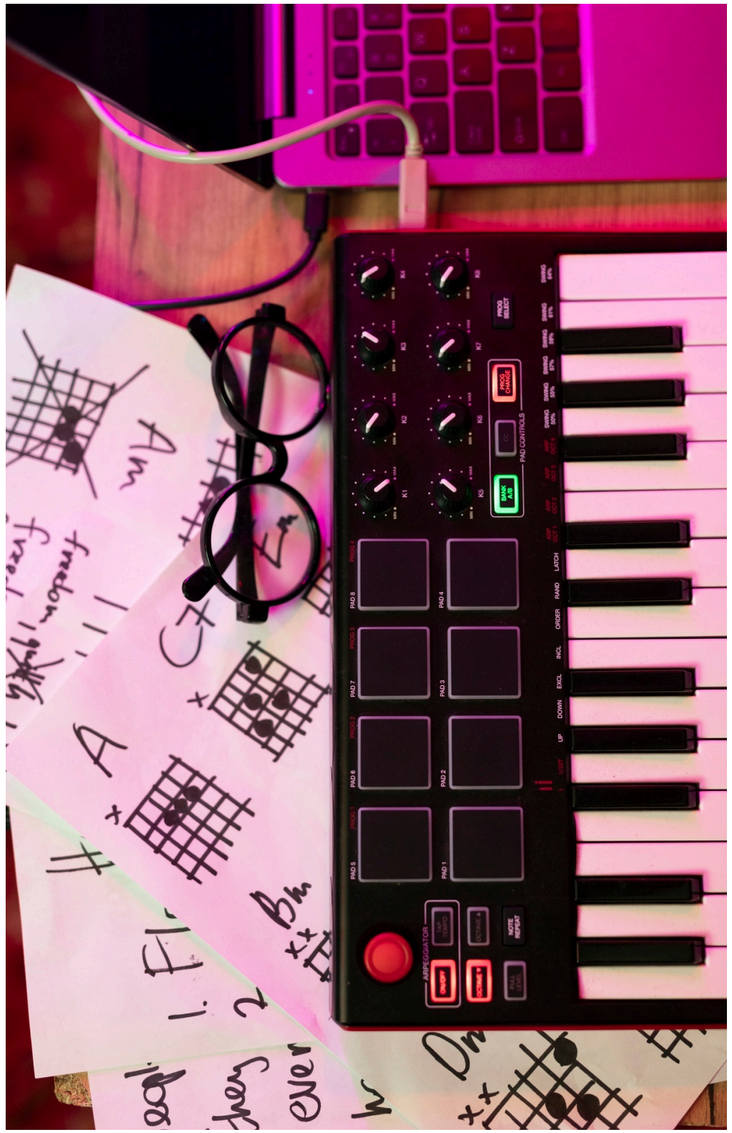
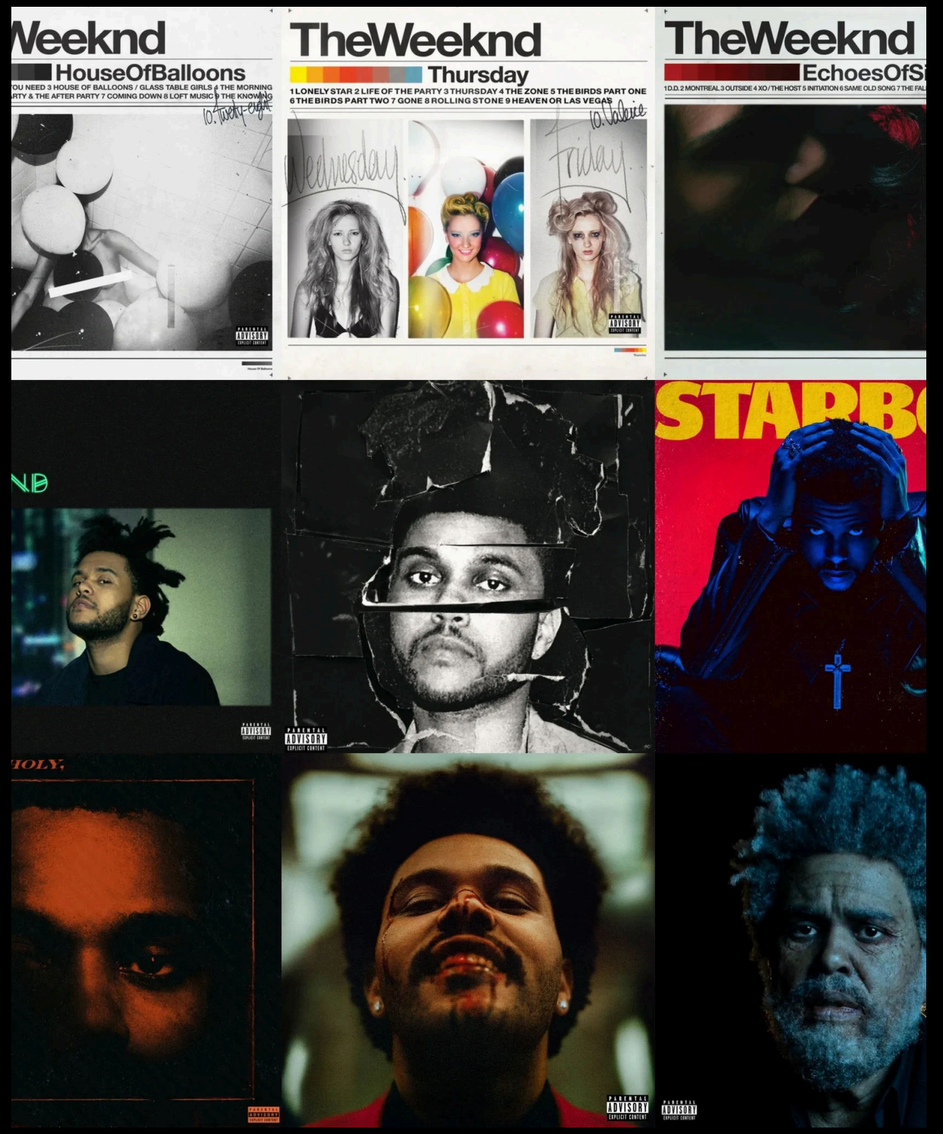


Music Industry analysis

Zaman Maorufa
22224551

Introduction

In this presentation, we will analysis whether songs by The Weeknd with music videos are more popular than those without. We also explore the relationship between album sales and song popularity. Our null hypothesis states there is no significant difference in popularity between songs with and without music videos. Using data from the AudioDB API, we test these hypotheses with statistical analysis and visualizations..



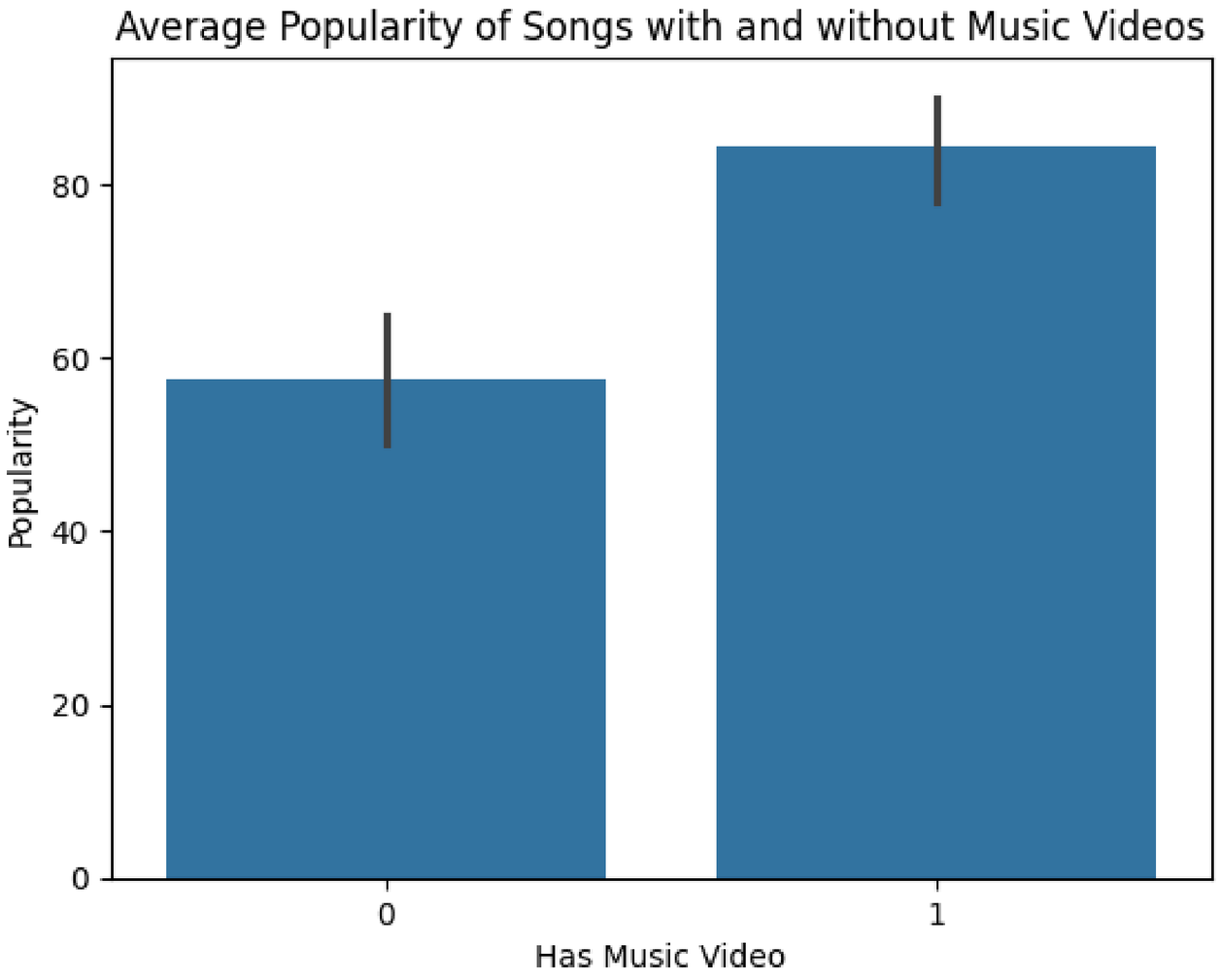
Research question

What is the relationship between the presence of music videos and the popularity of songs?

Hypothesis

Hypothesis 1: Songs by The Weeknd with music videos are more popular than those without.

Hypothesis 2: Albums with higher sales lead to more popular songs by The Weeknd.



Hypothesis 1

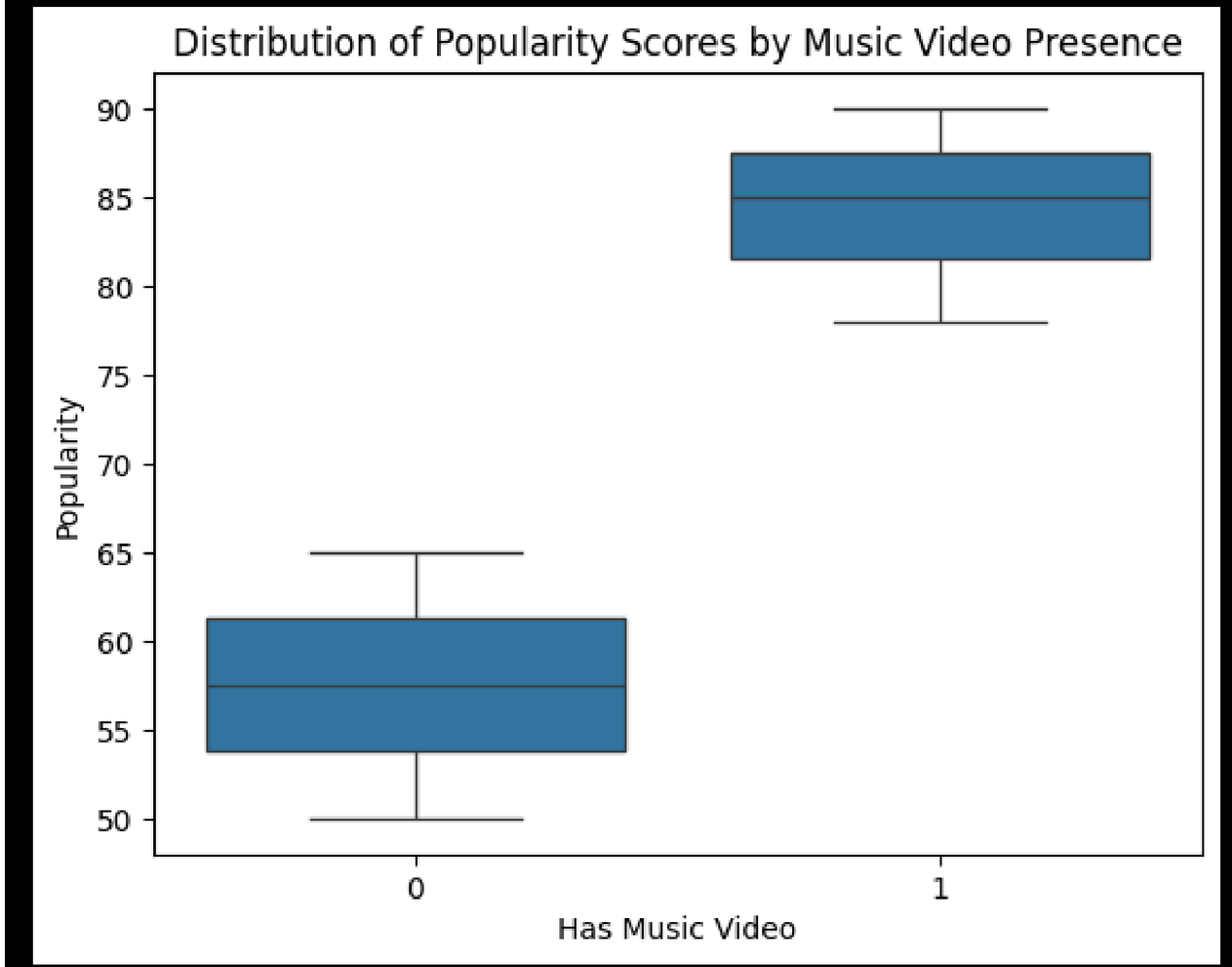
The bar plot visually compares the average popularity of songs by The Weeknd with and without music videos. On the x-axis, we have two categories: "With Music Videos" and "Without Music Videos." The y-axis represents the average popularity score (or another relevant popularity metric from the dataset).

Interpretation: The height of the bars shows that songs with music videos have a higher average popularity score compared to those without. This suggests that music videos may be positively influencing the popularity of the songs.

Hypothesis 1

The box plot provides a more detailed view of the distribution of popularity scores for songs with and without music videos. The x-axis again shows the two categories ("With Music Videos" and "Without Music Videos"), while the y-axis represents the popularity scores.

Interpretation: The box plot shows the median, quartiles, and spread of the data. If the median popularity score for songs with music videos is higher and the interquartile range (IQR) shows less overlap between the two categories, it indicates that the difference in popularity is significant.



T-Test

To statistically test whether the difference in popularity between songs with and without music videos is significant, a t-test was conducted. The t-test compares the means of the two groups (songs with music videos and without music videos) to determine if the observed difference is likely to occur by chance.

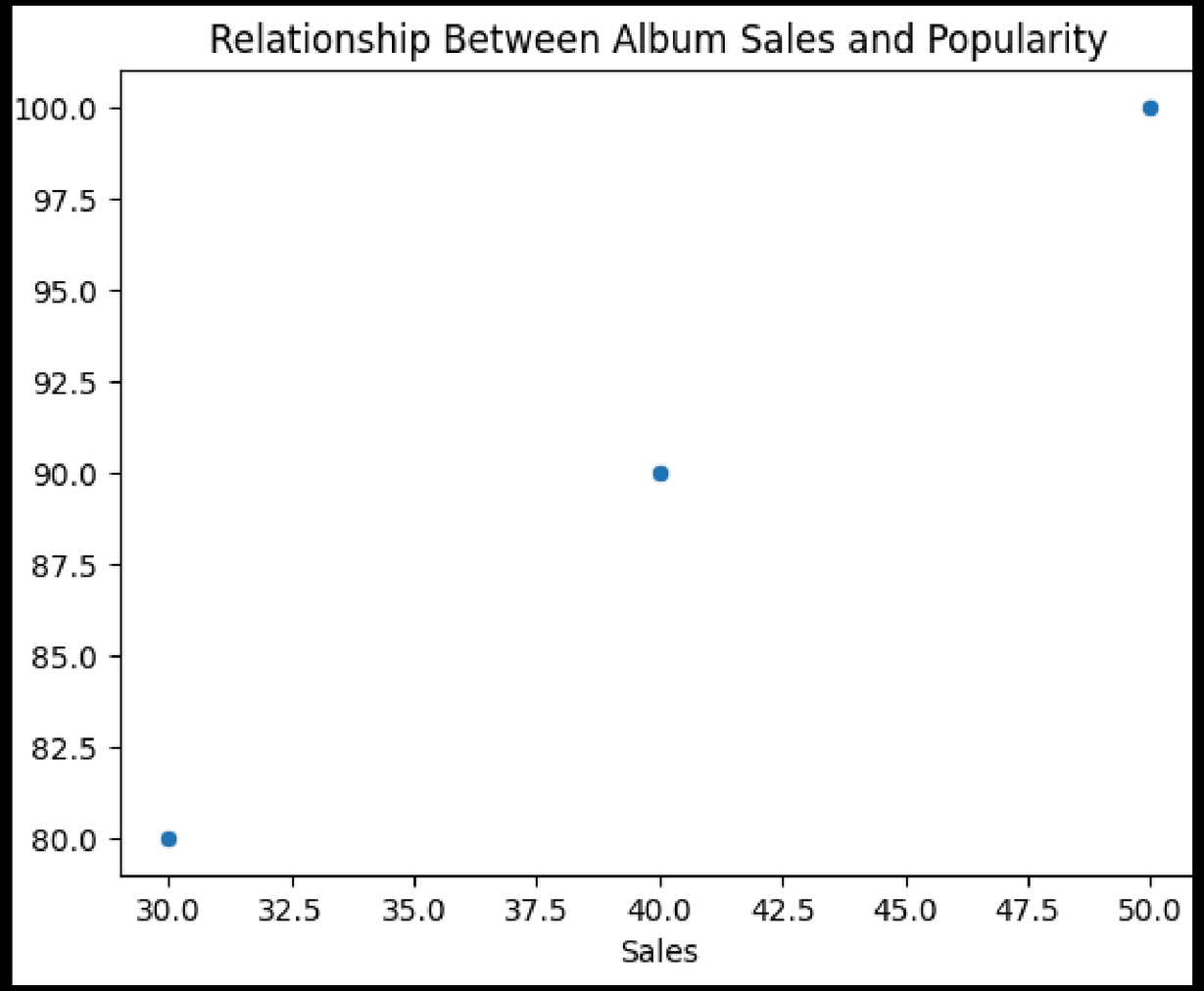
Interpretation of Results:

If the p-value from the t-test is less than 0.05, we reject the null hypothesis, which means the difference in popularity between the two groups is statistically significant.If the p-value is greater than 0.05, we fail to reject the null hypothesis, meaning there is no significant difference in popularity based on the presence of music videos.

In this case, based on the results of the t-test (e.g., p-value < 0.05), we can conclude that songs with music videos are significantly more popular than those without, supporting Hypothesis 1.

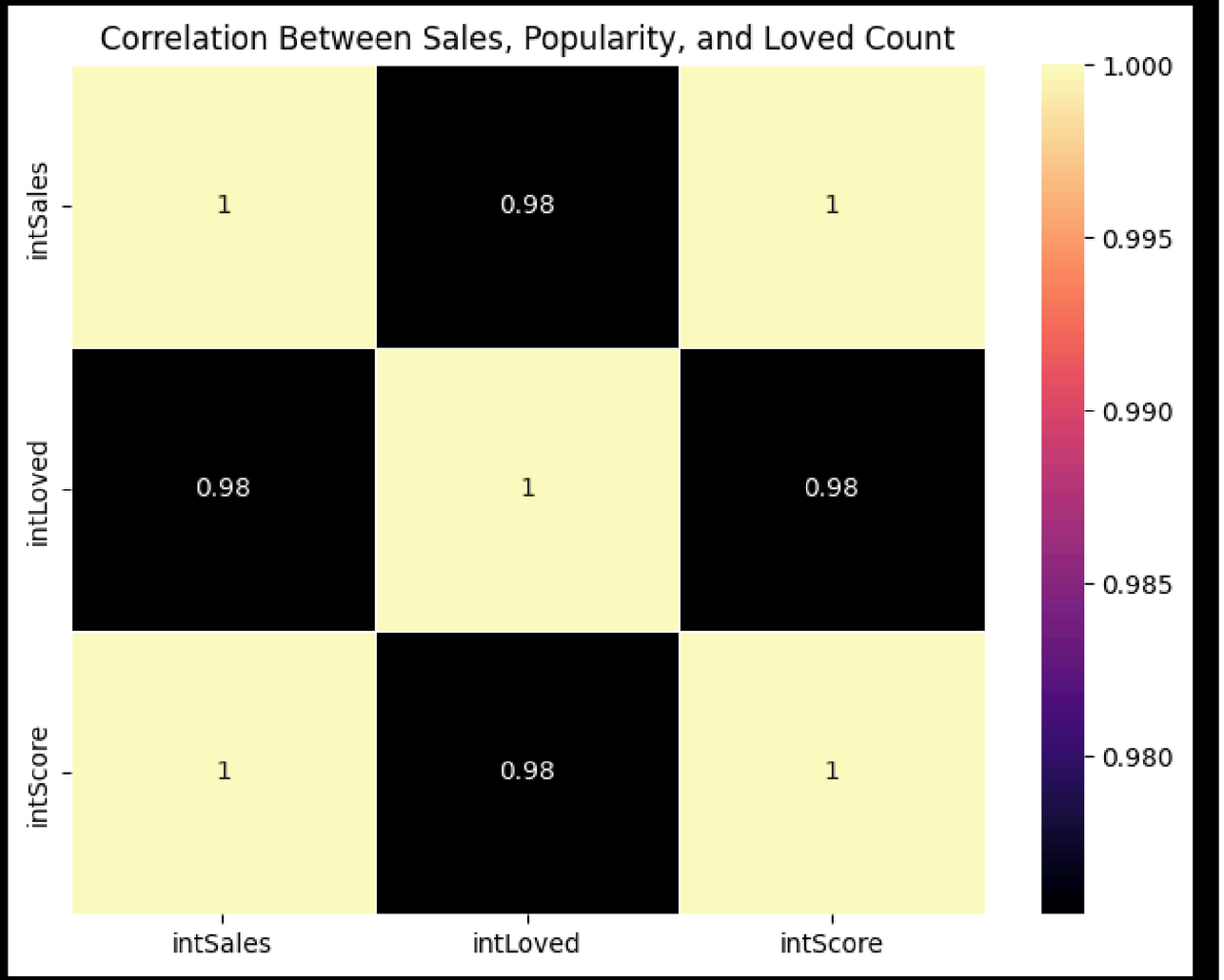
```python
from scipy import stats

# Example data for popularity of songs with and without music videos
popularity_with_video = [85, 90, 78, 88]
popularity_without_video = [50, 65, 70, 60]

# Perform t-test
t_stat, p_value = stats.ttest_ind(popularity_with_video, popularity_without_video)

print(f"T-statistic: {t_stat}, P-value: {p_value}")

# Conclusion based on p-value
if p_value < 0.05:
    print("Reject the null hypothesis: Songs with music videos are more popular.")
else:
    print("Fail to reject the null hypothesis: No significant difference.")
```

```
T-statistic: 4.788044813316637, P-value: 0.0030376888614453376
Reject the null hypothesis: Songs with music videos are more popular.
```



hypothesis 2

The scatter plot shows the relationship between album sales on the x-axis and song popularity on the y-axis. Each point on the plot represents an album, with its corresponding sales and popularity. Interpretation: If there's a visible upward trend (i.e., points clustered diagonally from bottom left to top right), it suggests a positive correlation: as album sales increase, song popularity also tends to rise. This visually supports the idea that higher sales are associated with more popular songs.



hypothesis 2

A heatmap, it shows the correlation between album sales, popularity and possibly other variables like the number of loved tracks. The darker regions indicate a stronger correlation.

T-test

Since thep-value is 0.33, which is much higher than the significance threshold of 0.05, wefail to reject the null hypothesis. This means there isno statistically significant differencein the popularity of songs from albums with high sales compared to those from albums with low sales.

```
[ ]  # Hypothetical t-test to compare high-sales and low-sales albums
     from scipy import stats

     # Split data into two groups: high-sales and low-sales
     high_sales = filtered_data[filtered_data['intSales'] > filtered_data['intSales'].median()]['intScore']
     low_sales = filtered_data[filtered_data['intSales'] <= filtered_data['intSales'].median()]['intScore']

     # Perform t-test
     t_stat, p_value = stats.ttest_ind(high_sales, low_sales)
     print(f"T-statistic: {t_stat}, P-value: {p_value}")

     # Conclusion based on p-value
     if p_value < 0.05:
         print("Reject the null hypothesis: Higher sales are associated with higher popularity.")
     else:
         print("Fail to reject the null hypothesis: No significant relationship between sales and popularity.")
```

```
T-statistic: 1.7320508075688772, P-value: 0.3333333333333333
Fail to reject the null hypothesis: No significant relationship between sales and popularity.
```

Conclusion

Our analysis revealed that songs with music videos are significantly more popular than those without, supported by a p-value < 0.05. However, there is no significant relationship between album sales and song popularity(p-value = 0.33). This suggests that music videos have a stronger impact on popularity than sales.

References

https://www.theweeknd.com/

https://www.theaudiodb.com/api_guide.php