

Analysis of Spotify Listening Habits Based on Weather Conditions

January 2025

Purpose of the Project

This project explores the relationship between music listening habits and weather conditions. By analyzing Spotify listening data alongside weather information, the goal is to uncover patterns in music preferences and how they align with various weather conditions such as sunny, rainy, and cloudy days. The project aims to understand whether weather conditions influence features like genre, energy, and tempo of the music.

Goals of the Project

The primary objectives are:

1. **Listening Preferences:** Analyzing how genres, energy levels, and tempo vary with different weather conditions.
 2. **Weather Data Dependency:** Evaluating whether music preferences are independent of or dependent on weather conditions using statistical methods.
 3. **Artist Preferences:** Identifying the relationship between frequently listened artists and weather conditions.
-

Hypotheses

- **Null Hypothesis (H_0):** Spotify listening habits are independent of weather conditions.
- **Alternative Hypothesis (H_1):** Spotify listening habits are dependent on weather conditions.

A Chi-Square Test of Independence was employed to evaluate the relationship between these variables.

Methodology

Data Collection

1. **Spotify Data:**
 - Extracted listening history, including song names, artists, genres, energy levels, tempo, and listening timestamps using the Spotify API.
2. **Weather Data:**
 - Collected weather information such as temperature, condition (sunny, rainy, cloudy), and humidity using the Visualcrossing API.

Data Processing

1. **Data Matching:**
 - Merged Spotify listening data with weather information based on timestamps.
2. **Data Cleaning:**
 - Removed incomplete records and standardized categories for weather conditions.

Statistical Testing

1. **Chi-Square Test of Independence:**
 - Categorical variables (e.g., weather conditions, artists) were analyzed to test for dependency.
 - Features like tempo and energy levels were binned into categories for analysis.

Visualization

- Distributions of genres, energy, tempo, and artist preferences were visualized for each weather condition.
 - Comparative bar charts and pie charts were used to highlight key patterns.
-

Findings

1. **Artist Preferences**

There was no noticeable trend in the listening history linked specifically to weather conditions. The same artists or similar music styles were chosen regardless of whether it was sunny, rainy, or cloudy.
 2. **Music Type vs Weather**

No consistent or significant relationship was observed between frequently listened artists/music types and different weather conditions. There was no discernible preference for a particular music type based on the weather.
 3. **Energy and Tempo Patterns**

There was no clear difference in the energy and tempo levels of songs among sunny, rainy, or cloudy days. While some days favored higher-tempo songs, such occurrences did not form a consistent pattern related to weather conditions.
 4. **Statistical Results**

Analyses (e.g., Chi-Square Test of Independence) indicated no statistically significant relationship between weather conditions and listening habits, suggesting that the null hypothesis (H_0 : weather conditions and listening habits are independent) could be accepted.
-

Limitations

1. **Dataset Scope:**

- Relies on personal listening data, limiting generalizability to broader populations.

2. **Binning Strategy:**

- Results are sensitive to the binning of continuous variables like tempo and energy.

3. **Weather Specificity:**

- Aggregated weather conditions may overlook nuanced effects of temperature or humidity.
-

Conclusion

Based on the analysis, no statistically significant relationship was found between Spotify listening habits and weather conditions. Variations in artist preferences, genre, energy, and tempo did not reveal a consistent pattern tied to different weather scenarios. Thus, these findings support the idea that environmental factors—at least in terms of weather—do not substantially influence musical choices. Future research could explore additional variables or incorporate more diverse datasets to further investigate these observations.

Acknowledgments

This project was completed as part of DSA210 – 29079 leveraging Python libraries such as pandas, matplotlib, seaborn, and scipy, and supported by data from Spotify and Visualcrossing APIs.