In [42]:
```python
import pandas as pd
pd.set_option('display.max.rows', 10)
```

In [43]:
```python
df = pd.read_csv(r"C:\Users\kallzz\Desktop\Data Analytics Stuff\Data Analyst - Boot
```

In [44]:
```python
df
```

Out[44]:

|     | Country | Region |
| --- | --- | --- |
| 0 | Afghanistan | ASIA (EX. NEAR EAST) |
| 1 | Albania | EASTERN EUROPE |
| 2 | Algeria | NORTHERN AFRICA |
| 3 | American Samoa | OCEANIA |
| 4 | Andorra | WESTERN EUROPE |
| ... | ... | ... |
| 222 | West Bank | NEAR EAST |
| 223 | Western Sahara | NORTHERN AFRICA |
| 224 | Yemen | NEAR EAST |
| 225 | Zambia | SUB-SAHARAN AFRICA |
| 226 | Zimbabwe | SUB-SAHARAN AFRICA |

227 rows × 2 columns

In [45]:
```python
# read_csv can be used to read text files but not in a proper way
df = pd.read_csv(r"C:\Users\kallzz\Desktop\Data Analytics Stuff\Data Analyst - Boot
```

In [46]:
```python
df
```

Out[46]:

|     | Country\tRegion |
| --- | --- |
| 0 | Afghanistan \tASIA (EX. NEAR EAST) |
| 1 | Albania \tEASTERN EUROPE |
| 2 | Algeria \tNORTHERN AFRICA |
| 3 | American Samoa \tOCEANIA ... |
| 4 | Andorra \tWESTERN EUROPE |
| ... | ... |
| 222 | West Bank \tNEAR EAST |
| 223 | Western Sahara \tNORTHERN AFRICA ... |
| 224 | Yemen \tNEAR EAST |
| 225 | Zambia \tSUB-SAHARAN AFRICA |
| 226 | Zimbabwe \tSUB-SAHARAN AFRICA |

227 rows × 1 columns

In [47]:
```python
# there is \t delimeter in text file

df = pd.read_csv(r"C:\Users\kallzz\Desktop\Data Analytics Stuff\Data Analyst - Boot
df
```

Out[47]:

|     | Country | Region |
| --- | --- | --- |
| 0 | Afghanistan | ASIA (EX. NEAR EAST) |
| 1 | Albania | EASTERN EUROPE |
| 2 | Algeria | NORTHERN AFRICA |
| 3 | American Samoa | OCEANIA |
| 4 | Andorra | WESTERN EUROPE |
| ... | ... | ... |
| 222 | West Bank | NEAR EAST |
| 223 | Western Sahara | NORTHERN AFRICA |
| 224 | Yemen | NEAR EAST |
| 225 | Zambia | SUB-SAHARAN AFRICA |
| 226 | Zimbabwe | SUB-SAHARAN AFRICA |

227 rows × 2 columns

In [48]:
```python
df = pd.read_table(r"C:\Users\kallzz\Desktop\Data Analytics Stuff\Data Analyst - Bc
df
```

Out[48]:

|     | Country | Region |
| --- | --- | --- |
| 0 | Afghanistan | ASIA (EX. NEAR EAST) |
| 1 | Albania | EASTERN EUROPE |
| 2 | Algeria | NORTHERN AFRICA |
| 3 | American Samoa | OCEANIA |
| 4 | Andorra | WESTERN EUROPE |
| ... | ... | ... |
| 222 | West Bank | NEAR EAST |
| 223 | Western Sahara | NORTHERN AFRICA |
| 224 | Yemen | NEAR EAST |
| 225 | Zambia | SUB-SAHARAN AFRICA |
| 226 | Zimbabwe | SUB-SAHARAN AFRICA |

227 rows × 2 columns

In [49]:
```python
# read_table can be used to read data from different types of file format but we ne
df = pd.read_table(r"C:\Users\kallzz\Desktop\Data Analytics Stuff\Data Analyst - Bc
df
```

Out[49]:                                      [

| | |
|---|---|
| **0** | { |
| **1** | "12 Strong": { |
| **2** | "Genre": "Action", |
| **3** | "Gross": "$453,173", |
| **4** | "IMDB Metascore": "54", |
| **...** | ... |
| **236** | "Rating": "PG13", |
| **237** | "Tomato Score": 12 |
| **238** | } |
| **239** | } |
| **240** | ] |

241 rows × 1 columns

In [50]:
```
df = pd.read_json(r"C:\Users\kallzz\Desktop\Data Analytics Stuff\Data Analyst - Boo
df
```

Out[50]:

| | 12 Strong | A Fantastic Woman (Una Mujer Fantástica) | All The Money In The World | Bilal: A New Breed Of Hero | Call Me By Your Name | Darkest Hour | D T |
|---|---|---|---|---|---|---|---|
| **0** | {'Genre': 'Action', 'Gross': '$453,173', 'IMDB... | {'popcornscore': 83, 'rating': 'R', 'tomatosco... | {'popcornscore': 71, 'rating': 'R', 'tomatosco... | {'popcornscore': 91, 'rating': 'PG13', 'tomato... | {'popcornscore': 87, 'rating': 'R', 'tomatosco... | {'popcornscore': 84, 'rating': 'PG13', 'tomato... | {'C 'A ' '$49 'I |
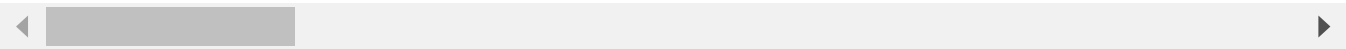
In [51]:
```
# to view entire row and column data in pandas
# pd.set_option('display.max.rows', 260)
# pd.set_option('display.max.columns', 50)
```

In [52]:
```
# read_excel by defaults reads the first sheet data
df = pd.read_excel(r"C:\Users\kallzz\Desktop\Data Analytics Stuff\Data Analyst - Bo
df
```

Out[52]:

| | Rank | CCA3 | Country | Capital |
|---|---|---|---|---|
| **0** | 36 | AFG | Afghanistan | Kabul |
| **1** | 138 | ALB | Albania | Tirana |
| **2** | 34 | DZA | Algeria | Algiers |
| **3** | 213 | ASM | American Samoa | Pago Pago |
| **4** | 203 | AND | Andorra | Andorra la Vella |
| **...** | ... | ... | ... | ... |
| **229** | 226 | WLF | Wallis and Futuna | Mata-Utu |
| **230** | 172 | ESH | Western Sahara | El AaiÃºn |
| **231** | 46 | YEM | Yemen | Sanaa |
| **232** | 63 | ZMB | Zambia | Lusaka |
| **233** | 74 | ZWE | Zimbabwe | Harare |

234 rows × 4 columns

In [53]:
```python
df.info()
# Provides information about non-null data and datatype and memory size
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 234 entries, 0 to 233
Data columns (total 4 columns):
 #   Column   Non-Null Count  Dtype
---  ------   --------------  -----
 0   Rank     234 non-null    int64
 1   CCA3     234 non-null    object
 2   Country  234 non-null    object
 3   Capital  234 non-null    object
dtypes: int64(1), object(3)
memory usage: 7.4+ KB
```

In [54]:
```python
df.head()
```

Out[54]:

| | Rank | CCA3 | Country | Capital |
|---|---|---|---|---|
| **0** | 36 | AFG | Afghanistan | Kabul |
| **1** | 138 | ALB | Albania | Tirana |
| **2** | 34 | DZA | Algeria | Algiers |
| **3** | 213 | ASM | American Samoa | Pago Pago |
| **4** | 203 | AND | Andorra | Andorra la Vella |

In [55]:
```python
df.head(10)
```

Out[55]:

| | Rank | CCA3 | Country | Capital |
|---|---|---|---|---|
| **0** | 36 | AFG | Afghanistan | Kabul |
| **1** | 138 | ALB | Albania | Tirana |
| **2** | 34 | DZA | Algeria | Algiers |
| **3** | 213 | ASM | American Samoa | Pago Pago |
| **4** | 203 | AND | Andorra | Andorra la Vella |
| **5** | 42 | AGO | Angola | Luanda |
| **6** | 224 | AIA | Anguilla | The Valley |
| **7** | 201 | ATG | Antigua and Barbuda | Saint Johnâ€™s |
| **8** | 33 | ARG | Argentina | Buenos Aires |
| **9** | 140 | ARM | Armenia | Yerevan |

In [56]:
```python
df.tail()
```

Out[56]:

| | Rank | CCA3 | Country | Capital |
|---|---|---|---|---|
| **229** | 226 | WLF | Wallis and Futuna | Mata-Utu |
| **230** | 172 | ESH | Western Sahara | El AaiÃºn |
| **231** | 46 | YEM | Yemen | Sanaa |
| **232** | 63 | ZMB | Zambia | Lusaka |
| **233** | 74 | ZWE | Zimbabwe | Harare |

In [57]:
```python
# to view a specific column data
df['Rank']
```

Out[57]:
```
0        36
1       138
2        34
3       213
4       203
       ...
229     226
230     172
231      46
232      63
233      74
Name: Rank, Length: 234, dtype: int64
```

In [58]:
```python
df.columns
```

Out[58]:
```
Index(['Rank', 'CCA3', 'Country', 'Capital'], dtype='object')
```

In [59]:
```python
df2 = df[['Rank', 'CCA3']]
df2
```

Out[59]:

| | Rank | CCA3 |
|---|---|---|
| **0** | 36 | AFG |
| **1** | 138 | ALB |
| **2** | 34 | DZA |
| **3** | 213 | ASM |
| **4** | 203 | AND |
| **...** | ... | ... |
| **229** | 226 | WLF |
| **230** | 172 | ESH |
| **231** | 46 | YEM |
| **232** | 63 | ZMB |
| **233** | 74 | ZWE |

234 rows × 2 columns

In [60]:
```python
# to view data based on index
df.loc[224]
```

Out[60]:
```
Rank                  43
CCA3                 UZB
Country       Uzbekistan
Capital         Tashkent
Name: 224, dtype: object
```

In [61]:
```python
# iloc is used to fetch the data based on integer based index(original index) even
df.iloc[224]
```

Out[61]:
```
Rank                  43
CCA3                 UZB
Country       Uzbekistan
Capital         Tashkent
Name: 224, dtype: object
```

In [ ]:

In [ ]: