**CS210 - Term Project Report**

19/01/2024

**Objective**:

Exploring how life events influence my music preferences. By pairing diary entries with my Spotify history, I'll analyze shifts in music taste before and after key events.

**Hypothesis**:

The events occurred in my life effected my music preferences.

**Methodology**:

**Data Collection and Cleaning**:

* **Diary Entries**:
  + I determined some specific dates that I thought can be influencial for my life during 2023.
* **Streaming History**:
  + The collection is done by data request from Spotify. It includes data from the beginning of 2023 until the date the data was requested. It is cleaned to avoid any biases and distruptive data. Songs streamed less than 60 seconds, songs streamed one time and one day, songs streamed between 00.00 and 07.00 excluded.

**Data Analysis**:

* **Analytical Approach**:
  + **Tool Selection**: I chose to use Dash and Plotly for data visualization due to their interactive capabilities and the dynamic nature of the graphs they produce. This allowed for a more in-depth and user-friendly exploration of the data.
* **Data Visualization**:
  + Implementation with Dash and Plotly:
    - **Data Preparation**: Before visualization, the data from Spotify and diary entries were formatted and merged into a coherent structure using pandas, a Python data analysis library.
    - **Interactive Dashboards**: Using Dash, I developed interactive web-based dashboards that enabled me to filter and view the data according to different parameters, such as time periods, music genres, and specific life events.
    - **Graph Types**: Several types of graphs were created using Plotly, including:
      * **Time Series Graphs**: To display changes in music preferences over time.
      * **Bar Charts**: To compare the frequency of different genres or artists before and after key life events.
      * **Scatter Plots**: To explore any correlations between the moods or themes of the music and the nature of the life events.
* **Statistical Analysis**:
  + **Correlation Analysis**: I conducted a correlation analysis to determine if there was a statistically significant relationship between life events and changes in music preferences.

**Results:**

* Graph of number of songs listened each day:

A graph showing a graph

Description automatically generated with medium confidence

**A graph with a dotted line

Description automatically generatedA graph showing a line

Description automatically generated with medium confidenceFebruary 27:** Breakup

Weekly Average of Valence

Weekly Average of Tempo

**A graph with a dotted line

Description automatically generated**A graph with a dotted line

Description automatically generated

Weekly Average of Energy

Weekly Average of Loudness

A graph with a dotted line

Description automatically generated

Weekly Average of Speechiness

Valence: A measure from 0.0 to 1.0 describing the musical positiveness conveyed by a track. Tracks with high valence sound more positive (e.g. happy, cheerful, euphoric), while tracks with low valence sound more negative (e.g. sad, depressed, angry).

According to the hypothesis test for Valence with %95 confidence:

* P-value for comparison between Target Date and One Week Before it: 1.004e-07
* P-value for comparison between Target Date and One Week After it: 1.052e-07
* Significant difference between Target Date and One Week Before it
* Significant difference between Target Date and One Week After it

According to the hypothesis test for Tempo with %95 confidence:

Tempo: The overall estimated tempo of a track in beats per minute (BPM). In musical terminology, tempo is the speed or pace of a given piece and derives directly from the average beat duration.

* P-value for comparison between Target Date and One Week Before it: 0.073
* P-value for comparison between Target Date and One Week After it: 0.23
* No significant difference between Target Date and One Week Before it
* No significant difference between Target Date and One Week After it

According to the hypothesis test for Energy with %95 confidence:

Energy: Energy is a measure from 0.0 to 1.0 and represents a perceptual measure of intensity and activity. Typically, energetic tracks feel fast, loud, and noisy.

* P-value for comparison between Target Date and One Week Before it: 0.617
* P-value for comparison between Target Date and One Week After it: 0.0001
* No significant difference between Target Date and One Week Before it
* Significant difference between Target Date and One Week After it

According to the hypothesis test for Loudness with %95 confidence:

Loudness: The overall loudness of a track in decibels (dB). Loudness is the quality of a sound that is the primary psychological correlate of physical strength (amplitude).

* P-value for comparison between Target Date and One Week Before it: 0.471
* P-value for comparison between Target Date and One Week After it: 2.488e-08
* No significant difference between Target Date and One Week Before it
* Significant difference between Target Date and One Week After it

According to the hypothesis test for Speechiness with %95 confidence:

Speechiness: Speechiness detects the presence of spoken words in a track. The more exclusively speech-like the recording the closer to 1.0 the attribute value.

* P-value for comparison between Target Date and One Week Before it: 3.414e-07
* P-value for comparison between Target Date and One Week After it: 0.035
* Significant difference between Target Date and One Week Before it
* Significant difference between Target Date and One Week After it

**May 21:** Elections

A graph showing a line

Description automatically generated with medium confidence

Weekly Average of Valence

A graph with a dotted line

Description automatically generated

Weekly Average of Tempo

A graph with a dotted line

Description automatically generated

Weekly Average of Energy

A graph showing a line

Description automatically generated with medium confidenceA graph with a dotted line

Description automatically generated

Weekly Average of Speechiness

Weekly Average of Loudness

According to the hypothesis test for Valence with %95 confidence:

* P-value for comparison between Target Date and One Week Before it: 0.005
* P-value for comparison between Target Date and One Week After it: 0.114
* Significant difference between Target Date and One Week Before it
* No significant difference between Target Date and One Week After it

According to the hypothesis test for Tempo with %95 confidence:

* P-value for comparison between Target Date and One Week Before it: 0.0003
* P-value for comparison between Target Date and One Week After it: 0.028
* Significant difference between Target Date and One Week Before it
* Significant difference between Target Date and One Week After it

According to the hypothesis test for Energy with %95 confidence:

* P-value for comparison between Target Date and One Week Before it: 0.00018
* P-value for comparison between Target Date and One Week After it: 0.0041
* Significant difference between Target Date and One Week Before it
* Significant difference between Target Date and One Week After it

According to the hypothesis test for Loudness with %95 confidence:

* P-value for comparison between Target Date and One Week Before it: 0.017
* P-value for comparison between Target Date and One Week After it: 0.011
* Significant difference between Target Date and One Week Before it
* Significant difference between Target Date and One Week After it

According to the hypothesis test for Speechiness with %95 confidence:

* P-value for comparison between Target Date and One Week Before it: 0.112
* P-value for comparison between Target Date and One Week After it: 0.014
* No significant difference between Target Date and One Week Before it
* Significant difference between Target Date and One Week After it

**August 19:** Rowing Races

A graph showing a line

Description automatically generated with medium confidenceA graph with a dotted line

Description automatically generatedA blue line on a white background

Description automatically generatedA graph showing a curve

Description automatically generated with medium confidenceA graph with a dotted line

Description automatically generated

Weekly Average of Loudness

Weekly Average of Energy

Weekly Average of Valence

Weekly Average of Tempo

Weekly Average of Speechiness

According to the hypothesis test for Valence with %95 confidence:

* P-value for comparison between Target Date and One Week Before it: 0.0005
* P-value for comparison between Target Date and One Week After it: 0.1404
* Significant difference between Target Date and One Week Before it
* No significant difference between Target Date and One Week After it

According to the hypothesis test for Tempo with %95 confidence:

* P-value for comparison between Target Date and One Week Before it: 0.7832
* P-value for comparison between Target Date and One Week After it: 0.0067
* No significant difference between Target Date and One Week Before it
* Significant difference between Target Date and One Week After it

According to the hypothesis test for Energy with %95 confidence:

* P-value for comparison between Target Date and One Week Before it: 1.478e-08
* P-value for comparison between Target Date and One Week After it: 0.0021
* Significant difference between Target Date and One Week Before it
* Significant difference between Target Date and One Week After it

According to the hypothesis test for Loudness with %95 confidence:

* P-value for comparison between Target Date and One Week Before it: 2.104e-07
* P-value for comparison between Target Date and One Week After it: 0.509
* Significant difference between Target Date and One Week Before it
* No significant difference between Target Date and One Week After it

According to the hypothesis test for Speechiness with %95 confidence:

* P-value for comparison between Target Date and One Week Before it: 1.27e-09
* P-value for comparison between Target Date and One Week After it: 0.0484
* Significant difference between Target Date and One Week Before it
* Significant difference between Target Date and One Week After it

**October 2:** Return to the University

A graph with a dotted line

Description automatically generatedA graph showing a line

Description automatically generated with medium confidenceA graph showing a line

Description automatically generated with medium confidenceA blue dotted line on a white background

Description automatically generatedA graph showing a line

Description automatically generated with medium confidence

According to the hypothesis test for Valence with %95 confidence:

* P-value for comparison between Target Date and One Week Before it: 0.763
* P-value for comparison between Target Date and One Week After it: 0.251
* No significant difference between Target Date and One Week Before it
* No significant difference between Target Date and One Week After it

According to the hypothesis test for Tempo with %95 confidence:

* P-value for comparison between Target Date and One Week Before it: 0.665
* P-value for comparison between Target Date and One Week After it: 0.475
* No significant difference between Target Date and One Week Before it
* No significant difference between Target Date and One Week After it

According to the hypothesis test for Energy with %95 confidence:

* P-value for comparison between Target Date and One Week Before it: 0.021
* P-value for comparison between Target Date and One Week After it: 0.037
* Significant difference between Target Date and One Week Before it
* Significant difference between Target Date and One Week After it

According to the hypothesis test for Loudness with %95 confidence:

* P-value for comparison between Target Date and One Week Before it: 0.192
* P-value for comparison between Target Date and One Week After it: 0.057
* No significant difference between Target Date and One Week Before it
* No significant difference between Target Date and One Week After it

According to the hypothesis test for Speechiness with %95 confidence:

* P-value for comparison between Target Date and One Week Before it: 0.001
* P-value for comparison between Target Date and One Week After it: 0.0669
* Significant difference between Target Date and One Week Before it
* No significant difference between Target Date and One Week After it

**Discussion:**

**Interpretation of Results:**

* **Impact of Events:**
  + The analysis revealed a significant shift in music preferences coinciding with specific events like elections and rowing races.
  + Contrary to expectations, other events such as a personal breakup and returning to university did not show a notable impact on music preferences. This could suggest that not all life events have a uniform effect on music choices or that the influence of such events is more nuanced and less directly observable.

**Limitations:**

* **Lyrics Analysis:**
  + One of the primary limitations of this study is the focus on numerical data (like genre or artist frequency) without considering the lyrical content of the music. Given that my preferences are heavily influenced by lyrics, this omission might have led to an incomplete understanding of the true nature of the shifts in music tastes.
* **Temporal Scope:**
  + The study's restriction to a single year (2023) also presents a limitation. Music preferences can evolve over longer periods, and a multi-year analysis might provide a more comprehensive and reliable understanding of the relationship between life events and music preferences. The one-year scope may not capture the full extent of gradual shifts in taste or delayed reactions to life events.

**Conclusion:**

This study contributes to the academic understanding of the relationship between life events and musical preferences. By analyzing personal diary entries alongside Spotify streaming data from 2023, it highlights the significant influence of specific external events like elections and sporting competitions on musical tastes. Conversely, it notes the lack of noticeable impact from personal and academic life events, challenging traditional assumptions about the universality of these influences.

The study, however, is limited by its focus on quantitative data and its one-year scope, potentially overlooking the nuanced role of lyrical content and long-term trends in music preference evolution.

In summary, while offering valuable insights into how external experiences shape music choices, this research underscores the need for future studies incorporating both qualitative aspects and extended timeframes to fully capture the complex interplay between life narratives and musical tastes.