

```
import requests

from bs4 import BeautifulSoup

import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

# Function to scrape movie data from IMDb
def scrape_imdb_data(url):
    response = requests.get(url)

    soup = BeautifulSoup(response.content, 'html.parser')

    movies = []

    for movie in soup.find_all('div', class_='lister-item-content'):
        title = movie.find('h3', class_='lister-item-header').find('a').text.strip()

        rating = float(movie.find('div', class_='inline-block ratings-imdb-
rating').find('strong').text.strip())

        year = int(movie.find('span', class_='lister-item-year').text.strip('('))

        movies.append({'Title': title, 'Rating': rating, 'Year': year})

    return movies

# URL of IMDb's Top Rated Movies page
url = 'https://www.imdb.com/chart/top/'

# Scrape the data
movies_data = scrape_imdb_data(url)

# Convert the scraped data into a DataFrame
df = pd.DataFrame(movies_data)

# Display the first few rows of the DataFrame
print(df.head())
```

```
# Check basic statistics of numerical columns
print(df.describe())

# Data visualization

# Histogram of movie ratings
plt.figure(figsize=(10, 6))
sns.histplot(df['Rating'], bins=20, kde=True)
plt.title('Distribution of Movie Ratings')
plt.xlabel('Rating')
plt.ylabel('Frequency')
plt.show()

# Box plot of movie ratings
plt.figure(figsize=(10, 6))
sns.boxplot(df['Rating'])
plt.title('Boxplot of Movie Ratings')
plt.xlabel('Rating')
plt.show()

# Line plot of average movie ratings over the years
avg_rating_year = df.groupby('Year')['Rating'].mean().reset_index()
plt.figure(figsize=(10, 6))
sns.lineplot(x='Year', y='Rating', data=avg_rating_year)
plt.title('Average Movie Ratings Over the Years')
plt.xlabel('Year')
plt.ylabel('Average Rating')
plt.show()
```