GameSeeker – A Personalized Game Recommendation Engine

Keer Feng

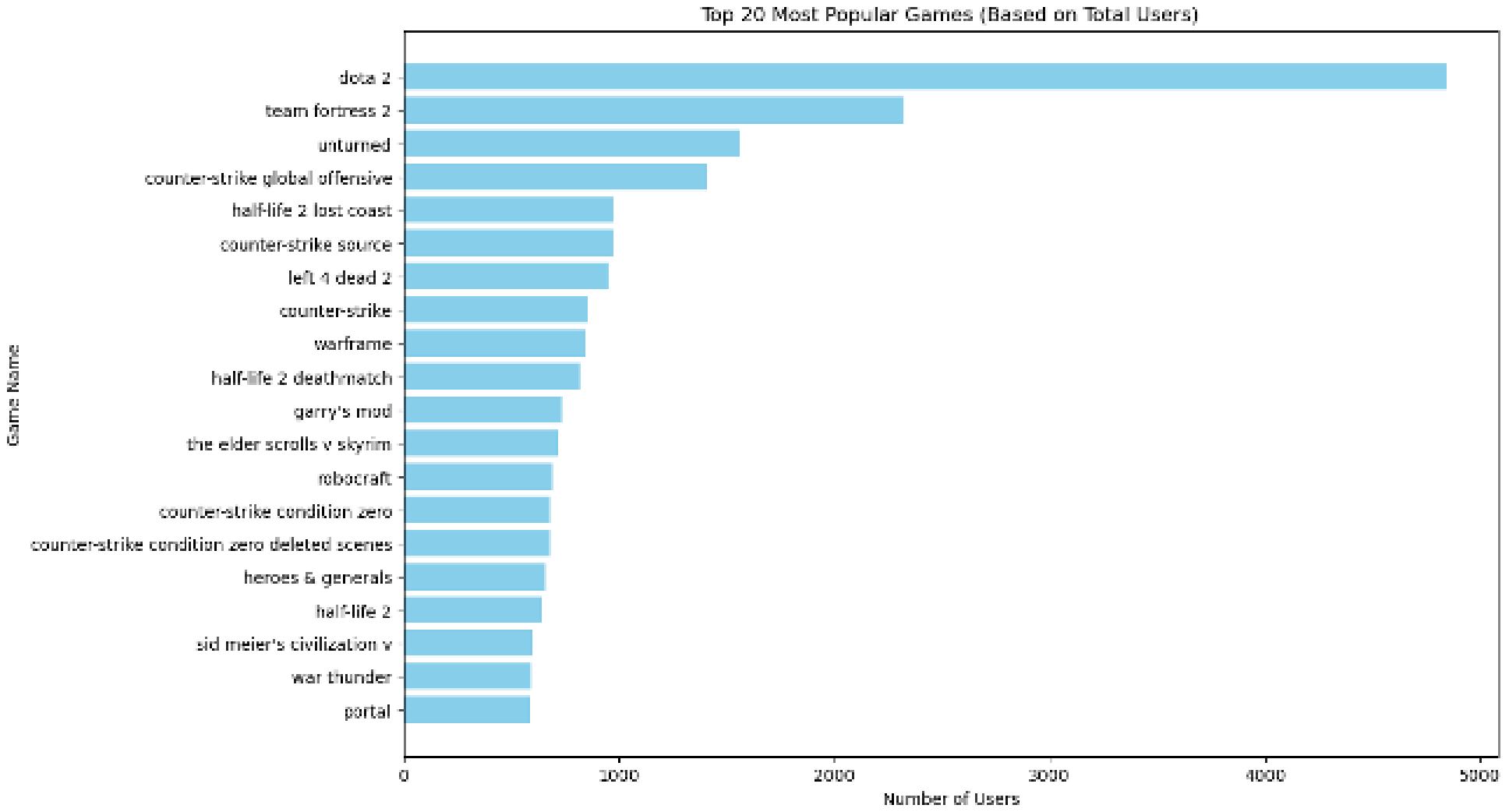
Game Recommendation System: Problem and Opportunity

Subject Area: Improving game recommendation systems for platforms like Steam, Epic Games, and PlayStation Store.

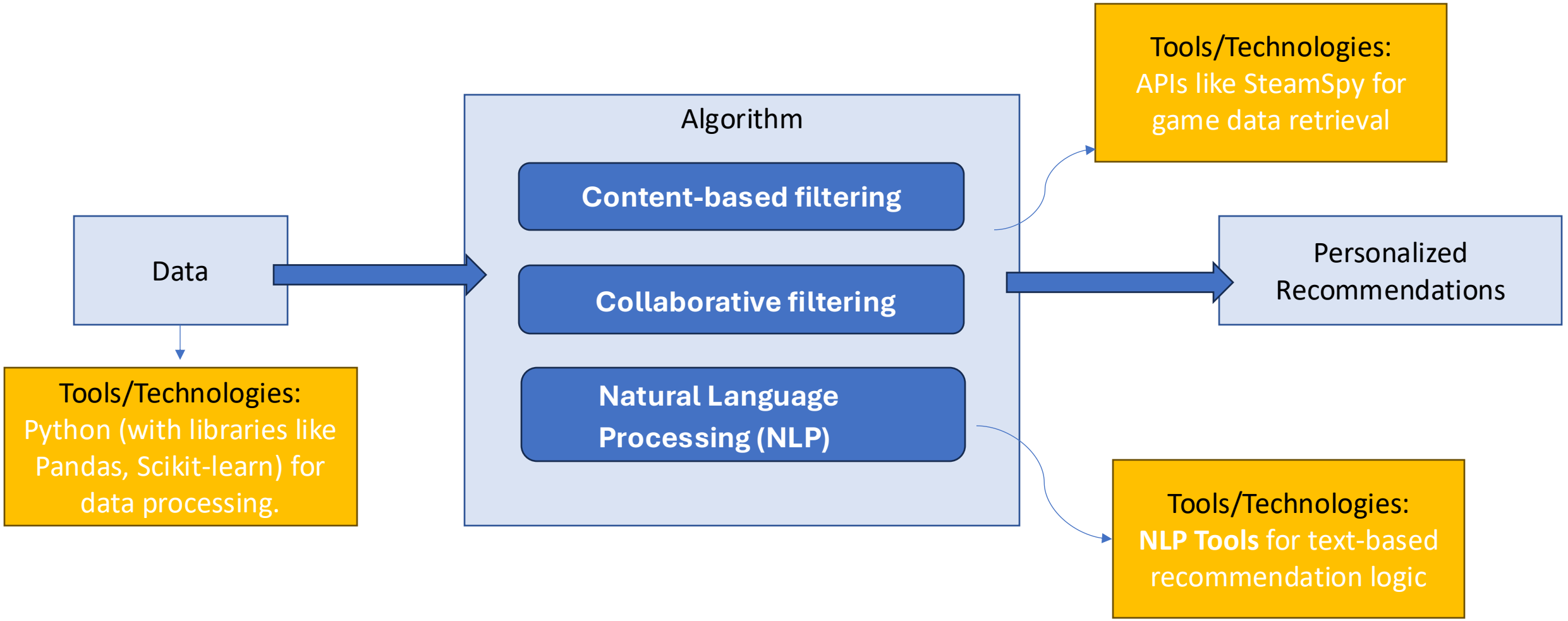
Goal: Build a system that offers **personalized recommendations**

Challenge: Recommending games that match individual preferences (art style, gameplay, story, etc.), instead of just popular titles.

Problem with the Current Approach



Proposed Data Science Approach for Game Recommendations



Potential Impact of Game Recommendation System

Impact on Stakeholders:

- **Players**: Will find games that truly match their interests, making the discovery process more enjoyable and efficient.
- **Developers**: Especially indie developers, can reach players who are genuinely interested in their specific game types, boosting visibility and sales.
- **Platforms**: Enhanced recommendation systems can improve user satisfaction, engagement, and retention, which directly translates to increased revenue.

Datasets Overview:

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.

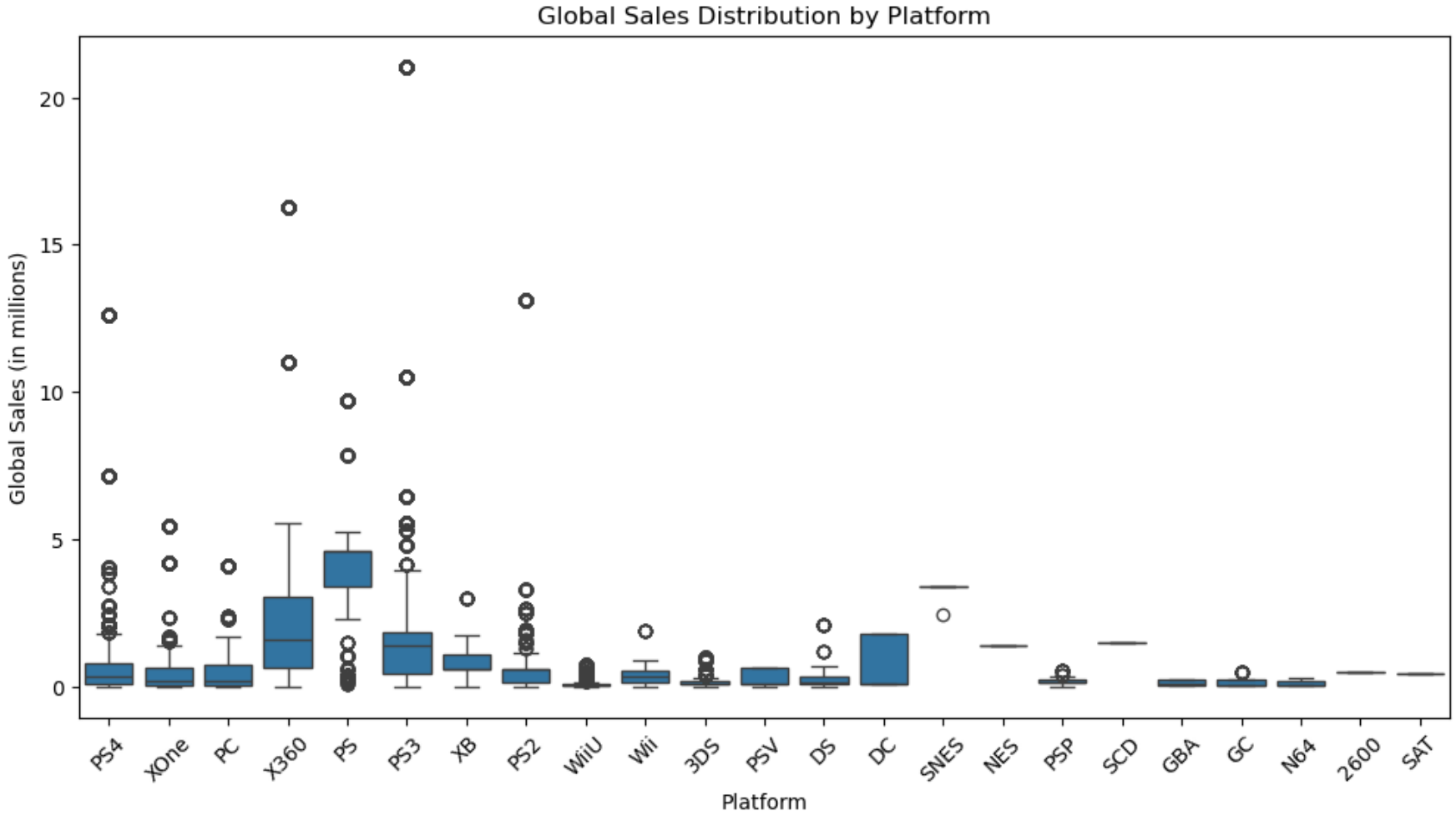
	Name	Platform	Year_of_Release	Genre	Publisher	NA_Sales	EU_Sales	JP_Sales	Other_Sales	Global_Sales	Critic_Score	Critic_Count	User_Score	User_Count	Developer	Rating
0	Wii Sports	Wii	2006.0	Sports	Nintendo	41.36	28.96	3.77	8.45	82.53	76.0	51.0	8.0	322.0	Nintendo	E
1	Super Mario Bros.	NES	1985.0	Platform	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

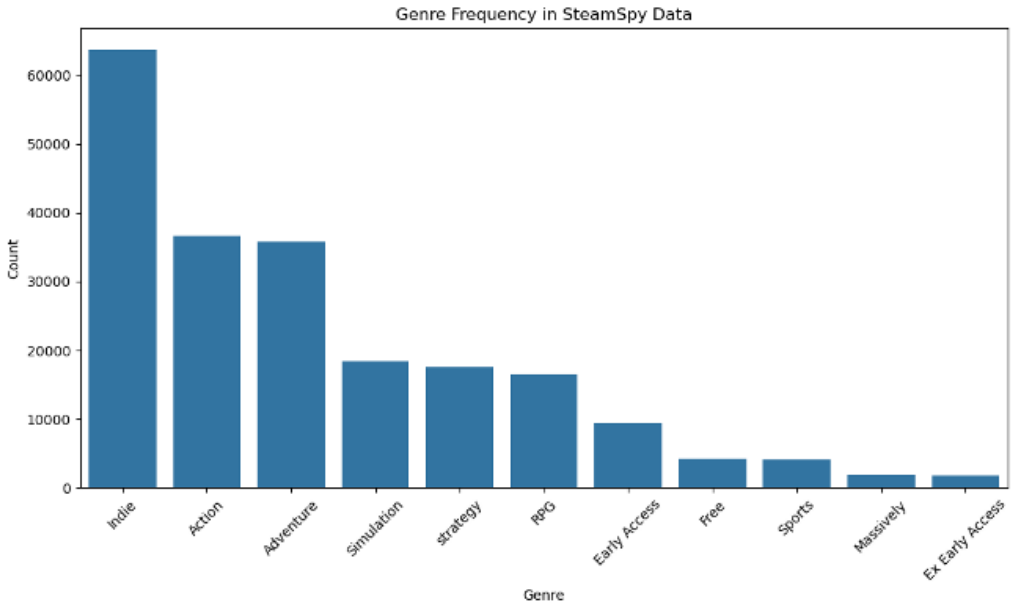
SteamSpy Data: Aggregated statistics across genres and game popularity

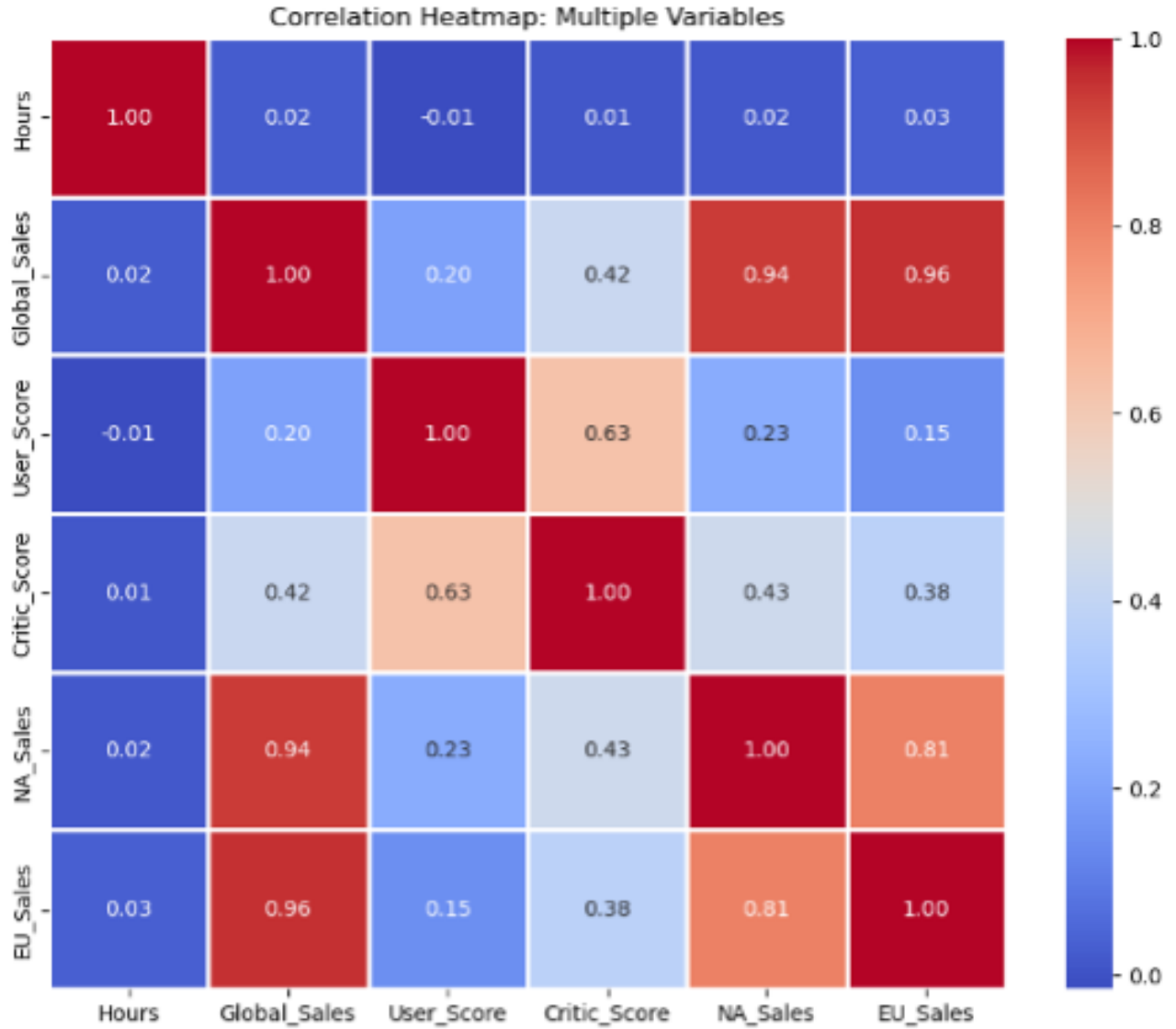
	#	Game	Release date	Price	Score rank(Userscore / Metascore)	Owners	Playtime (Median)	Developer(s)	Publisher(s)	Genre
0	63	Dota 2	Jul 9, 2013	Free	N/A (N/A/90%)	200,000,000 .. 500,000,000	22:47 (11:16)	Valve	Valve	Action
1	385	Counter-Strike: Global Offensive	Aug 21, 2012	Free	N/A (N/A/83%)	100,000,000 .. 200,000,000	12:25 (05:28)	Valve	Valve	Action
2	4	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	Action
3	121	Apex Legends	Nov 4, 2020	Free	N/A (N/A/88%)	50,000,000 .. 100,000,000	09:57 (03:00)	Respawn	Electronic Arts	Action
4	4184	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	Action

Data Quality Concerns:

- **Missing Data**
- **Duplicates & Inconsistencies**
- **Outliers**




Genre Frequency in SteamSpy Data
Count
60000
50000
40000
30000
20000
10000
0
Indie
Action
Adventure
Simulation
strategy
RPG
Early Access
Free
Sports
Massively
Ex Early Access
Genre


Correlation Heatmap: Multiple Variables
Hours
Global_Sales
User_Score
Critic_Score
NA_Sales
EU_Sales
1.00 0.02 -0.01 0.01 0.02 0.03
0.02 1.00 0.20 0.42 0.94 0.96
-0.01 0.20 1.00 0.63 0.23 0.15
0.01 0.42 0.63 1.00 0.43 0.38
0.02 0.94 0.23 0.43 1.00 0.81
0.03 0.96 0.15 0.38 0.81 1.00
1.0
0.8
0.6
0.4
0.2
0.0

Next Steps: Data Processing, Feature Engineering, and Modeling

- Data Processing: Handle missing values, outliers and refine the Model for real-time recommendations.
- Feature Engineering: Create new features, such as engagement scores and genre-based features.
- Modeling: Develop a baseline recommendation model using collaborative filtering and content-based filtering techniques.