# Predicting Success of Steam Games

MICHAL TRNĚNÝ

MASARYK UNIVERSITY, 2016

### Task

Investigate how well can be post-release success of games predicted from basic information about games (available before release)

No data regarding previews, early reviews, user activity on social networks...



### Steam

Largest platform for selling PC games

Since 2004

Over 10,000 games in total



### Steam Charts

Tracks how many players are in game every hour

Monthly statistics

Since July 2012

Not very accurate measure of success (?)

(however, Steam Charts is the best source with sufficient history)



### SteamSpy

Tracks number of owners

Since April, 2015

### steamspy Rocket League



Manually collected data for about 1,400 games

No access to history

Still not a good measure of success?

### Store | Hub | SteamDB | Site Developer: Psyonix, Inc. Publisher: Psyonix, Inc. Genre: Action, Indie, Racing, Sports Languages: English, French, Italian, German, Spanish, Dutch, Portuguese, Japanese,

Korean, Russian, Turkish Tags: Multiplayer (1476), Racing (1430), Soccer (1229), Sports (1142), Competitive (974), Team-Based (943), Football (847), Online Co-Op (800), Action (694), Co-op (620), Funny (593), Fast-Paced (585), Local Multiplayer (534), Local Co-Op (530), Great Soundtrack (519), Split Screen (502), 4 Player Local (473), Singleplayer (467), Indie (362), Casual (343)

Category: Single-player, Multi-player, Co-op, Shared/Split Screen, Cross-Platform Multiplayer, Steam Achievements, Full controller support, Steam Trading Cards, Steam Cloud, Stats Release date: Jul 7, 2015

Price: \$19.99

Score rank: 87% Userscore: 93% Old userscore: 93% Metascore: 86% Owners: 4,327,767 ± 53,251 **(**)2016-04-21

Players in the last 2 weeks: 1,239,904 ± 28,629 (28.65%)

Players total: 4,153,044 ± 52,178 (95.96%)

Peak concurrent players yesterday: 56,289

YouTube stats: 3,998,144 views and 28,875 comments for videos uploaded last week,

over 50 new videos uploaded yesterday

Playtime in the last 2 weeks: 08:32 (average) 03:23 (median)

Playtime total: 62:51 (average) 18:20 (median)

OWNERS CHANGES AUDIENCE (2 WEEKS) AUDIENCE TOTAL CCU (DAILY) CCU (HOURLY) TWITCH (DAILY)

TWITCH (HOURLY) YOUTUBE REVIEWS MORE LIKE THIS RELATED GEOGRAPHY PLAYTIME (TOTAL)

PLAYTIME (2 WEEKS)

### Owners data:



Q Search

### Data

About 4,700 games from July, 2012 – July, 2016

Completely omitted Early Access titles and Free-to-play titles

Info such as genres, price, release date, descriptions, languages, features

+ Thumbnail and screenshots of every game

Full image, rPlotter

### Screenshots

K-means used to extract 5 most dominant colors

First screenshot and thumbnail

and thumbhail







Reduced image, rPlotter



Reduced image, frequent



Full image, rPlotter









Reduced image, frequent







Full image, rPlotter

Reduced image, rPlotter



Reduced image, frequent





## Subsetting

Pick only games whose publisher has a history of at least 2 games

Covers 30-40 % games

### Notable attributes

Previous games: how many, max and min players of any game, Gini index

Developer, Publisher (limited to max 53, rest is "other")

text descriptions

Price

HW requirements: disk space, RAM, GPU

Languages

Number of colors on thumbnail and first screenshot + dominant thumbnail color

### Results

Training, validation, test set (60 %, 20 %, 20 %)

Classification (10 classes)	Random Forest	Regression (values 0-15)	SVM
		Correlation coefficient	0.80
		MAE	1.24
		RMSE	1.67
Accuracy	46 % / 44 %	within +-1 from actual	53 % / 52 %
within +-1 from actual	87 % / 84 %	within +-2 from actual	82 % / 77 %

Regression (SVM, polynomial kernel)

