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

```
# 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")
```

```
# Question 1: What was the most frequent movie duration in the 1990s? Save an approximate answer as an integer called duration (use 1990 as the decade's start year).
movies=netflix_df[netflix_df["type"] == "Movie"] #created data set that shows the only netfix Movies
movies_1990 = movies[(movies['release_year'] >= 1990) & (movies['release_year'] < 2000)] #created new data set from previously created data set where only movies from the 1990's are selected
duration_1 = movies_1990['duration'].mode() #selected the duration from my new data set that showed up the most using mode function
duration = duration_1.iloc[0] #selected the top row as that is the most frequent duration of movies from the 1990's
print(duration)
print(type(duration)) #made sure saved as an integer

#Question 2: A movie is considered short if it is less than 90 minutes. Count the number of short action movies released in the 1990s and save this integer as short_movie_count.
action_movies_1990s = movies_1990[movies_1990["genre"] == "Action"] #selected data set from above that was already created to show data of only Movies from 1990's. Refined the data to only include movies who's genre was action.
short_movie_count = 0
for label, row in action_movies_1990s.iterrows():
    if row["duration"] < 90:
        short_movie_count = short_movie_count + 1
    else:
        short_movie_count = short_movie_count #created if/else case statement that went through each row of my new data set starting at 0. if the movies duration was less than 90, it would add +1 to teh count. If anything else, it would keep the count where it was at.
print(short_movie_count)
```

```
94
<class 'numpy.int64'>
7
```