360i Data Engineering Challenges

This takehome challenge involves pulling data from an API, creating a database, and running an analysis. We should be able to replicate your work - please provide any code, notes, and/or documents you used while analyzing the data in a zip file package. This helps us follow your thought process. This take home assignment is property of 360i and is not to be shared without permission.

Intro

The API we'll be working with is the Spotify API. We recommend the `spotipy` library to work with the API in Python. We ask that you munge the data into a SQLite database. `sqlite3` is part of the Python standard library. Please attach the SQLite db file as part of your submission. Finally, answer some questions using SQL queries and `pandas`. Feel free to add visualizations using the tool of your choice.

Spotipy API

To use the spotipy python API, you need to use a spotify "app". If you have a spotify account you can set one up on their website. Otherwise feel free to use ours:

```
os.environ['SPOTIPY_CLIENT_ID'] = '00b7317977ad4c0d971af8274f1aa790'
os.environ['SPOTIPY_CLIENT_SECRET'] = '6efbf45fe72d435f9739d0c0f4c26db5'
os.environ['SPOTIPY_REDIRECT_URI'] = 'https://360i.com/'

base_url = 'https://api.spotify.com'
scope = 'playlist-read-private'
# spotify:playlist:
rap_caviar = '5yolys8XG4q7YfjYGl5Lff'

token = util.prompt_for_user_token('Puffer Fish',scope=scope)

spotify = Spotify(auth=token)
```

Documentation for the python API can be found here.

ETL

Please build a database as described below using the Spotify "Rap Caviar" playlist. You will be making two tables in a database:

- tracks
- artists

Fill the `tracks` table with songs from the Spotify "Rap Caviar" playlist.

`tracks` table minimum fields:

- name
- popularity
- duration_ms
- artist_name

For songs with multiple artists, you can just use the first one listed.

Fill the `artists` table with all of the artists in the tracks.

`artists` table minimum fields: - id - name - popularity - followers

Analysis

Basic questions:

- How many songs are in the playlist?
- What are the top 5 tracks by artist follower count?
- Which song is the longest?
- What is the relationship between track and artist popularity?

If you find that you have some time and want to investigate ideas outside of the questions above, we'd be delighted but it's not expected. We understand that you've got a busy schedule. Below are some possible directions to go in.

- Add additional fields to the tables of your choosing
- Pick another playlist and add to the track and artist tables.
- Compare features across playlists
- What features are predictive of track popularity?