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.

You work for a production company that specializes in nostalgic styles. You want to do some research on movies released in the 1990's. You'll delve into Netflix data and perform exploratory data analysis to better understand this awesome movie decade!

You have been supplied with the dataset `netflix_data.csv`, along with the following table detailing the column names and descriptions. Feel free to experiment further after submitting!

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

```python
# Importing pandas and matplotlib
import pandas as pd
import matplotlib.pyplot as plt

# Read in the Netflix CSV as a DataFrame
netflix_df = pd.read_csv("netflix_data.csv")
netflix_df.head()
```

	...	↑↓	...	↑↓	...	↑↓	...	↑↓	director	...	↑↓	cast	...	↑↓	country	...	↑↓	date_added	...	↑↓	relea...
0	s2		Movie		7:19		Jorge Michel Grau		Demián Bichir, Héctor Bonilla, Oscar Serrano,...		Mexico		December 23, 2016	
1	s3		Movie		23:59		Gilbert Chan		Tedd Chan, Stella Chung, Henley Hii, Lawren...		Singapore		December 20, 2018	
2	s4		Movie		9		Shane Acker		Elijah Wood, John C. Reilly, Jennifer Connelly...		United States		November 16, 2017	
3	s5		Movie		21		Robert Luketic		Jim Sturgess, Kevin Spacey, Kate Bosworth, A...		United States		January 1, 2020	
4	s6		TV Show		46		Serdar Akar		Erdal Beşikçioğlu, Yasemin Allen, Melis Birka...		Turkey		July 1, 2017	

Rows: 5

```python
# filter the data for movies released in the 1990s
import pandas as pd
import matplotlib.pyplot as plt

# Read in the NETFLIX CSV as a DataFrame
netflix_df = pd.read_csv("netflix_data.csv")
# subset and filtering dataframe column by values
netflix_90s_df = netflix_df[(netflix_df["release_year"] >= 1990) & (netflix_df["release_year"] <= 1999)]
# sort release_year by ascending
# the most frequent movie by duration in 1990s
netflix_90s_df.sort_values(by="duration", ascending=False)
```

			title			director			cast		
3387	s5431	Movie	Schindler's List			Steven Spielberg			Liam Neeson, Ben Kingsley, Ralph Fiennes, C
1785	s2860	Movie	Hum Aapke Hain Koun			Sooraj R. Barjatya			Madhuri Dixit, Salman Khan, Mohnish Bahl, R
4746	s7682	Movie	Wyatt Earp			Lawrence Kasdan			Kevin Costner, Dennis Quaid, Gene Hackman
2415	s3861	Movie	Magnolia			Paul Thomas Anderson			Jeremy Blackman, Tom Cruise, Melinda Dillo.
2985	s4774	Movie	Pardes			Subhash Ghai			Shah Rukh Khan, Amrish Puri, Mahima Chau.
2192	s3461	Movie	Kuch Kuch Hota Hai			Karan Johar			Shah Rukh Khan, Kajol, Rani Mukerji, Farida J
3704	s5954	Movie	Taal			Subhash Ghai			Anil Kapoor, Aishwarya Rai Bachchan, Aksha
3160	s5078	Movie	Raja Hindustani			Dharmesh Darshan			Aamir Khan, Karisma Kapoor, Suresh Oberoi,
4436	s7130	Movie	Together For Eternity			Sooraj R. Barjatya			Salman Khan, Karisma Kapoor, Saif Ali Khan,
1786	s2861	Movie	Hum Saath-Saath Hain			Sooraj R. Barjatya			Salman Khan, Karisma Kapoor, Saif Ali Khan,
240	s344	Movie	Agneepath			Mukul Anand			Amitabh Bachchan, Mithun Chakraborty, Da.
4490	s7224	Movie	Trimurti			Mukul Anand			Anil Kapoor, Shah Rukh Khan, Jackie Shroff, .
2450	s3913	Movie	Mann			Indra Kumar			Aamir Khan, Manisha Koirala, Sharmila Tagor
2124	s3354	Movie	Khalnayak			Subhash Ghai			Sanjay Dutt, Madhuri Dixit, Jackie Shroff, An.
3483	s5571	Movie	Shiva			Ram Gopal Varma			Nagarjuna Akkineni, Amala, Rajendranath Zu
1120	s1761	Movie	Dil Se			Mani Ratnam			Shah Rukh Khan. Manisha Koirala. Preitu Zint

Rows: 184

```python
# find the most frequent movie duration
import pandas as pd
import matplotlib.pyplot as plt

# Read in the NETFLIX CSV as a DataFrame
netflix_df = pd.read_csv("netflix_data.csv")
# subset and filtering dataframe column by values
netflix_90s_df = netflix_df[(netflix_df["release_year"] >= 1990) & (netflix_df["release_year"] <= 1999)]
# sort release_year by ascending
netflix_90s_df = netflix_90s_df.sort_values(by="release_year", ascending=True)
# the most frequent movie duration
netflix_90s_df = netflix_90s_df.sort_values(by="duration", ascending=False)
# create a histogram
plt.figure(figsize=(20, 16))
plt.hist(netflix_90s_df["duration"], bins=10, alpha=0.7)

# Add titles and labels
plt.title("Histogram of Movie Durations in the 90s", fontsize=16)
plt.xlabel("Duration (minutes)", fontsize=14)
plt.ylabel("Frequency", fontsize=14)

# Show the plot
plt.show()

# Find the duration of the movie "Schindler's List"
duration = netflix_90s_df[(netflix_90s_df["type"] == "Movie") & (netflix_90s_df["title"] == "Schindler's List")]
duration = 100
```

Histogram of Movie Durations in the 90s