

Spotify History Data Report

The reason why I chose spotify data in this project is to find out the frequency of listening to music and understanding how my music interest has changed in the last 3 years.

Data source:

I sent a request to Spotify for my music history data and after 20 days, they sent me all the data for the last eight years and it contains every song started and time that song played. And I gathered the necessary parts of the document in a csv file.

Data analysis

I use different types of EDA, visualization and machine learning models. Firstly I combined the different documents into one csv file and also dropped the unnecessary information. And then I added some numerical columns for visualization and testing, training part. In the test training part I use hyperparameters to find the accuracies between music played time and year, month and day. In addition correlation of music played time found with other numerical items.

Findings

42.532 songs started on spotify in the last 3 years.

Use of shuffle rate is 12 percent.

Most listened 10 song:

KHRBR: 2172 min

Chicago Freestyle (feat. Giveon): 979 min

Nude: 844 min

Gerçekleşebilir: 800 min

Bad and Boujee (feat. Lil Uzi Vert): 784 min

Zehirli Melodiler: 781 min

MAFIA: 779 min

Hala Bir Umut Var: 741 min

Believe Me: 733 min

Holy Key (feat. Big Sean, Kendrick Lamar & Betty Wright): 728 min

Most listened 3 album:

Music To Be Murdered By: 3586 min

Certified Lover Boy: 3039 min

The Marshall Mathers LP2: 2250 min

Top 5 artist:

Eminem: 11506 min

Drake: 7008 min

Ati242: 4794 min

UZI: 2976 min

Ceg: 2779 min

most skipped song before 10 second:

Nude: 150 times

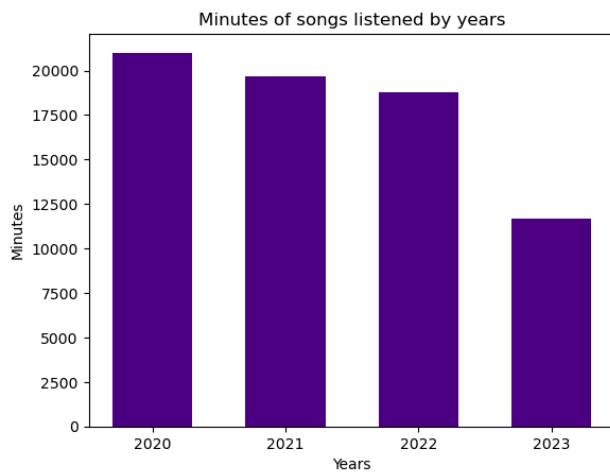
Gamzedeyim Deva Bulmam: 95 times

Aynen Öyle: 88 times

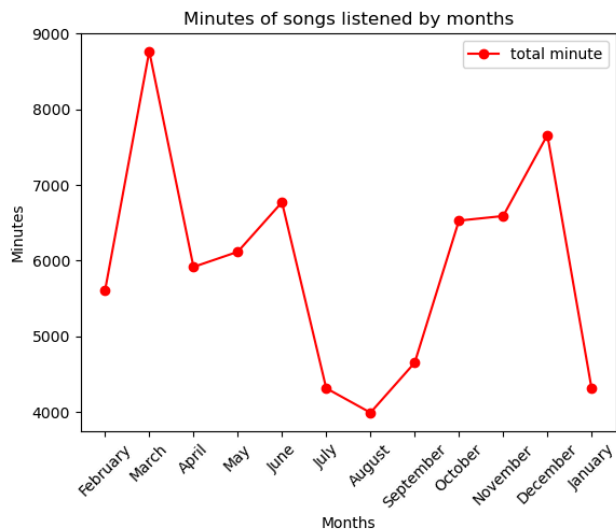
Shawty: 86 times

ne deve ne kush: 85 times

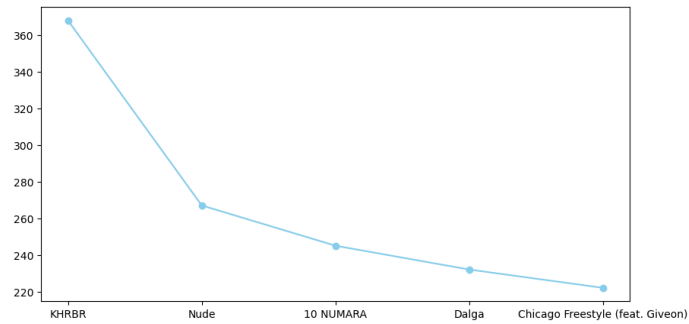
Minutes of songs listened by years:



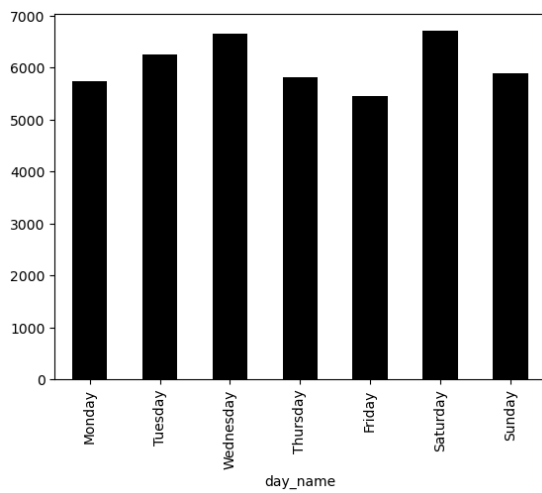
Minutes of songs listened by month:



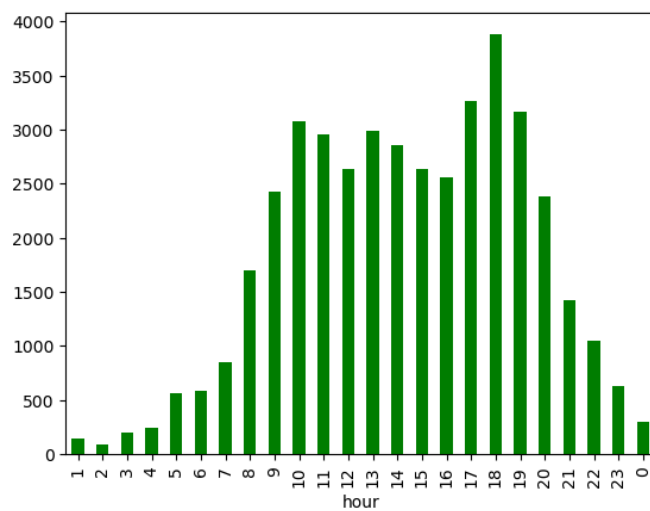
Most listened songs until the end:



Number of music played by day:



Number of music played by hour:



Correlation of played with year, month, day and hour:

year 0.021783
month 0.051492
day -0.008669
hour -0.036466

Test Accuracy = 0.011755025273304338

Classification Accuracy = 0.011755025273304338

Limitations and future work

If Spotify sent more detailed document, I might make a more exhaustive analysis. Also I don't have enough information for machine learning models and it limited my work.

For future work, I can improve the visualization by using more diverse types of bars and plots. And I might add detailed predictions using machine learning models.