



# Board games recommender

Matěj Bukáček, Petr Janík, Jakub Kraus, Michal Salášek,  
2023

# Goal & Motivation

- Develop a recommender system for board games
  - users who like to play board games and want to try new ones
- Possible monetization:
  - e-shop
  - referral links to e-shops

# Data

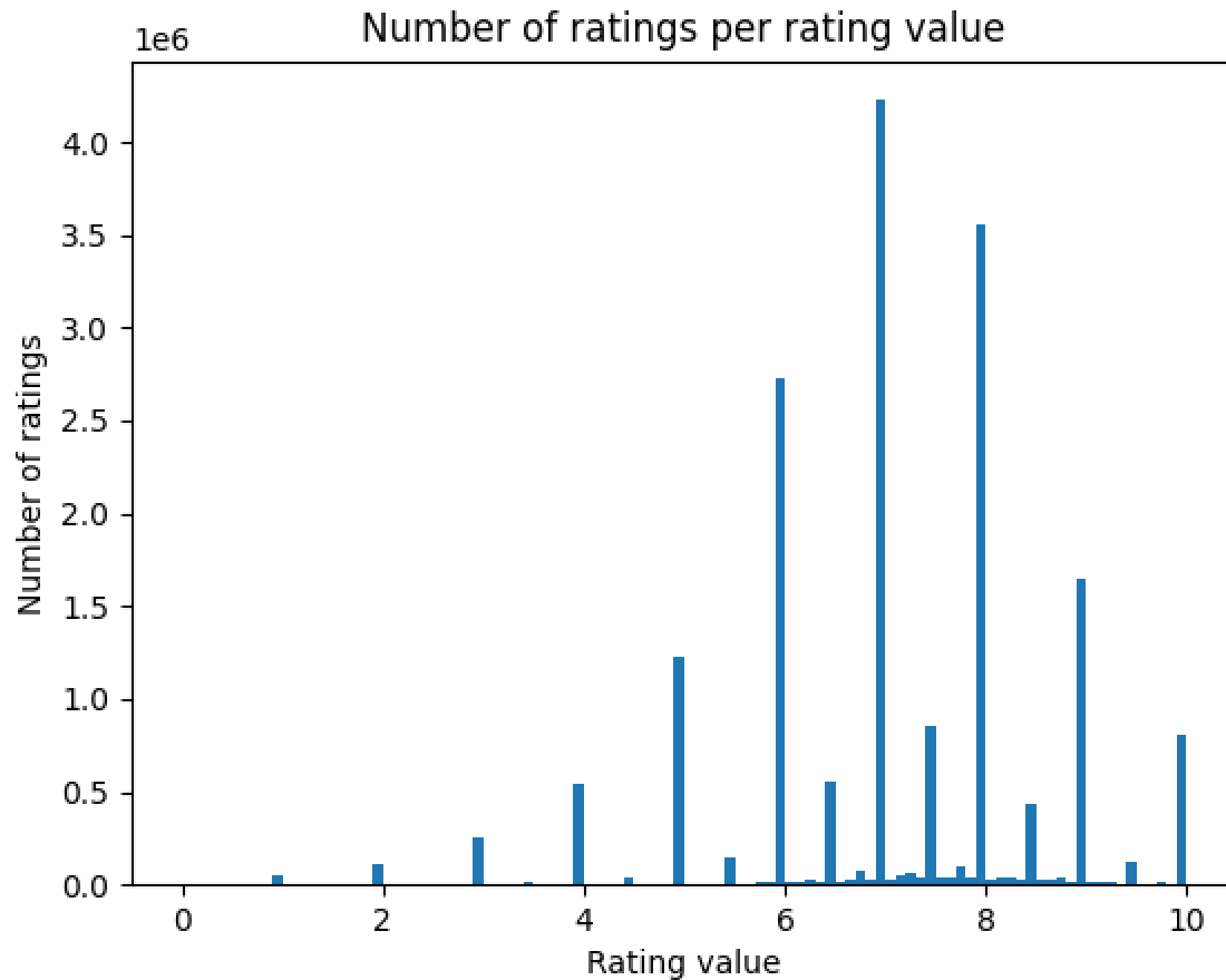
- from [kaggle](#)
- games , mechanics , subcategories , themes
- user\_ratings
  - BGGId - BoardGameGeek game ID
  - Rating - Raw rating given by user
  - Username - User giving rating

# Data preprocessing

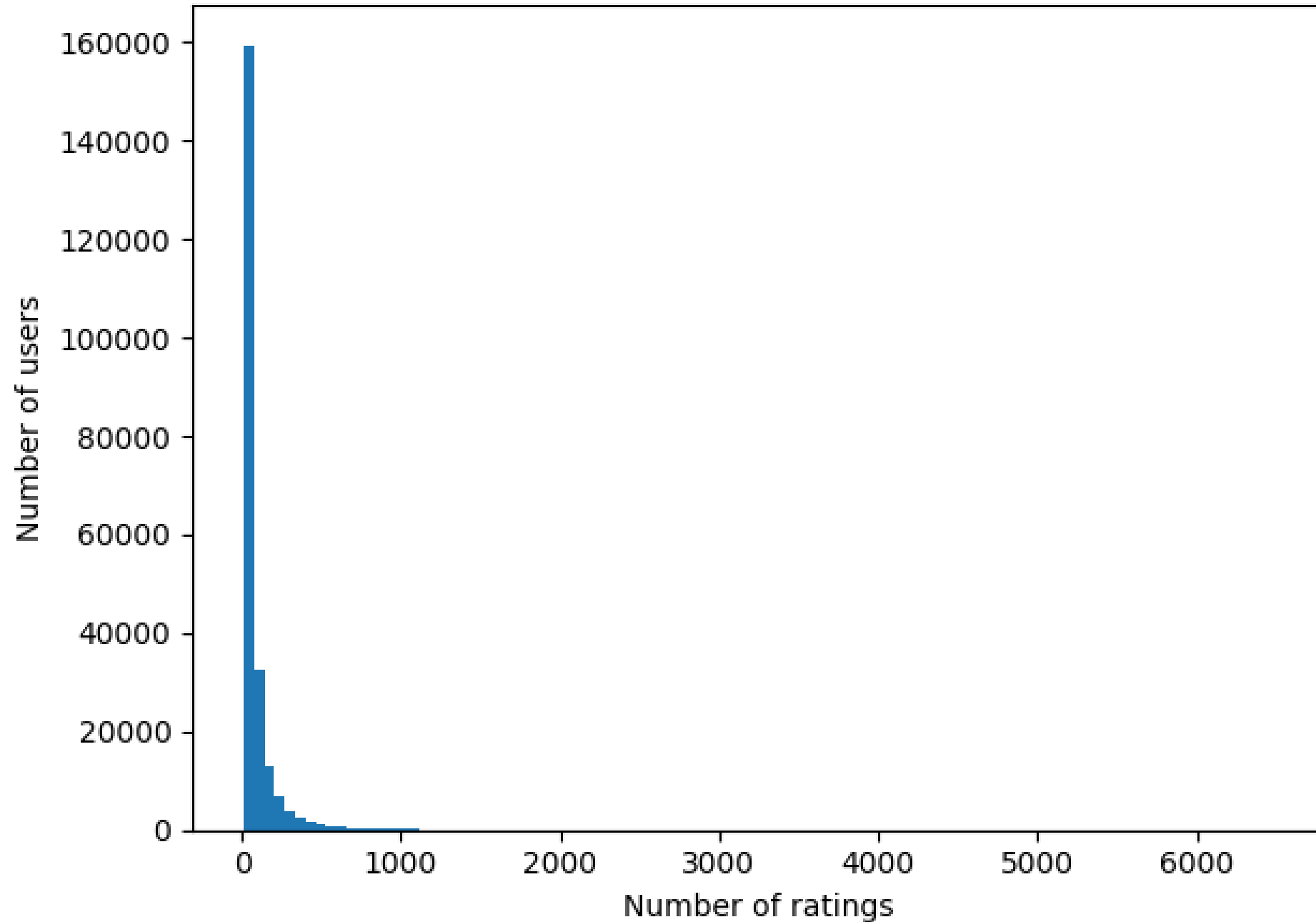
- removed duplicate ratings
- removed users with less than 10 ratings
- removed games with less than 10 ratings

# Data analysis

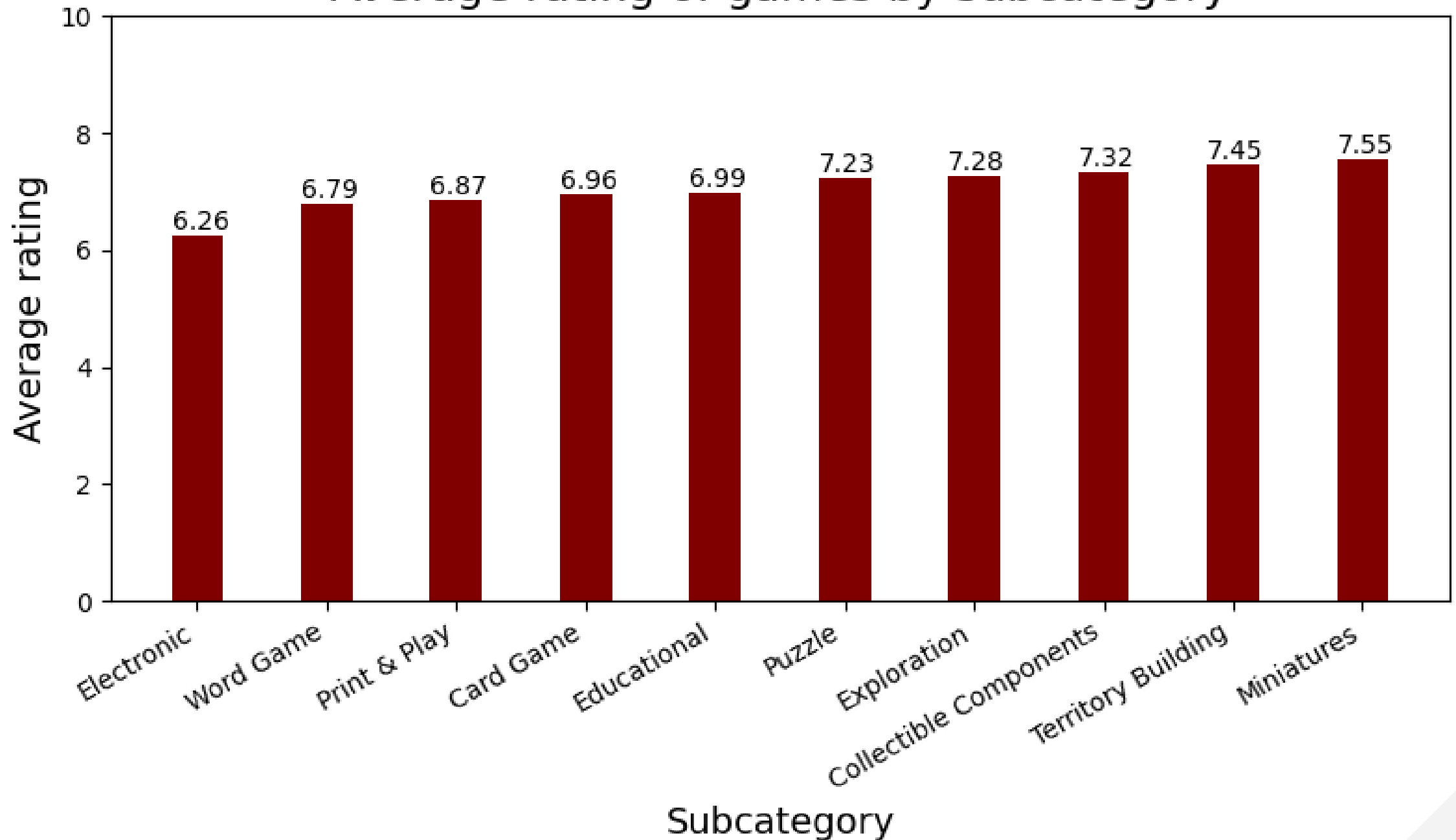
- **18 340 284** user ratings
- **224 557** users, **21 919** games, **157** mechanics, **217** themes, **10** subcategories
- density of user ratings matrix: **0.37%**
- ratings per user:
  - average: **81.67**
  - median: **39**
  - maximum: **6478**
- average rating: **7.10** (range **[0, 10]**)



Number of users per number of ratings

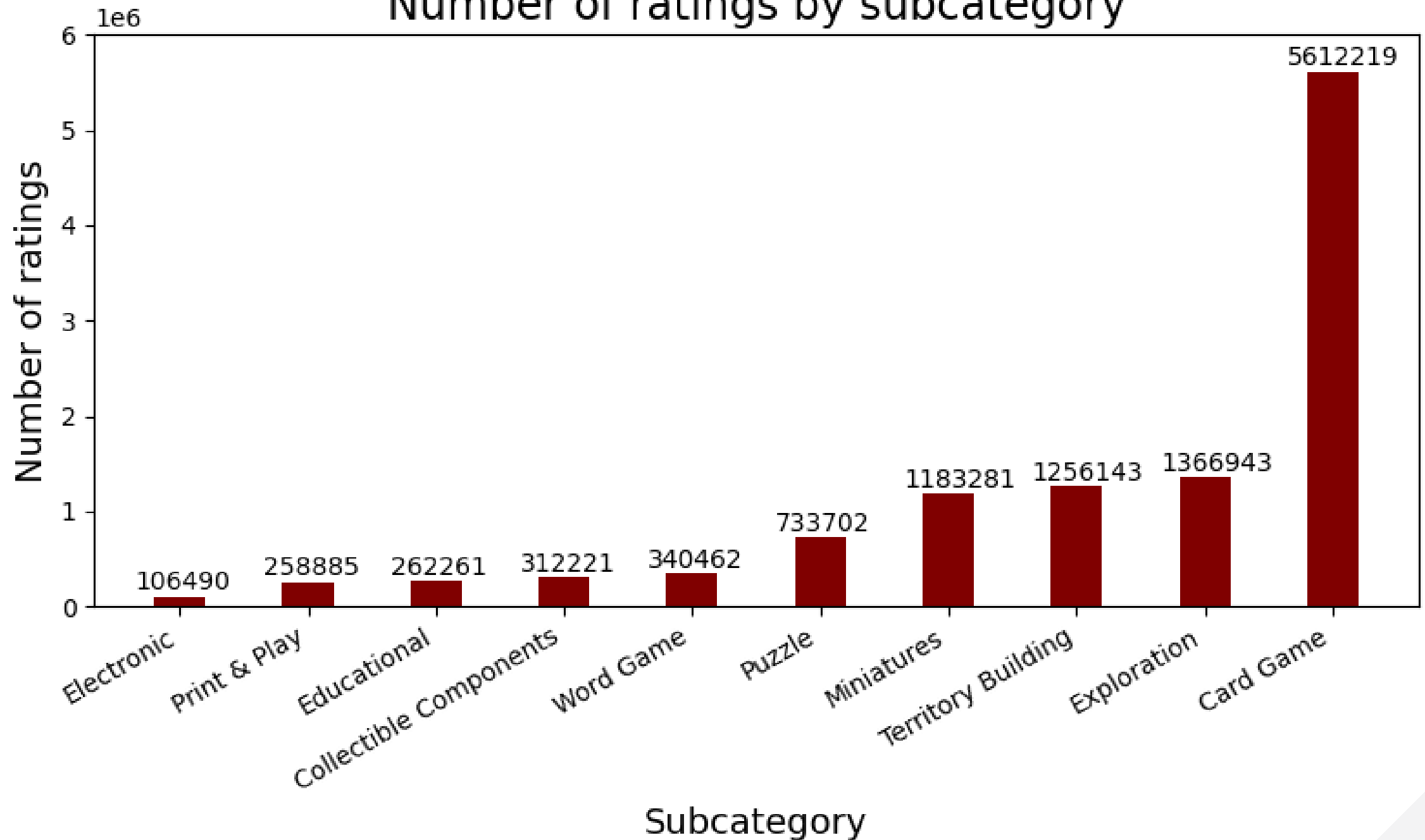


# Average rating of games by subcategory





# Number of ratings by subcategory



# TF-IDF

- most of relevant data are binary flags + some numerical values and Description
- pipeline:
  - convert binary flags and numerical values to text
  - concatenate with description
  - compute TF-IDF matrix
  - compute cosine similarities
  - choose rows with rated games
  - sort games based on similarity score

# Problems

- during development:
  - converting everything to text (so that library function could be used)
  - mapping index in matrix to index in database
- finished product:
  - reimplementations

# Evaluation

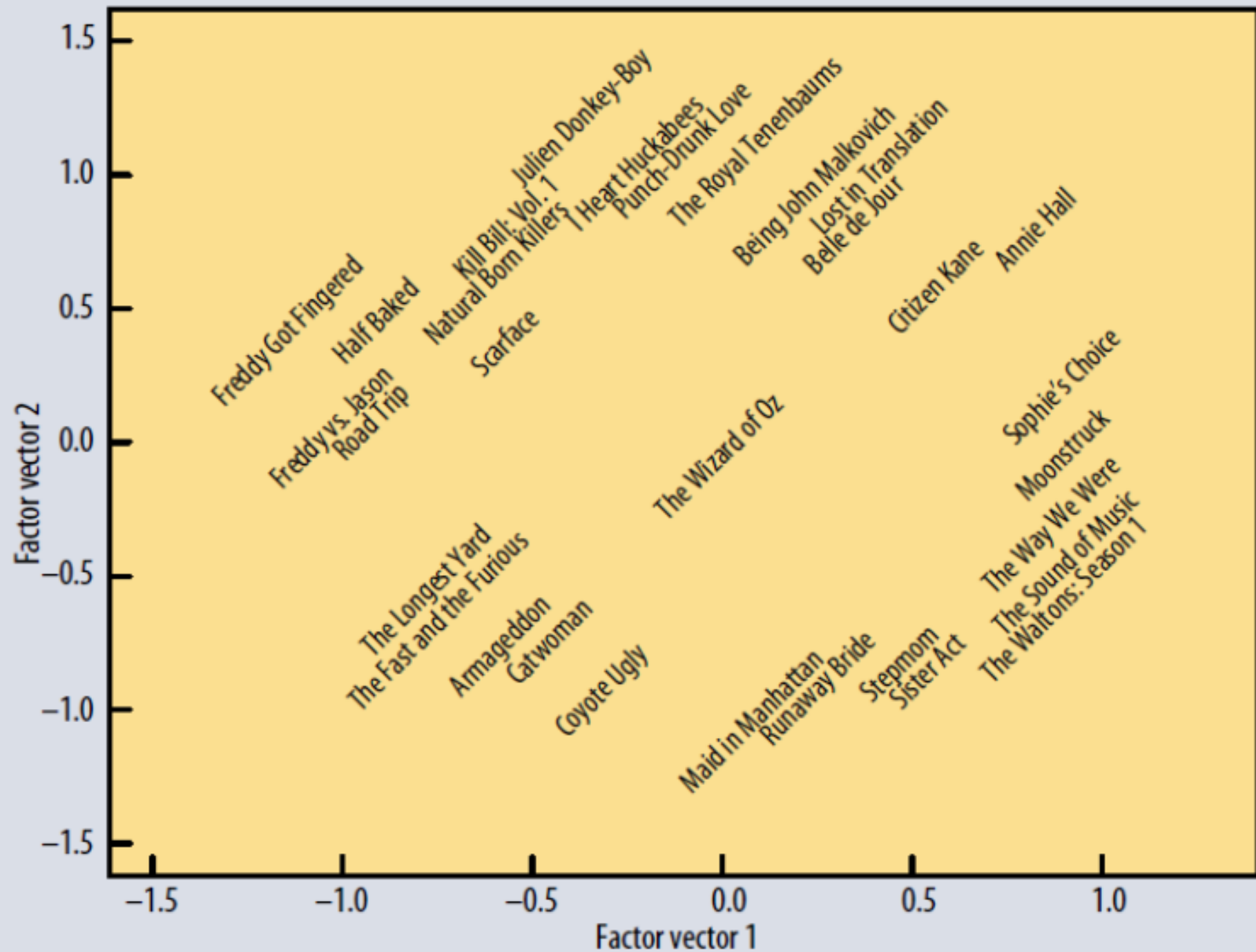
- no exact metric, just by feedback
- mostly positive feedback, but problems when game has too many reimplementations  
(those are very similar, so they get high score)
- examples of recommendations at the end

# Latent factors

- **idea:**
  - we are trying to model "taste" of users and "features" of items
  - matrices of user / item latent factors
- **approach:** minimize squared errors (+ regularization)
- **method:** stochastic / mini-batch gradient descent

$$\hat{r}_{ui} = q_i^T p_u$$

$$\min_{q,p} \sum_{(u,i) \in T} (r_{ui} - q_i^T p_u)^2 + \lambda(||q_i||^2 + ||p_u||^2)$$

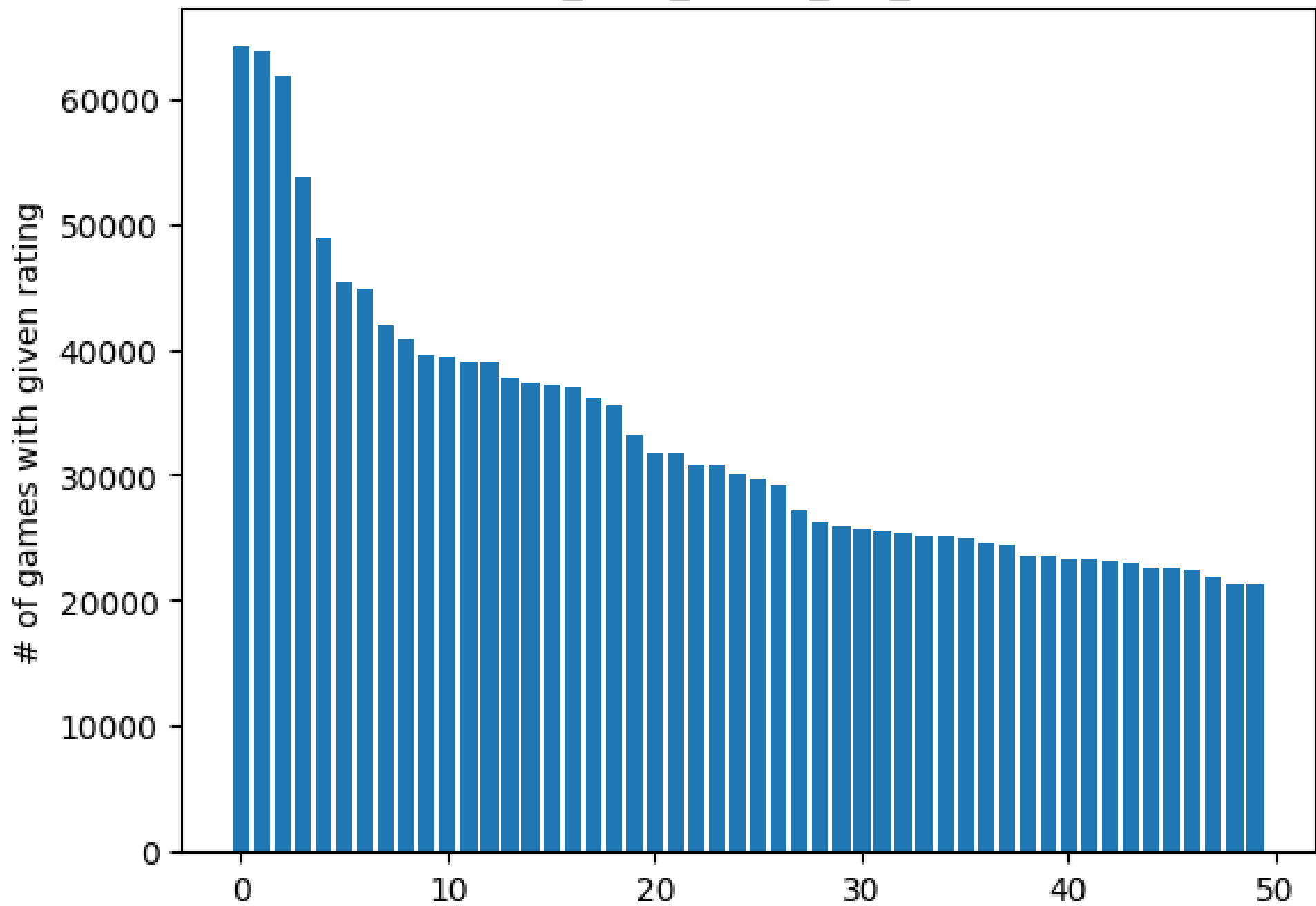


# Pipeline I

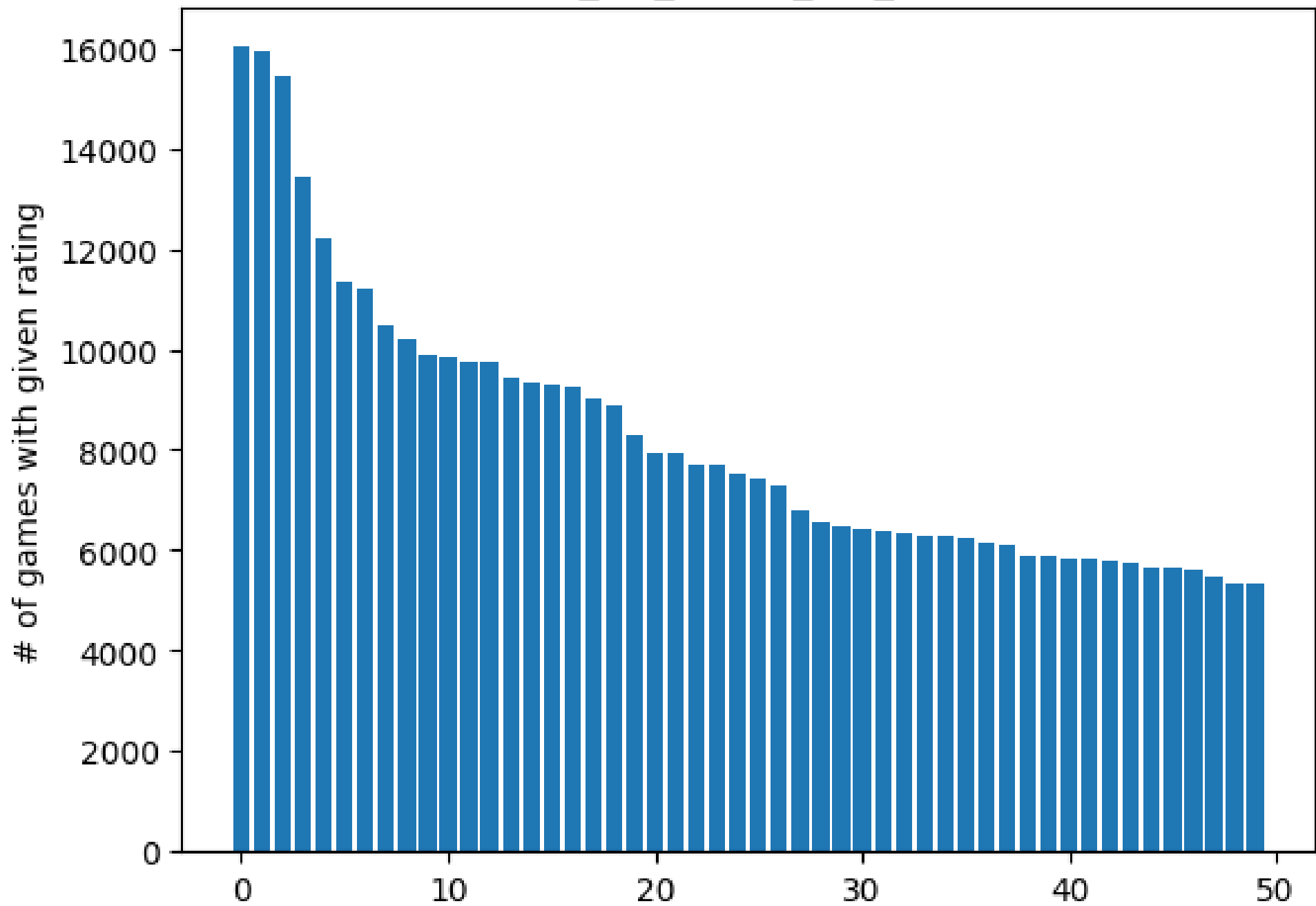
- split dataset into **train**, **validation** and **test** set
  - **idea**: all games should be present in all three datasets



X\_train\_BGGId\_top\_50



X\_val\_BGGId\_top\_50



# Pipeline II

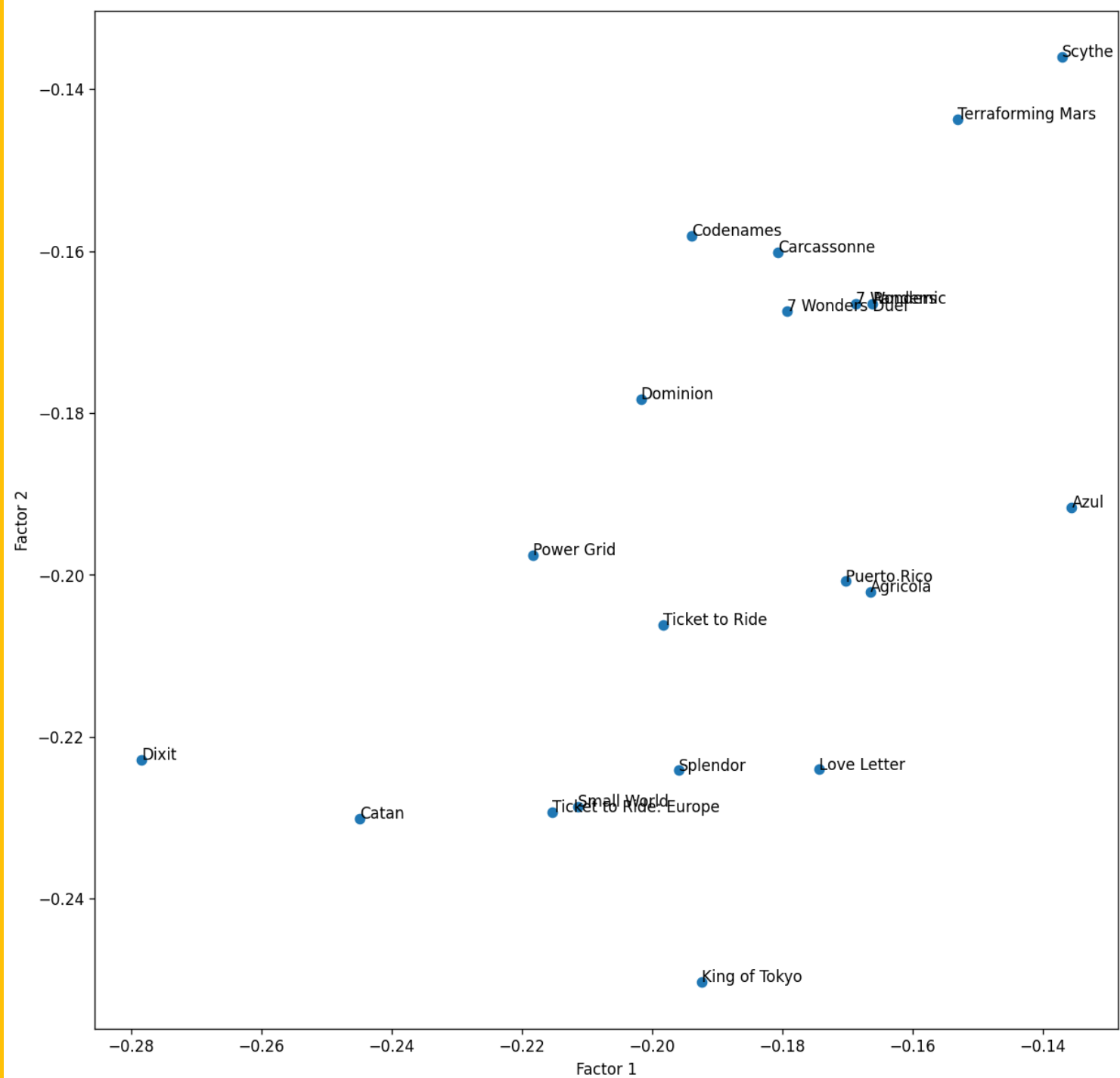
- use training set (user ratings) to **train** the model
  - adjust matrices of user / item factors using **gradient descent**
- use validation set to prevent overfitting
  - compute **RMSE**
  - apply early stopping if necessary
- use test set to **evaluate** the trained model (RMSE)

# Initial results I

- **stochastic** gradient descent too **slow**
  - necessary to reduce data
- adopt **mini-batch** gradient descent
  - allows to train on full data with more epochs / hyper-parameter tuning

# Initial results II

- predictions not quite "reasonable" (no obvious pattern)
  - 2 latent factors  $\rightarrow$  2D plane (similar to PCA)  $\rightarrow$  **find features** in similar / "opposite" clusters
  - compare RMSE of latent factors model with **baselines**



# RMSE comparison

- global mean: 1.530
- user mean: 1.376
- item mean: 1.316
- global mean + item/user bias: 1.230
- latent factors: 1.19
- latent factors with global effects: 1.19

# New user

- approximate **user factors** from ratings
- item factors matrix is constant
- least squares
- simple approach (systems of equations)
- examples of recommendations at the end



# Experience report

- necessary to implement gradient descent myself
- necessary to make mini-batch
- additional effort to confirm algorithms are implemented correctly
- computing recommendations for new user initially unclear

# Memory based CF

- **pipeline:**
  - ratings matrix (users  $\times$  games) - get users that rated the same games as me
  - unrated game - keep users who rated it
  - find k most similar users to me
  - get mean rating

# Qualitative evaluation

- only on local device (not integrated into the app)
- low amount of feedback - testing in Postman
- not very intuitive, much novelty and unexpected recommendations
- stick to category - RPGs recommend RPGs

POST

http://localhost:8000/recommendations/item-based-cf

Params

Authorization

Headers (9)

Body

Pre-request Script

Tests

Settings

☐ none☐ form-data☐ x-www-form-urlencoded☒ raw☐ binary☐ GraphQL

JSON



```
1 {  
2   "offset": 0,  
3   "limit": 3,  
4   "ratings": [  
5     {  
6       "gameId": 823, // The Lord of the Rings  
7       "value": 10  
8     },  
9     {  
10      "gameId": 7467, // The Lord of the Rings: The Two Towers Board Game  
11      "value": 9  
12    },  
13    {  
14      "gameId": 77423, // The Lord of the Rings: The Card Game  
15      "value": 9  
16    },  
17    {  
18      "gameId": 175755, // The Lord of the Rings: Journey to Mordor  
19      "value": 10  
20    }  
21  ]  
22 }
```

Pretty

Raw

Preview

Visualize

JSON



```
1 {
2   "games": [
3     {
4       "id": 5867,
5       "name": "10 Days in Europe",
6       "yearPublished": 2002,
7       "difficulty": 1,
8       "avgRating": 6,
9       "minPlayers": 2,
10      "maxPlayers": 4,
11      "playtime": 30,
12      "minAge": 10,
13      "image": "https://cf.geekdo-images.com/7jXx06W2pr66YKWrZ7Dywg__original/img/yIUeQ-GifItuTGgLWGWhyA1cus=/0x0/filters:format(png)/pic4895073.png"
14    },
15    {
16      "id": 194789,
17      "name": "10 Minute Heist: The Wizard's Tower",
18      "yearPublished": 2017,
19      "difficulty": 1,
20      "avgRating": 6,
21      "minPlayers": 2,
22      "maxPlayers": 5,
23      "playtime": 10,
24      "minAge": 14,
25      "image": "https://cf.geekdo-images.com/KFvJSZokVUTkLpZCpmEzHA__original/img/kK7wVB1_FLA12GU4HI-IZrNFWEQ=/0x0/filters:format(jpeg)/pic3584961.jpg"
26    },
27    {
28      "id": 155122,
29      "name": "1066, Tears to Many Mothers",
30      "yearPublished": 2018,
31      "difficulty": 2,
32      "avgRating": 7,
33      "minPlayers": 1,
34      "maxPlayers": 2,
35      "playtime": 60,
36      "minAge": 12,
37      "image": "https://cf.geekdo-images.com/1UM58z00q18nocyVi548vQ__original/img/COBaRqsMDsILepjsTkdbF5Avkn4=/0x0/filters:format(jpeg)/pic4200779.jpg"
38    }
39  ],
40  "totalNumberOfGames": 3593
41 }
```

POST

http://localhost:8000/recommendations/item-based-cf

Params

Authorization

Headers (9)

Body

Pre-request Script

Tests

Settings

☐ none

☐ form-data

☐ x-www-form-urlencoded

☒ raw

☐ binary

☐ GraphQL

JSON



```
1  {
2    "offset": 0,
3    "limit": 3,
4    "ratings": [
5      {
6        "gameId": 3955, // BANG
7        "value": 10
8      },
9      {
10       "gameId": 463, // Magic: The Gathering
11       "value": 10
12     },
13     {
14       "gameId": 77423, // The Lord of the Rings: The Card Game
15       "value": 9
16     },
17     {
18       "gameId": 6366, // Dungeons & Dragons: The Fantasy Adventure Board Game
19       "value": 10
20     }
21   ]
22 }
```

Pretty

Raw

Preview

Visualize

JSON



```
1 {
2   "games": [
3     {
4       "id": 7865,
5       "name": "10 Days in Africa",
6       "yearPublished": 2003,
7       "difficulty": 1,
8       "avgRating": 6,
9       "minPlayers": 2,
10      "maxPlayers": 4,
11      "playtime": 30,
12      "minAge": 10,
13      "image": "https://cf.geekdo-images.com/B1Bji6N6r1kUvI_Glp7U1w__original/img/3Y_svR8FUml-yicbCq3bwBvSYKc=/0x0/filters:format(jpeg)/pic1229634.jpg"
14    },
15    {
16      "id": 7866,
17      "name": "10 Days in the USA",
18      "yearPublished": 2003,
19      "difficulty": 1,
20      "avgRating": 6,
21      "minPlayers": 2,
22      "maxPlayers": 4,
23      "playtime": 30,
24      "minAge": 10,
25      "image": "https://cf.geekdo-images.com/RVv508Pwn5NiCFwAfCmlbw__original/img/H5J1XXYwmFZCGu9XU7gf8iGMGTy=/0x0/filters:format(jpeg)/pic4306793.jpg"
26    },
27    {
28      "id": 422,
29      "name": "1835",
30      "yearPublished": 1990,
31      "difficulty": 3,
32      "avgRating": 7,
33      "minPlayers": 3,
34      "maxPlayers": 7,
35      "playtime": 360,
36      "minAge": 16,
37      "image": "https://cf.geekdo-images.com/7UvvhAg807X1IITYIRALfag__original/img/1Vu52eKoaONv177opE7uJzwmqE8=/0x0/filters:format(jpeg)/pic202312.jpg"
38    }
39  ],
40  "totalNumberOfGames": 3202
41 }
```

# Encountered problems

- performance - not usable in real-time, not deployed
- implementation - several new technologies (also programming on Windows)
- memory - creation of ratings matrix on local device



# User interface

Search all games by name

🔍

piranha

×



### Piranha

From 2004

Rating: 5 / 10

Difficulty: 1 / 5

5+

2 - 6 players

20 minutes

- DETAILS
- BUY
- RATE



### Piranha Panic

From 2005

Rating: 4 / 10

Difficulty: 1 / 5

5+

2 - 4 players

15 minutes

- DETAILS
- BUY
- RATE



### Piranha Pedro

From 2004

Rating: 6 / 10

Difficulty: 1 / 5

8+

2 - 6 players

30 minutes

- DETAILS
- BUY
- RATE

Search all games by name

🔍 piranha

×



### Piranha

From 2004

Rating: 5 / 10

Difficulty: 1 / 5



5+



2 - 6 players



20 minutes

DETAILS

BUY

RATE

### Rate Piranha Panic

How much did you enjoy Piranha Panic?



CANCEL

SAVE

### Piranha Pedro

From 2004

Rating: 6 / 10

Difficulty: 1 / 5



8+



2 - 6 players



30 minutes

DETAILS

BUY




RATE

Search my ratings by name



RATE GAMES

RESET RATINGS

Name	Image	Your rating	Rating given at
UNO		★★☆☆☆☆☆☆	08/05/2023, 17:21:11
Catan		★★★★★★★☆☆	08/05/2023, 17:20:33
Piranha Panic		★★★★★★★☆☆	08/05/2023, 17:20:21
Rows per page: 3 ▾ 1-3 of 3 < >			

## TF-IDF



## CATAN: 3D Edition

From 2021

Rating: 7 / 10

Difficulty: 2 / 5



10+



3 - 4 players



120 minutes

[DETAILS](#)[BUY](#)[RATE](#)

## Catan: 25th Anniversary Edition

From 2020

Rating: 7 / 10

Difficulty: 2 / 5



10+



3 - 6 players



120 minutes

[DETAILS](#)[BUY](#)[RATE](#)

## Catan: 25 Jahre Jubiläums-Edition

From 2020

Rating: 7 / 10

Difficulty: 0 / 5



10+



3 - 4 players



120 minutes

[DETAILS](#)[BUY](#)[RATE](#)

## Latent factors



### Frontier: The Card Game

From 2017

Rating: 8 / 10

Difficulty: 1 / 5



12+



1 - 4 players



90 minutes

[DETAILS](#)[BUY](#)[RATE](#)

### Magi Kitchen

From 2014

Rating: 6 / 10

Difficulty: 1 / 5



8+



4 - 8 players



15 minutes

[DETAILS](#)[BUY](#)[RATE](#)

### Celestial Rainbows

From 2013

Rating: 5 / 10

Difficulty: 1 / 5



8+



2 - 8 players



15 minutes

[DETAILS](#)[BUY](#)[RATE](#)



## Top-rated



## Gloomhaven

From 2017

Rating: 8 / 10

Difficulty: 3 / 5



14+



1 - 4 players



120 minutes

[DETAILS](#)[BUY](#)[RATE](#)

## Pandemic Legacy: Season 1

From 2015

Rating: 8 / 10

Difficulty: 2 / 5



13+



2 - 4 players



60 minutes

[DETAILS](#)[BUY](#)[RATE](#)

## Brass: Birmingham

From 2018

Rating: 8 / 10

Difficulty: 3 / 5



14+



2 - 4 players



120 minutes

[DETAILS](#)[BUY](#)[RATE](#)

## Most rated



## Pandemic

From 2008

Rating: 7 / 10

Difficulty: 2 / 5



8+



2 - 4 players



45 minutes

[DETAILS](#)[BUY](#)[RATE](#)

## Carcassonne

From 2000

Rating: 7 / 10

Difficulty: 1 / 5



7+



2 - 5 players



45 minutes

[DETAILS](#)[BUY](#)[RATE](#)

## Catan

From 1995

Rating: 7 / 10

Difficulty: 2 / 5



10+



3 - 4 players



120 minutes

[DETAILS](#)[BUY](#)[RATE](#)



[DETAILS](#)[BUY](#)[RATE](#)[DETAILS](#)[BUY](#)[RATE](#)[DETAILS](#)[BUY](#)[RATE](#)

## Random picks



### Monikers: Something Something

From 2016

Rating: 8 / 10

Difficulty: 1 / 5



18+



4 - 20 players



60 minutes

[DETAILS](#)[BUY](#)[RATE](#)

### Das Jagdspiel

From 1954

Rating: 5 / 10

Difficulty: 1 / 5



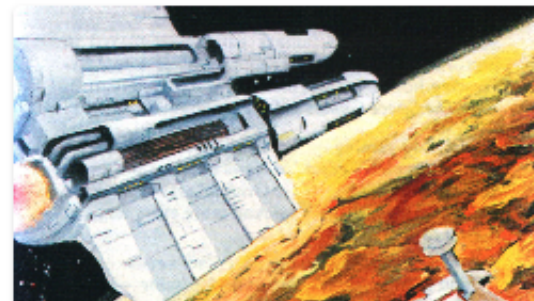
8+



2 - 8 players



45 minutes

[DETAILS](#)[BUY](#)[RATE](#)

### Hot Spot

From 1979

Rating: 5 / 10

Difficulty: 1 / 5



10+



2 - 2 players



60 minutes

[DETAILS](#)[BUY](#)[RATE](#)

# Thank you



Try it out: <https://pv254-board-games-recommender.vercel.app/>

Code: <https://github.com/petr7555/pv254-board-games-recommender>