



home

about us

contact

GAMESEEKER-

A PERSONALIZED

GAME RECOMMENDATION ENGINE



Keer Feng



PROBLEM STATEMENT, PROPOSED SOLUTION, AND IMPACT

Problem:

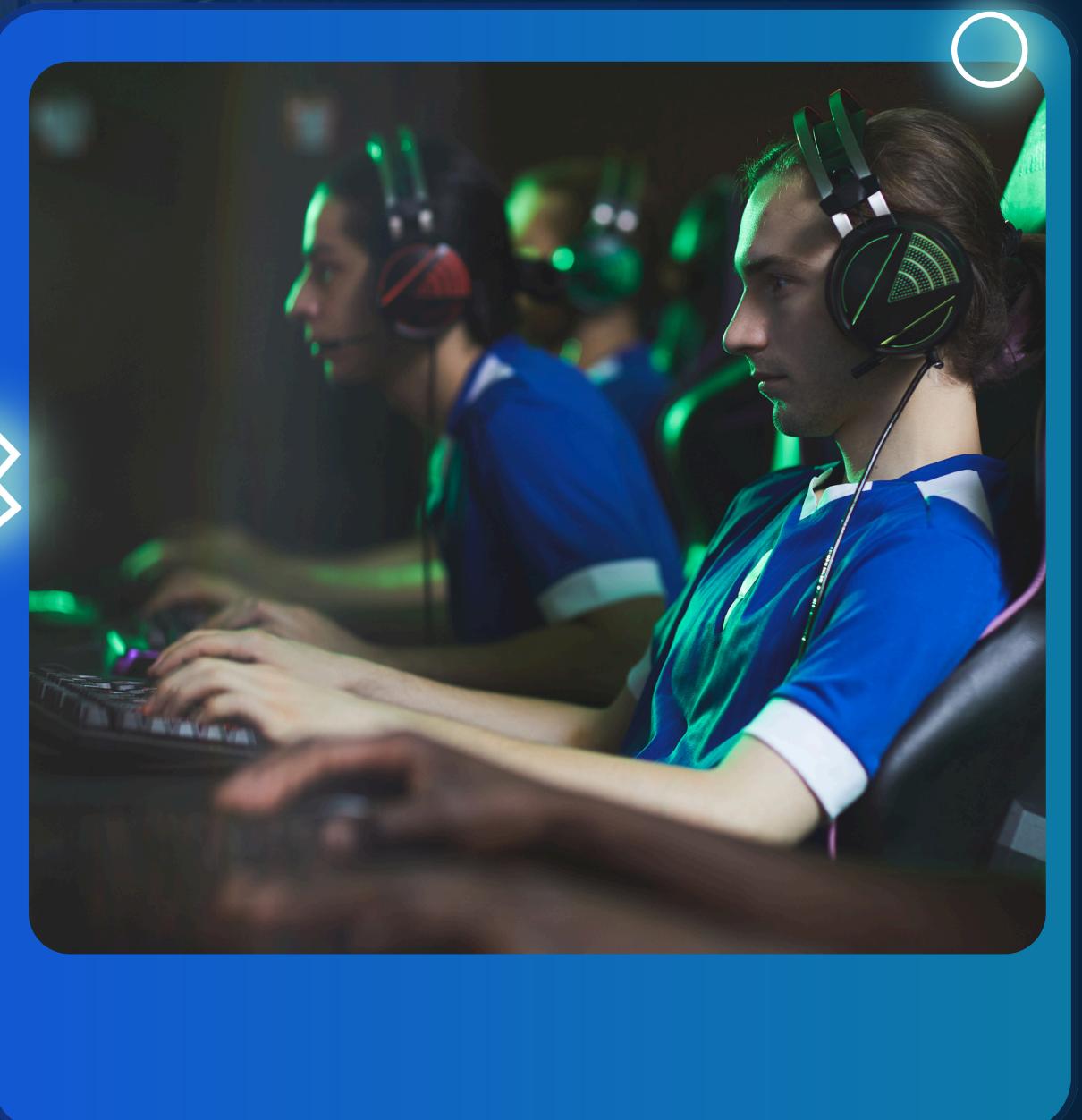
- Players face difficulty discovering games suited to their preferences due to overwhelming options on digital platforms.

Proposed Solution:

- Build a personalized game recommendation engine using player behavior and review sentiment analysis.

Impact Estimate:

- Enhance player satisfaction, increase engagement with lesser-known games, and help indie developers reach target audiences.





DATASET SOURCES AND PREPROCESSING

Steam-200k Dataset: 200,000 Steam games, including playtime, ratings, and user behavior data.

	user id	Name	purchase	hours	0
0	151603712	The Elder Scrolls V Skyrim	purchase	1.0	0
1	151603712	The Elder Scrolls V Skyrim	play	273.0	0
2	151603712	Fallout 4	purchase	1.0	0
3	151603712	Fallout 4	play	87.0	0
4	151603712	Spore	purchase	1.0	0

Video Games Sales Dataset: Provides global sales figures and regional performance.

	Name	Platform	Year_of_Release	Genre	Publisher	NA_Sales	EU_Sales	JP_Sales	Other_Sales	Global_Sales	Critic_Score	Critic_Count	User_Score	User_Count	Developer	Rating
0	Wii Sports	Wii	2006.0	Sports	Nintendo	41.36	28.96	3.77	8.45	82.53	76.0	51.0	8.0	322.0	Nintendo	E
1	Super Mario Bros.	NES	1985.0	Platform	Nintendo	29.08	3.58	6.81	0.77	40.24	NaN	NaN	NaN	NaN	NaN	NaN
2	Mario Kart Wii	Wii	2008.0	Racing	Nintendo	15.68	12.76	3.79	3.29	35.52	82.0	73.0	8.3	709.0	Nintendo	E
3	Wii Sports Resort	Wii	2009.0	Sports	Nintendo	15.61	10.93	3.28	2.95	32.77	80.0	73.0	8.0	192.0	Nintendo	E
4	Pokemon Red/Pokemon Blue	GB	1996.0	Role-Playing	Nintendo	11.27	8.89	10.22	1.00	31.37	NaN	NaN	NaN	NaN	NaN	NaN

SteamSpy Data: Aggregated statistics across genres and game popularity

#	Game	Release date	Price	Score rank/Userscore / Metascore)	Owners	Playtime (Median)	Developer(s)	Publisher(s)	Genre
0	Dota 2	Jul 9, 2013	Free	N/A (N/A/90%) 200,000,000..500,000,000	22:47 (11:16)		Valve	Valve	Action
1	Counter-Strike: Global Offensive	Aug 21, 2012	Free	N/A (N/A/83%) 100,000,000..200,000,000	12:25 (05:28)		Valve	Valve	Action
2	Grand Theft Auto V	Apr 14, 2015	29.98	N/A (N/A/96%) 50,000,000..100,000,000	11:26 (03:14)	Rockstar North	Rockstar Games	Rockstar Games	Action
3	Apex Legends	Nov 4, 2020	Free	N/A (N/A/88%) 50,000,000..100,000,000	09:57 (03:00)	Respawn	Electronic Arts	Electronic Arts	Action
4	Unturned	Jul 7, 2017	Free	N/A (N/A) 50,000,000..100,000,000	43:28 (45:04)	Smartly Dressed Games	Smartly Dressed Games	Smartly Dressed Games	Action





DATASET SOURCES AND PREPROCESSING

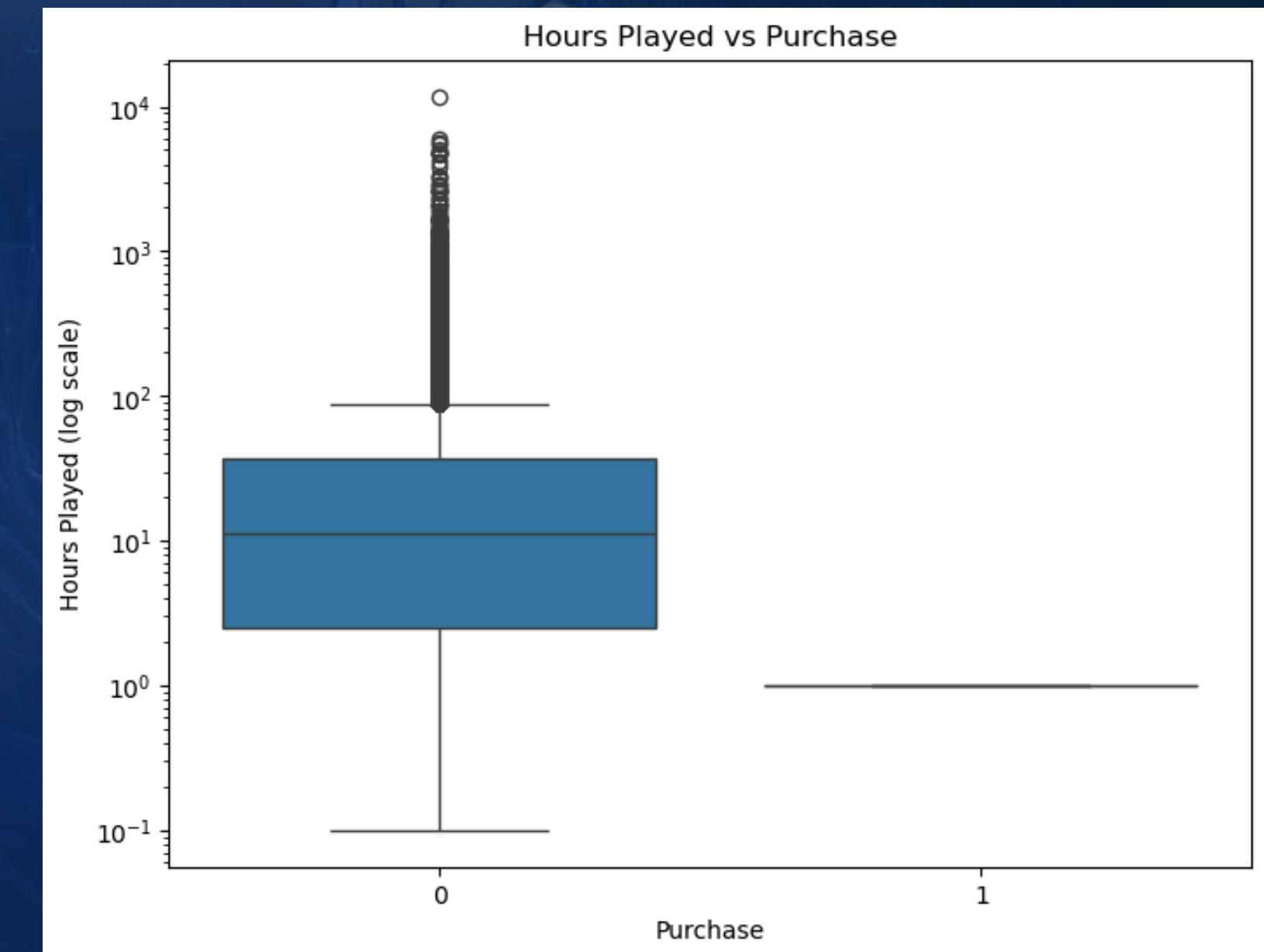
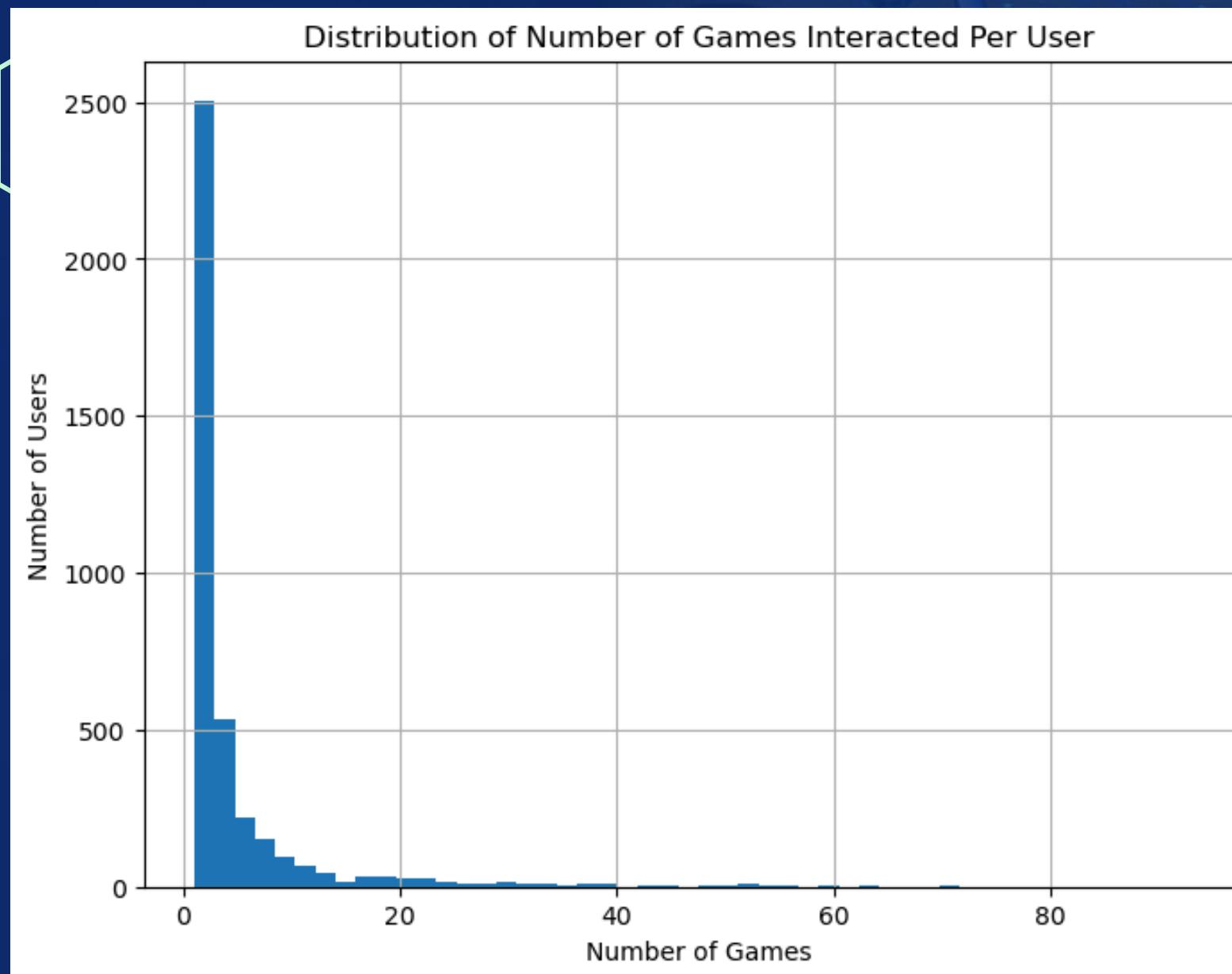
Preprocessing Steps:

- Cleaned missing values
- Standardized columns across datasets for merging.
- Converted 'Price' from text to numeric, handled 'free' games as 0.
- One-hot encoded Genre_x into multiple genre features.
- Aggregated user interaction data to game level for consistent analysis.



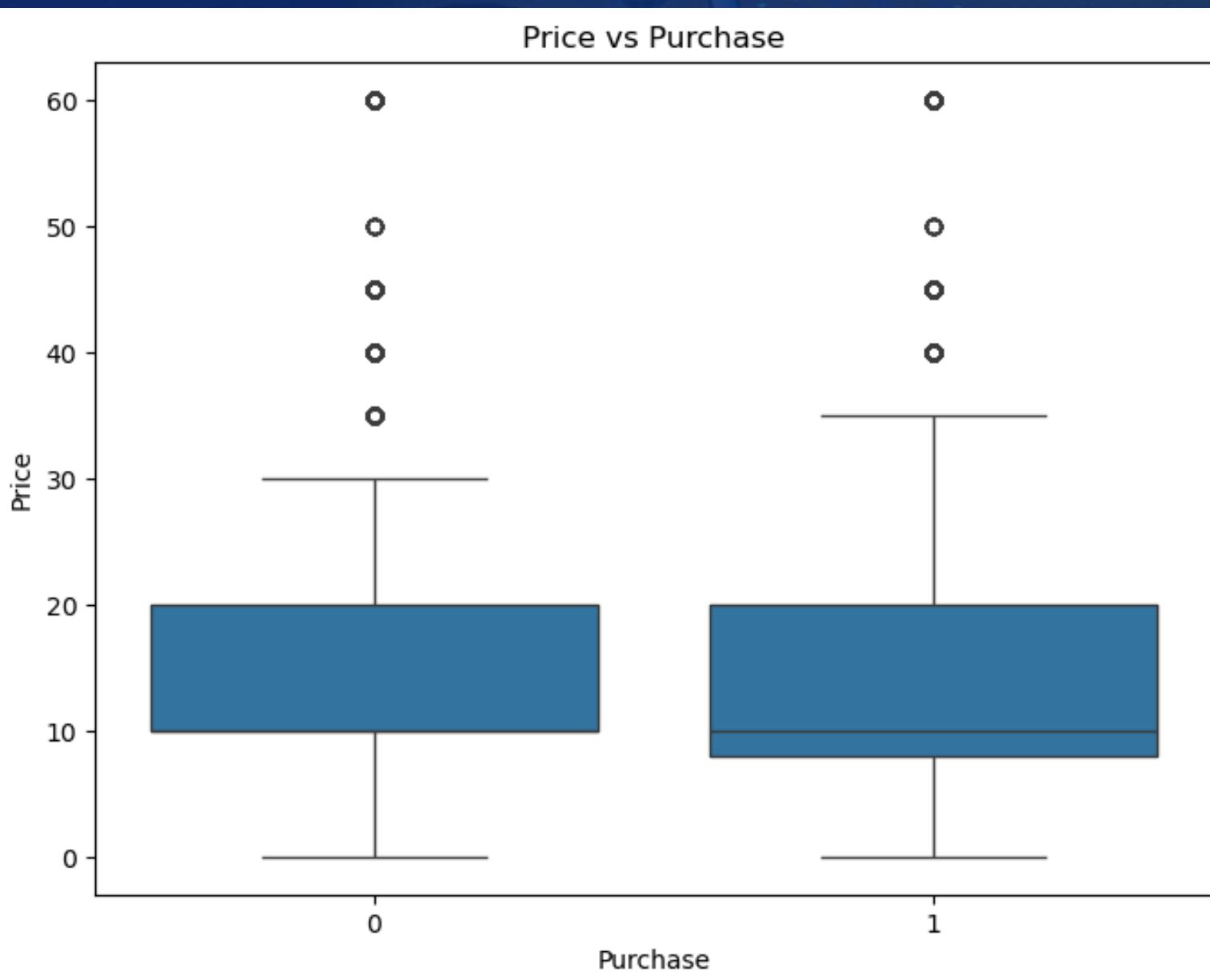
[home](#)[about us](#)[contact](#)

EXPLORATORY DATA ANALYSIS HIGHLIGHTS





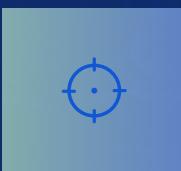
EXPLORATORY DATA ANALYSIS HIGHLIGHTS



Question	Answer
Does lower price guarantee purchase?	✗ No, but lower prices slightly help
Are expensive games purchased?	✓ Yes, but much less often
Is price alone a strong feature?	✗ No – needs to combine with other features

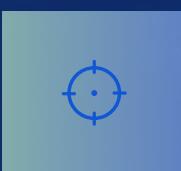
[home](#)[about us](#)[contact](#)

MODEL COMPARISON & INTERPRETATION



Models Tested (Classifier)

- Logistic Regression – baseline model
- Decision Tree
- Random Forest ✓ (best performer)



Performance Metrics

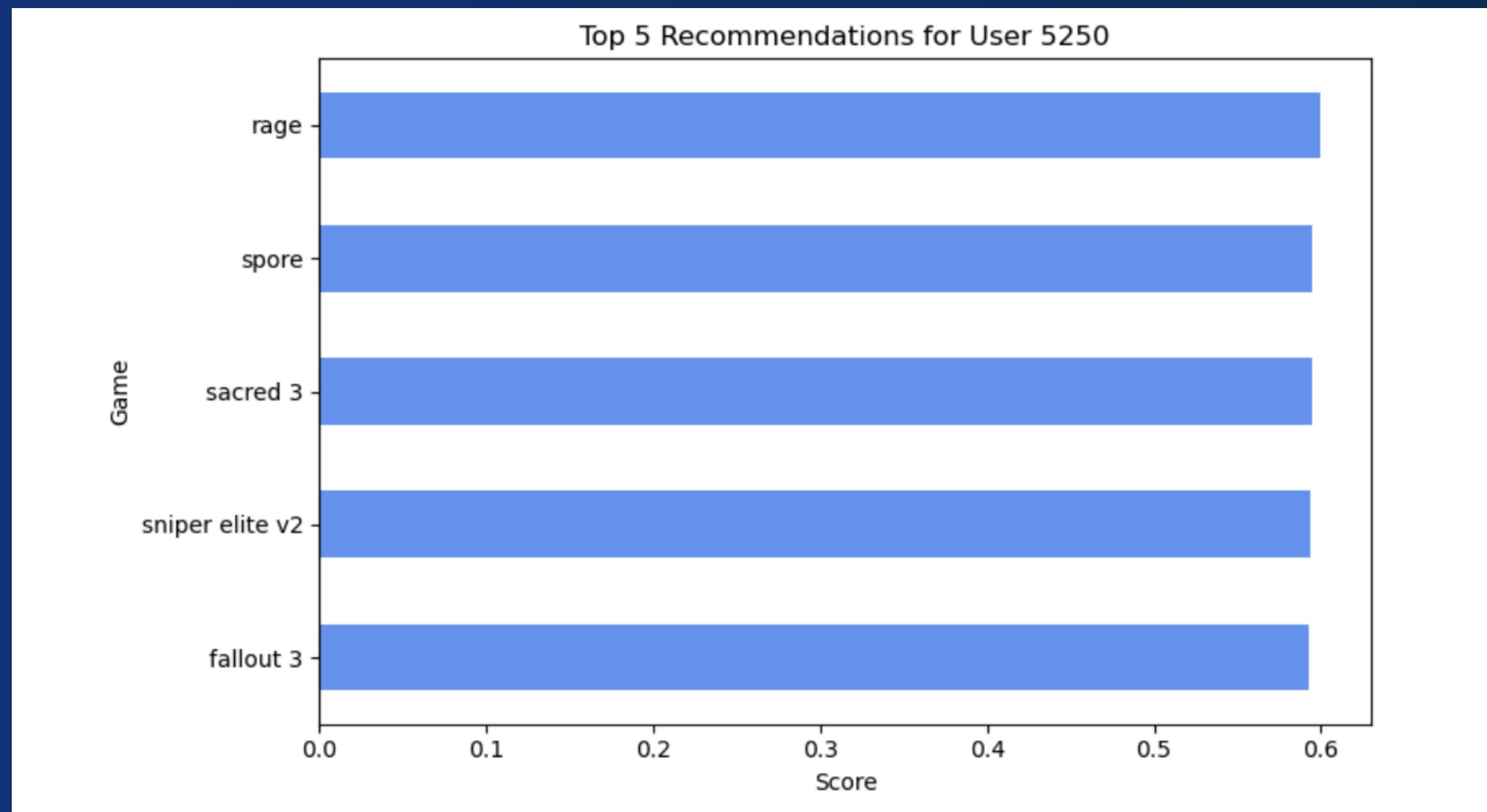
Model	Accurac	Precisio	Recall	F1 Score	ROC AUC
	y	n			
Logistic Reg.	0.90	0.85	1.00	0.92	0.87
Decision Tree	0.996	0.993	1.00	0.996	0.996
Random Forest	0.996	0.993	1.00	0.996	0.996

MODEL PERFORMANCE & RECOMMENDATION STRATEGY

- 
- Tree-based models generalize well and outperform linear baseline
 - RF chosen for its stability, scalability, and interpretability
 - Hybrid recommender blends both strategies, mitigating weaknesses like cold-start or overspecialization

RECOMMENDED GAMES FOR USER 5250

These are the top 5 games our system suggests, based on what they played and what similar users enjoyed.





ROADMAP FOR ADVANCED MODELING

- Expand feature set with TF-IDF review analysis
 - Handle class imbalance and improve model robustness
 - Optimize models with Decision Trees, Random Forests, ensembles
 - Evaluate adding more review data (pros: richer features; cons: extra cleaning)
 - Build hybrid recommendation system (content + collaborative filtering)



home

about us

contact



THANK YOU!