Hello Team!

Taylor Swift has become one of the biggest pop stars in the world. She recently began her ‘Eras Tour’, her first tour since 2018! As a result, Taylor Swift and her music has become a frequent topic in the office as there are many Swifties employed here. Some of these conversations have resulted in heated and lengthy debates surrounding Taylor Swift's music and which songs and albums are most popular. I am hoping you could use this Spotify dataset of songs from Taylor Swift’s studio 9 albums to answer the questions below and help settle some of these office debates.

1. During what time span, did Taylor’s songs reached peak popularity? When did her songs reach lowest popularity?
2. Songs like All Too Well (the 10 minute version) are quite popular with many of her fans. Many of the office Swifties love Tayor Swift because of her song story-telling abilities; this often results in songs which are longer in length. Please describe the relationship between song duration and popularity.
3. Some critics have argued that Taylor Swift’s popularity has increased since her transition from the country to pop music genre, does the data support this? There is a divide among office Swifties about which genre resulted in more popular music. Is Taylor’s most popular music from a the country or pop (or alternative) genre?
4. Overall, what attribute(s) makes for the most popular and least popular Taylor Swift music?

I look forward to hearing back from you and seeing what you come up with!
Best,

Your Database Team Leader

Question 1: During what time span (preferably years), did Taylor's songs reached peak popularity? When did her songs reach lowest popularity?

```
SELECT YEAR(a.date) AS year, MAX(s.popularity) AS max_popularity,
AVG(s.popularity) AS mean_popularity
FROM songs s
LEFT JOIN albums a on s.release_date_id = a.id
GROUP BY YEAR(a.date);
```

	year	max_popularity	mean_popularity
1	2006	59	50.1333
2	2010	64	49.7273
3	2012	72	60.5000
4	2014	82	54.4211
5	2017	78	71.8667
6	2019	80	72.1111
7	2020	65	62.6471
8	2021	76	65.5349

Taylor Swift's music reached the peak of its popularity during the year 2019 when her average popularity index reached a value of 72.1111, while the lowest point in time for her music was during the year 2006 when the average popularity index of her music was 50.1333. This is not surprising as 2006 was the very beginning of Taylor's career. Interestingly, the maximum popularity index for one of her songs is 82, which was during the year 2014. This song could potentially be an outlier, because although it is her most popular, the overall popularity of her music during this year is less popular by comparison.

Question 2: Songs like 'All Too Well (the 10 minute version)' are quite popular with many of her fans. Many of the office Swifties love Tayor Swift because of her song story-telling abilities; this often results in songs which are longer in length. Please describe the relationship between song duration and popularity.

```
SELECT name AS song_name, popularity,
      (length/60000) AS length_in_minutes
FROM songs
                        ORDER BY popularity LIMIT 10;
```

	song_name	popularity	length_in_minutes
1	I Wish You Would - Voice Memo	0	1.7856
2	Blank Space - Voice Memo	0	2.1864
3	I Know Places - Voice Memo	0	3.6056
4	Back To December	43	4.8840
5	Never Grow Up	44	4.8413
6	Innocent	44	5.0378
7	Mine - POP Mix	45	3.8424
8	A Perfectly Good Heart	46	3.6691
9	Long Live	47	5.2993
10	Tied Together with a Smile	47	4.1351

```sql
SELECT name AS song_name, popularity, (length/60000) AS length_in_minutes
FROM songs
ORDER BY popularity DESC LIMIT 10;
```

	song_name	popularity	length_in_minutes
1	Blank Space	82	3.8638
2	Lover	80	3.6884
3	Shake It Off	80	3.6533
4	You Need To Calm Down	78	2.8560
5	Delicate	78	3.8709
6	Look What You Made Me Do	77	3.5309
7	Cruel Summer	77	2.9738
8	ME! (feat. Brendon Urie of Panic! At The Disco)	77	3.2167
9	Paper Rings	76	3.7067
10	Getaway Car	76	3.8938

```sql
SELECT a.name, AVG(s.popularity) AS mean_popularity, SUM((s.length/60000)) AS
total_minutes
FROM songs s LEFT JOIN albums a on a.id = s.album_id
GROUP BY a.name;
```

	name	mean_popularity	total_minutes
1	Taylor Swift	50.1333	53.4928
2	Speak Now (Deluxe Package)	49.7273	101.1889
3	Red (Deluxe Edition)	60.5000	90.6748
4	1989 (Deluxe)	54.4211	68.7610
5	reputation	71.8667	55.7551
6	Lover	72.1111	61.8566
7	folklore (deluxe version)	62.6471	67.1401
8	evermore (deluxe version)	65.4706	69.0815
9	Fearless (Taylor's Version)	65.5769	106.5419

It seems like the (top 10) most popular Taylor Swift songs tend to have a song length of approximately 3.5 - 3.8 minutes. The (top 10) least popular Taylor songs seems to have song a longer duration in general, approximately 3.8 - 4.8 minutes long. However, it is worth noting that the least popular songs have a much larger range (lengths from 1.7 - 5.2) while the range of the most popular songs has less variability (lengths from 2.8 - 3.89).
Overall, songs that are most popular also are moderate in length. Songs that are least popular have a wider range of duration, but tend to be longer in length.

Question 3: Some critics have argued that Taylor Swift's popularity has increased since her transition from the country to pop music genre, does the data support this? There is a divide among office Swifties about which genre resulted in more popular music. Is Taylor's most popular music from a the country or pop (or alternative) genre?

```
SELECT YEAR(a.date) AS year, MAX(s.popularity) AS max_popularity,
AVG(s.popularity) AS mean_popularity
FROM songs s LEFT JOIN albums a on s.release_date_id = a.id
GROUP BY YEAR(a.date);
```

	year	max_popularity	mean_popularity
1	2006	59	50.1333
2	2010	64	49.7273
3	2012	72	60.5000
4	2014	82	54.4211
5	2017	78	71.8667
6	2019	80	72.1111
7	2020	65	62.6471
8	2021	76	65.5349

```
SELECT g.name, MAX(s.popularity) AS max_popularity, AVG(s.popularity) AS
mean_popularity
FROM songs s LEFT JOIN genres g on s.genre_id = g.id
GROUP BY g.name;
```

	name	max_popularity	mean_popularity
1	country	76	57.4353
2	pop	82	65.5769
3	alternative	72	64.0588

Taylor's first pop album ('1989') was released in the year 2014 after her first 3 country albums. From 2012 to 2014 Taylor's average average popularity index decreased from approximately 60.5 to 54.4. Although, Taylor released her most popular song, with a rating of 82, during this year (2014).

However, after Taylor released two additional albums under the pop genre ('reputation' and 'Lover'), she reached her highest average popularity of 72.1 during the year 2019.

In my opinion it seems like Taylor Swift's music took a small dip in popularity when she transitioned from country to pop, but as she continued to release more pop music, her music reached the height of its popularity.

Overall, it seems that the most popular music Taylor has released is from the pop, alternative and country genre in descending order based on the mean popularity of songs from each genre.

Question 4: Overall, what attribute(s) makes for the most popular and least popular Taylor Swift music?

```
SELECT  s.name as song_name, s.popularity, g.name AS genre FROM songs s
LEFT JOIN genres g on s.genre_id = g.id
ORDER BY s.popularity DESC LIMIT 10;
```

	song_name	popularity	genre
1	Blank Space	82	pop
2	Lover	80	pop
3	Shake It Off	80	pop
4	You Need To Calm Down	78	pop
5	Delicate	78	pop
6	Look What You Made Me Do	77	pop
7	Cruel Summer	77	pop
8	ME! (feat. Brendon Urie of Panic! At The Disco)	77	pop
9	Paper Rings	76	pop
10	Getaway Car	76	pop

```
SELECT  s.name as song_name, s.popularity, g.name AS genre FROM songs s
LEFT JOIN genres g on s.genre_id = g.id
ORDER BY s.popularity LIMIT 10;
```

	song_name	popularity	genre
1	I Wish You Would - Voice Memo	0	pop
2	Blank Space - Voice Memo	0	pop
3	I Know Places - Voice Memo	0	pop
4	Back To December	43	country
5	Never Grow Up	44	country
6	Innocent	44	country
7	Mine - POP Mix	45	country
8	A Perfectly Good Heart	46	country
9	Long Live	47	country
10	Tied Together with a Smile	47	country

```
SELECT  s.name as song_name, s.popularity, a.name AS album FROM songs s
LEFT JOIN albums a on s.album_id = a.id
ORDER BY s.popularity DESC LIMIT 10;
```

	song_name	popularity	album
1	Blank Space	82	1989 (Deluxe)
2	Lover	80	Lover
3	Shake It Off	80	1989 (Deluxe)
4	You Need To Calm Down	78	Lover
5	Delicate	78	reputation
6	Look What You Made Me Do	77	reputation
7	Cruel Summer	77	Lover
8	ME! (feat. Brendon Urie of Panic! At The Disco)	77	Lover
9	Paper Rings	76	Lover
10	Getaway Car	76	reputation

```sql
SELECT  s.name as song_name, s.popularity, a.name AS album FROM songs s
LEFT JOIN albums a on s.album_id = a.id
ORDER BY s.popularity LIMIT 10;
```

	song_name	popularity	album
1	I Wish You Would - Voice Memo	0	1989 (Deluxe)
2	Blank Space - Voice Memo	0	1989 (Deluxe)
3	I Know Places - Voice Memo	0	1989 (Deluxe)
4	Back To December	43	Speak Now (Deluxe Package)
5	Never Grow Up	44	Speak Now (Deluxe Package)
6	Innocent	44	Speak Now (Deluxe Package)
7	Mine - POP Mix	45	Speak Now (Deluxe Package)
8	A Perfectly Good Heart	46	Taylor Swift
9	Long Live	47	Speak Now (Deluxe Package)
10	Tied Together with a Smile	47	Taylor Swift

```sql
SELECT name AS song_name, popularity, (length/60000) AS minutes, danceability,
acousticness, energy, tempo
FROM songs
ORDER BY popularity DESC LIMIT 10;
```

	song_name	popularity	minutes	danceability	acousticness	energy	tempo
1	Blank Space	82	3.8638	0.76	0.103	0.703	95.997
2	Lover	80	3.6884	0.359	0.492	0.543	68.534
3	Shake It Off	80	3.6533	0.647	0.0647	0.8	160.078
4	You Need To Calm Down	78	2.8560	0.771	0.00929	0.671	85.026
5	Delicate	78	3.8709	0.75	0.216	0.404	95.045
6	Look What You Made Me Do	77	3.5309	0.766	0.204	0.709	128.07
7	Cruel Summer	77	2.9738	0.552	0.117	0.702	169.994
8	ME! (feat. Brendon Urie of Panic! At The Disco)	77	3.2167	0.61	0.033	0.83	182.162
9	Paper Rings	76	3.7067	0.811	0.0129	0.719	103.979
10	Getaway Car	76	3.8938	0.562	0.00465	0.689	172.054

```sql
SELECT name AS song_name, popularity, (length/60000) AS minutes, danceability,
acousticness, energy, tempo
FROM songs
ORDER BY popularity LIMIT 10;
```

	song_name	popularity	minutes	danceability	acousticness	energy	tempo
1	I Wish You Would - Voice Memo	0	1.7856	0.781	0.717	0.357	118.317
2	Blank Space - Voice Memo	0	2.1864	0.675	0.801	0.234	127.296
3	I Know Places - Voice Memo	0	3.6056	0.592	0.829	0.128	78.828
4	Back To December	43	4.8840	0.525	0.113	0.676	141.95
5	Never Grow Up	44	4.8413	0.715	0.829	0.308	124.899
6	Innocent	44	5.0378	0.553	0.202	0.604	133.989
7	Mine - POP Mix	45	3.8424	0.696	0.00461	0.768	121.05
8	A Perfectly Good Heart	46	3.6691	0.483	0.00349	0.751	156.092
9	Long Live	47	5.2993	0.412	0.0426	0.682	203.959
10	Tied Together with a Smile	47	4.1351	0.479	0.525	0.578	146.165

```sql
SELECT a.name AS album, AVG(s.popularity) AS mean_popularity,
ROUND(AVG((s.length/60000)), 4) AS mean_song_length_in_minutes,
     ROUND(AVG(s.danceability), 4) AS mean_danceability,
ROUND(AVG(s.acousticness), 4) AS mean_acousticness
FROM songs s
LEFT JOIN albums a on s.album_id = a.id
GROUP BY a.name
ORDER BY AVG(s.popularity) DESC LIMIT 10;
```

	album	mean_popularity	mean_song_length_in_minutes	mean_danceability	mean_acousticness
1	Lover	72.1111	3.4365	0.6582	0.3337
2	reputation	71.8667	3.7170	0.6579	0.1385
3	Fearless (Taylor's Version)	65.5769	4.0978	0.551	0.2141
4	evermore (deluxe version)	65.4706	4.0636	0.5268	0.7941
5	folklore (deluxe version)	62.6471	3.9494	0.5419	0.7176
6	Red (Deluxe Edition)	60.5000	4.1216	0.6334	0.1488
7	1989 (Deluxe)	54.4211	3.6190	0.6332	0.2446
8	Taylor Swift	50.1333	3.5662	0.5453	0.183
9	Speak Now (Deluxe Package)	49.7273	4.5995	0.559	0.2265

```sql
SELECT a.name AS album, AVG(s.popularity) AS mean_popularity,
ROUND(AVG(s.energy), 4) AS mean_energy,
      ROUND(AVG(s.tempo), 4) AS mean_tempo
FROM songs s
LEFT JOIN albums a on s.album_id = a.id
GROUP BY a.name
ORDER BY AVG(s.popularity) DESC LIMIT 10;
```

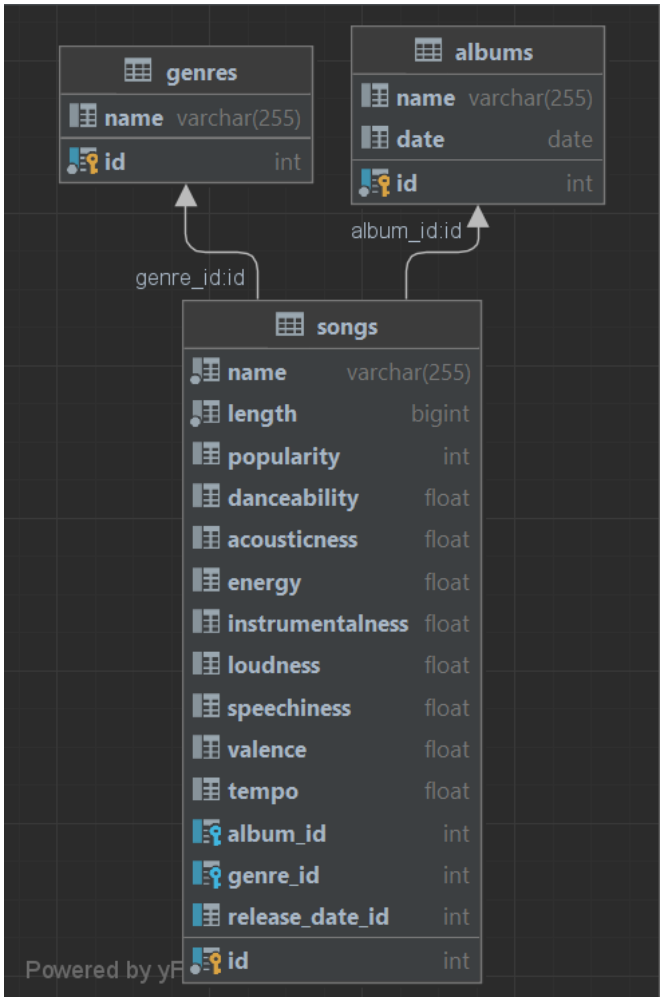
	album	mean_popularity	mean_energy	mean_tempo
1	Lover	72.1111	0.5452	119.9727
2	reputation	71.8667	0.5829	127.5401
3	Fearless (Taylor's Version)	65.5769	0.6391	131.2372
4	evermore (deluxe version)	65.4706	0.4941	120.7073
5	folklore (deluxe version)	62.6471	0.4158	119.8844
6	Red (Deluxe Edition)	60.5000	0.6008	110.2965
7	1989 (Deluxe)	54.4211	0.6248	127.0331
8	Taylor Swift	50.1333	0.6643	126.0538
9	Speak Now (Deluxe Package)	49.7273	0.6594	132.8357

In general, the most popular songs and albums from Taylor Swift were released under the pop genre, specifically from the ‘Lover’ album. The least popular songs and albums were released under the country genre, specifically from the ‘Speak Now’ album. Additionally, the majority of popular songs seem to have a slower tempo (approximately 95 -120 beats per minute) compared to least popular songs (which have a tempo of approximately 120-150 beats per minute).

Otherwise, there doesn’t seem to be a conclusive trend which separates the most and least popular songs as there is either high variability or virtually no variability among the other attributes when compared across the high and low ends of the popularity index.

In my database I decided to separate the genre, album release date, and album name from the rest of the attributes because this information was very repetitive and I wanted to normalize the tables as much as possible and minimize data redundancy. All continuous decimal filed were made to be float data types and all strings were made to be varchar data types.

ERD Diagram



albums Table

Field Name	Logical Field Name	Data Type	Key
id	Identification Number	INT	PRI
name	Album Name	VARCHAR	

genres Table

Field Name	Logical Field Name	Data Type	Key
id	Identification Number	INT	PRI
name	Genre Name	VARCHAR	

songs Table

Field Name	Logical Field Name	Data Type	Key
id	Unique Identification Number	INT	PRI

name	Song Name	VARCHAR	
length	Length in Miliseconds	BIGINT	
popularity	Popularity Index (0-100)	INT	
danceability	Danceability Index (0-1)	FLOAT	
acousticness	Acousticness (0-1)	FLOAT	
energy	Energy (0-1)	FLOAT	
instrumentalness	Instrumentalness (0-1)	FLOAT	
loudness	Loudness	FLOAT	
spechiness	Spechiness (0-1)	FLOAT	
valence	Valence (0-1)	FLOAT	
tempo	Tempo	FLOAT	
album_id	Album Identification Number	INT	MUL
genre_id	Genre Identification Number	INT	MUL
release_date_id	Release Date Identification Number	INT	MUL