

START GAME

PLAY

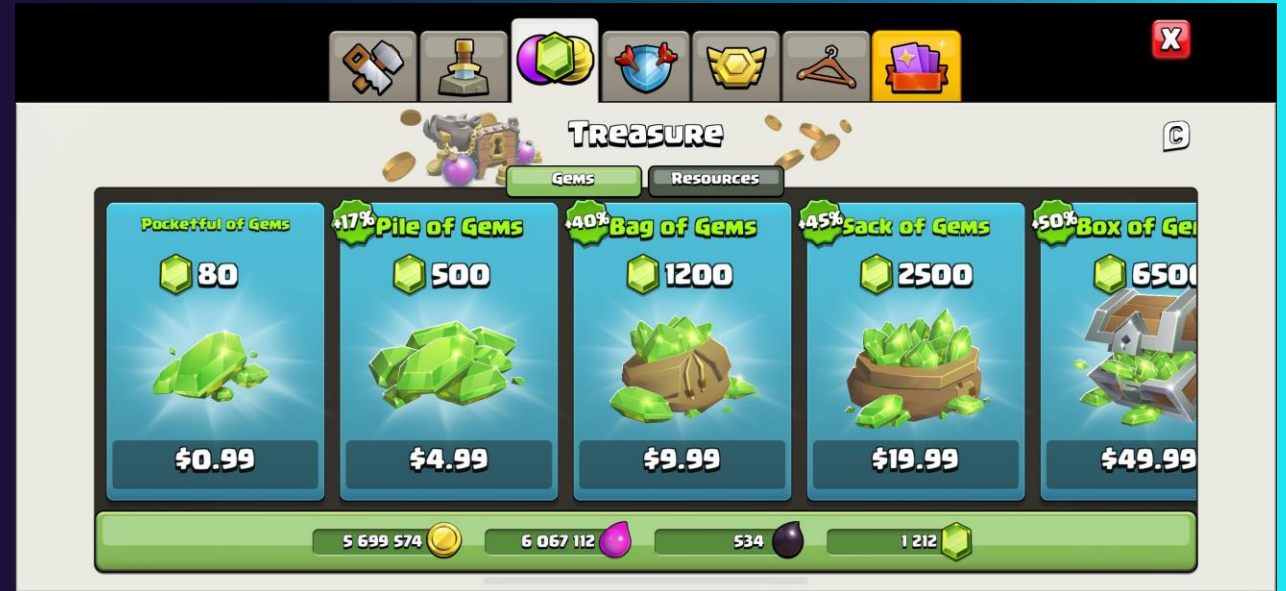
In-Game Purchase Predictions

A Deep Dive into Player Behavior



Understanding in-game Purchases

- ❑ Free Games = Money
- ❑ Cosmetics, pay to win, etc.
- ❑ Makes the game more fun

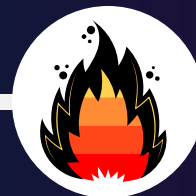
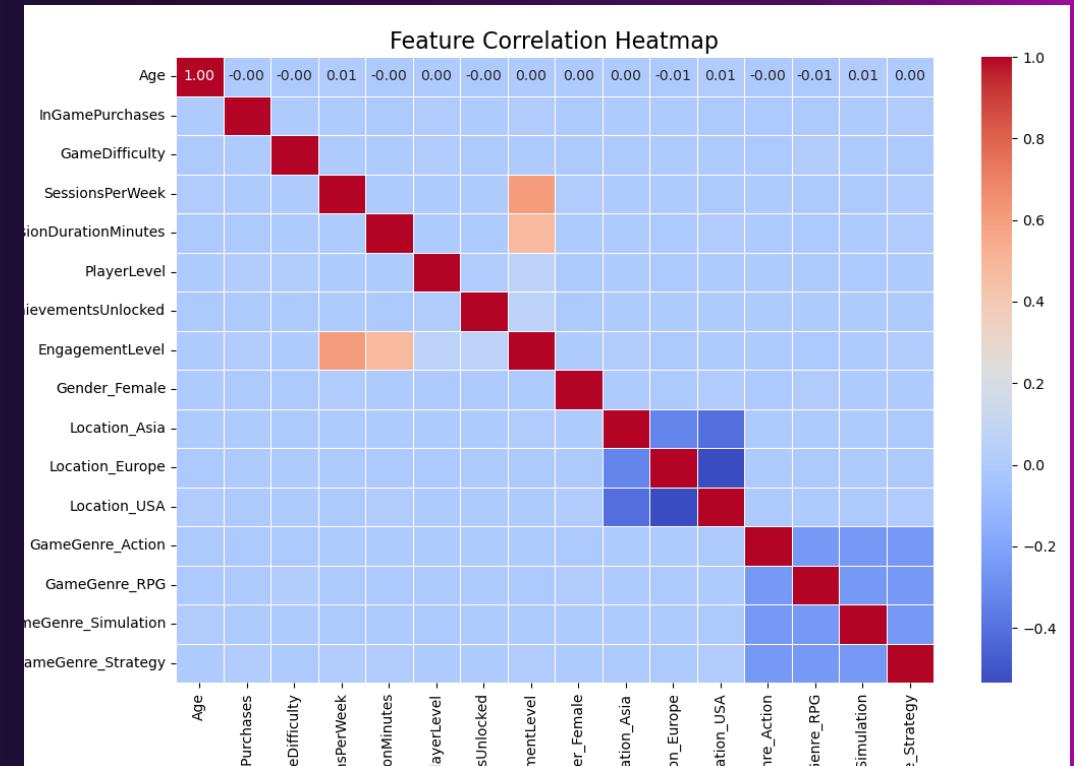


Buy Now

Our Data Set & EDA

Data columns (total 13 columns):

#	Column	Non-Null Count	Dtype
---	-----	-----	-----
0	PlayerID	40034 non-null	int64
1	Age	40034 non-null	int64
2	Gender	40034 non-null	object
3	Location	40034 non-null	object
4	GameGenre	40034 non-null	object
5	PlayTimeHours	40034 non-null	float64
6	InGamePurchases	40034 non-null	int64
7	GameDifficulty	40034 non-null	object
8	SessionsPerWeek	40034 non-null	int64
9	AvgSessionDurationMinutes	40034 non-null	int64
10	PlayerLevel	40034 non-null	int64
11	AchievementsUnlocked	40034 non-null	int64
12	EngagementLevel	40034 non-null	object



Finances

Estimated New Annual Revenue for Client

Revenue Source	Incremental Purchases/ Yr**	Average Cost per Purchase *	Total Revenue
In App Purchases - Established Game	6000	\$12.28	\$73,680
Total New Annual Revenue			\$73,680

Estimated Cost of Consulting Project

Team Members	Consultant	Hourly Rate***	Total Hours	Total Cost
Machine Learning Data Scientist	A. Baraban	\$155.00	80	\$12,400
Statistical Analyst	R. Hough	\$149.00	80	11,920
Data Modeler	C. Wanko	\$141.00	80	11,280
Business Insight and Analytics Manager	M. Nicholas	\$162.50	80	13,000
Total Consultant Cost			240	\$48,600
Data and Licensing Cost				500
Administrative and Supplies Costs				2,000
Infrastructure Cost (4% of Consultants)				1,944
TOTAL COST TO CLIENT				\$53,044

Rate of Return Year 1	39%
5 Year Rate of Return	137%

* - Avg. Cost per In-App purchase derived from: <https://www.statista.com/statistics/1215270/iap-gaming-app-transactions-value-genre/>

** - Estimate of incremental 10% on base of 5,000/month or 60,000/year.

Note that 33% of App Store and Google Play apps garner btwn 1,000 and 10,000 downloads per month.

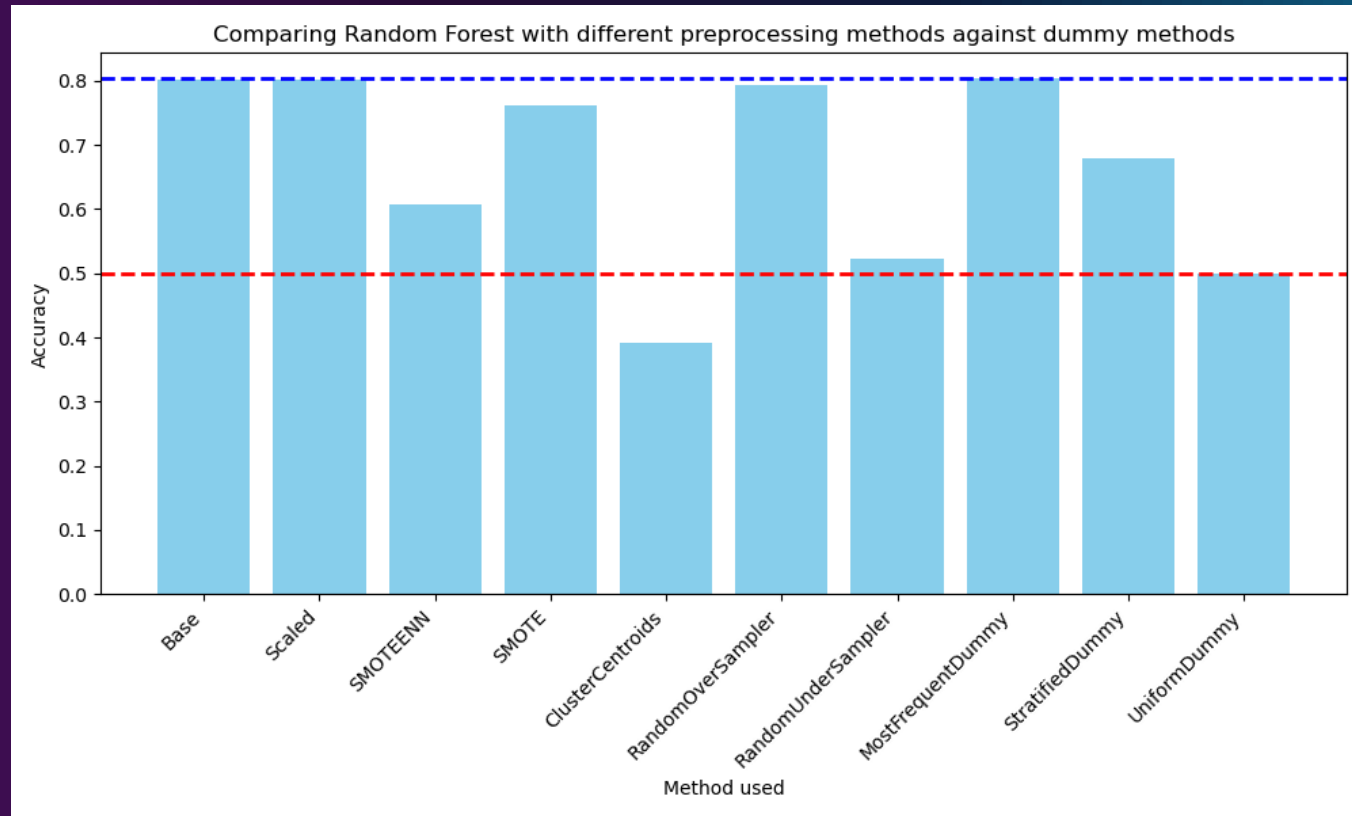
***- Estimated Hourly rates are derived from <https://www.mastersofbusinessanalytics.com/careers/>

Calculation = Estimated Annual Salary/2,080 Hours/year * 1.3 to include employer benefits x 2.5 Fee upcharge

		Revenue				
	Investment	Year 1	Year 2	Year 3	Year 4	Year 5
Cash Flow	-\$53,044	\$73,680	\$73,680	\$73,680	\$73,680	\$73,680



Initial analysis



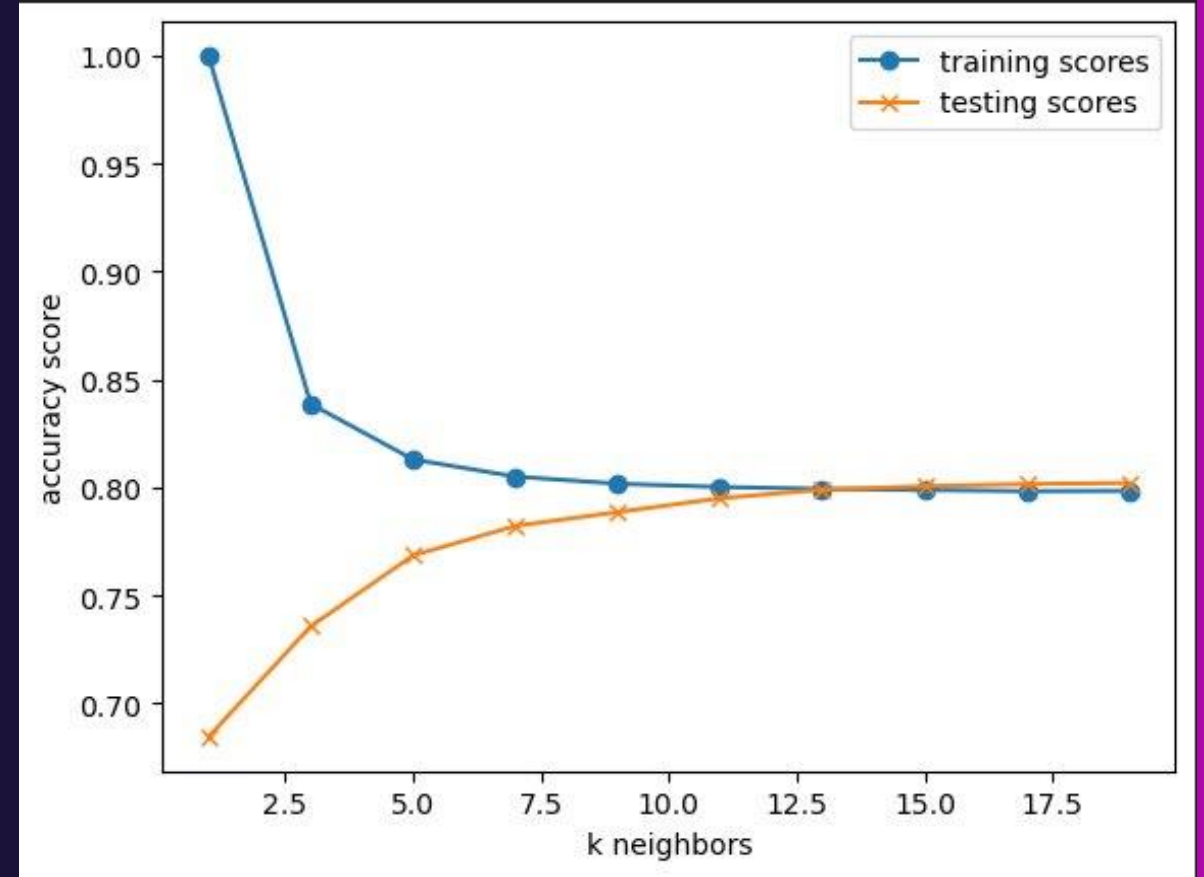
Griefing

When the data becomes too toxic...



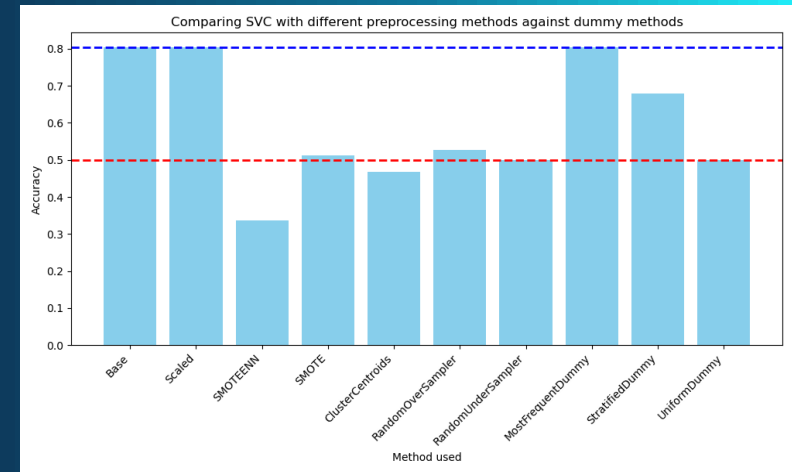
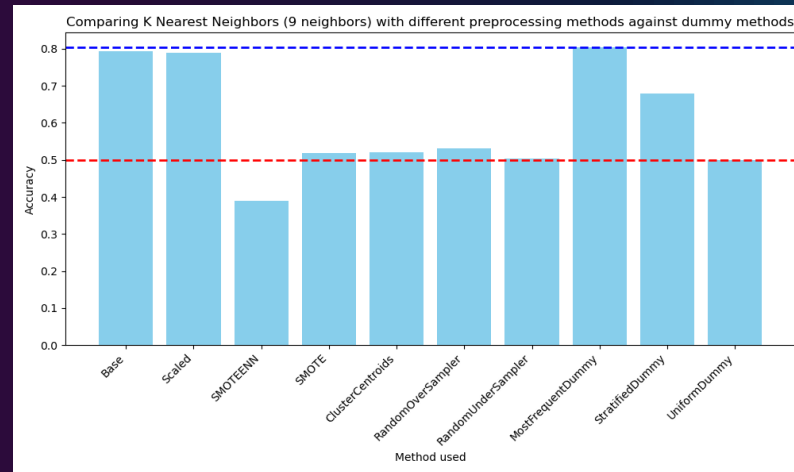
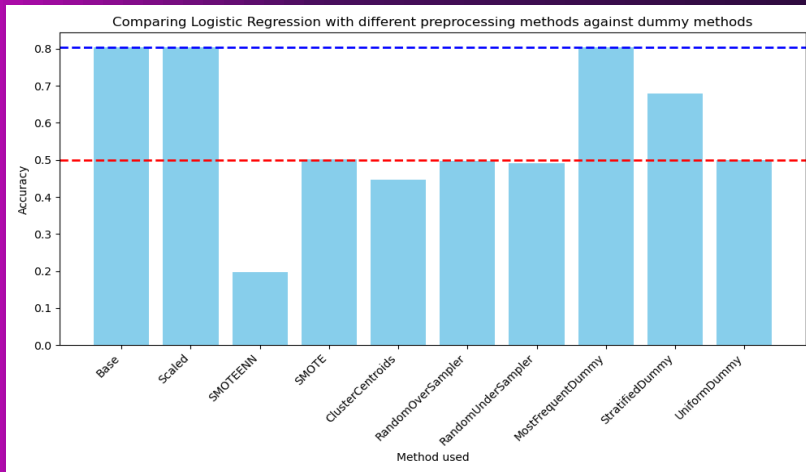
Test/Train values for KNN

Test scores  train scores

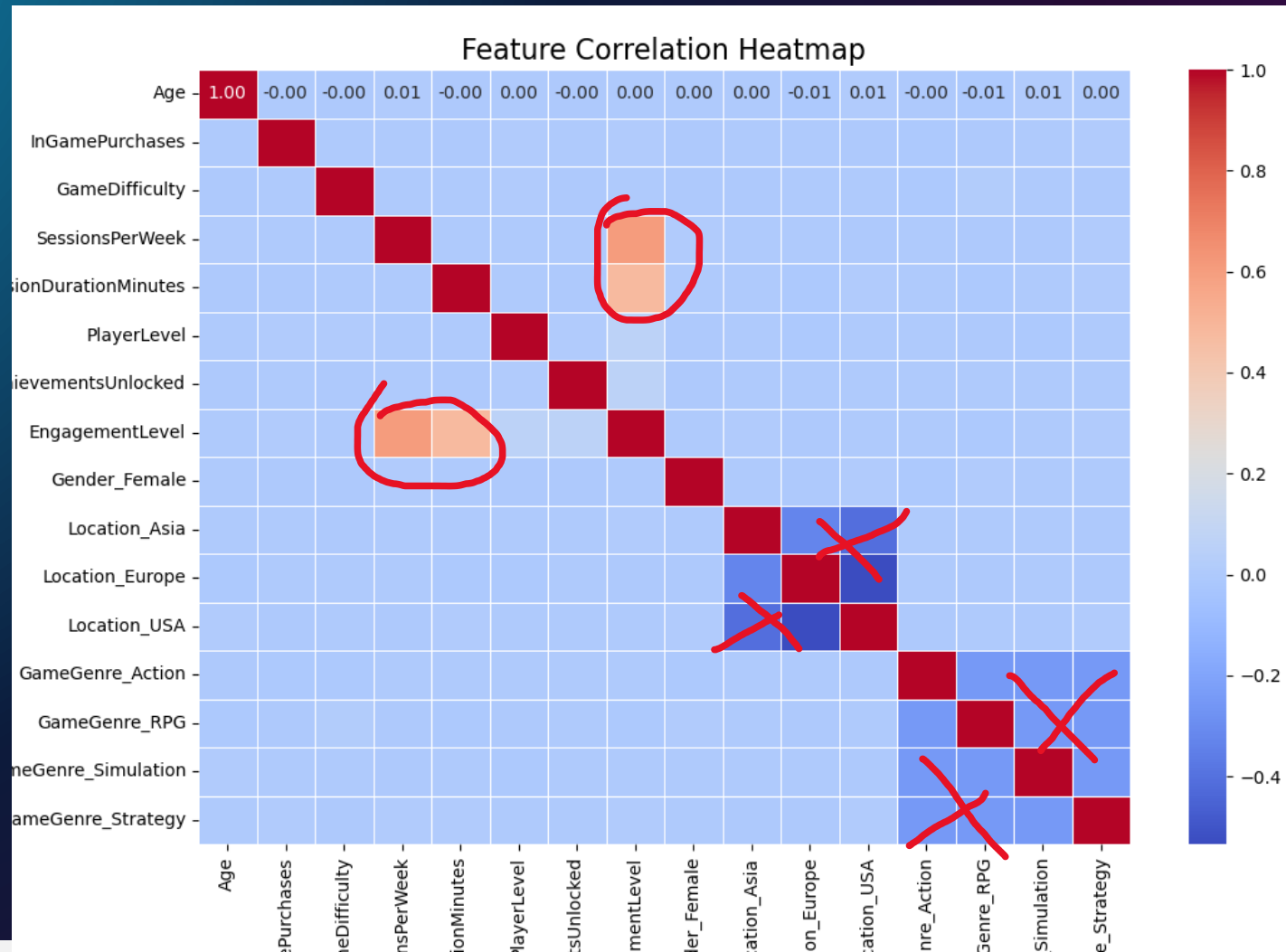


Models Aren't Learning

Look at the blue line...

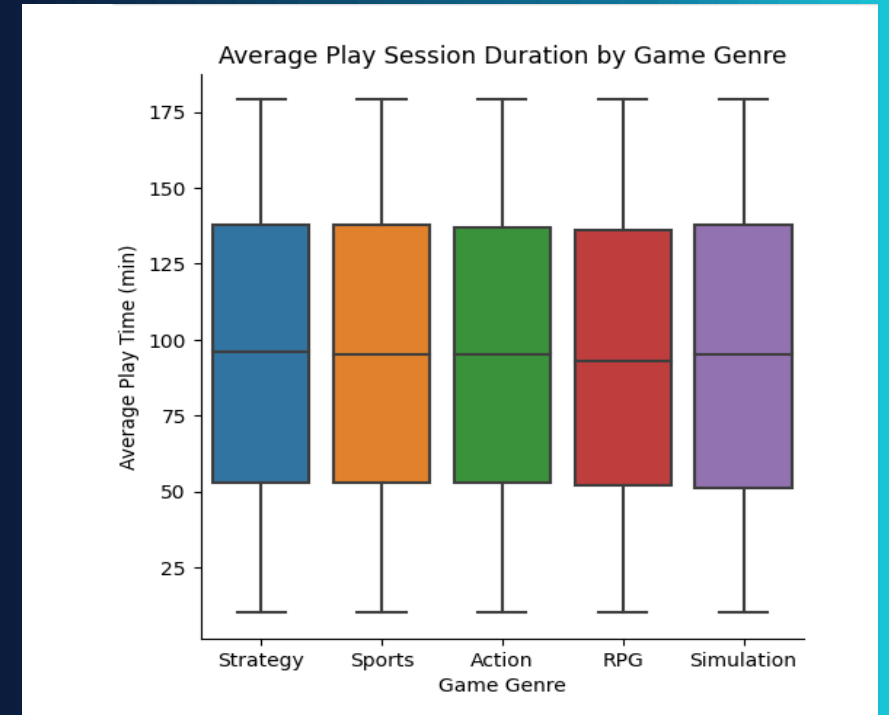
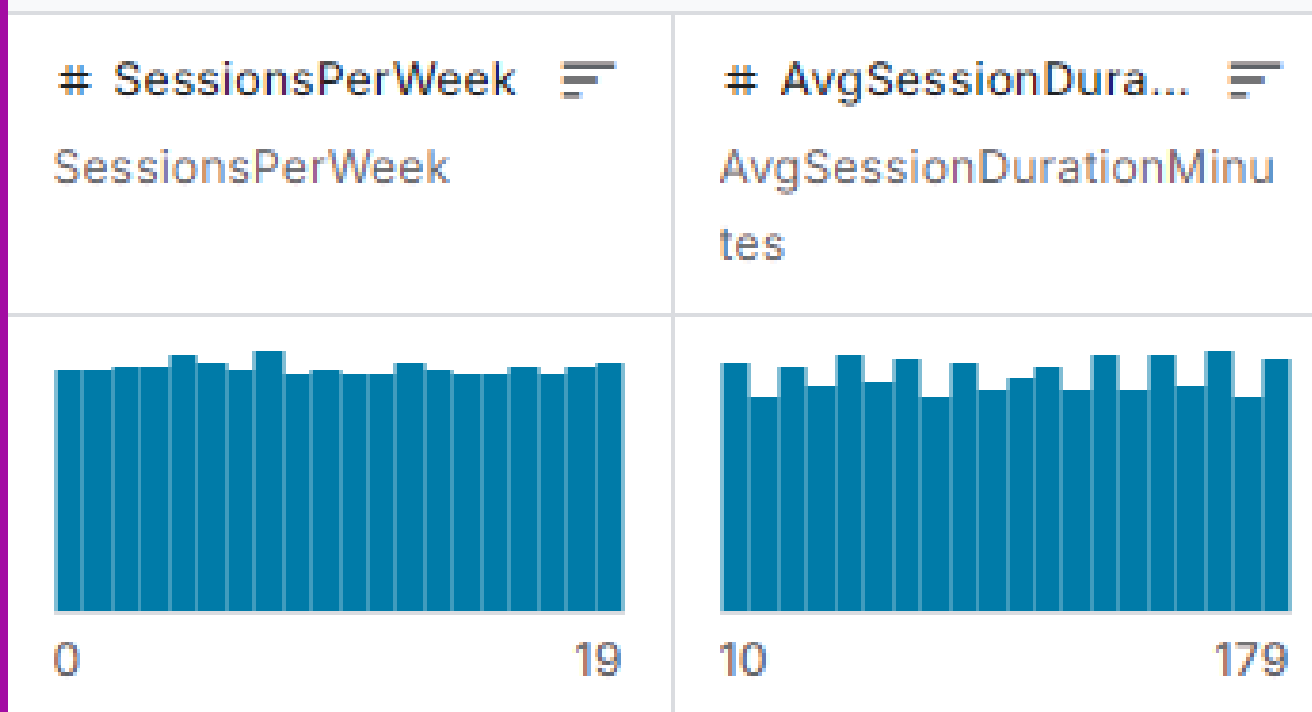


Correlations



Correlations of one hot encoded columns are denoted by red X, disregard them

Distributions



The response...



Comment on This is Synthetic Data, Isn't It?



Posted 2d ago · In [Predict Liver Disease: 1700 Records Dataset](#)

I didn't design the data card, Kaggle is putting the data source in the bottom, but sure we did take your feedback into consideration and added a section on the top.

For the nature of the data, it's not your friend who 'just' discovered this, we were talking with data providers about this matter and we were waiting for them to give us permission to share all the information about it, for example now we still waiting for them to launch their website before we cite them on the data cards.

We understand your frustration and have listened to your concerns. We recognize that this is your first interaction on Kaggle in three years, and similarly, for your friend, his first in seven years on the platform. We would love to see your contributions here on Kaggle, whether through sharing synthetic or real data, participating in competitions, or engaging with other categories like notebooks. Contributing to the community is important, rather than simply consuming free content and complaining about it.

Another option you may not be aware of is the ability to report content. If you come across something on Kaggle that you find inappropriate or concerning, you can click "Report" and submit your complaint. Kaggle moderators will review it, and if the content needs to be removed, they will take the necessary action.

Show less





**GAME
OVER**

The text "GAME OVER" is rendered in a bold, yellow, pixelated font with a digital glitch effect. Each letter has a cyan shadow and is surrounded by small, colorful geometric shapes (circles, triangles, squares, and crosses) in yellow, cyan, and magenta. The background is a dark blue gradient with a cyan-to-blue gradient on the right side.

PLAY AGAIN ?

YES

NO

Appendix

<https://github.com/baraban7890/In-App-Purchase-Analysis>