# Music and Smoking Habits Analysis

## Author: Emir Emre AYDIN

Music has always been an important role in my daily life, shaping my emotions and behaviors in profound ways. Since October 4th, I have been tracking my smoking habits with the goal of understanding how these behaviors affects with my music preferences and external events. This project seeks the patterns in smoking habits influenced by specific artists or even life events like exams. By combining personal smoking data, Spotify listening history, and external event markers, I aim to explore whether music plays a role in triggering certain behaviors. Through data cleaning, statistical analysis, and visualization, the project offers a structured approach to identifying meaningful correlations. Tools like Python, Spotify's Request History, and statistical libraries have been employed to extract insights and visualize the findings effectively. Ultimately, this analysis is a journey into the intersection of personal habits and music, with the hope of better understanding *what drives certain behaviors and how they can be shaped or mitigated.*

The first step in the project involved *cleaning and reformatting Spotify data* to ensure consistency and usability. Using the Request History data, the raw dataset was transformed to extract meaningful features like listening time, duration, artist names, and dates. From this, two distinct datasets were created: the first focused on daily listening averages to observe overall trends, and the second detailed how much time was spent listening to specific artists on each day. These datasets provided the foundation for analyzing correlations between music preferences and smoking habits, enabling a deeper understanding of both general trends and artist-specific influences. This preprocessing step was critical for aligning the music data with smoking behavior for further analysis.

## Separate Analysis of Music and Smoking

### A) Smoking Habbits Analysis

Following the preprocessing, I conducted separate analyses to explore the patterns and distributions in smoking habits. I began by visualizing the data using histograms and time series plots, which provided a clear picture of the overall trends and daily variations. To assess whether the data followed a normal distribution, I performed a normality assessment using Q-Q plots, skewness, and kurtosis.


Histogram of Daily Smoking Habits
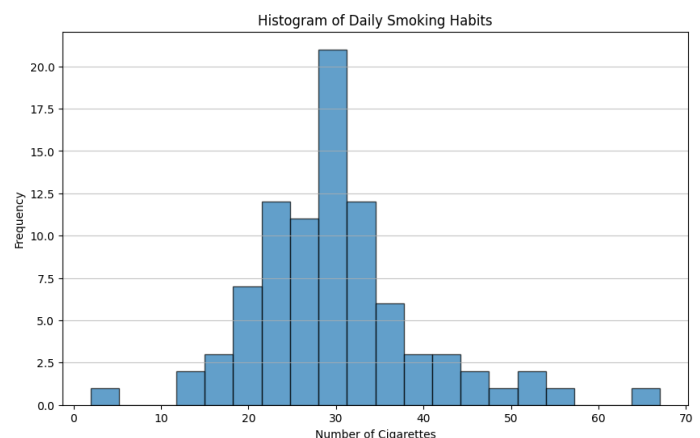
The results of the normality assessment can be summarized as follows:

1. **Q-Q Plot Test**:

   • The data points generally followed the diagonal line, which is indicative of normality.

   • However, deviations were observed at the tails, suggesting potential outliers or a departure from strict normality in the extremes.

2. **Skewness**:

   • The skewness value was **0.856**, indicating a moderate right skew in the data.

   • A positive skew suggests that the smoking habits were slightly more concentrated towards lower values, with a tail extending to higher values.
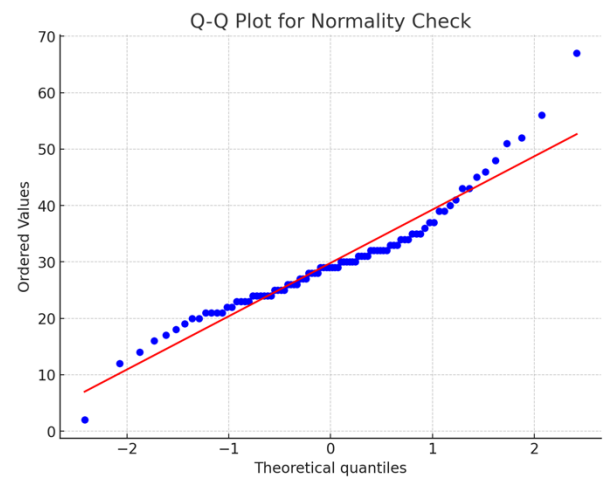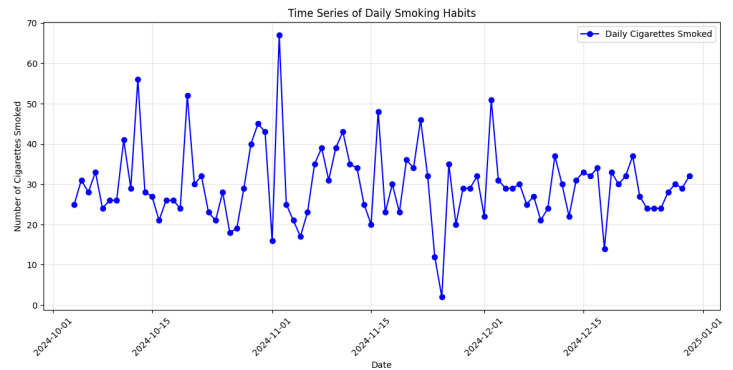


These findings suggest that while the data appears approximately normal, there are noticeable deviations at the extremes, which may affect certain statistical analyses. As a results.

This analysis examines the impact of special days on smoking habits, highlighting how different types of days influence cigarette consumption. According to the data, "Happy and Relaxation" as socializing days, such as parties or birthdays, show the highest average cigarette consumption, suggesting that social events may trigger increased smoking. In contrast, "Stressful/Exam and Preparation" days have moderate consumption levels, indicating that smoking might be used as



a coping mechanism during stressful periods. "Regular Days," on the other hand, exhibit the lowest consumption, suggesting that daily routines are less likely to provoke smoking. This analysis emphasizes the significant influence of special days on smoking habits, providing valuable insights into behavioral triggers.
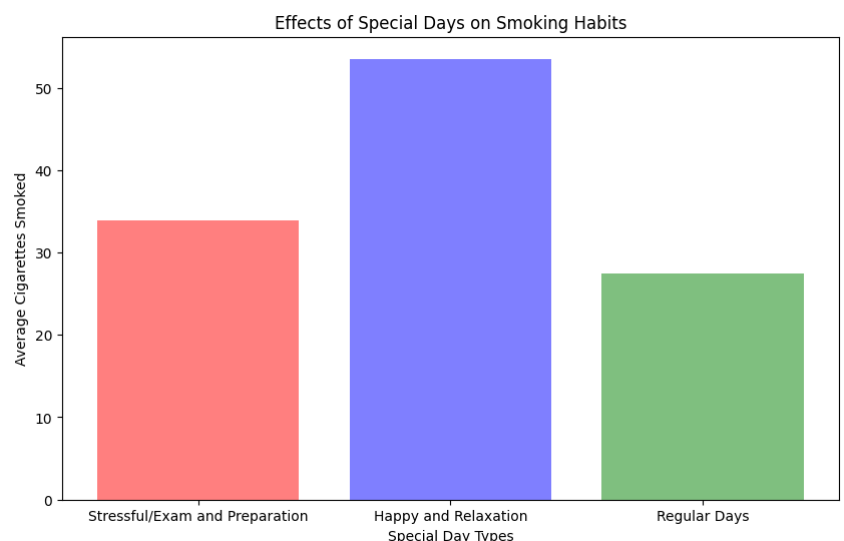
The smoking data offers a comprehensive view of daily cigarette consumption, including information on special days, such as stressful events or celebrations, and their influence on smoking habits. Time series and distribution analyses revealed significant variability in smoking patterns, with peaks corresponding to specific triggers like exams or social events. This dataset forms the foundation for exploring behavioral trends and identifying correlations with external factors, such as music preferences, to understand the underlying drivers of smoking habits.

## B) Music Data Analysis

After analyzing smoking habits, I shifted my focus to the music data to explore how listening behaviors might influence or reflect smoking patterns. By leveraging metrics such as daily listening averagesand artist-specific trends, the goal was to uncover meaningful relationships between music and smoking. This transition illustrates the combination of two behavioral datasets to better understand the habits between music and daily habits.

I began my analysis of the music data by creating a dataset that showcased the artists I listened to the most over the study period. This dataset ranked the top 50 artists based on the total number of plays, offering a detailed overview of my music preferences. The results highlighted distinct listening patterns, with certain artists dominating my daily plays, reflecting



Top 50 Most Listened Musicians

their prominence in my music habits. To ensure the accuracy and reliability of this dataset, I compared the findings with my Spotify Wrapped summary for the year, which provides an official overview of listening trends. The comparison revealed a strong alignment between my self-collected data and Spotify's analysis, confirming that the dataset I built accurately represents my listening habits. By establishing this dataset, I could move forward with exploring deeper correlations between music preferences and smoking behaviors, knowing that the core data was both consistent and trustworthy.
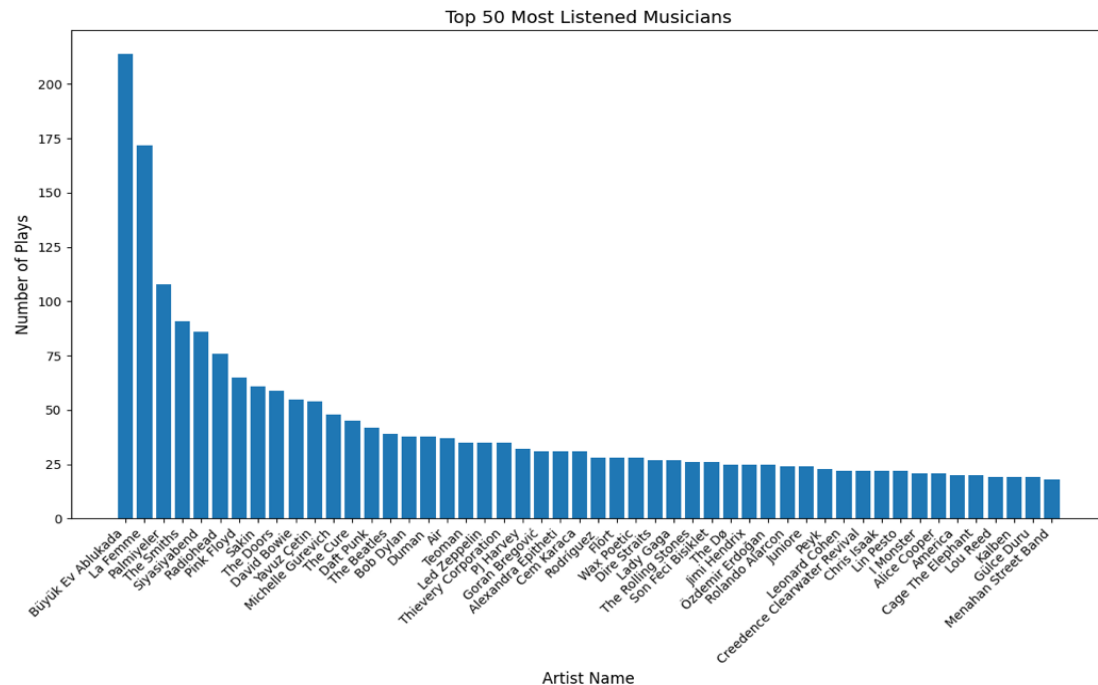
My next analysis explored the average daily listening trends across the days of the week, aiming to understand how listening habits varied depending on the day. The data revealed distinct patterns, with clear peaks and troughs that aligned with the structure of the week. The highest average listening count occurred on Saturdays, with an impressive 126.1 plays, followed by Fridays at 113.2 plays and Sundays at 100.4 plays. These findings suggest that listening habits peak toward the weekend, likely reflecting more leisure time and a greater inclination to relax or engage with music.

However, Thursday had the lowest average listening count at 64.6 plays, indicating a notable dip in listening activity during the mid-week. Monday and Tuesday exhibited moderate levels of activity, with averages of 77.3 and 78 plays, respectively, reflecting a more balanced consumption pattern as people transition into their weekly routines. Wednesday showed a sharp increase in listening activity, with an average of 99.4 plays, perhaps marking the midpoint of the week when individuals might seek more engagement with music to recharge.

These results highlight how *listening habits are influenced by the structure of the week*, with increased activity on days traditionally associated with relaxation or social activities. The significant drop on Thursday suggests it could be a busy day for other activities, leaving less time for music. This analysis provides valuable insights into the temporal patterns of music consumption, offering a foundation for exploring how these habits intersect with other behaviors, such as smoking or external life events. Understanding these trends can help uncover broader behavioral rhythms and their implications.
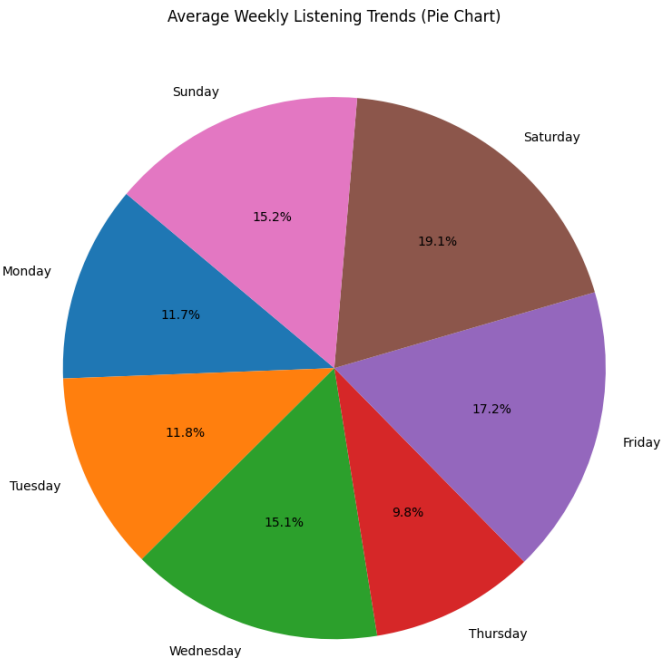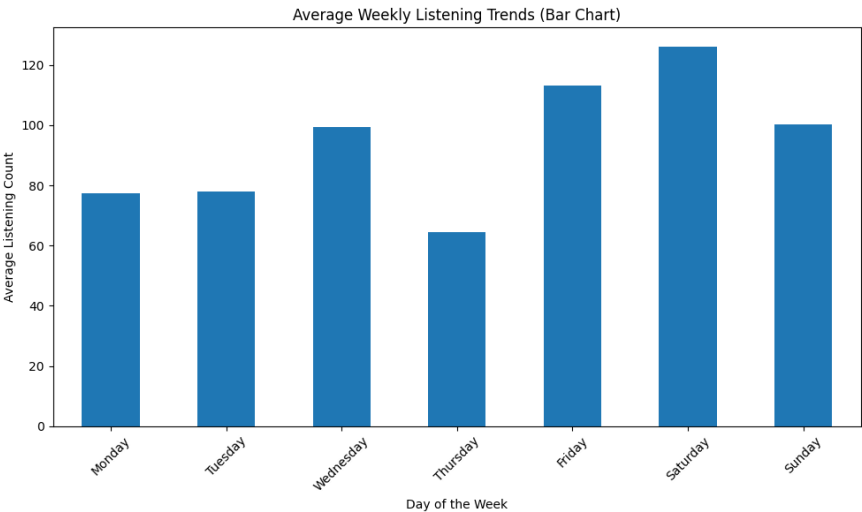
In conclusion, the analysis of music data revealed clear patterns in listening habits, both in terms of artist preferences and daily trends. The most listened-to artists, such as Büyük Ev Ablukada and La Femme, *aligned with personal preferences and the Spotify Wrapped summary, validating the data's accuracy*. Weekly trends highlighted a significant increase in music consumption during weekends, with Saturday showing the highest average listening counts, reflecting how leisure time influences listening behaviors. These findings provide a solid foundation for exploring how music consumption interacts with other behaviors, such as smoking habits, offering insights into the potential role of music as a behavioral trigger or coping mechanism.

## Combined Analysis of Music Habbits and Smokng Habitts

**Introduction**

This project explores the intriguing intersection of music preferences and smoking habits, aiming to uncover patterns and correlations that offer deeper insights into behavioral triggers. By analyzing daily smoking data alongside detailed music consumption, such as artist preferences and listening trends. The analysis is grounded in a combination of statistical assessments, time series visualizations, and network mapping, which collectively provide a comprehensive understanding of how music and smoking habits interact. The statistically significant correlation (Pearson coefficient = 0.57, p-value = 3.75e-06) *suggests a meaningful relationship*, further supported by observations from artist connections and special day analyses. This integrated approach not only highlights the influence of music.

The comparison of daily smoking habits and music listening trends provides an insightful look into how these two behaviors interact over time. The time series plot illustrates that cigarette consumption tends to remain relatively stable, with a steady baseline punctuated by
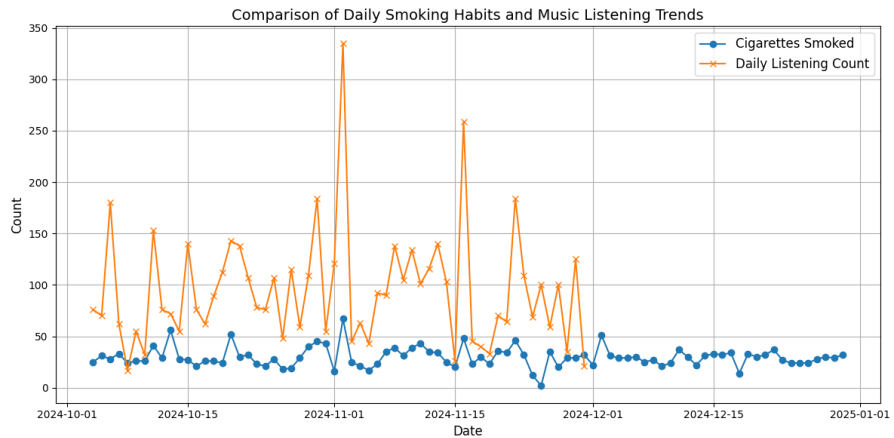


occasional spikes and dips. These fluctuations in smoking are likely driven by external factors such as stressful events, exam periods, or social activities. In contrast, music listening counts demonstrate far greater variability, with distinct peaks on specific days. These spikes often align with weekends or periods of relaxation, suggesting that music consumption is more influenced by leisure time and personal mood.

When viewed together, *the plot reveals intriguing overlaps where higher smoking levels coincide with peaks in music listening,* hinting at a possible behavioral connection. These moments may indicate that music serves as a coping mechanism during heightened smoking activity or that

both behaviors are influenced by shared triggers such as stress or celebratory events. While the trends do not perfectly mirror each other, the variability in music listening and the steadiness of smoking habits provide a foundation for exploring potential correlations. This analysis underscores the importance of integrating multiple behavioral datasets to identify patterns and gain a deeper understanding of the interplay between music preferences and lifestyle choices.

The scatter plot comparing daily cigarettes smoked with daily music listening counts reveals an interesting pattern, and the Pearson correlation coefficient of **0.57** indicates a *moderate positive correlation between the two variables.* This suggests that as the number of cigarettes smoked increases, so does the number of music listens, albeit not perfectly. The p-value of **3.75e-06** confirms that this relationship is statistically significant, *meaning that the observed correlation is unlikely to have occurred by chance.* While the correlation is moderate, it points to a possible



connection where increased smoking behavior may coincide with higher music listening, potentially due to shared underlying factors, such as stress or a desire to relax. The scatter plot itself shows that the majority of data points are clustered around lower cigarette consumption and moderate listening counts, with a few notable outliers representing higher smoking days associated with higher listening activity. This indicates that extreme values in either behavior may sometimes occur simultaneously, but overall, the general trend suggests that these two habits may be linked under certain circumstances. The next steps in the analysis would involve exploring the causal or behavioral factors contributing to this correlation, potentially looking into external triggers such as specific days, moods, or events, and further dissecting how they influence both smoking and music preferences.
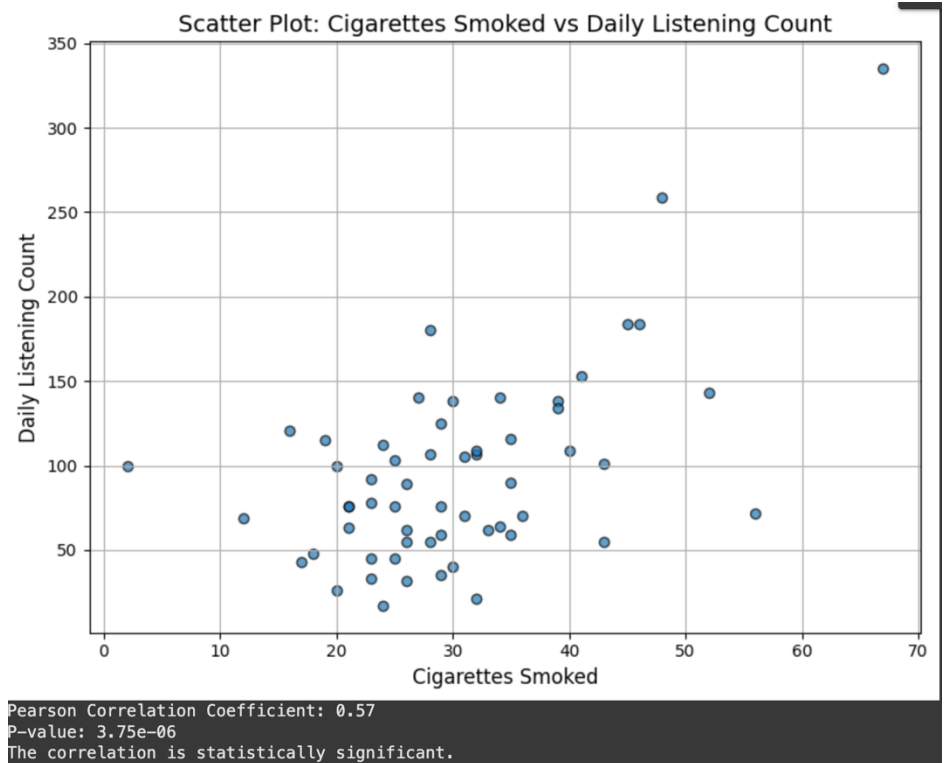
The correlation matrix heatmap reinforces the insights gained from the previous scatter plot. The matrix visually confirms a moderate positive correlation between daily listening count and cigarettes smoked, with a correlation coefficient of **0.57**. This aligns with the findings from the scatter plot, where an increase in daily cigarette consumption corresponds with higher music listening counts. The heatmap's color gradient further emphasizes this moderate correlation, with the value of 0.57 highlighted in the matrix, indicating a meaningful, albeit not perfect, relationship between



the two variables. These consistent results across both the scatter plot and the heatmap solidify the conclusion that there is a statistically significant connection between smoking habits and music consumption, suggesting that these two behaviors may be influenced by similar factors or triggers. The next step in this analysis would be to investigate potential external factors contributing to this relationship, such as stress levels, social settings, or emotional states.
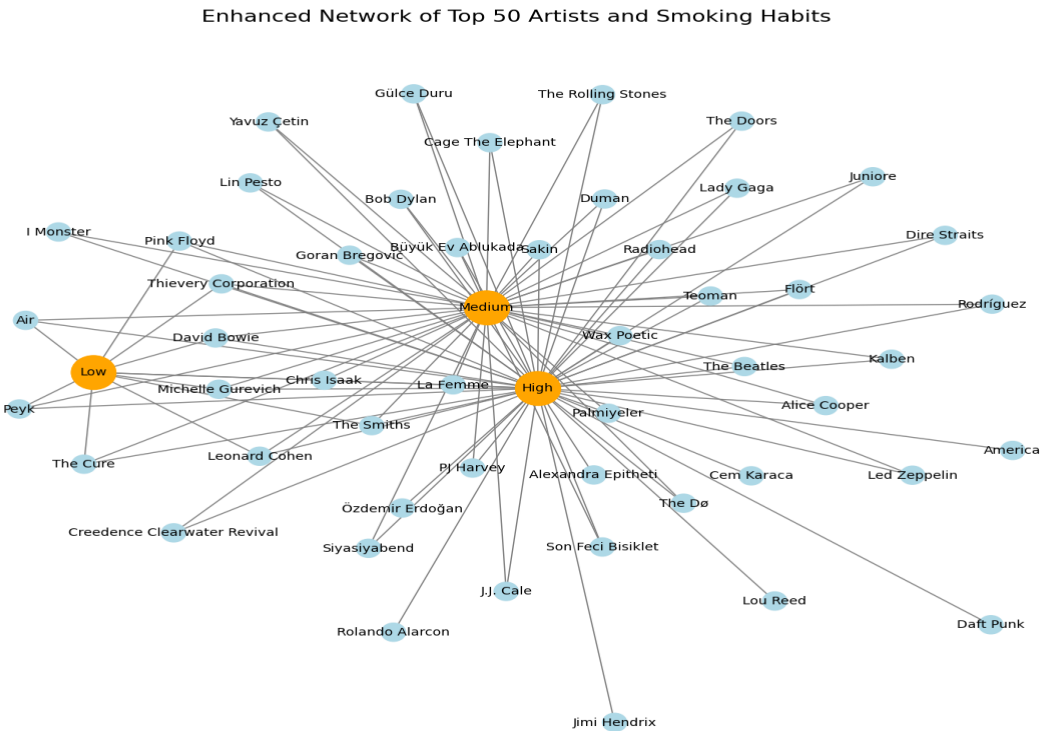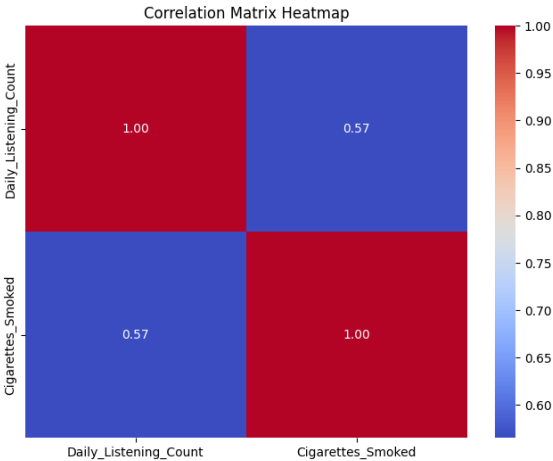


The enhanced network visualization of *the top 50 artists and their connections to smoking* habits reveals interesting patterns in how smoking levels correlate with musical preferences. The

network is divided into three smoking levels: **Low(X<10)**, **Medium(X<20)**, and **High(X>=20)**, with artists connected to these levels based on listening habits and smoking data. This section explores the key findings and insights derived from the network, offering a more detailed examination of the connections between music preferences and smoking behaviors.

**Observations:**

1.      **Central Smoking Levels**:

•       The **Medium** and **High** smoking levels are the most densely connected, indicating that a significant number of popular artists are linked to these two smoking categories. This suggests that a large portion of listening to these artists corresponds with moderate to high levels of smoking. Artists such as **Radiohead**, **Bob Dylan**, and **David Bowie** are strongly associated with the **Medium** smoking level, implying that their music may resonate more with individuals who have moderate smoking habits.

•       **Low** smoking levels, on the other hand, are connected to fewer artists, including **Peyk** and **Air**. This could indicate that individuals with lower smoking habits are less likely to engage with the music of these specific artists, potentially reflecting a niche or more restrained listening audience.

2.      **Diverse Artist Connections**:

•       The **High** smoking level, characterized by higher daily cigarette consumption, displays a broader range of musical influences. Artists such as **Pink Floyd**, **Jimi Hendrix**, **The Rolling Stones**, and **Led Zeppelin** are strongly connected to this category, which may reflect a cultural association between rock and alternative genres and smoking behaviors. These genres are often linked to lifestyle choices or behaviors that may include smoking as a social or cultural activity.

•       The variety in artist connections under **High** smoking habits also highlights how smoking behavior might be tied to particular lifestyle choices or subcultures that celebrate or normalize smoking.

3.      **Focused Connections for Low Smoking Levels**:

•       **Low** smoking habits are associated with fewer, but more specific, artists. The artists in this category, like **Peyk** and **Air**, could be appealing to a more niche or health-conscious audience that maintains lower smoking levels. This suggests a more selective group of listeners who engage with artists who might represent a different set of values, such as wellness or tranquility, contrasting with the cultural associations tied to higher smoking levels.


**Insights:**

1.     **Potential Cultural Trends**:

•The clustering around certain smoking levels (e.g., **High**) suggests that smoking habits are culturally linked with specific music genres. For example, rock and alternative genres—often associated with rebellious —are more strongly linked to higher smoking levels, which could reflect broader societal trends where smoking is seen as part of the rock and roll culture.

•The music genre preferences tied to smoking habits suggest that certain artists may be appealing to individuals who use smoking as a coping mechanism or as part of a lifestyle choice. For instance, genres that align with relaxation or socializing might be more linked to higher smoking habits.

2.     **Targeted Patterns**:

•The network graph emphasizes the dominant smoking levels that shape musical preferences, providing a foundation to explore how certain behaviors are linked to specific cultural or demographic patterns. For example, the significant representation of **High** smoking habits among classic rock artists may point to an audience demographic that is more likely to engage in behaviors associated with a more indulgent or relaxed lifestyle.

•The graph also suggests that understanding the connections between music and smoking can reveal underlying psychological, social, or behavioral factors that influence both habits. By investigating these patterns further, we can gain deeper insights into how music preferences correlate with lifestyle choices and how these can be influenced by or associated with broader cultural trends.


In conclusion, this network analysis not only highlights the relationship between smoking habits and music listening preferences but also suggests potential cultural and behavioral links that can provide valuable insights into understanding these intertwined habits. Further exploration into these connections could uncover additional nuances in the lifestyle choices of listeners and their engagement with specific music genres.


**Combined Analysis of Music Habits and Smoking Habits**

This project delves into the complex relationship between daily smoking habits and music listening preferences, aiming to uncover correlations and behavioral patterns that may shed light on how these two aspects of daily life influence each other. By analyzing smoking data alongside detailed music consumption metrics, such as artist preferences, daily listening counts, and weekly trends, the study seeks to explore whether music acts as a trigger or coping mechanism in

relation to smoking. The correlation between these behaviors is assessed through statistical methods like Pearson's correlation coefficient, which revealed a moderate positive relationship (0.57), indicating that higher smoking levels may coincide with more frequent music listening. This statistical significance is further corroborated by the heatmap analysis, which confirmed the correlation and provided a clear visual representation of the relationship.

The time series analysis of smoking habits versus music listening trends highlights a steady baseline in smoking behavior, with notable peaks that often correspond to specific events such as stressful periods or social occasions. In contrast, music listening exhibits greater variability, with marked spikes on weekends and relaxation days, suggesting that music consumption is heavily influenced by leisure and personal mood. When overlaid, these patterns show some intriguing overlaps where higher smoking levels correspond with peaks in music listening, hinting at a potential behavioral connection. This suggests that for some individuals, both music and smoking may be triggered by similar external factors like stress or socializing.

The scatter plot and the Pearson correlation analysis reinforce these findings, showing a moderate correlation between daily smoking counts and music listening. The network analysis of top 50 artists and smoking levels further deepens the insights, revealing that smoking habits—particularly the **High** category—are linked with iconic rock and alternative artists like **Pink Floyd**, **The Rolling Stones**, and **Jimi Hendrix**, reflecting a cultural connection between certain genres and smoking behaviors. On the other hand, **Low** smoking levels were associated with fewer, more niche artists, suggesting a more selective audience. These patterns support the hypothesis that music genres, particularly those linked with relaxed or rebellious lifestyles, may play a role in influencing smoking habits.

In conclusion, this combined analysis highlights the intriguing intersection between music preferences and smoking habits, uncovering both cultural and behavioral connections. The moderate correlation and network mappings suggest that external triggers, such as social settings or emotional states, may simultaneously influence both music listening and smoking behavior. Further research into these connections could reveal deeper insights into how lifestyle choices, music genres, and psychological factors are intertwined, providing valuable knowledge that could influence both music-related and health-related interventions.