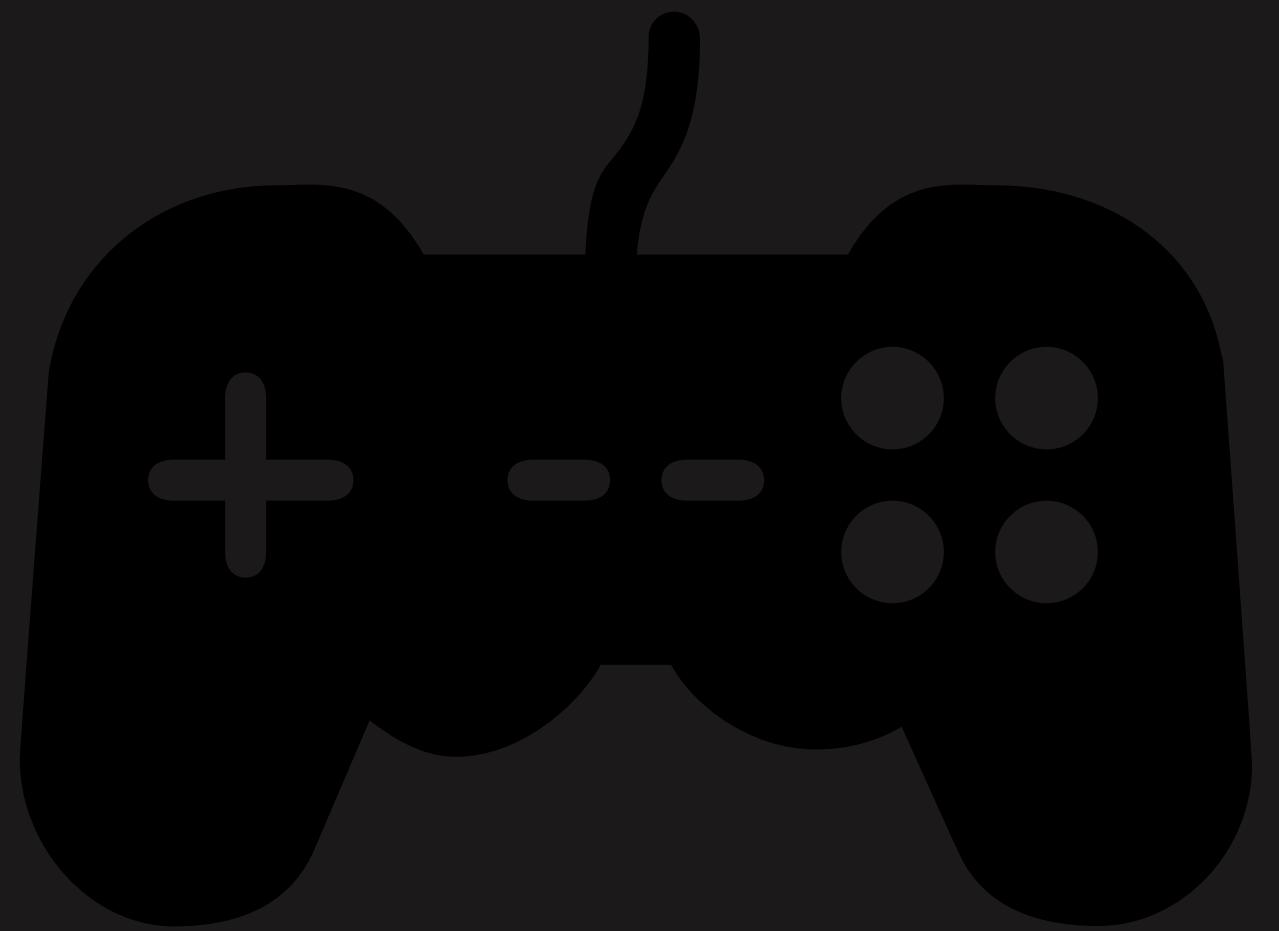


# STEAM GAME RECOMMENDER

Matthew Buggey  
Sprint 1

# Project Overview

- This project's goal is to create a game recommender system for the Steam Game Store.
- Specifically this project aims to better predict future successful 'Indie' (independent developer) titles.



# Project Overview

- Indie games are powerful potential sellers that are difficult to predict popularity for as Indie studios are created and disbanded constantly.
- Additionally many Indie studios lack funding compared to longtime game development companies and thus do not have the funds or notoriety to properly market their products.
- Despite this the gaming space has seen a surge in Indie game popularity and development, it is not uncommon anymore to have a game you've never heard of seemingly randomly jump to the top of the sales charts



# Brief Overview of the Steam Store

The screenshot shows the Steam Store homepage with a festive winter theme. A large banner at the top left announces the "STEAM WINTER SALE" from December 21 to January 4 at 10 AM Pacific. Below this, a section titled "Winners of the 2023 Steam Awards!" displays various award categories and their winners, each with a small thumbnail and a discount offer. To the right, there's a "PLAYERS LIKE YOU LOVE" section based on game history, featuring "DEATHLOOP" and "SHADOW OF WAR" with significant discounts. The main navigation bar at the top includes links for "Your Store", "New & Noteworthy", "Categories", "Points Shop", "News", "Labs", and a search bar.

**Winners of the 2023 Steam Awards!**

The community has spoken and the votes have been counted. These are your winners!

[LEARN MORE](#)

Award Category	Winner Game	Discount	Original Price	Sale Price
Game of the Year	Baldur's Gate 3	-10%	CDN\$ 79.99	CDN\$ 71.99
VR Game of the Year	Labyrintine	-60%	CDN\$ 12.99	CDN\$ 5.19
Labor of Love	Red Dead Redemption II	-67%	CDN\$ 79.99	CDN\$ 26.39
Best Game on Steam Deck	Hogwarts Legacy	-50%	CDN\$ 79.99	CDN\$ 39.99
Better With Friends	LETHAL COMPANY	-	CDN\$ 12.99	CDN\$ 12.99
Outstanding Visual Style	ATOMIC HEART	-50%	CDN\$ 79.99	CDN\$ 39.99
Most Innovative Gameplay	STARFIELD	-30%	CDN\$ 89.99	CDN\$ 62.99
Best Game You Suck At	SIFU	-50%	CDN\$ 43.99	CDN\$ 21.99
Best Soundtrack	THE LAST OF US PART II	-33%	CDN\$ 79.99	CDN\$ 53.59
Outstanding Story-Rich Game	Baldur's Gate 3	-10%	CDN\$ 79.99	CDN\$ 71.99
Sit Back and Relax	DAVE THE TREASURE	-20%	CDN\$ 25.99	CDN\$ 20.79

**PLAYERS LIKE YOU LOVE**

BASED ON THE GAMES YOU'VE PLAYED

-50% CDN\$ 64.99 CDN\$ 32.49	-70% CDN\$ 64.99 CDN\$ 16.49	-85% CDN\$ 59.99 CDN\$ 8.99
-80% CDN\$ 79.99 CDN\$ 15.99	-75% CDN\$ 52.99 CDN\$ 13.24	-80% CDN\$ 54.99 CDN\$ 10.99
-67% CDN\$ 39.99 CDN\$ 13.19	-40% CDN\$ 64.99 CDN\$ 38.99	-75% CDN\$ 49.99 CDN\$ 12.49
-80% CDN\$ 79.99 CDN\$ 15.99	-50% CDN\$ 77.99 CDN\$ 38.99	-67% CDN\$ 39.99 CDN\$ 13.19

# Brief Overview of the Steam Store

The screenshot displays the Steam Store's Home page with a dark theme. On the left, a sidebar lists various games under categories like 'GAMES AND SOFTWARE' and 'UNCATEGORIZED'. The main area features a 'WHAT'S NEW' section with cards for recent news and events, including 'PUBG: BATTLEGROUNDS Weekly Bans Notice', 'The Return of the Azure Dragon - Beryl M762', 'Dev Blog #4 Dying Light 2', 'GTA Online: The Chop Shop Now Available', 'Yuletide 2023', 'Dead by Daylight's Bone Chill 2023 Event', and 'Bonus Special Drops for December PUBG: BATTLEGROUNDS'. Below this is a 'Recent Games' section showing thumbnails for titles like 'Baldur's Gate 3', 'JEDI SURVIVOR', 'LETHAL COMPANY', 'The Outer Worlds', 'DEAD BY DAYLIGHT', 'OUTER WILDS', 'PIZZA TOWER', 'ESCAPE SIMULATOR', 'WE WERE HERE EXPEDITIONS: THE FRIENDSHIP', 'Wallpaper ENGINE', and 'BREATH & FRESH'. At the bottom, a 'Play Next' section shows recommended games like 'HALO: THE MASTER CHIEF COLLECTION', 'BIOSHOCK REMASTERED', 'ALIEN ISOLATION', 'MAD MAX', 'RUST', 'OUTLAST', 'HELLTAKER', 'WAR THUNDER AIR SUPERIORITY', 'STYX: MASTER OF SHADOWS', and 'THE T.A.L.O. PRINCIPLE'.

# Data Science Solution

## Collaborative Filtering

Collaborative filtering in the recommendation department focuses on the relationship between users and items

Steam is unique in that it acts both as a store for games and a community hub for players

An incredible amount of unique user-generated content is available on demand for almost every title

CUSTOMER REVIEWS

Overall Reviews:  
**Overwhelmingly Positive** (195,161 reviews)

REVIEW TYPE ▾ PURCHASE TYPE ▾ LANGUAGE ▾ DATE RANGE ▾ PLAYTIME ▾ DISPLAY AS: Summary

Filters Your Languages

Showing 185,679 reviews that match the filters above ( **Overwhelmingly Positive** )

MOST HELPFUL REVIEWS IN THE PAST 30 DAYS

Kiru Senpai Recommended 214.5 hrs on record (157.7 hrs at review time)  
3 reviews   
POSTED: 18 DECEMBER, 2023  
Hour 1: Help I'm lost.  
Hour 10: HELP IM LOST.  
Hour 157: HELP ME IM ♫♫♫♫♫ LOST.  
10/10 Still lost while writing this review.

Was this review helpful?  
 Yes No Funny Award

1,539 people found this review helpful  
654 people found this review funny

skli Recommended 29.3 hrs on record (9.0 hrs at review time)  
17 products in account 1 review   
POSTED: 9 DECEMBER, 2023  
who cares if a furry made it, it's better than the new cod

Was this review helpful?  
 Yes No Funny Award

8,530 people found this review helpful  
1,695 people found this review funny

DeadbutDreaming Recommended 33.9 hrs on record (14.7 hrs at review time)  
248 products in account 17 reviews   
POSTED: 21 DECEMBER, 2023  
For every like this review gets, I will eat one spoonful of marinara sauce.

Was this review helpful?  
 Yes No Funny Award

RECENTLY POSTED

Masterwuu3 1.8 hrs   
POSTED: 2 JANUARY  
Product received for free  
A gift from the gods

JuiyYT 6.3 hrs   
POSTED: 2 JANUARY  
W game

bumer 12.2 hrs   
POSTED: 2 JANUARY  
nerf the outside monster pls pls slsp slsp sfojdkghdfsgjkhajfeaofnhrdjrfk

Slipknot 4.5 hrs   
POSTED: 2 JANUARY  
nbjdk

nightminer64 9.3 hrs

# Estimated Impact

## High End Outliers

Among Us is the most recent indie release to gain notoriety

Revenue information is private, however it is estimated from player counts that sales increased from \$3 million in 2019 to around \$50 million in 2020.



Source:  
<https://helplama.com/among->

# Estimated Impact

## Other Examples



New potential fad game

Older extremely popular indie title

# Estimated Impact

## Other Examples



Very old yet popular game that receives infrequent updates but has strong replayability

Story based game only good for one playthrough

# Dataset

Sourced from Kaggle

Contains three csv files and one JSON file which contains metadata related to each game entry



CSVs are as follows:

Games	Users	Recommendations	Games_metadata
Descriptive data about each unique game. Including release date, number of reviews, rating and software compatibility.	Descriptive data about unique users. How many games said user owns, and how many reviews said user has posted.	Individual reviews by users on games, includes user playtime, whether they recommend the game or not and whether their review received any awards from other users.	A JSON file which can be appended to games.csv. Contains individual game descriptions along with the tags used to label the game ex: Indie, Sci-Fi, Fantasy, etc

# Dataset

Two main finds:

There is a disparity between user\_reviews  
a cell that came preloaded on games.csv  
and num\_reviews, my mapped count of  
reviews from recommendations.csv

	user_reviews	num_reviews
0	2199	1686
1	21	0
2	3722	6943
3	873	877
4	8784	4850
...	...	...
50867	101	0
50868	29458	0
50869	1128	0
50870	82	0
50871	144	0



Rating system works on both a positive ratio and review quantity method.

Positive %	10-49 reviews	50-499 reviews	500+ reviews
95%-100%			Overwhelmingly Positive
90%-94%	Positive	Very Positive	
85%-89%	80%-100%	80%-100%	Very Positive
80%-84%			80%-94%
75%-79%		Mostly Positive	
70%-74%		70%-79%	
65%-69%			
60%-64%			
55%-59%			
50%-54%		Mixed	
45%-49%		40%-69%	
40%-44%			
35%-39%			
30%-34%			
25%-29%		Mostly Negative	
20%-24%		20%-39%	
15%-19%			
10%-14%			
5%-9%	Negative	Very Negative	Overwhelmingly Negative
0%-4%	0%-19%	0%-19%	0%-19%

Seemingly how Steam ratings work - chart by @runevision Sep 2020

# Next Steps

- Find an ideal way to hot encode the rating system
- Determine whether user\_reviews or num\_reviews is the ideal metric
- Begin a rough grouping model for games based on tag features
- Research inputting a user profile into my model to test recommendations

