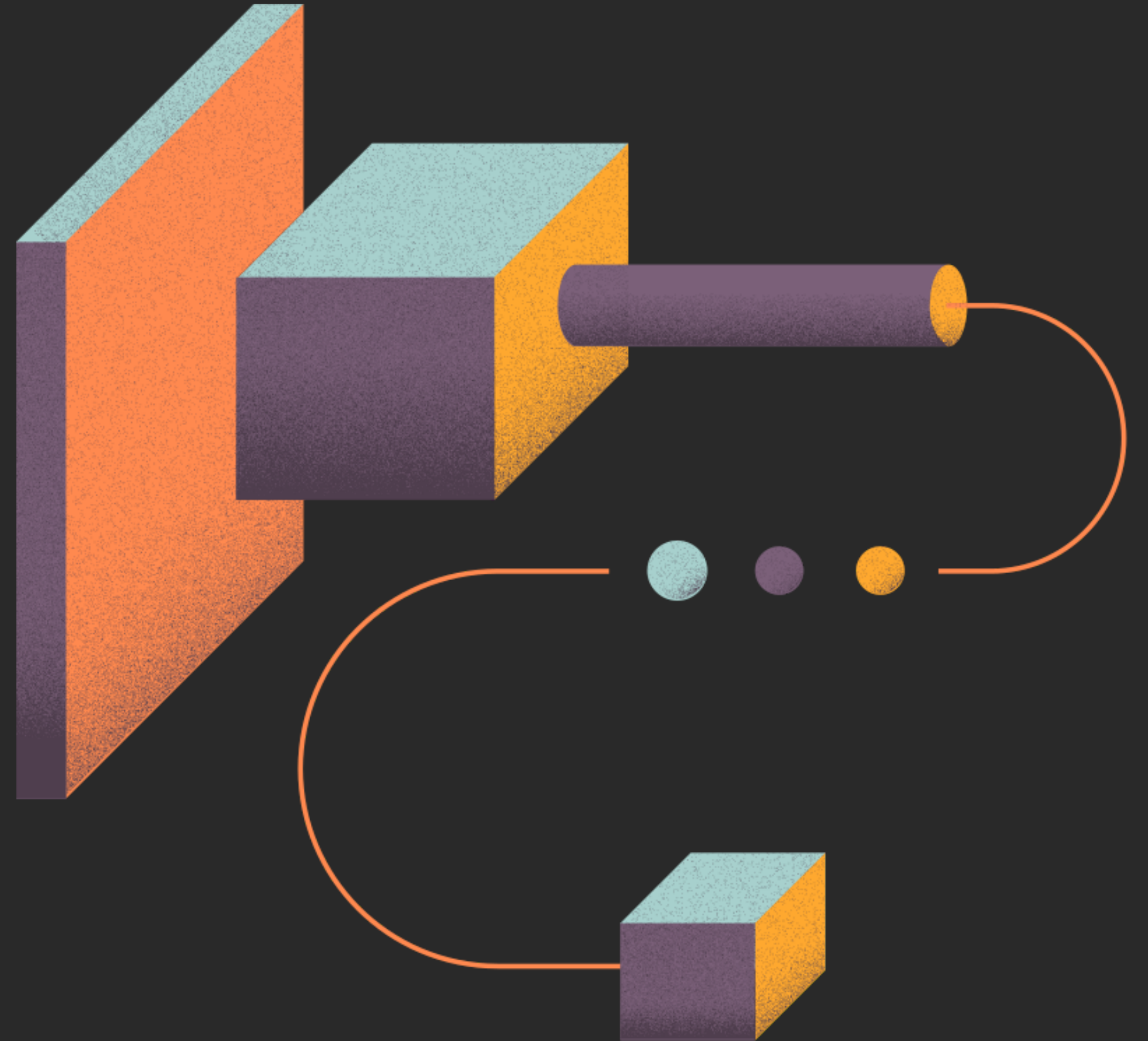




MUSIC RECOMMENDATION BY ETL, ML AND POWER BI ANALYSIS



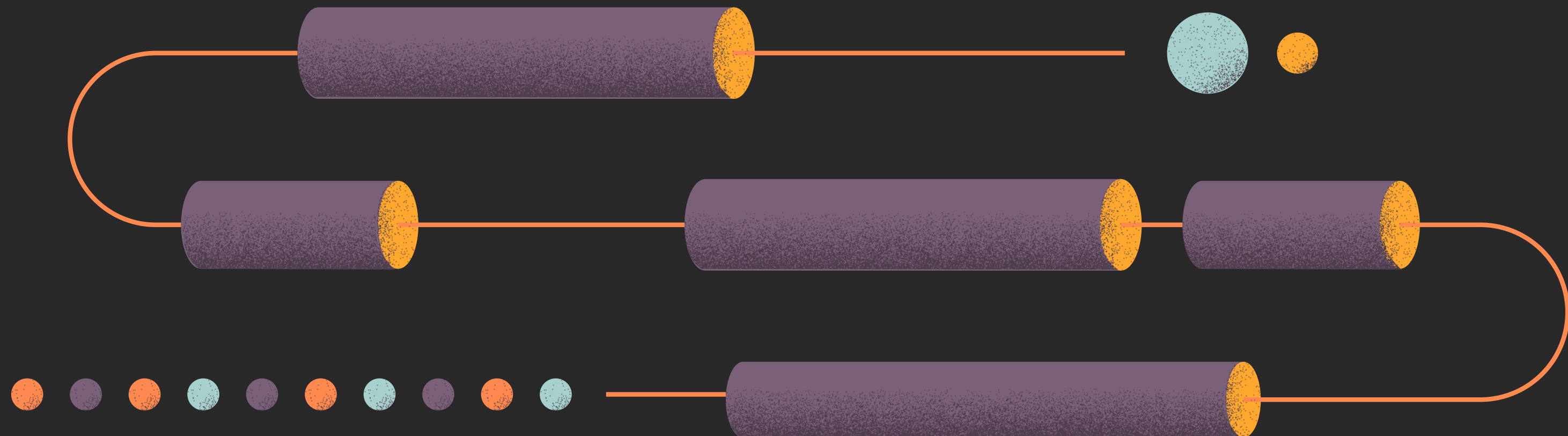


WELCOME EVERYONE

Welcome Message

Presenting to you a recommendation system using ETL, ML and Power BI analytics and visualization

GET STARTED





LIST OF CONTENT

01

Opening Slide

Presentations are communication tools that can be used as demonstrations.

04

About Us

Team information

05+

System Report

Project Details and explanation

12

End Slide

Ending note



ALL ABOUT THE TEAM



Muhammad Hamza

CT-19057

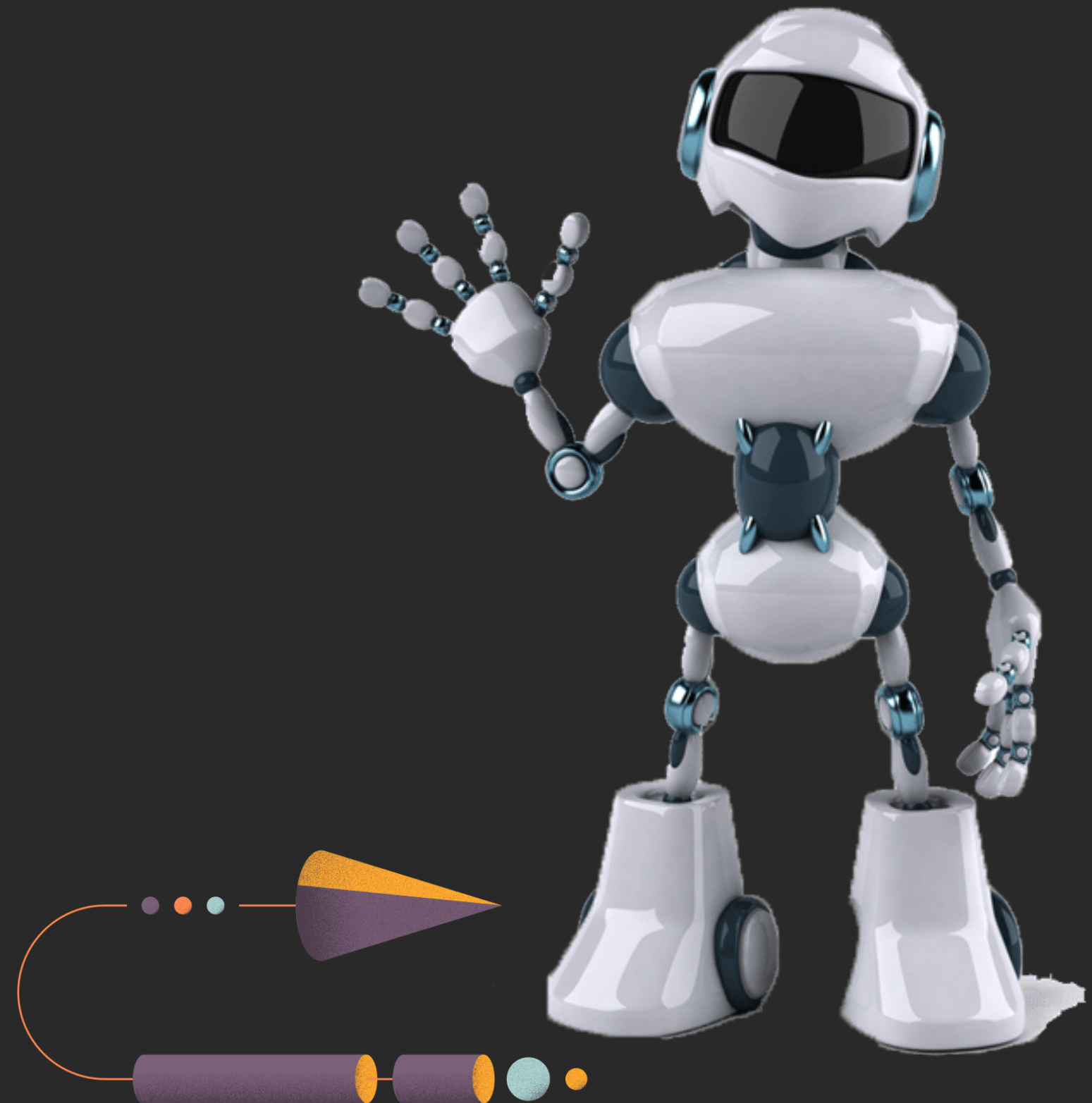


Fashad Ahmed Siddique

CT-19043

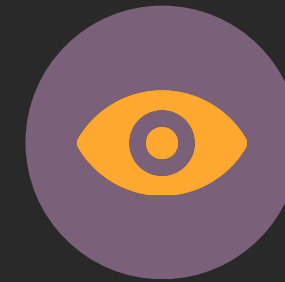
Why Are We Here ?

To present our Music Recommendation System using ETL (extract, transform, load), Power BI visualization and machine learning algorithms.





PROBLEM STATEMENT



Our Vision

Improve user engagement, increase revenue, and differentiate our music streaming service from the competition by providing a competent music recommendation system.

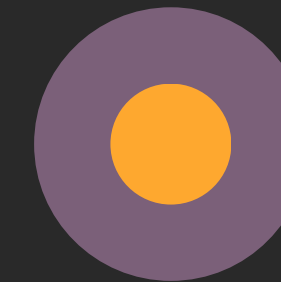


A problem well stated is a problem half solved.

John Dewey

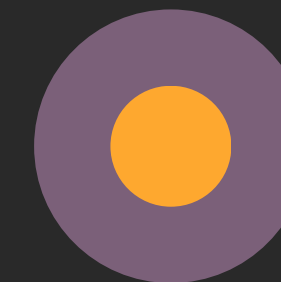


SYSTEM FLOW



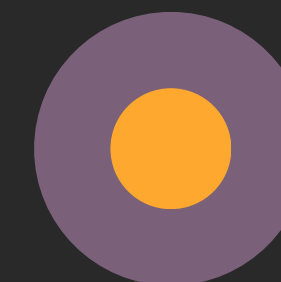
ETL PIPELINE with SPOTIFY API

Built an ETL pipeline that extract data directly from the spotify user.



Data Visualization with POWER BI

Using the loaded data from the pipeline to present useful analytics to the audience.



Recommendations using various ML algos

Using the pipeline data to infer useful features and then process them separately to produce computations to suggest similar song choices with respect to genres they listen to.



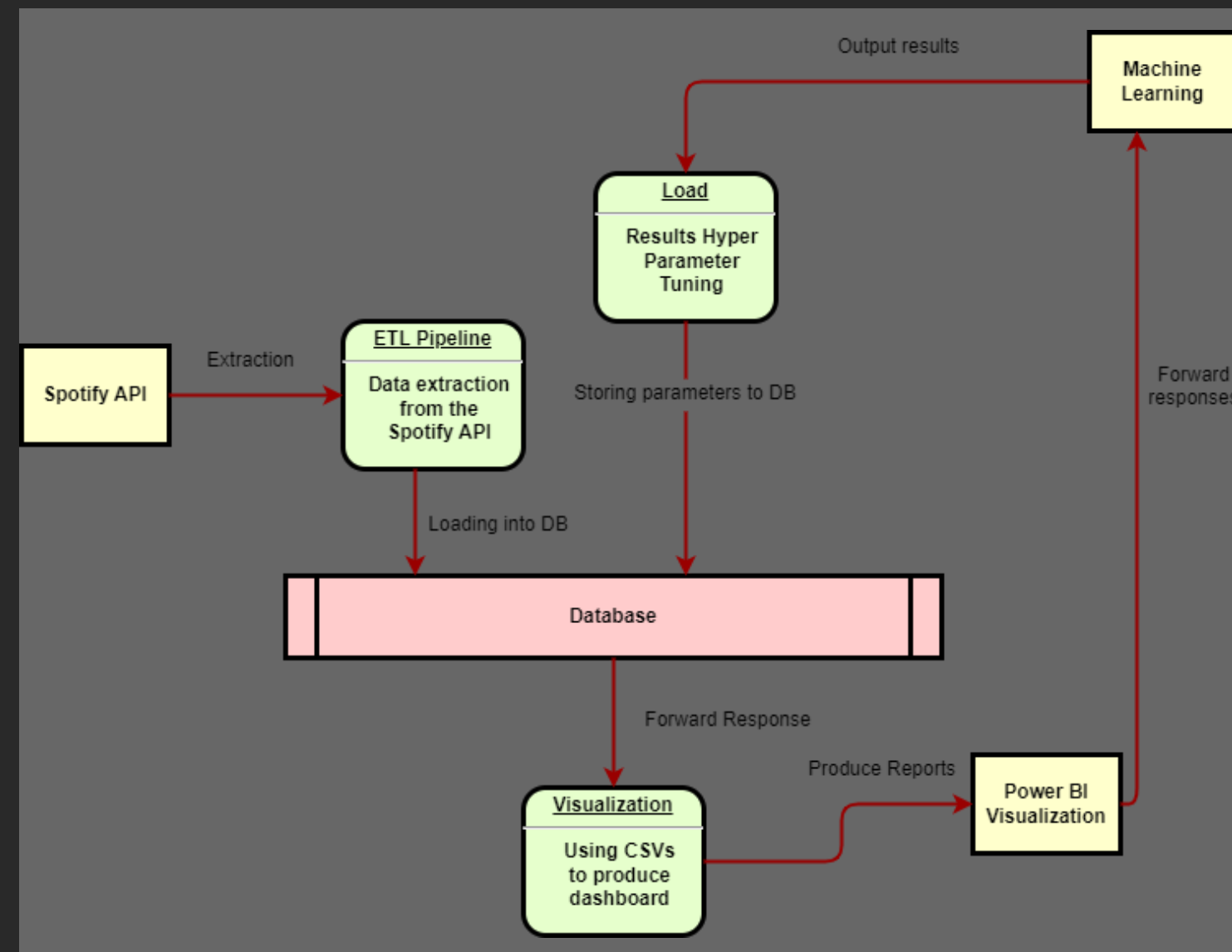
METHODOLOGY DIAGRAM

Step One

Using ETL pipeline to extract all the data from the source and pulling other data from secondary repositories

Step Four

Using the loaded data in the machine learning algorithms to compute the features and produce valuable results as a recommendation for the user



Step Two

Loading the data into the dataabase

Step Thre

Using the loaded data in a visualization tool to infer useful analytics and producing useful. insights



SOLUTION & CODE

ETL Code

```
HEADERS = {
    "Accept": "application/json",
    "Content-Type": "application/json",
    "Authorization": "Bearer {token}".format(token=TOKEN)
}

# Convert time to Unix timestamp in milliseconds
today = datetime.datetime.now()
yesterday = today - datetime.timedelta(days=99)
yesterday_unix_timestamp = int(yesterday.timestamp()) * 1000

req = requests.get("https://api.spotify.com/v1/me/player/recently-played?after={time}".format(time=yesterday_unix_timestamp), headers = headers)

data = req.json()

song_names = []
artist_names = []
played_at_list = []
timestamps = []

# Extracting only the relevant bits of data from json
for song in data["items"]:
    song_names.append(song["track"]["name"])
    artist_names.append(song["track"]["album"]["artists"][0]["name"])
    played_at_list.append(song["played_at"])
    timestamps.append(song["played_at"][0:10])

song_dict = {
    "song_name": song_names,
    "artist_name": artist_names,
    "played_at": played_at_list,
    "timestamp": timestamps
}
```

ML code

```
from sklearn.cluster import KMeans
from sklearn.preprocessing import StandardScaler
from sklearn.pipeline import Pipeline

cluster_pipeline = Pipeline([('scaler', StandardScaler()), ('kmeans', KMeans(n_clusters=10))])
X = genre_data.select_dtypes(np.number)
cluster_pipeline.fit(X)
genre_data['cluster'] = cluster_pipeline.predict(X)

# Visualizing the Clusters with t-SNE

from sklearn.manifold import TSNE

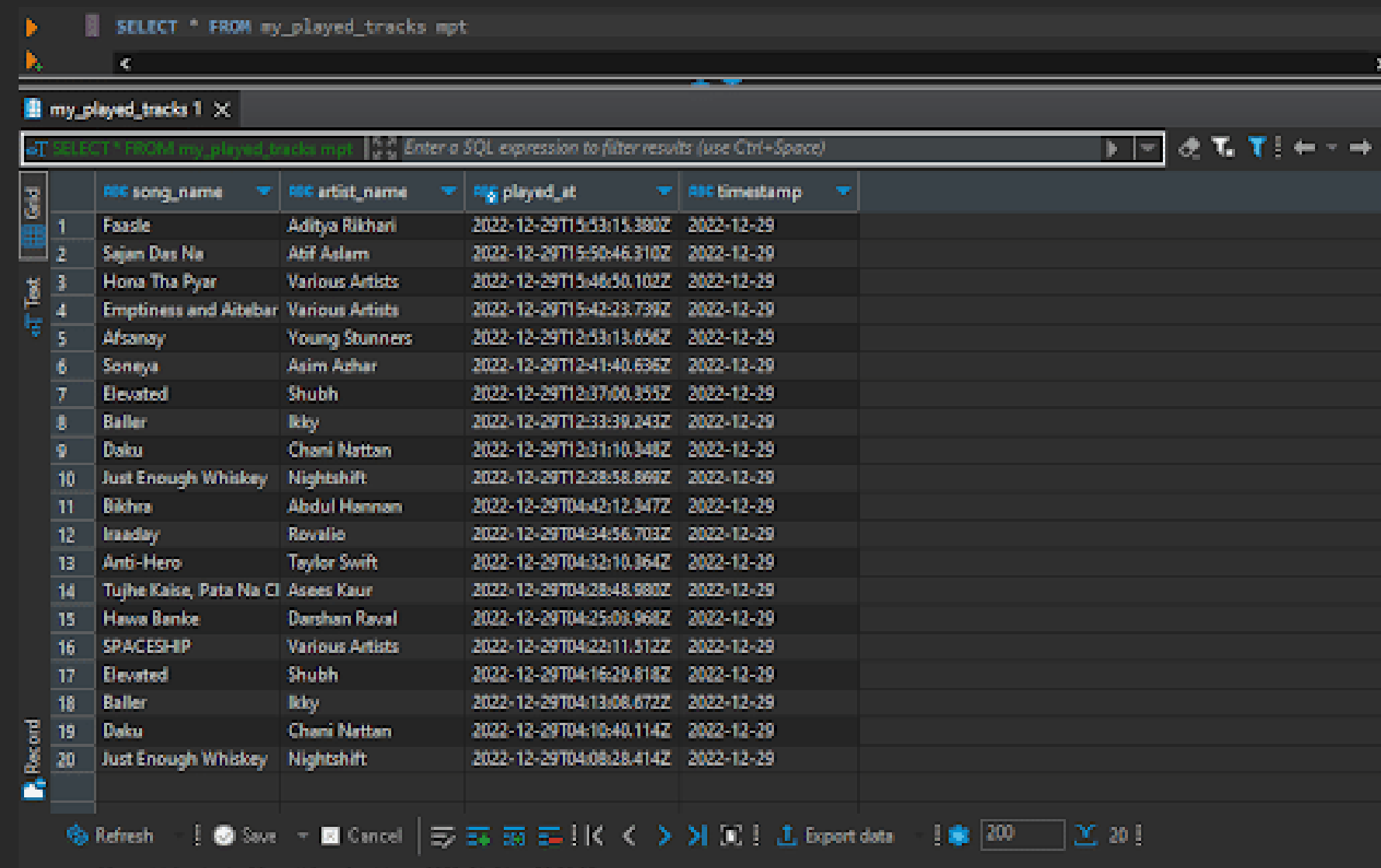
tsne_pipeline = Pipeline([('scaler', StandardScaler()), ('tsne', TSNE(n_components=2, verbose=1))])
genre_embedding = tsne_pipeline.fit_transform(X)
projection = pd.DataFrame(columns=['x', 'y'], data=genre_embedding)
projection['genres'] = genre_data['genres']
projection['cluster'] = genre_data['cluster']

fig = px.scatter(
    projection, x='x', y='y', color='cluster', hover_data=['x', 'y', 'genres'])
fig.show()
```



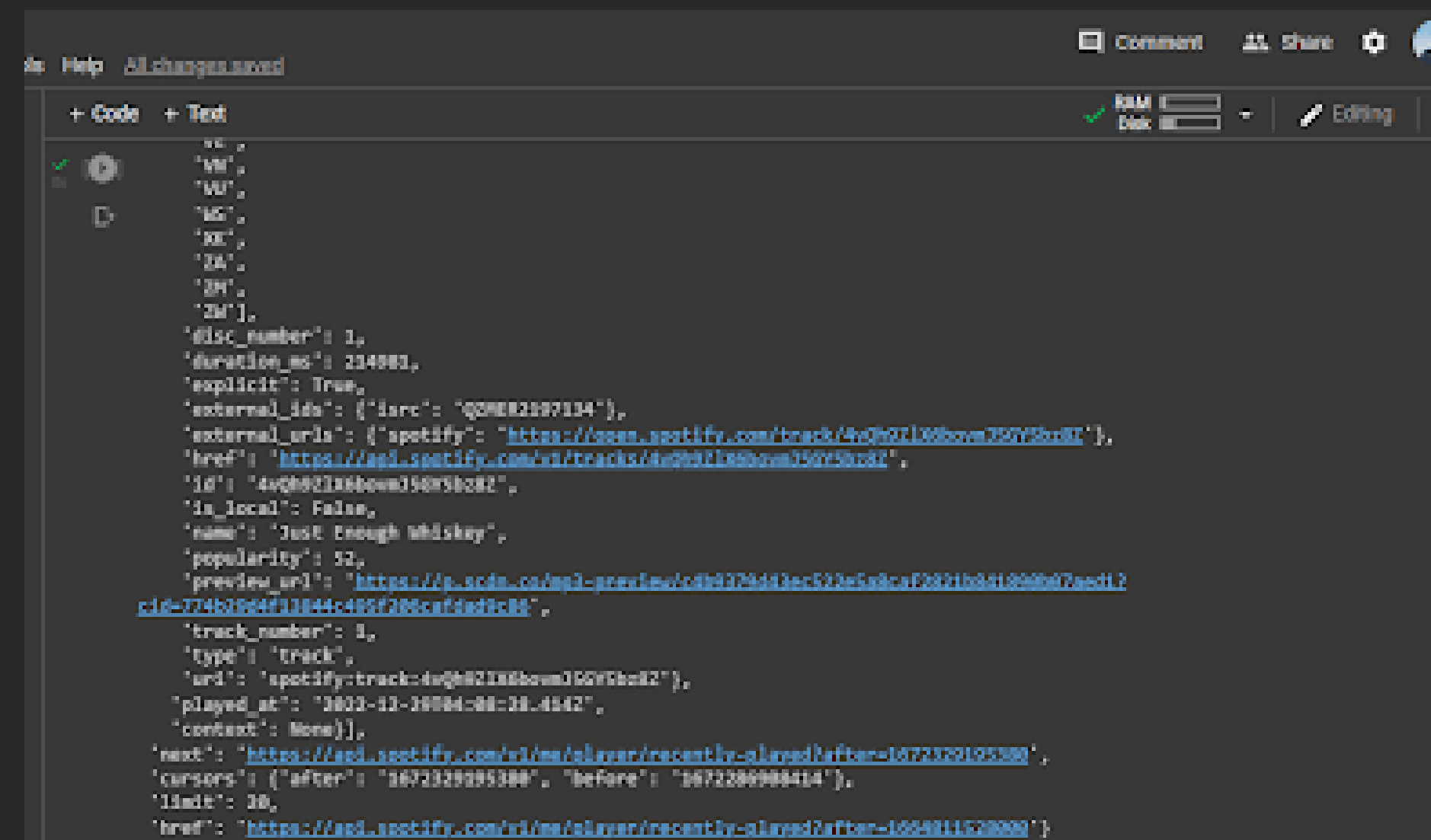

SOLUTION & CODE

Database view



	song_name	artist_name	played_at	timestamp
1	Faale	Aditya Rikhan	2022-12-29T15:53:15.380Z	2022-12-29
2	Sejan Das Na	A6f Aslam	2022-12-29T15:50:46.310Z	2022-12-29
3	Hona Tha Pyar	Various Artists	2022-12-29T15:46:30.102Z	2022-12-29
4	Emptiness and Aitebar	Various Artists	2022-12-29T15:42:23.750Z	2022-12-29
5	Afsanay	Young Stunners	2022-12-29T12:53:13.656Z	2022-12-29
6	Sonaya	Asim Ashar	2022-12-29T12:41:40.656Z	2022-12-29
7	Elevated	Shubh	2022-12-29T12:37:00.355Z	2022-12-29
8	Baller	Ikky	2022-12-29T12:33:39.243Z	2022-12-29
9	Daku	Chani Nattan	2022-12-29T12:31:10.348Z	2022-12-29
10	Just Enough Whiskey	Nightshift	2022-12-29T12:28:58.869Z	2022-12-29
11	Bikha	Abdul Hannan	2022-12-29T04:42:12.347Z	2022-12-29
12	Inaaday	Rovafio	2022-12-29T04:34:56.703Z	2022-12-29
13	Anti-Hero	Taylor Swift	2022-12-29T04:32:10.364Z	2022-12-29
14	Tujhe Kaisa, Pata Na Ci	Asees Kaur	2022-12-29T04:28:48.980Z	2022-12-29
15	Hawa Banke	Darshan Raval	2022-12-29T04:25:09.968Z	2022-12-29
16	SPACESHIP	Various Artists	2022-12-29T04:22:11.512Z	2022-12-29
17	Elevated	Shubh	2022-12-29T04:16:29.818Z	2022-12-29
18	Baller	Ikky	2022-12-29T04:13:08.672Z	2022-12-29
19	Daku	Chani Nattan	2022-12-29T04:10:40.114Z	2022-12-29
20	Just Enough Whiskey	Nightshift	2022-12-29T04:08:28.414Z	2022-12-29

Colab view



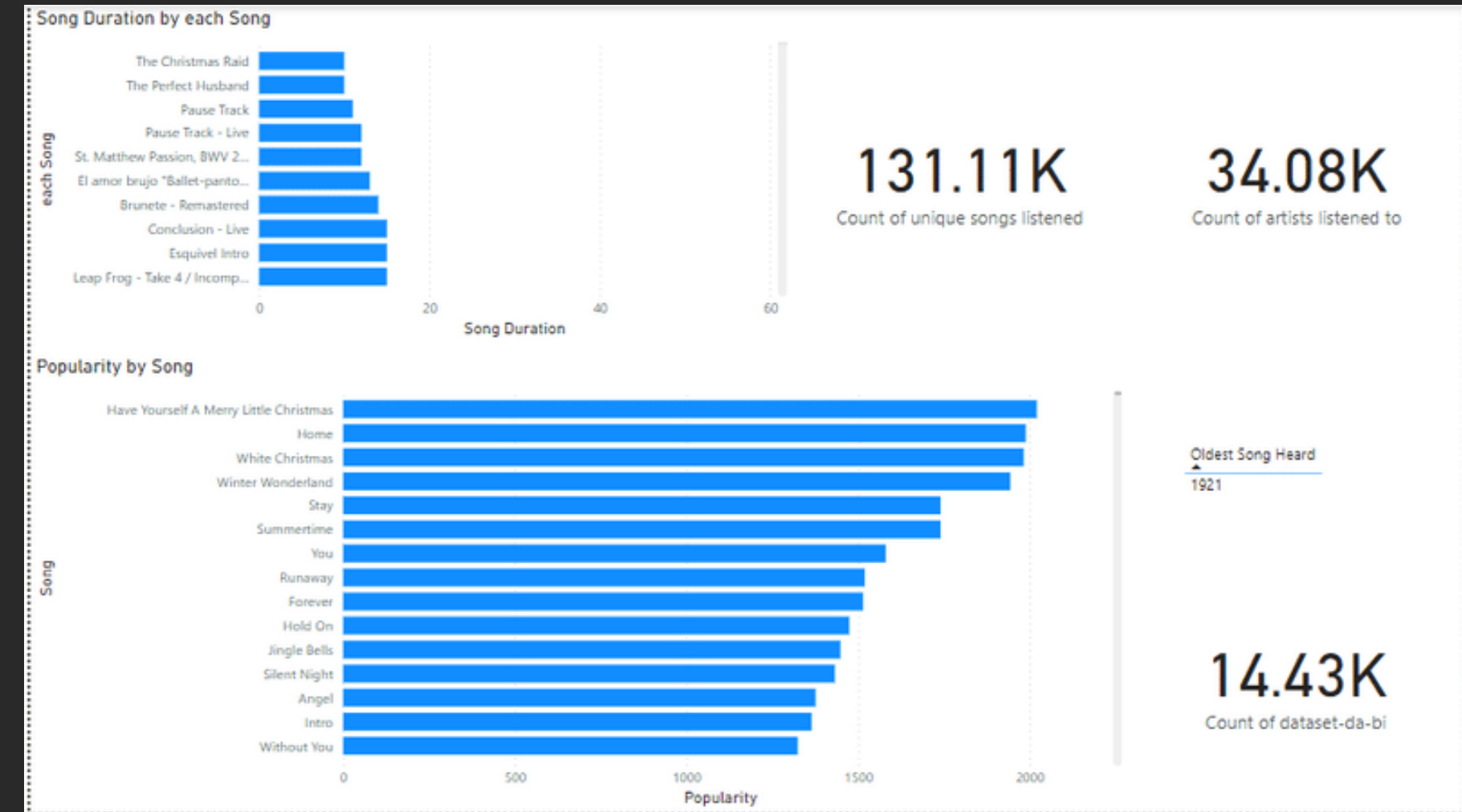
```
{
  "uri": "spotify:track:4xQh021X86owm3G0Y5b082",
  "track_number": 1,
  "type": "track",
  "uri": "spotify:track:4xQh021X86owm3G0Y5b082",
  "played_at": "2022-12-29T04:08:28.414Z",
  "context": {
    "next": "https://api.spotify.com/v1/me/player/recently-played?after=1672329195380",
    "cursors": {
      "after": "1672329195380",
      "before": "1672286988414"
    },
    "limit": 20,
    "href": "https://api.spotify.com/v1/me/player/recently-played?after=1669811528888"
  }
}
```

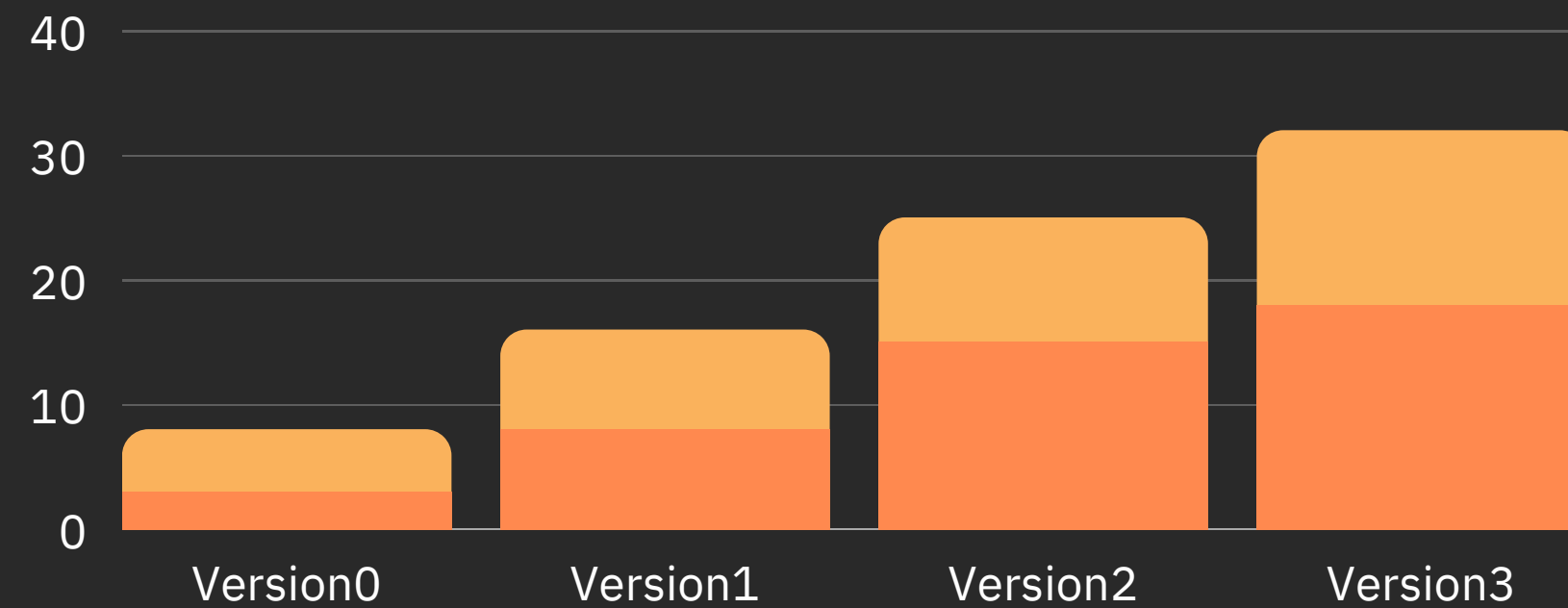
SOLUTION & CODE

Raw API view

```
{
  "items": [
    {
      "track": {
        "album": {
          "album_type": "single",
          "artists": [
            {
              "external_urls": {
                "spotify": "https://open.spotify.com/artist/3ozYqVCLohfpXIhalkhM8D"
              },
              "href": "https://api.spotify.com/v1/artists/3ozYqVCLohfpXIhalkhM8D",
              "id": "3ozYqVCLohfpXIhalkhM8D",
              "name": "Aditya Rikhari",
              "type": "artist",
              "uri": "spotify:artist:3ozYqVCLohfpXIhalkhM8D"
            }
          ]
        },
        "available_markets": [
          "US"
        ],
        "external_urls": {
          "spotify": "https://open.spotify.com/album/2xfAmgKBQroCYbDn118me3"
        },
        "href": "https://api.spotify.com/v1/albums/2xfAmgKBQroCYbDn118me3",
        "id": "2xfAmgKBQroCYbDn118me3",
        "name": "The Christmas Raid",
        "type": "album",
        "uri": "spotify:album:2xfAmgKBQroCYbDn118me3"
      },
      "artists": [
        {
          "external_urls": {
            "spotify": "https://open.spotify.com/artist/3ozYqVCLohfpXIhalkhM8D"
          },
          "href": "https://api.spotify.com/v1/artists/3ozYqVCLohfpXIhalkhM8D",
          "id": "3ozYqVCLohfpXIhalkhM8D",
          "name": "Aditya Rikhari",
          "type": "artist",
          "uri": "spotify:artist:3ozYqVCLohfpXIhalkhM8D"
        }
      ],
      "duration_ms": 180000,
      "explicit": false,
      "href": "https://api.spotify.com/v1/tracks/2xfAmgKBQroCYbDn118me3",
      "id": "2xfAmgKBQroCYbDn118me3",
      "name": "The Christmas Raid",
      "type": "track",
      "uri": "spotify:track:2xfAmgKBQroCYbDn118me3"
    }
  ],
  "next": null,
  "previous": null
}
```

Power BI Dashboard view





FUTURE WORK



Producing Better Versions

Coming up with better versions with the best methodologies to keep the system upto date and produce better results.



Enhancing Accuracy

As the usage increases, more and more data will be aded into the datasets ultimately helping increase the accuracy of the system.





**THANKS FOR
YOUR TIME**

