PROJECT: INVESTIGATING NETFLIX MOVIES

Netflix! What started in 1997 as a DVD rental service has since exploded into one of the largest entertainment and media companies.

Given the large number of movies and series available on the platform, it is a perfect opportunity to flex your exploratory data analysis skills and dive into the entertainment industry. Our friend has also been brushing up on their Python skills and has taken a first crack at a CSV file containing Netflix data. They believe that the average duration of movies has been declining. Using your friends initial research, you'll delve into the Netflix data to see if you can determine whether movie lengths are actually getting shorter and explain some of the contributing factors, if any.

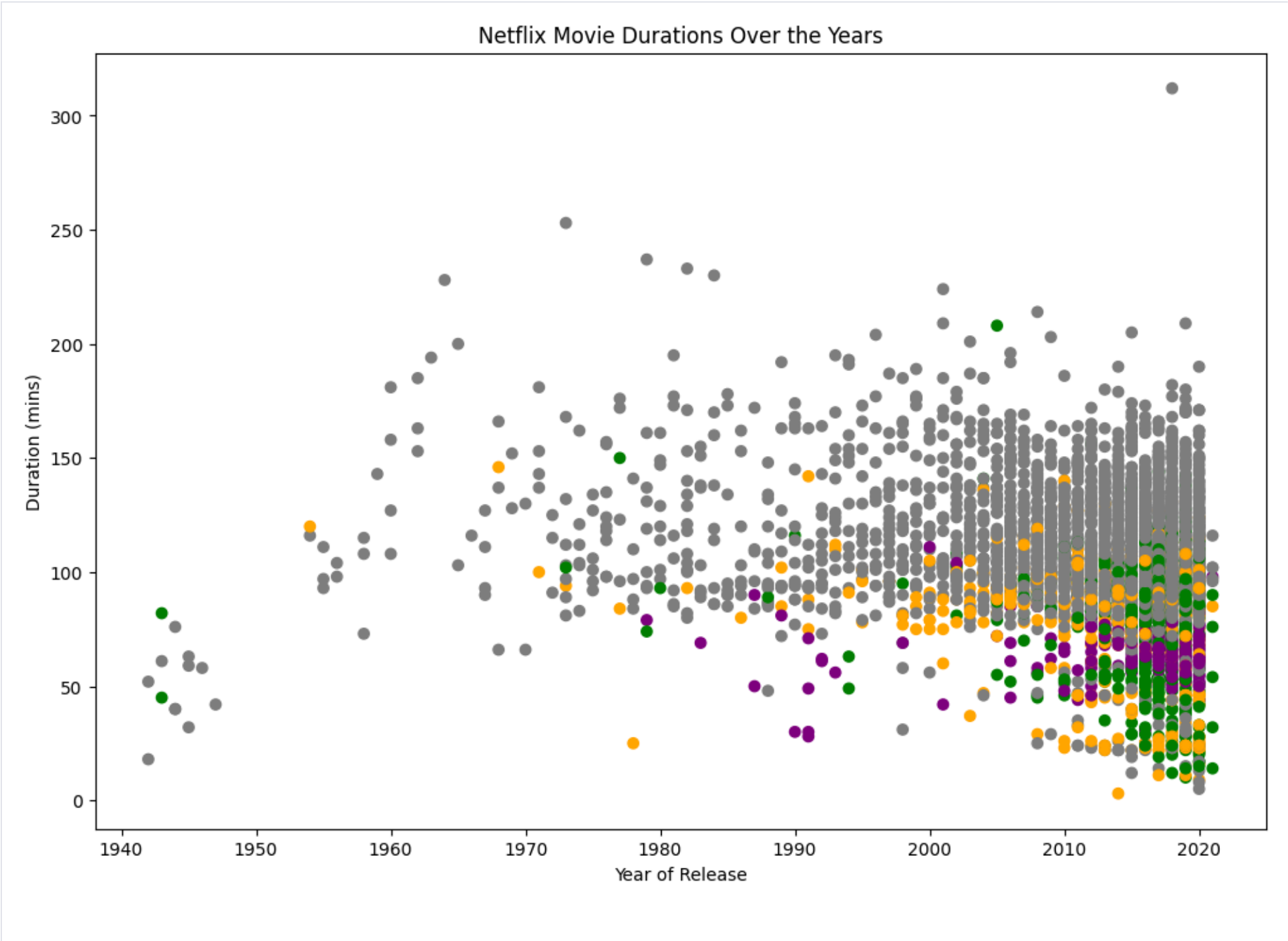
You have been supplied with the dataset `netflix_data.csv` , along with the following table detailing the column names and descriptions:

The data

netflix_data.csv

Column	Description
`show_id`	The ID of the show
`type`	Type of show
`title`	Title of the show
`director`	Director of the show
`cast`	Cast of the show
`country`	Country of origin
`date_added`	Date added to Netflix
`release_year`	Year of Netflix release
`duration`	Duration of the show in minutes
`description`	Description of the show
`genre`	Show genre

```
Data loaded successfully.
Movies filtered from the dataset.
Relevant columns selected.
Identified 420 short movies (duration < 60 mins).
Colors assigned based on movie genres.
First 10 color assignments: 1      grey
2      grey
3      grey
4      grey
6      grey
7      grey
8      grey
9      grey
10     grey
13    green
Name: color, dtype: object
```


Netflix Movie Durations Over the Years
Duration (mins)
0
50
100
150
200
250
300
1940
1950
1960
1970
1980
1990
2000
2010
2020
Year of Release

```
Scatter plot created for movie durations.
Conclusion: Trend analysis required to determine duration changes.
```