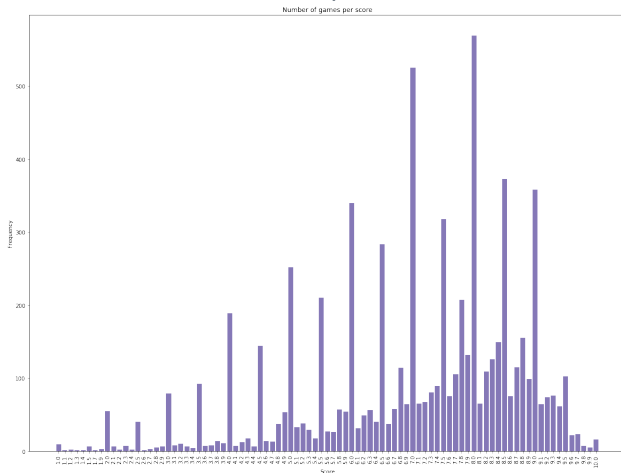




## Game count per year per platform.



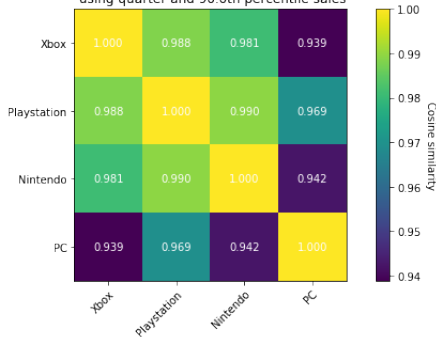
## Distribution of scores for all platforms.



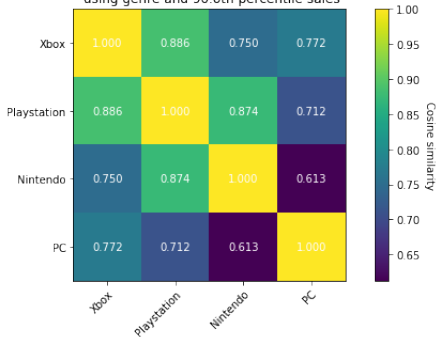


## Heatmaps of the cosine similarities of genre and 90th percentile sales per platform.

Heatmap of the cosine similarity between platforms using quarter and 90.0th percentile sales



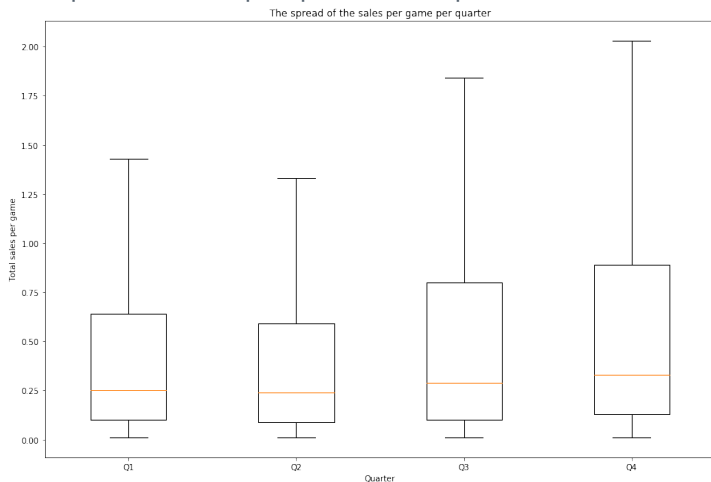
Heatmap of the cosine similarity between platforms using genre and 90.0th percentile sales



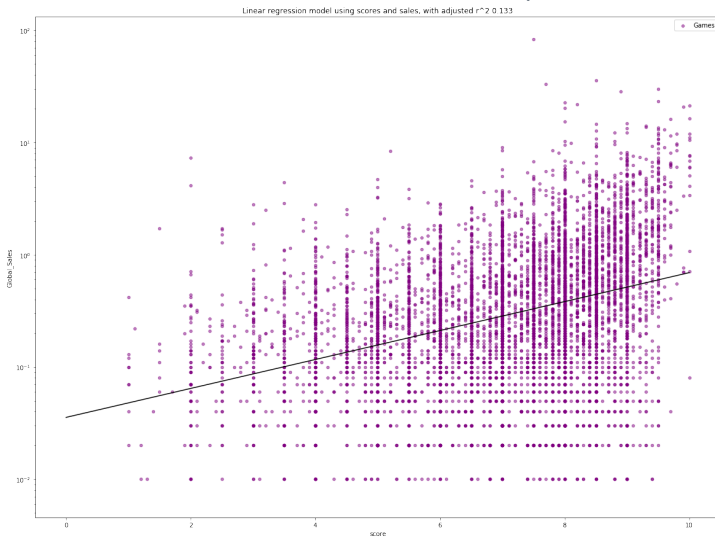




## The spread of sales per quarter for all platforms.



## Correlation between scores and sales for all platforms.



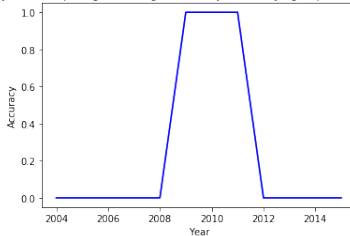


- ▶ OLS: ordinary least squares
- ▶ Use data of four years before for estimating the next year
- ▶ Dummy variables for categorical variables
- ▶ Twee modellen:
  - ▶  $Sales = score * X_0 + genre * X_1$
  - ▶  $Sales = score * X_0 + genre * X_1 + genre * score * X_2$

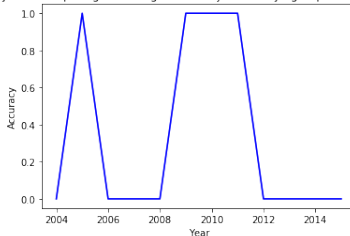


## Accuracy for predicting the top 1 big hit, for the two different models.

Accuracy with comparing to the big hit of the year we trying to predict on PlayStation

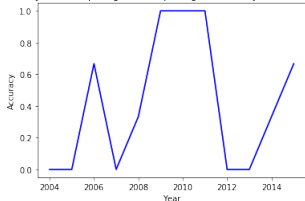


Accuracy with comparing to the big hit of the year we trying to predict on PlayStation

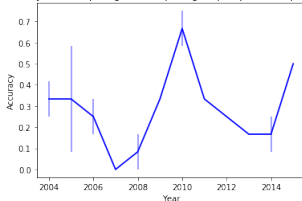


### Accuracy for predicting the top 3 big hits, for the two different models.

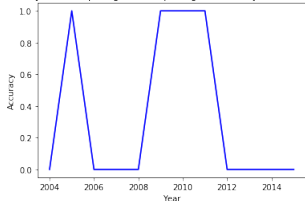
Accuracy with comparing to the top 3 big hits of the year on PlayStation



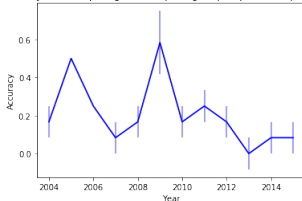
Accuracy with comparing to the top 3 big hit per quarter on playstation



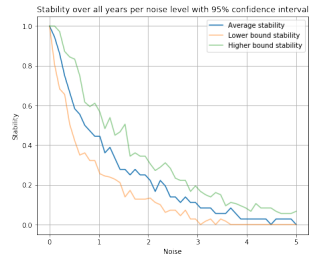
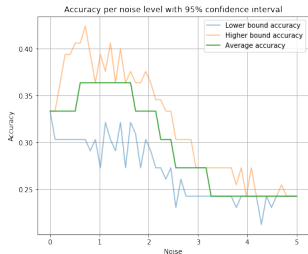
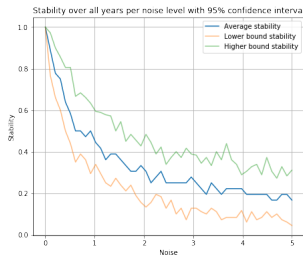
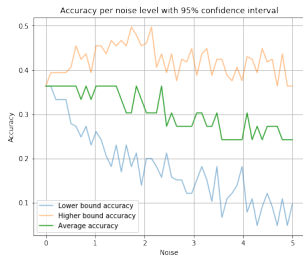
Accuracy with comparing to the top 3 big hits of the year on PlayStation



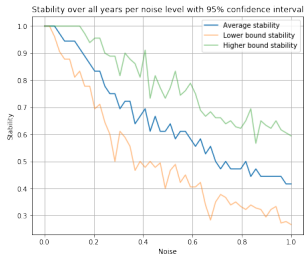
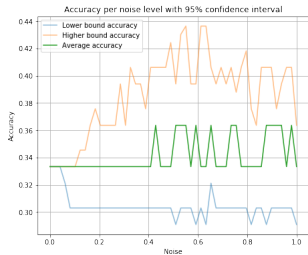
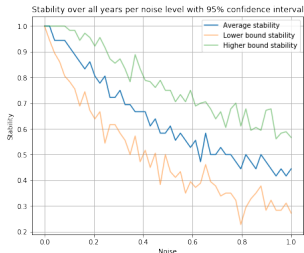
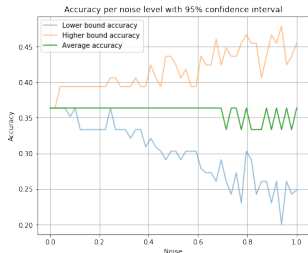
Accuracy with comparing to the top 3 big hit per quarter on playstation



## Accuracy and stability test with different noises, for all years.



## Zoomed in version, noise between 0 and 1.

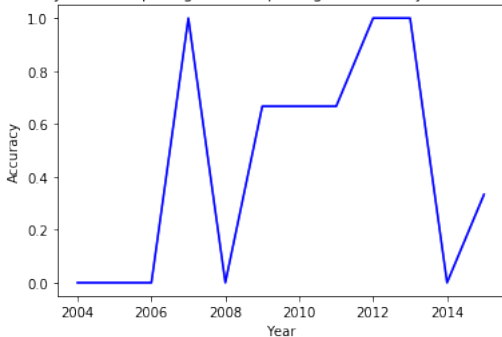




## Accuracy for predicting the top 3 big hits

- Model:  $Sales = genre * X$ .

Accuracy with comparing to the top 3 big hits of the year on PlayStation







- ▶ Find optimal time period to maximise quality.
- ▶ Does the average accuracy change when you swap the scores between genres?
- ▶ Explore the correlations and differences between consoles.
- ▶ Analyse the best and worst situations of each model and see what their strengths and weaknesses are.
- ▶ View the data as a time series, and project the changes over time.
- ▶ Improve these models by adding more information, such as development budget, development time and team size.



- ▶ We've proved our hypothesis.
- ▶ Better than random guesses, but not perfect.
- ▶ Score and genre model works better.
- ▶ Don't buy shares based on these models.

# Conclusion



Questions?