



Music Recommendation System



By: Theodore Chidi-Maha

Problem Context

- People's lives are too fast-paced and have too many distractions
 - They don't have time to sift through Spotify's extensive song library
- Causing them to be dissatisfied with the product and cancel their subscriptions



[This Photo](#) by Unknown Author is licensed under [CC BY-SA](#)

Problem to Solve

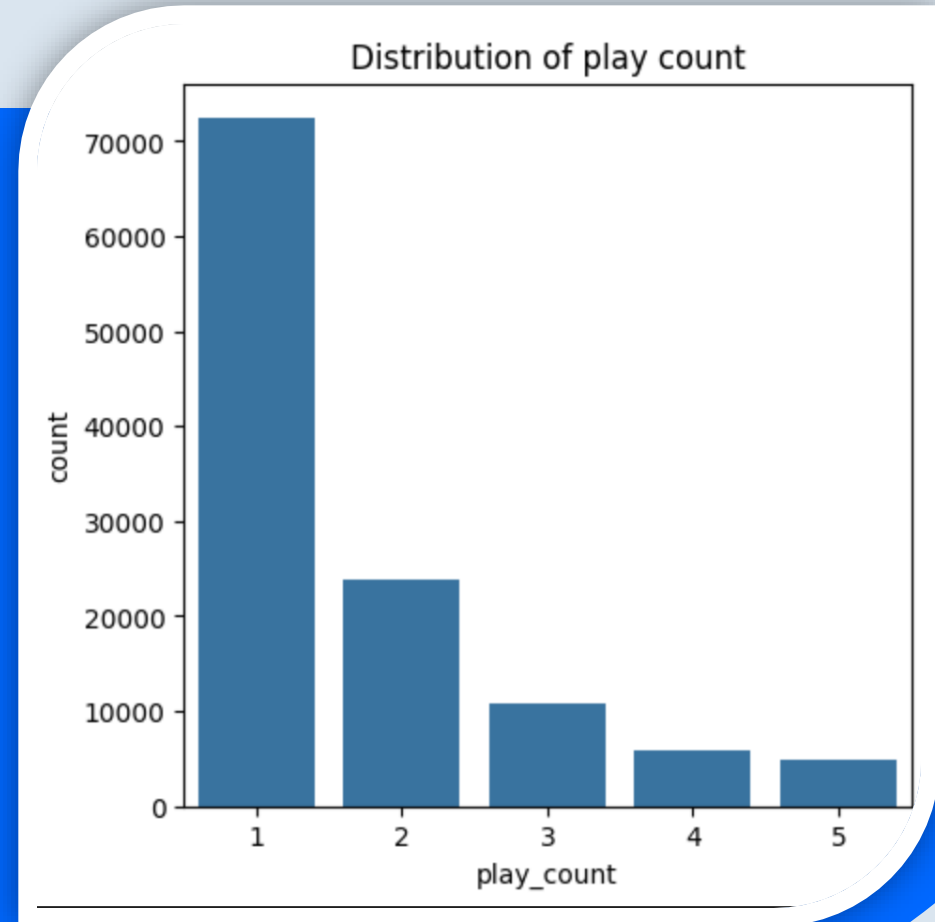
- Spotify relies on users continually engaging with their product to generate revenue.
- **we need a system that will enable users to engage with the platform more efficiently.**
- By discovering songs, they will enjoy even on a limited timeframe.



Data Insights

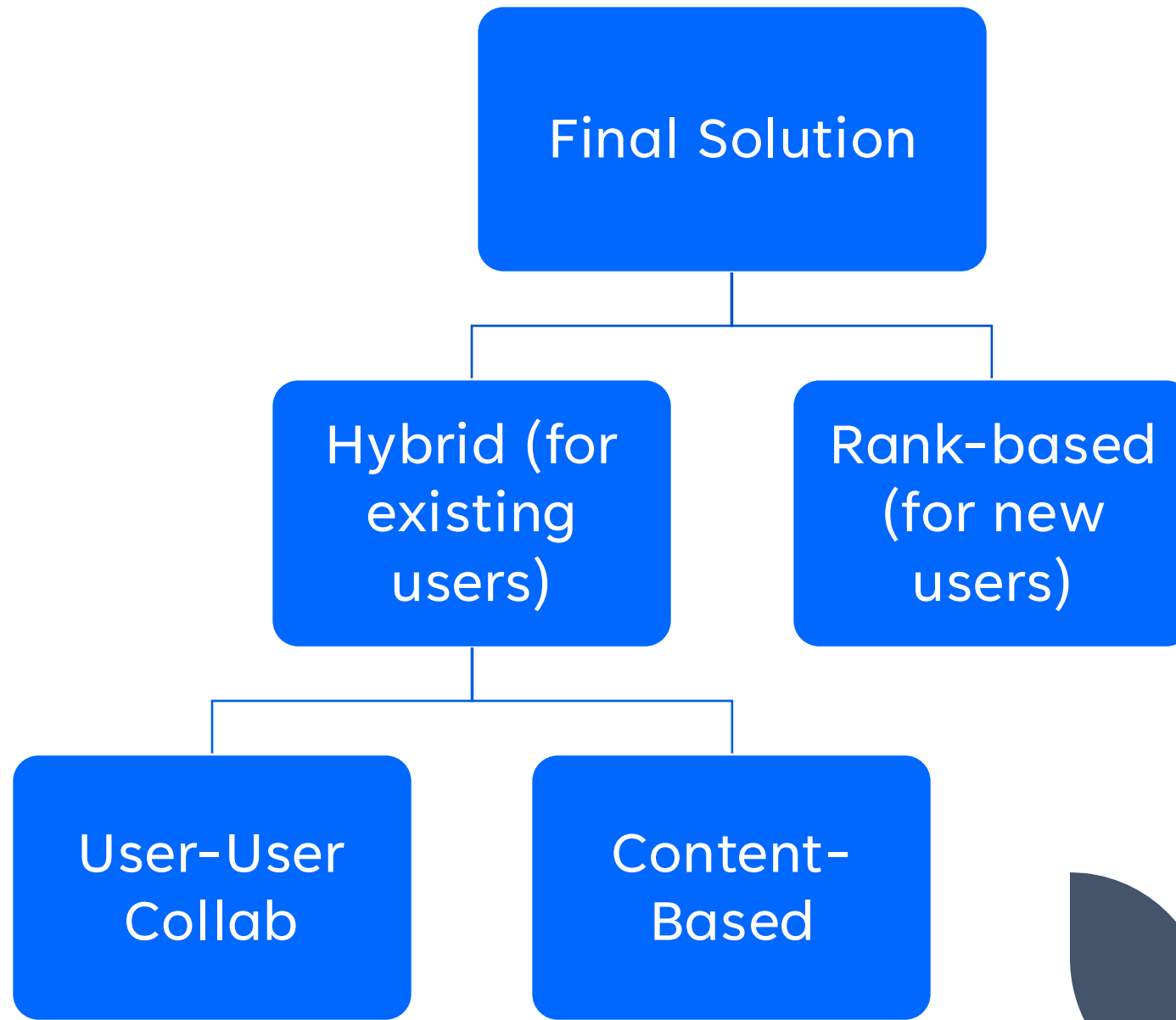
- Data exploration led to a few key takeaways:
- Song play counts is the best tool to identify how users rate the songs
- More notably, **there is a bias of 61% of users playing a song only 1 time**

* This data comes from a filtered dataset



Recommendation models

Systems	Insights
Rank-Based	<ul style="list-style-type: none">• Recommends songs based on popularity.• Useful only for brand new user.
User-User Collab	<ul style="list-style-type: none">• Used users' similarity• Had the best perform both in metrics and in practice.
Item-Item Collab	<ul style="list-style-type: none">• Used items' similarity• worse than user-user.
Content-Based	<ul style="list-style-type: none">• Used meta-data (similar titles, artists, and album names)• gave rather middling results
Matrix Factor	<ul style="list-style-type: none">• Used latent factors in user-item interactions.• Good performance in metrics; Poor performance in practice
Cluster-Based	<ul style="list-style-type: none">• <i>Unsupervised</i> grouping of users• Poor performance both in metrics and in practice.



Algorithm performance metrics

Systems	RMSE	Precision (relevant recommendations)	Practical Error Margin
<u>Hybrid</u>	<u>0.9003</u>	<u>50.9%</u>	<u>2%</u>
User-User Collab	1.0521	43.2%	2%
Item-Item Collab	1.0328	40.8%	2%
Content	*N/A	40%	*N/A
Matrix Factor	1.0083	43.2%	28.5%
Cluster	1.0532	40.8%	45%

- adding a 5-star ratings system for songs
- Add song Genres to dataset
- Also, Add User data [Age, country, etc.]
- Performing quarterly re-analysis of data and models

Recommendations/ Action items

Risks/Challenges

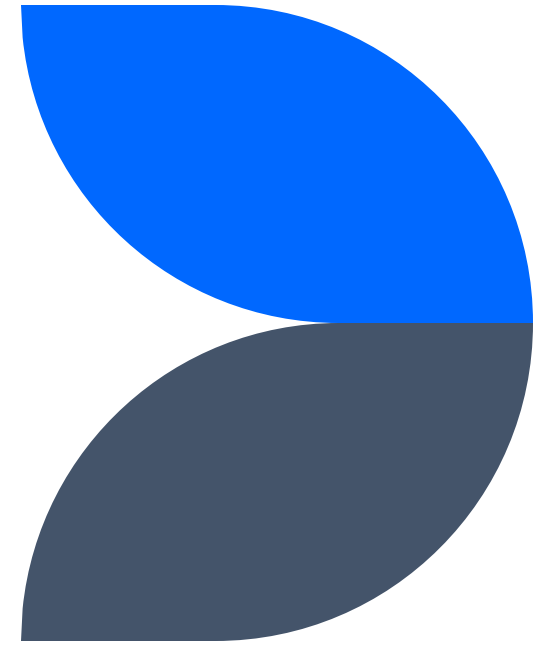
Risks:

- Increasing the amount of personal data we collect, will require us to increase data security

Challenges:

- Data concerns may cause user to desire the ability to opt-out of the recommendation system.
- We also need to be careful of the content we recommend to our users (ex. recommend explicit songs to minors)

Thank You



APPENDIX



Datasets Description

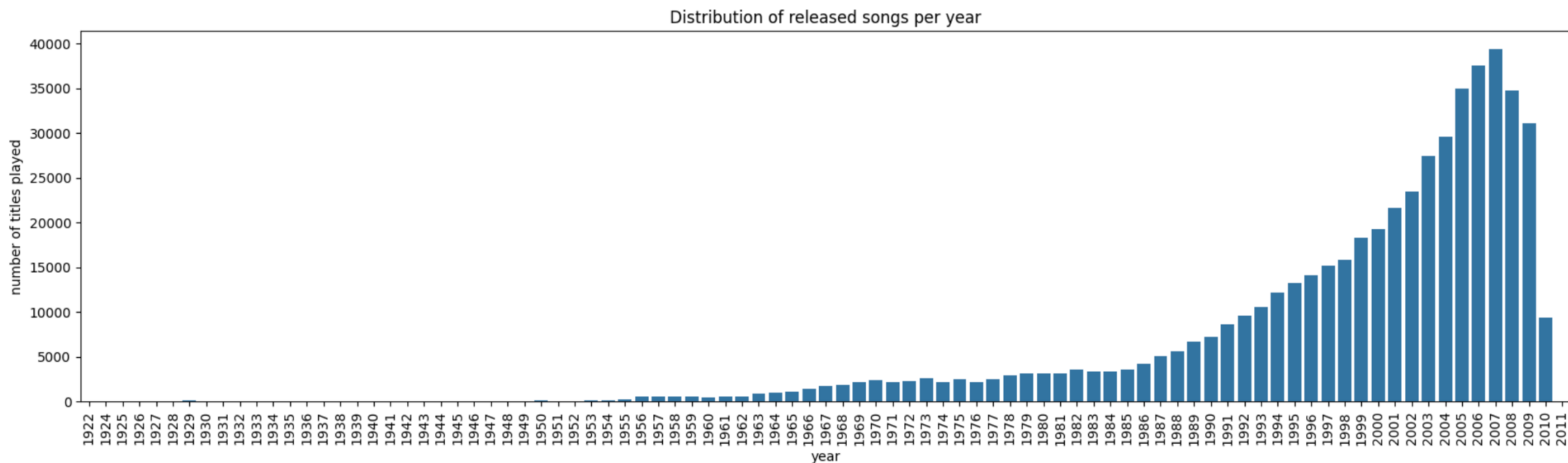
song_data:

- song_id: A unique id given to every song
- title: Title of the song
- Release: Name of the released album
- Artist_name: Name of the artist
- year: Year of song release

count_data:

- user_id: A unique id given to the user
- song_id: A unique id given to the song
- play_count: Number of times the song was played

Additional data trends (Yearly Releases)



Additional data trends (Yearly play counts)

