

Please prepare a 20-30 minute presentation summarizing your approach and findings for the task items. **The presentation should be aimed at an audience that would include both technical experts and non-technical business stakeholders.** Be prepared to discuss your methodology and rationale in detail, as well as your findings. We want to assess your technical, analytical and communication skills.

How you define and plan the analysis is up to you, and you are free to use any tools you want to complete the task.

Finally, be prepared to share any code or queries used, if you used a notebook, please send that to us in advance.

Optional items

If you find you complete the work quickly, here are some optional items to address.

- Based on your investigation into the data, what other use cases might you propose the data could be used for? Based on your understanding of the business domain, what additional types of information would you look to incorporate?
- An executive likes your analysis, but asks you to now integrate lyrics data, which currently exists as 1 text file per track. How would you incorporate/structure this dataset to flexibly perform analysis on how lyrics affect playlist performance?

About the Data

Download link: <https://drive.google.com/file/d/17UI4mfA2nLItslWGRD8C6cweY2TQvOke/view>

The data set contains simulated data for a sample of Spotify playlists, but the patterns within are similar to real data. The following table describes the fields available in the data set.

Column Name	Description
playlist_uri	The key, Spotify uri of the playlist
owner	Playlist owner, Spotify username
streams	Number of streams from playlist today
stream30s	Number of streams over 30 seconds from playlist today
dau	Number of Daily Active Users, i.e. users with a stream over 30 seconds from playlist today

wau	Number of Weekly Active Users, i.e. users with a stream over 30 seconds from playlist in past week
mau	Number of Monthly Active Users, i.e. users with a stream over 30 seconds from playlist in past month
mau_previous_months	Number of Monthly Active Users in the month prior to this one
mau_both_months	Number of users that were active on the playlist both this and the previous month
users	Users streaming at least one stream of any time duration (including skipped streams) in past month
skippers	Number of users who skipped more than 90 percent of their streams today
owner_country	Country of the playlist owner
n_tracks	Number of tracks in playlist
n_local_tracks	Change in number of local file tracks on playlist since yesterday
n_artists	Number of unique artists in playlist
n_albums	Number of unique albums in playlist
monthly_stream30s	Number of streams over 30 seconds this month
monthly_owner_stream30s	Number of streams over 30 seconds by playlist owner this month
tokens	List of playlist title tokens, stopwords and punctuation removed
genre_1	No. 1 Genre by weight of playlist tracks, from Gracenote metadata
genre_2	No. 2 Genre by weight of playlist tracks, from Gracenote metadata

genre_3	No. 3 Genre by weight of playlist tracks, from Gracenote metadata
mood_1	No. 1 Mood by weight of playlist tracks, from Gracenote metadata
mood_2	No. 2 Mood by weight of playlist tracks, from Gracenote metadata
mood_3	No. 3 Mood by weight of playlist tracks, from Gracenote metadata