



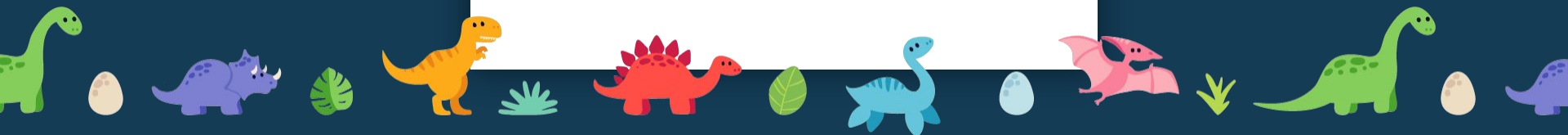
Steam Game Platform Recommendation System

Presented By EurekaGo!



Agenda

- About Steam
- Business Objective
- Data Description
- Modeling
- Recommendation

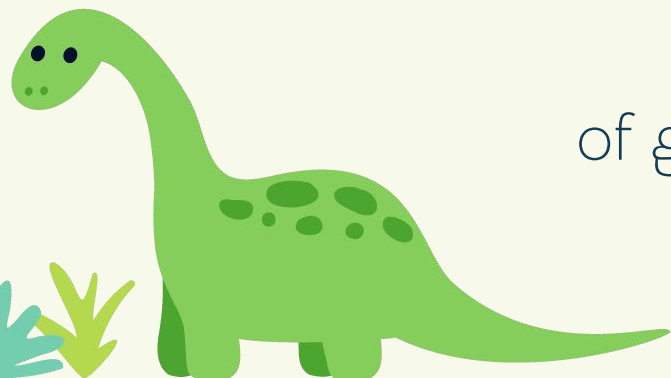




34,000
games on the site

95,000,000
monthly active users

18%
of global PC games sales



Features on Steam



Trading and Gifting



Read, Rate & Discuss





Business Initiative

Overall Objective is to **reacquire gamers**, augmenting the **gamer engagement** and **immersion**

- Identify inactive gamers and further stratify among them

Create game recommendation system that target each stratification

- Identify focal topics that gamers have been talking about

Refine our recommendation system with the focal topics





Evaluation Metrics

- **Click Through Rate:**
$$\frac{\text{Number of Clicks}}{\text{Number of Emails Sent}}$$

Emails (Inactive gamers) and Pop-up windows (Active gamers)
- **Conversion Rate:**
$$\frac{\text{Number of Purchases}}{\text{Number of Gamers Clicking on Promotion Link}}$$

Buy or not, Play or not
- **Game Play Time:** count(times a gamer play games per week)

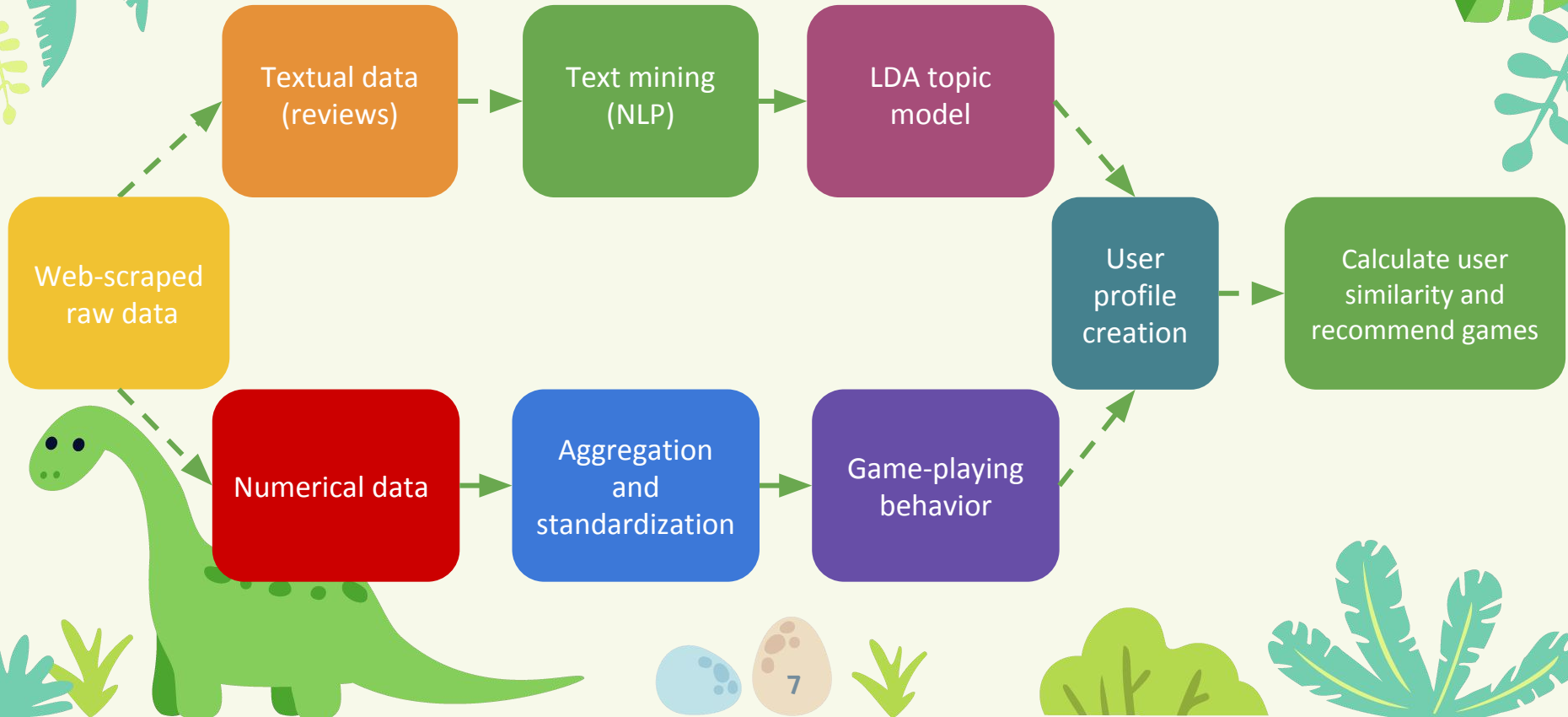
Track if the gamers churn easily after they re-activate
- **Social Engagement:** count(uses of social interactive features per week)

Track if gamers engage by leaving reviews, clicking 'recommend', 'votes up' and 'votes funny' buttons





Analytics Process



Data description

REVIEW-LEVEL TABLE

GameID
SteamID
NumGamesOwned
NumReviewsGiven
PlayTime
.....
WriteEarly

18 columns

GAME-LEVEL TABLE

AppID
AppName
Genre
Developer
Price
.....
IsFree

15 columns

91,925 Games - 6,742 Games

1,817,436 Reviews

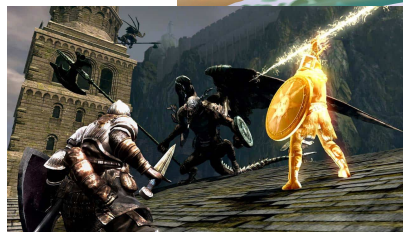
1,091,552 Distinct User

MOST POPULAR GENRES #TOP3



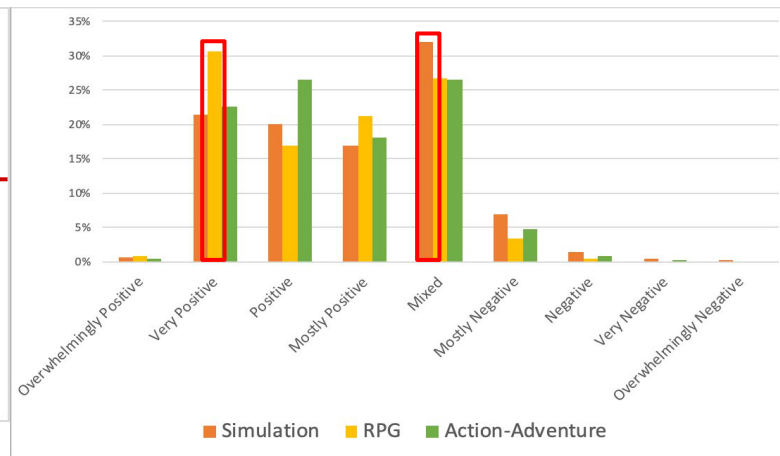
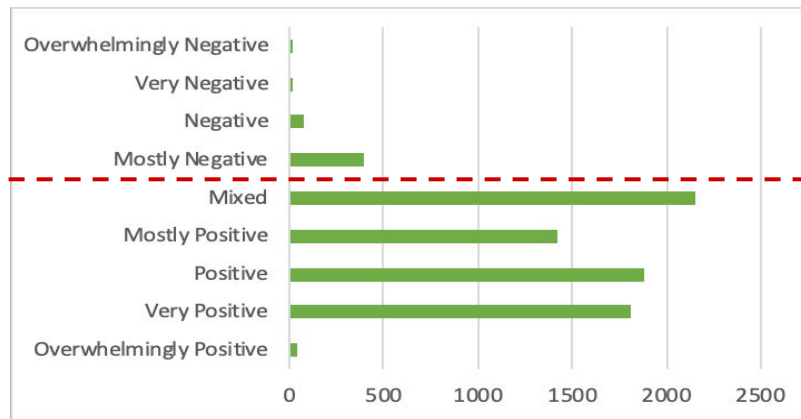
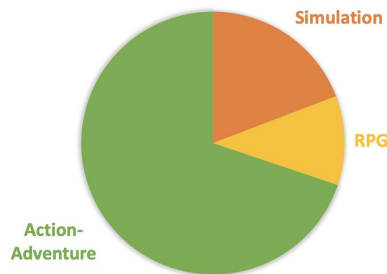
ACTION & ADVENTURE

SIMULATION

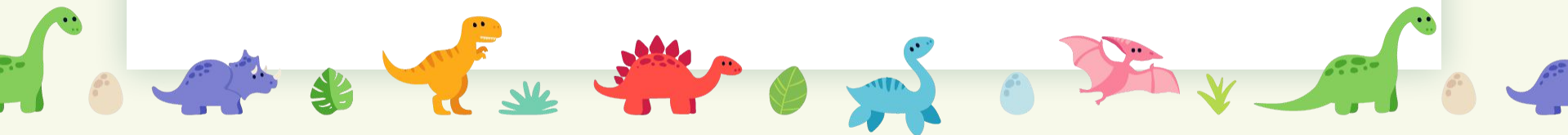
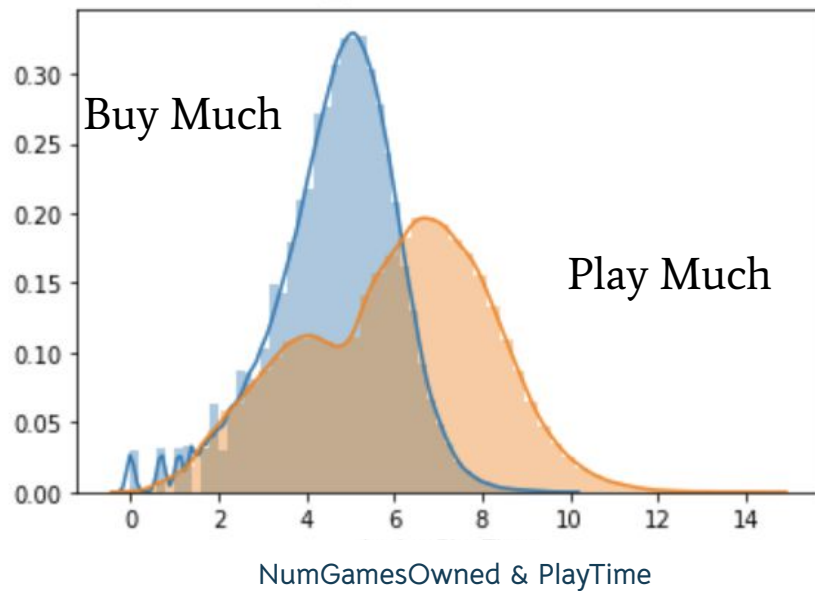


RPG
(ROLE-PLAY GAME)

EDA

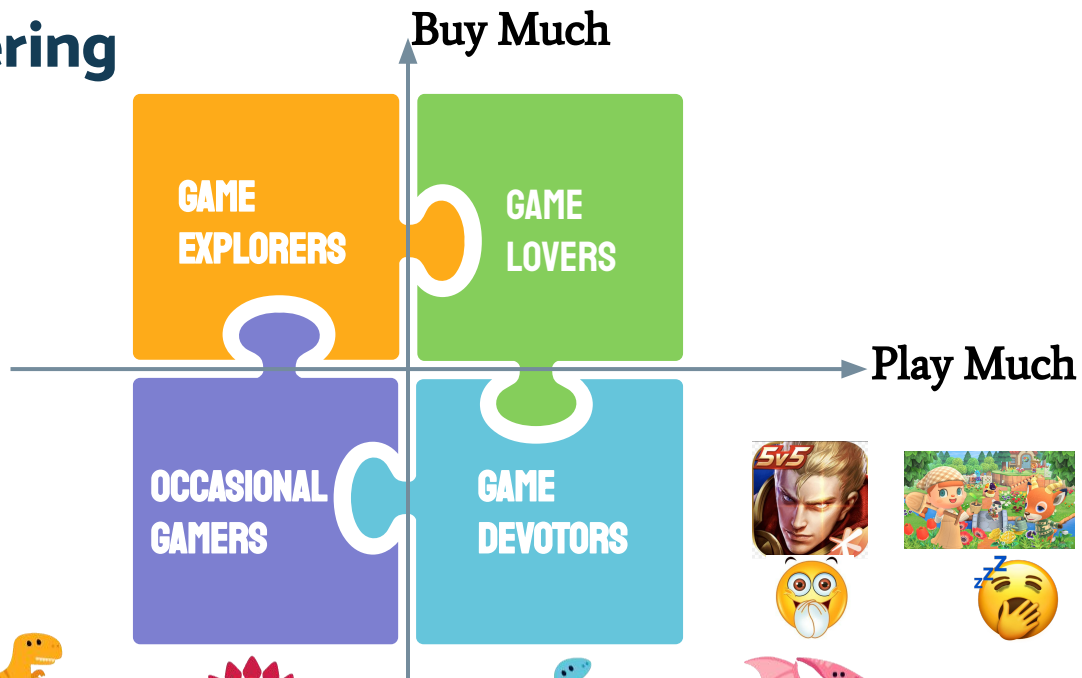


EDA





User Clustering





User Segmentation -- Who's Churning?



Active User

Actively playing games in 2019, 2020

Playtime per year > mean

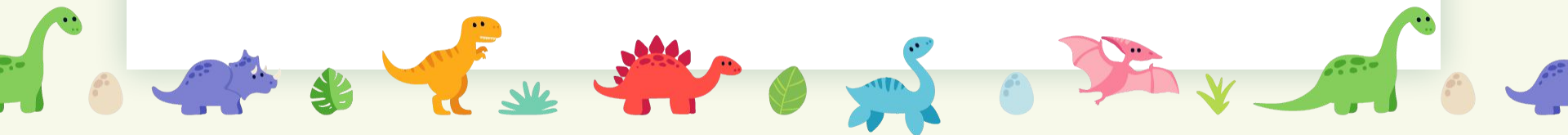


Inactive User

We want to reacquire!!

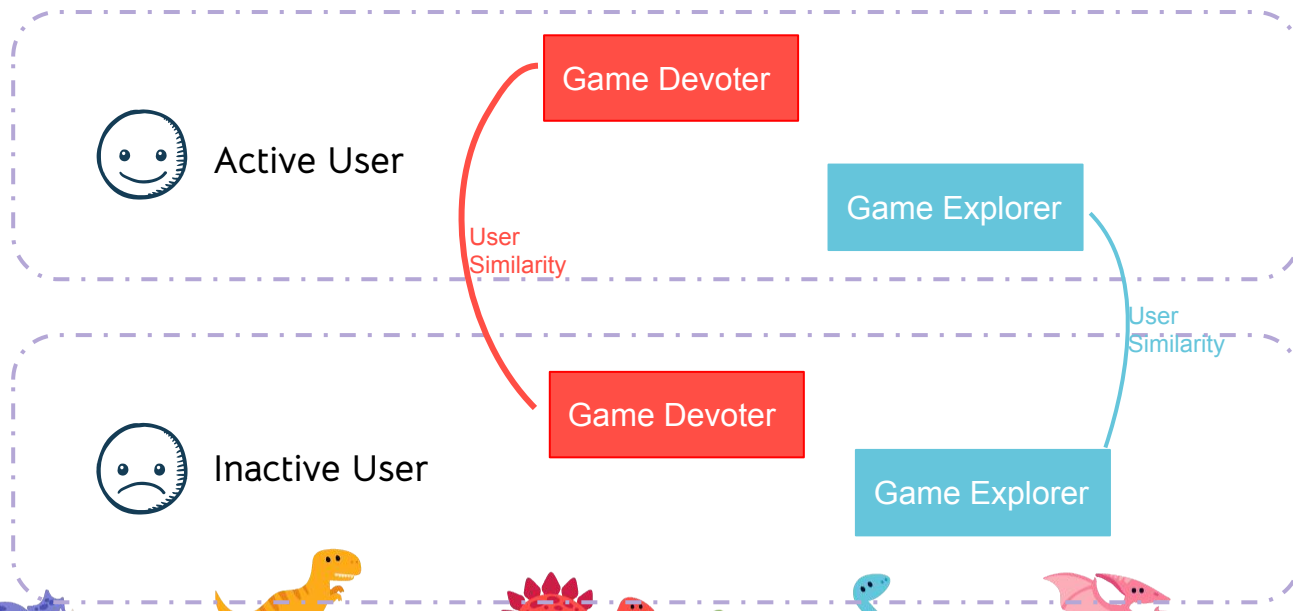
Actively played game in 2019 but STOPPED playing in 2020

Playtime per year > mean



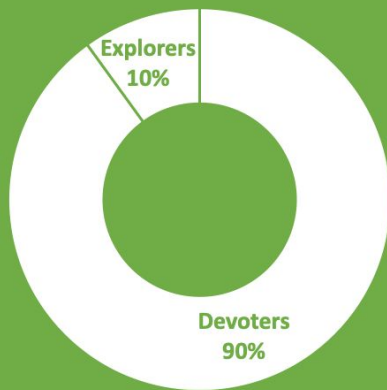


Matching Active and Inactive Gamers

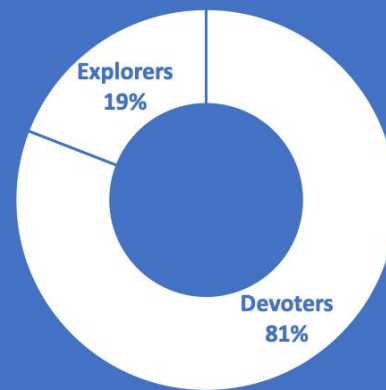


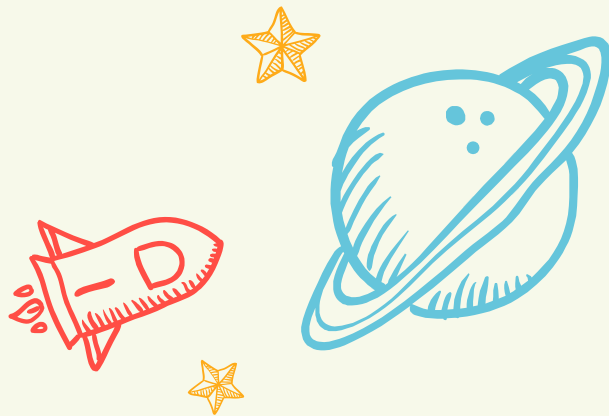
Distribution of 85,508 Users

OF 71,612 INACTIVE USERS

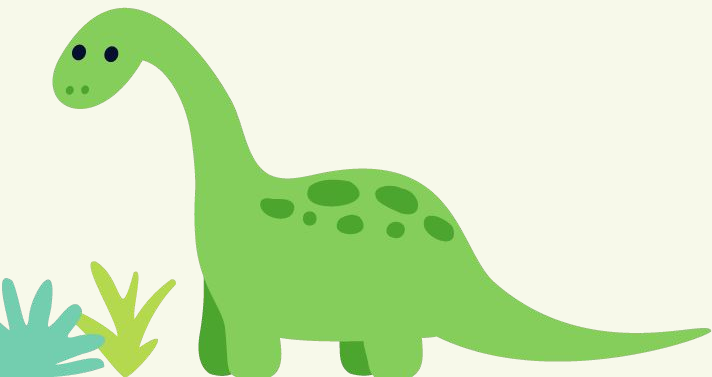


OF 13,800 ACTIVE USERS

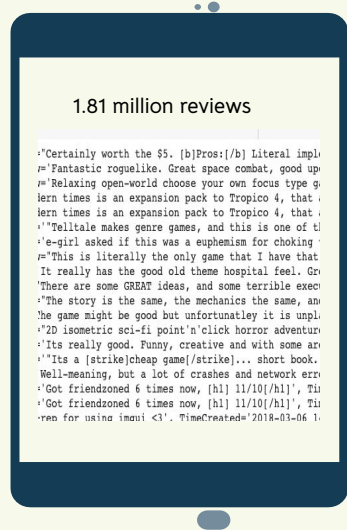




Model development



LDA Topics Model



"It was able to make me emotional and really get into the story"

GAME STORY AND U/UX

Topic 1

Voice act ,Great stori
Art style , Good story
Story line , stori great

SENTIMENT EXPRESSION

Topic 3

Super hot, super fun,
Simple fun, refund pls
Bring back, fun love

"Well, I need a refund guys"

'Super fun quirky game that I ran through pretty quickly.'

GAME STRATEGY

Topic 2

Click click, boss fight,
Level design, final
boss, Dungeon crawler

TACTICS

Topic 5

Early access, ca
not_wait, single
player, saint now,
fun friends

PERFORMANCE & STABILITY

Topic 4

Bug bug, every time,
lok forward, go back,
long time, full price

"Saw a few bugs getting the game into full screen"

GAME CHARACTER & CONTENT

Topic 6

Star war, main
character, dark soul,
replay value, visual
novel, cick adventure

"The main character has no special abilities to defeat enemies"



User Profile

Play time percentage by genre

Review percentage by topic

SteamID	NumGamesOwned	NumReviewsGiven	Action-Adventure	RPG	Simulation	RecommendedPercent	SteamAgeProxy	topic_0	topic_1	topic_2	topic_3	topic_4	topic_5
76561197960295681	309	4	1.0	0.0	0.0	1.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0
76561197960312016	58	2	0.0	0.0	1.0	1.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0
76561197960331413	391	4	0.0	1.0	0.0	1.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0
76561197960331786	1129	28	0.788	0.212	0.0	1.0	17.0	1.0	0.0	0.0	0.0	0.0	0.0
76561197960359040	315	7	1.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0
76561197960370005	154	1	1.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0
76561197960389353	1026	51	0.467	0.515	0.018	0.786	2619.0	0.143	0.0	0.071	0.429	0.071	0.286
76561197960409453	552	5	1.0	0.0	0.0	1.0	261.0	1.0	0.0	0.0	0.0	0.0	0.0
76561197960433625	1017	5	1.0	0.0	0.0	1.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0
76561197960438384	352	5	1.0	0.0	0.0	1.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0
76561197960446373	543	2	0.0	1.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0
76561197960459273	793	7	0.252	0.748	0.0	0.0	10.0	0.0	0.0	0.0	0.5	0.0	0.5
76561197960521661	980	10	1.0	0.0	0.0	1.0	118.0	0.0	0.5	0.0	0.5	0.0	0.0
76561197960543469	263	1	1.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0
76561197960566229	229	5	0.374	0.626	0.0	1.0	57.0	0.5	0.0	0.0	0.0	0.0	0.5
76561197960646352	178	1	1.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0
76561197960690083	103	1	1.0	0.0	0.0	1.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0
76561197960723156	223	8	0.0	1.0	0.0	1.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0
76561197960759629	205	2	0.0	1.0	0.0	1.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0
76561197960790283	214	2	1.0	0.0	0.0	1.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0

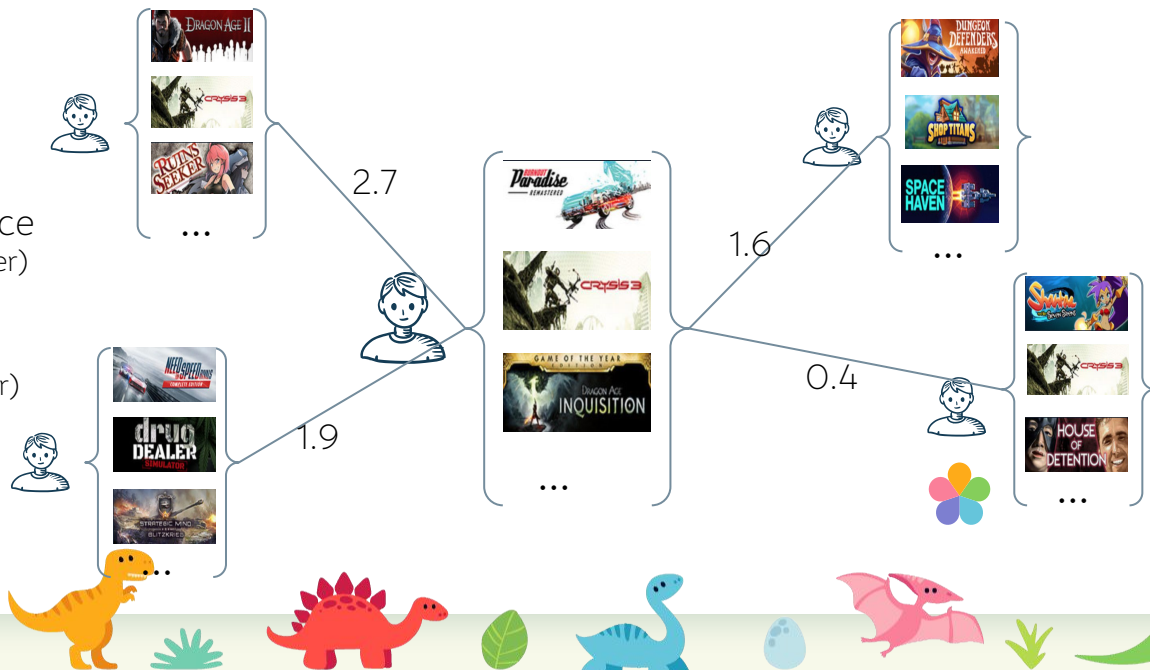
only showing top 20 rows



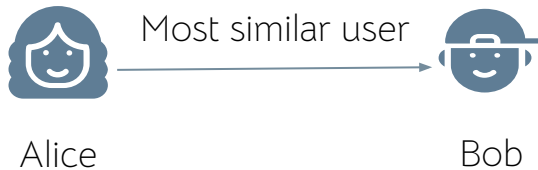
Recommendation System

Euclidean distance
(the smaller the closer)

Cosine similarity
(the higher the closer)



Show Case



User segment

Inactive Action-Adventure
Devotor

Active Action-Adventure
Devotor

Game list held



Scaling up plan

May expect...

1. Fast changing user activeness
2. Change of user preference
3. More incoming users
4. More new games released

Plan...

1. Update user profile frequently
2. ETL pipeline (model automation)
3. Talent acquisition



Thanks!

Any questions?

