Capstone Proposal

Project Manager: Yeonjae Zhang.
Title: Data Scientist.
Date: 17th May, 2022

Future Music Trends

Business Understanding

I listen to music. There are hundreds of music trends. Depending on the trends, we tend to choose the music to listen to. Music trends are the most important factor in music money. Most people spend their time enjoying music online. That caused them to leave their data traces online. With the collected data, I can analyze music trends and I can even predict future music trends. The prediction will help experts in the music industry. I will work on the project for the music lovers.

Data Understanding

The data of spotify daily top 200 songs with genres from 2017 - 2021 is offered by spotify. The data has 317,262 rows. It is an unorganized dataframe. Each row includes rank, title, genres, and date in a column. So I have to extract the information to organize my dataset.

Data Preparation

I will separate the information by creating appropriate columns. # is used to separate the information in the rows. First, I will use the split function to separate the columns. Second, I will create new columns and save the information to the columns. Third, I will count genres in the top 200 songs daily. Finally, I will separate the dataset into each genre with the counts and dates.

Modeling

I will train the data with the time series model. I will figure out trends, seasonality, and residuals. Also, I have to plot ACF and PACF to identify Auto Regressive and Moving Average. I will train the SARIMA Model to predict future song trends.

Evaluation

I will report MSE and AIC to evaluate the model on training and test dataset. Because I don't want to get wrong predictions on music trends. If I don't report the evaluation score. Anyone can't trust my model resutls.

Deployment

I will compare with each genre SARIMA model prediction. And I will choose most counted genre as trends. This prediction will help musicians, music investors, and other music lovers.

Tools

Pandas, Statsmodels, Matplotlib, and Seaborn will be used for this project.