Business Project -Proposing building a regression model for Warner Music Group

Calvin Yu

Client Problem/Need

- Warner Music group has received too many demos during the covid-19. How can they know which demo can become popular?

Data Science Solution Path

- Do a Exploratory analysis and create visualization for all the song attributes which includes keys, mode, valence, tempo , and so on

- Evaluating the relationship between each feature and the song profit

Impact Hypothesis

- By better understanding how each feature of the song contributes to its profit, Warner Music Group can identify what songs would have the most potential to make the most profits

Scoping considerations

- Assumption :
    - Popular songs tend to share similar features

- Measure of Success :
    - proven that there are relationships between the features and the profit
    - How well does the prediction model do (Future)

- Risk :
   - Melody cannot be defined

Methodology

- Data:
    - Spotify global 2019 most-streamed tracks (US)

    - Spotify-Data 1921-2020
        - Audio features of songs from 1921-2020

    - Top 200 most streamed songs on spotify 2020 (US)

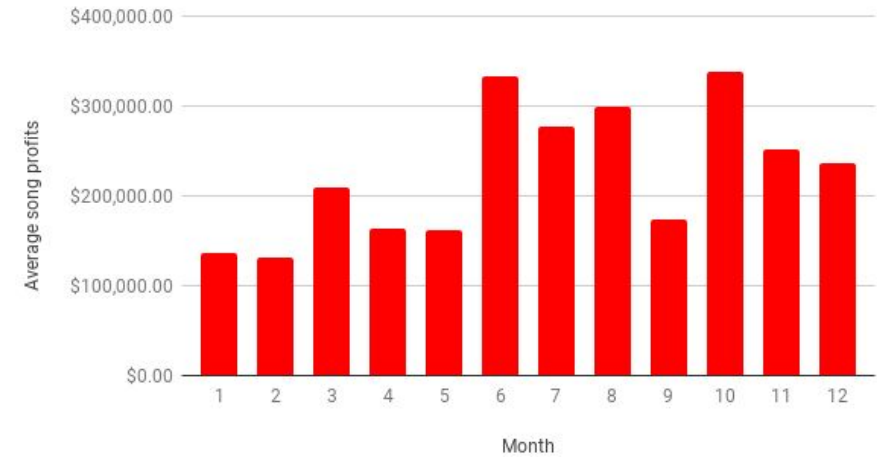# Methodology Cont.

- Exploratory Data Analysis

    - Create visualizations on every feature on how it relates to the profits

    - Categorical on Excel , Numerical on Tableau

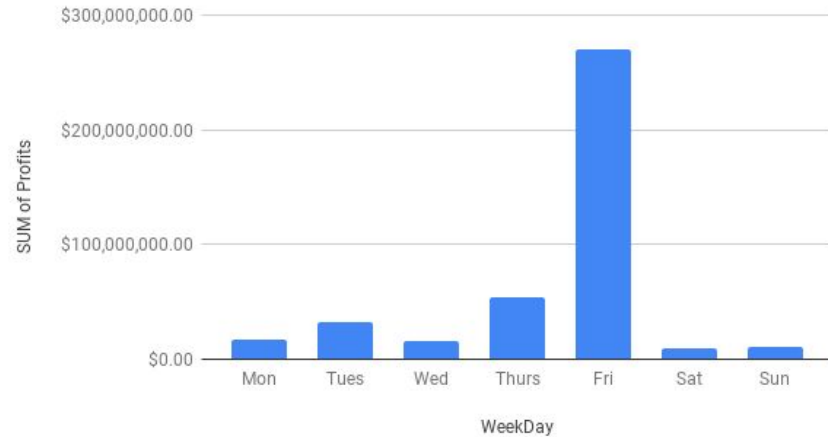# Results

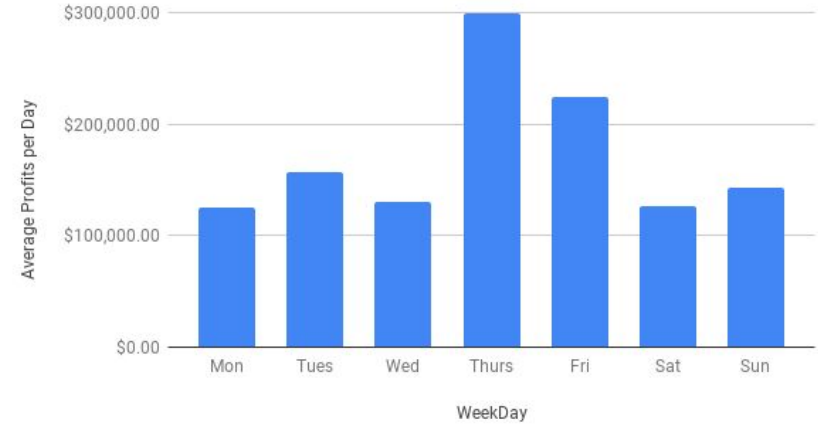SUM of Profits vs. Month
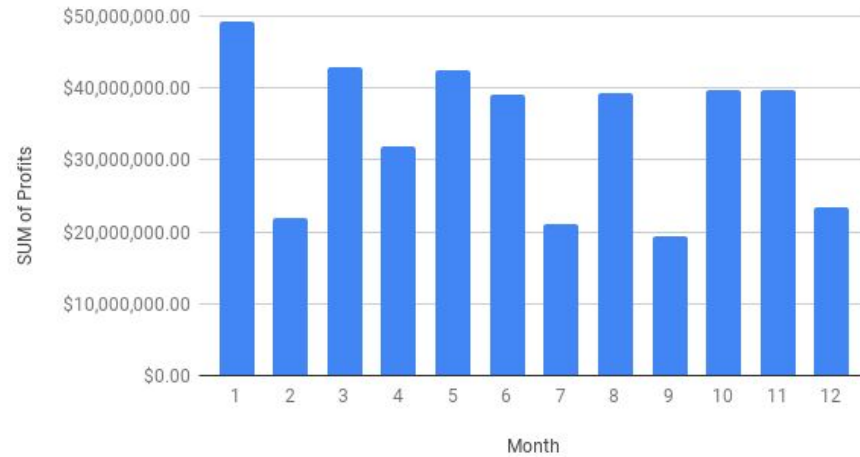
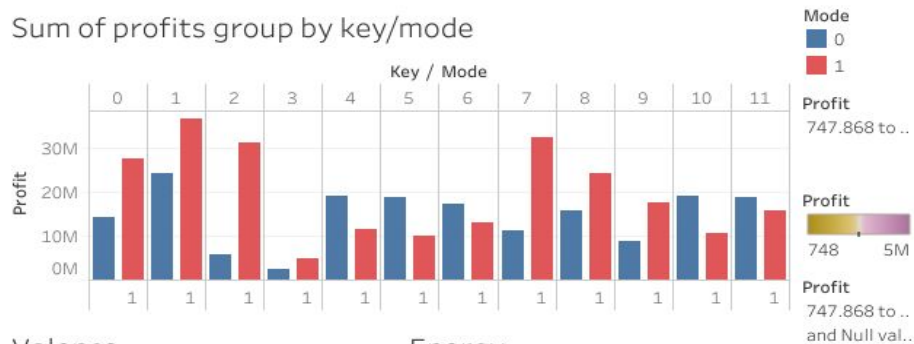Average song profits vs. Month
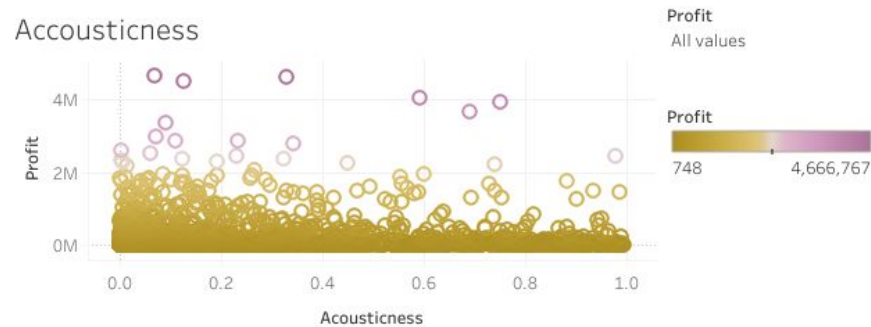
Results

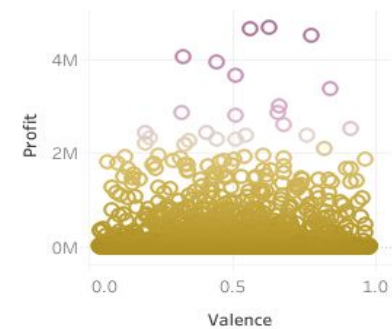SUM of Profits vs. WeekDay

Average Profits per Day vs. WeekDay

Sum of profits group by key/mode

Key / Mode

| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |

Mode
0
1

Profit
747.868 to ..

Profit
748   5M

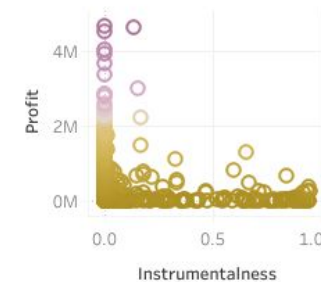Profit
747.868 to ..
and Null val..

Accousticness

Profit
All values

Profit
748   4,666,767

Valence
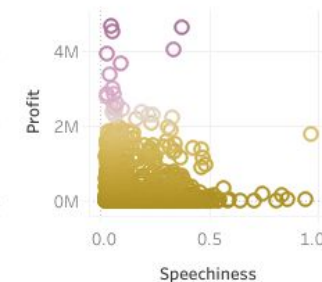
Energy
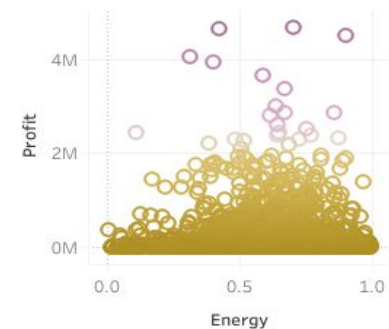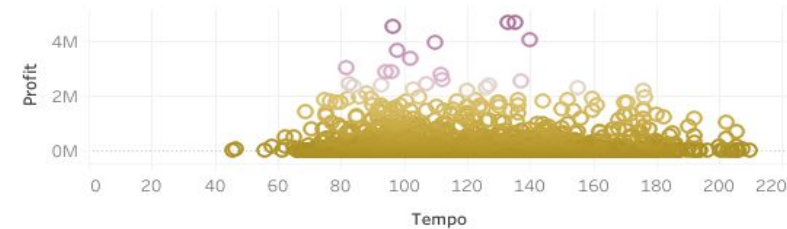
instrumentalness
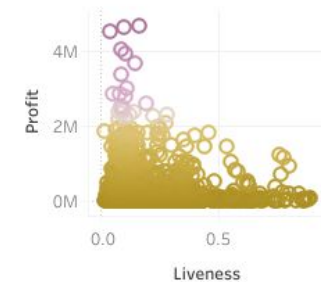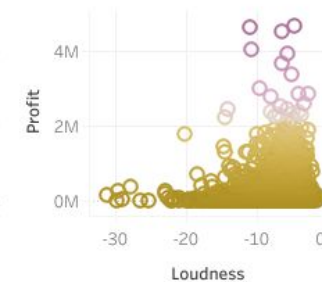
speechiness

Tempo

liveness

loudness

Conclusion

- Insights for me:
    - Popular songs follow a trend

- Insights for the Business :

    How to filter the demos based on its attributes

- Recommendation:
    - Create a regression model to predict the profits

Future Steps

- Build a prediction model that takes all the song attributes to generate an estimated profits.

- Generate a program which after uploading the song, it will automatically generate the values of each features