

Spotify Top 200 Global Analysis (R)

Overview

This project focuses on performing **data cleaning**, **exploratory data analysis (EDA)**, and **clustering (K-Means & Hierarchical)** on Spotify's weekly Top 200 Global dataset.

The goal is to explore relationships between popularity, stream counts, song characteristics, and trends while generating meaningful insights.

Tools & Technologies

- R (tidyverse, ggplot2, dplyr, cluster)
- Data Cleaning
- Exploratory Data Analysis (EDA)
- Data Visualization
- K-Means & Hierarchical Clustering

Steps Performed

1. Loading the dataset
2. Cleaning missing or inconsistent values
3. Exploratory data analysis
4. Stream distribution analysis
5. Country/genre-based insights
6. K-Means clustering
7. Hierarchical clustering
8. Visualization of results

Key Insights

- Songs with high stream counts form distinct clusters.
- Genres and countries show meaningful grouping patterns.
- Clear weekly trends and fluctuations exist within the dataset.

Files

- spotify_analysis.R
- charts/ (visualizations)
- data.csv

Author

Busenur Durak

GitHub: github.com/busenur-durak

LinkedIn: [linkedin.com/in/busenur-durak](https://www.linkedin.com/in/busenur-durak)

Spotify Top 200 Global Analizi (R)

Proje Özeti

Bu proje, Spotify Top 200 Global haftalık müzik verileri üzerinde **veri temizleme, keşifsel veri analizi (EDA) ve kümeleme (K-Means & Hiyerarşik Clustering)** çalışmaları yapmak amacıyla geliştirilmiştir. Amaç; popülerlik, stream sayıları, şarkı özellikleri ve trend davranışları arasındaki ilişkileri incelemek ve veriden anlamlı içgörüler üretmektir.

Kullanılan Araçlar & Teknolojiler

- R (tidyverse, ggplot2, dplyr, cluster)
- Veri Temizleme
- EDA (Keşifsel Veri Analizi)
- Veri Görselleştirme
- K-Means & Hiyerarşik Kümeleme Analizi

Yapılan Adımlar

1. Veri setinin yüklenmesi
2. Eksik/veri hatalarının temizlenmesi
3. Keşifsel veri analizi
4. Stream sayılarının dağılımı
5. Tür ve ülke bazlı analizler
6. K-Means ile kümeleme
7. Hiyerarşik kümeleme
8. Bulguların görselleştirilmesi

Önemli Bulgular

- Stream sayıları en popüler kümelerde belirgin şekilde yoğunlaşmıştır.
- Benzer müzik türleri ve ülkeler belirli segmentlerde gruplanmıştır.
- Trendlerde haftalık olarak anlamlı dalgalanmalar görülmüştür.

Dosyalar

- spotify_analysis.R
- charts/ → grafikler
- data.csv → analiz edilen veri

 **Yazar**

Busenur Durak

GitHub: github.com/busenur-durak

LinkedIn: linkedin.com/in/busenur-durak