

# Analisis Visualisasi Data Netflix Movies and Shows Tahun 2022

Muhammad Mitchell Tri Octaviano Syaifullah







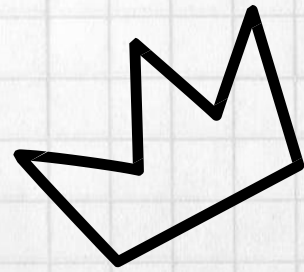
# Introduction

Netflix adalah sebuah platform layanan streaming film yang berbasis di Amerika. Didirikan pada tahun 1997, kini Netflix sudah berkembang sangat pesat dan hampir tersedia di seluruh negara.

Terdapat berbagai macam film yang tersedia dengan berbagai judul, genre, imdb score, dan lainnya. Dengan dilakukannya visualisasi data terhadap Movies dan Show yang ada di Netflix, diharapkan dapat mempermudah user untuk mencari tontonan yang tepat sesuai dengan keinginannya.

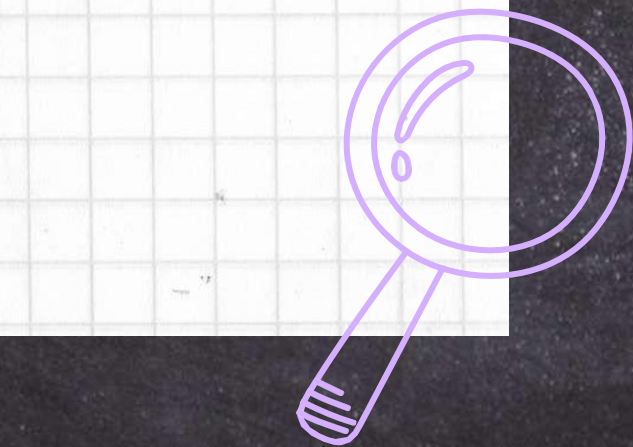






# Dataset

Pada project kali ini, saya menggunakan dataset Netflix Movies and Shows yang bersumber dari kaggle dan diolah menggunakan Google Collabs. Data ini terdiri dari 11 kolom yang berisi Judul, Tipe, Tahun Rilis, Durasi, Genre, Negara, imdb id, imdb score, dan imdb votes.



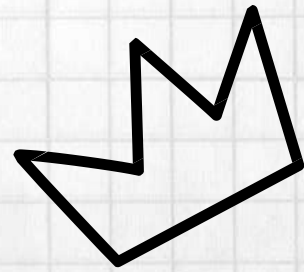


# Praprocessing Data

Pada Dataset Netflix Movies and Shows, sebelum saya melakukan visualisasi data, saya melakukan praproses data terlebih dahulu. Dari gambar disamping, terlihat masih ada missing value pada kolom title, imdb\_id, imdb\_score, dan imdb\_votes. Untuk kolom title dan imdb\_id saya mengganti nilai missing value dengan frekuensi. Namun, untuk imdb\_score dan imdb\_votes saya mengganti nilai missing value dengan rata-rata dikarenakan kolom imdb\_score dan imdb\_votes merupakan tipe data numerik

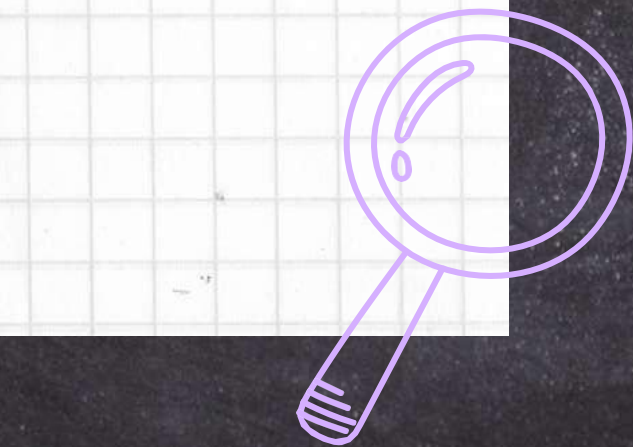
```
df.isnull().sum()
title      1
type       0
release_year 0
runtime    0
genres     0
production_countries 0
imdb_id    444
imdb_score 523
imdb_votes 539
dtype: int64
```





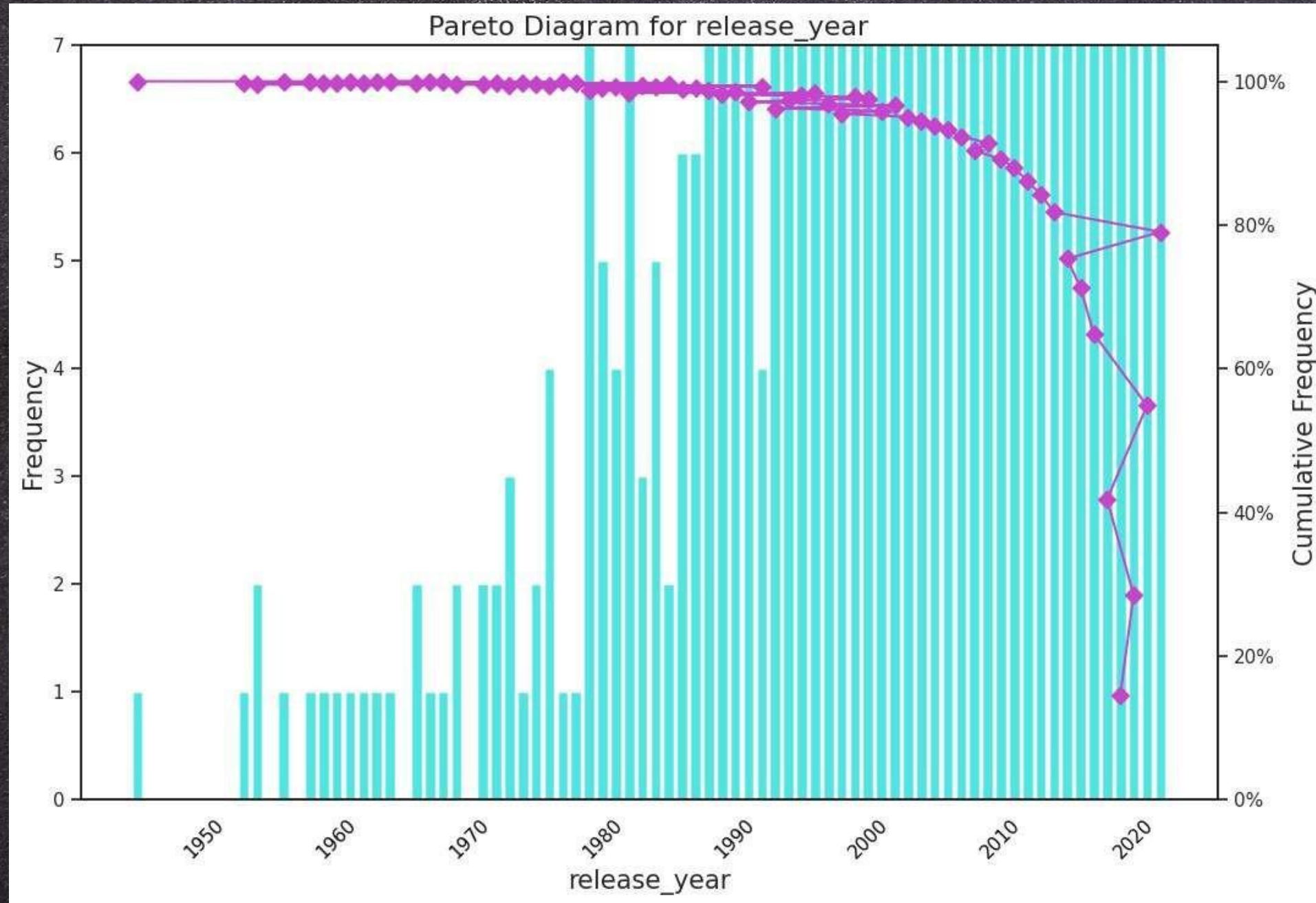
# Process

Saya melakukan analisa data Netflix untuk ditampilkan beberapa visualisasi, seperti Pareto Diagram untuk Tahun Rilis, Top 7 Genre dalam Netflix, imdb score terhadap Show dan Movies, Bar Chart perbandingan banyaknya film dengan durasi, serta Donat Chart perbandingan banyaknya Movies dan Show.





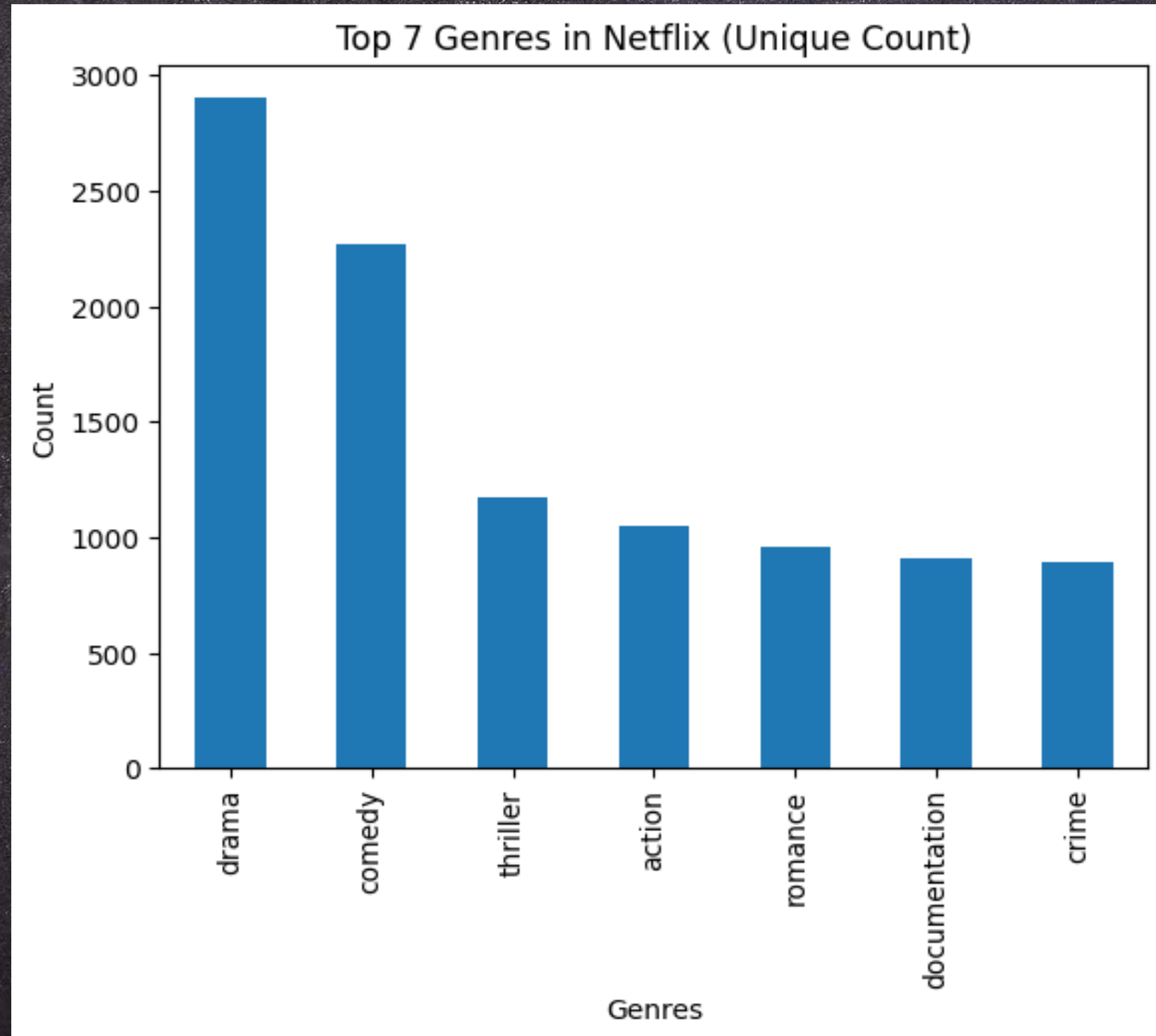
# Pareto diagram



Pada pareto diagram disamping, kita bisa melihat distribusi tahun rilis dari netflix movies and show yang dimana penyebaran data netflix movies and show kebanyakan di rentang tahun 2010-2020. Dan pada tahun 2019 merupakan tahun rilis dari netflix movies and show yang terbanyak.



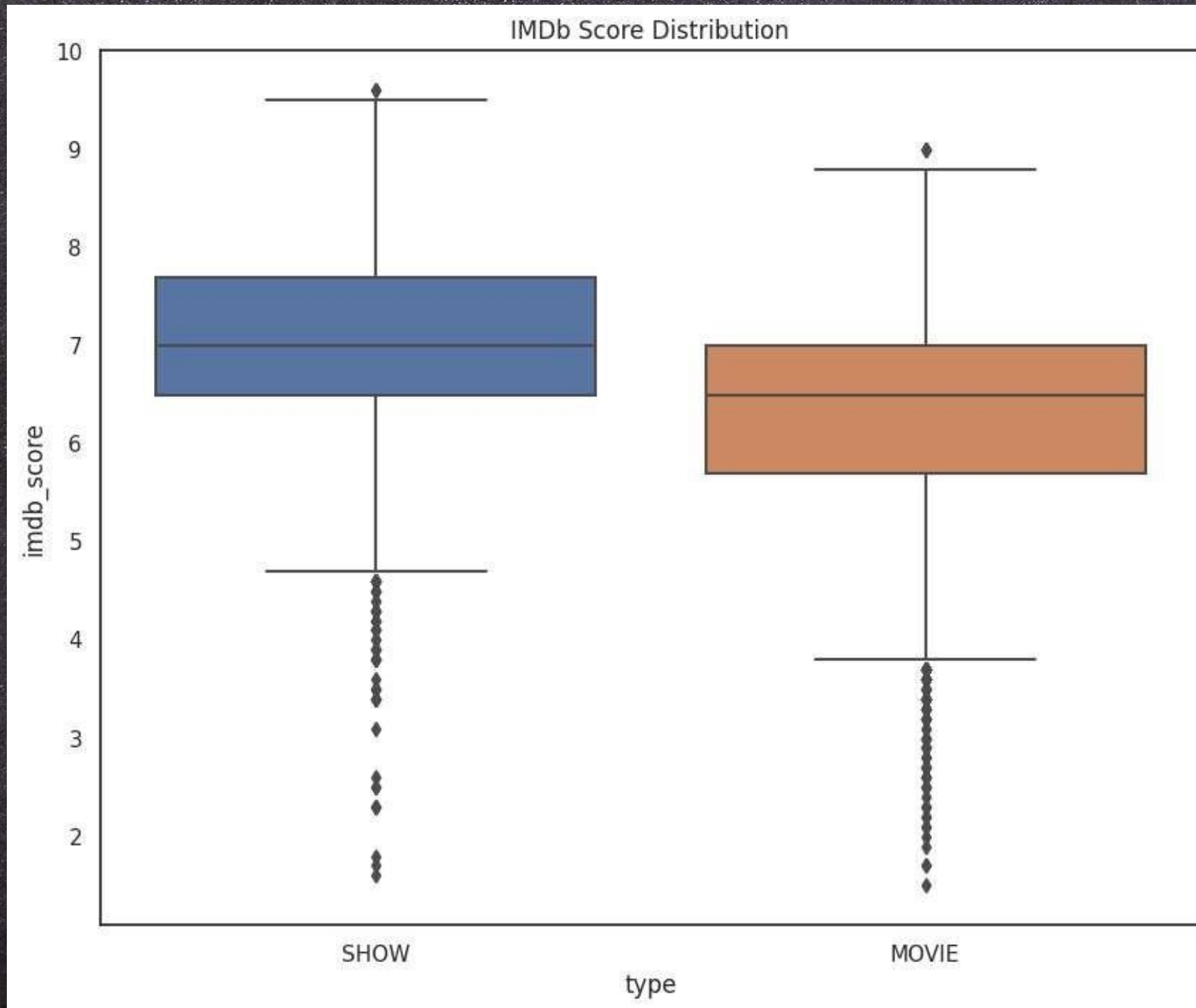
# Top 7 genre in Netflix



Pada data diagram batang di samping, terdapat 7 kategori genre yang terdapat pada film netflix yang sering di tonton dan yang paling sering di tonton yaitu genre drama dengan jumlah sekitar lebih dari 2800 film.



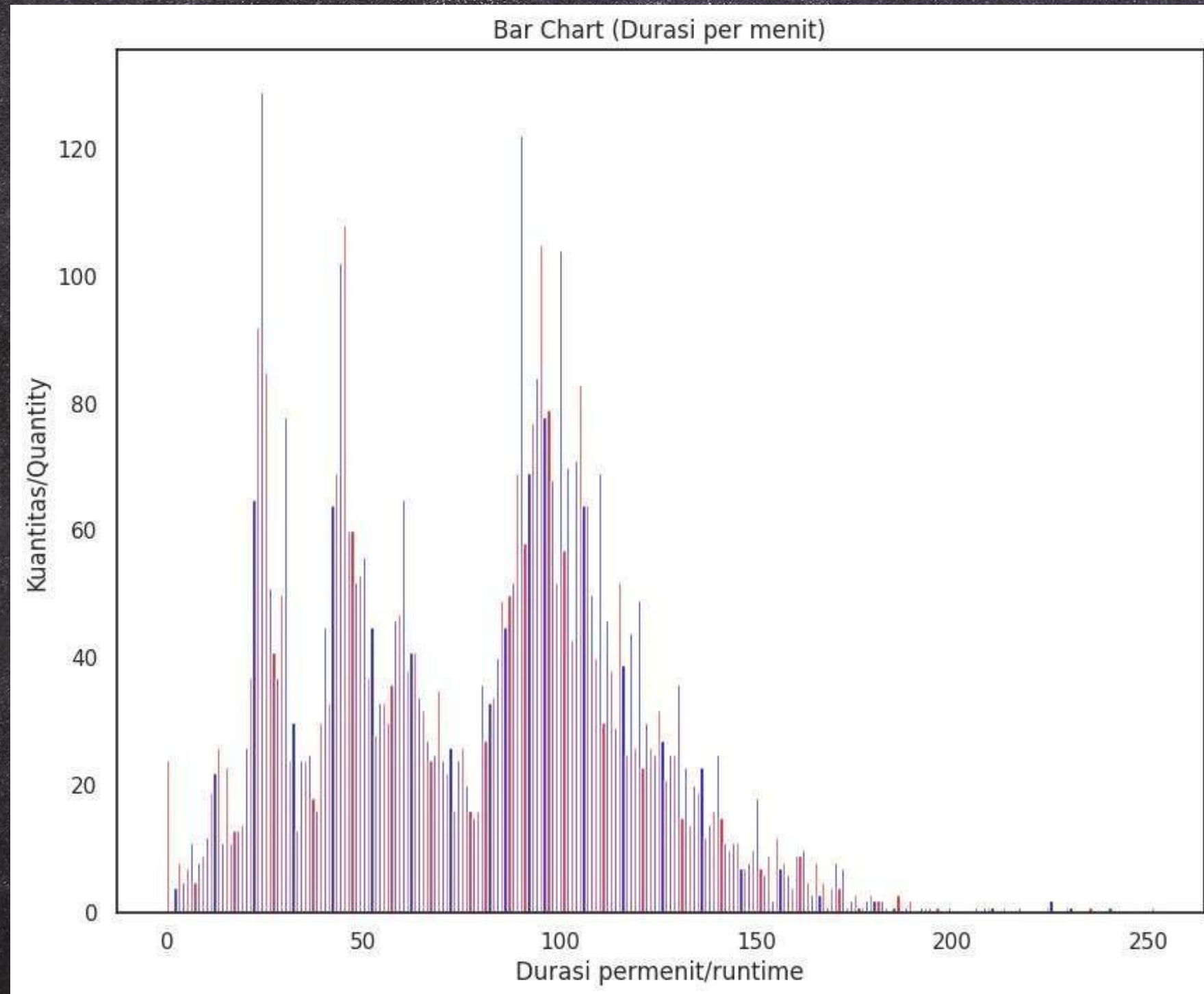
# Box whisker Plot



Pada boxplot disamping, kita bisa melihat distribusi imdb score dari masing-masing type film netflixnya. Pada type show, distribusi imdb score berada pada rentang 6,5 - 7,5. Sedangkan, pada type movie, distribusi imdb score berada pada rentang 5,5 - 7.



# Barchart

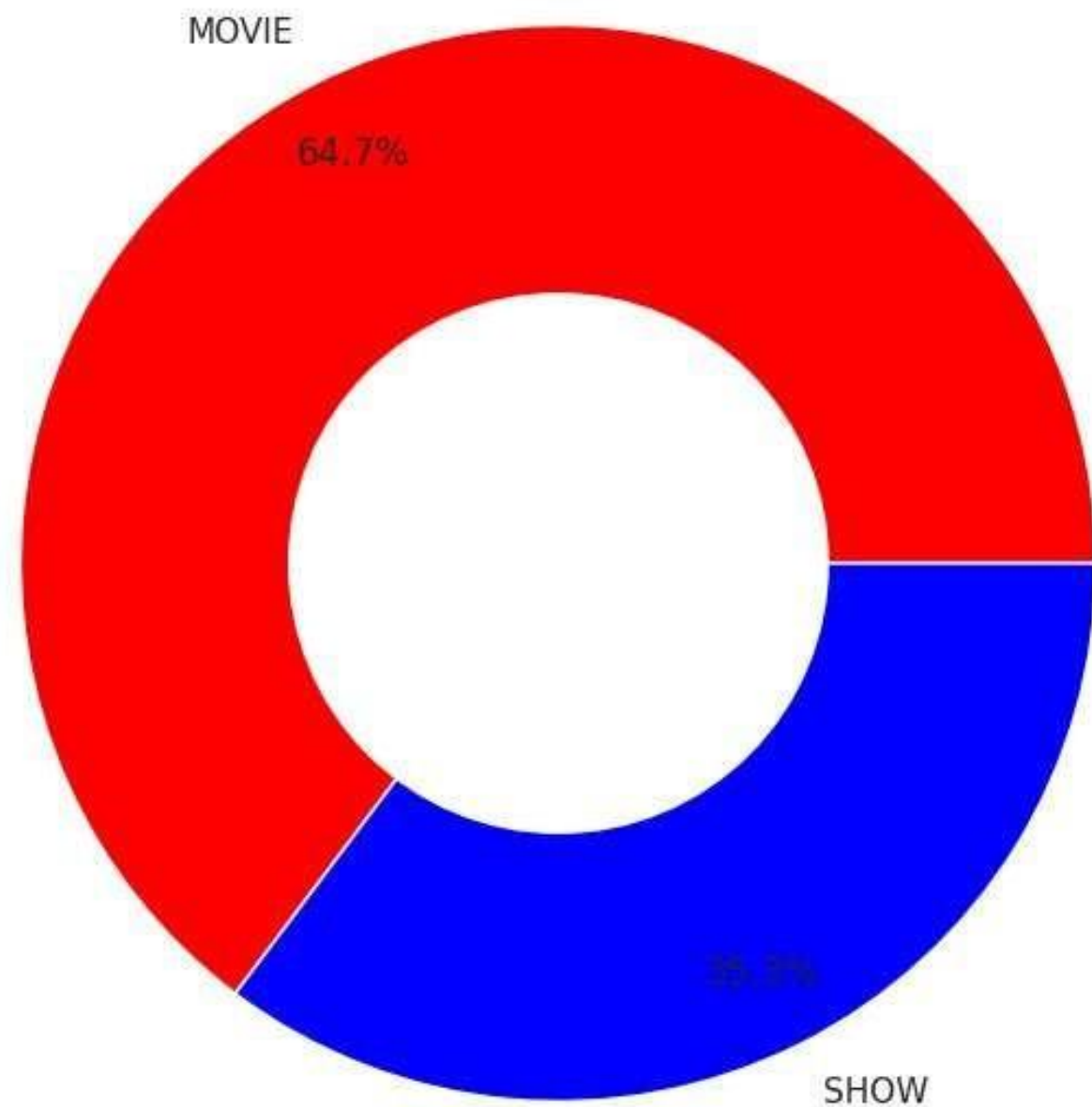


Pada Barchart disamping, kita bisa melihat durasi waktu dari setiap movie and show yang dimana penyebaran data netflix movies and show kebanyakan pada durasi 0-150 menit. Lalu pada durasi sekitar 25 menit merupakan durasi dengan film terbanyak yang melebihi 120 film baik movie maupun show.




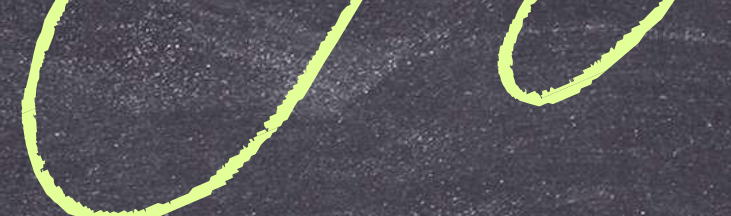
# Donuts Chart

Donat Chart for Percentage of Type Netflix Show and Movies

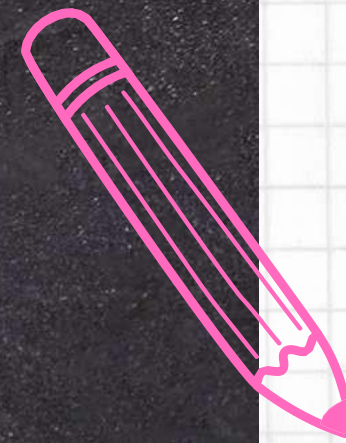
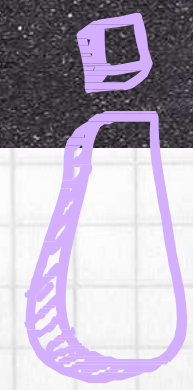


Pada Donat chart disamping, kita bisa melihat dari 5000 data yang ada, tipe film movie merupakan data terbanyak sebesar 64.7% yang ada di Netflix dan show lebih sedikit dari movie yaitu sebesar 35.3%.







# Kesimpulan



Kita bisa mengetahui bahwa dari 5000 lebih data Film-film di Netflix pada tahun 2022, film dengan tipe Movie mendominasi Netflix dibandingkan film dengan tipe Show. Lalu film dengan genre comedy merupakan film terbanyak ditonton baik tipe movie maupun show. Jadi daya jual Netflix lebih mendominasi pada Movie yang bergenre Comedy. Hal ini menunjukkan bahwa banyak penonton netflix tertarik dengan alur cerita yang kuat dan emosional.





THANK

YOU

