

Caitlin Sisilli

Homework 4

Question 1)

a)

```
import numpy as np
```

```
import pandas as pd
```

```
frame= pd.read_table('Movies.csv',sep=',')
```

```
print(frame.shape)
```

Output:

```
(50, 5)
```

b)

```
import numpy as np
```

```
import pandas as pd
```

```
frame= pd.read_table('Movies.csv',sep=',')
```

```
three=frame.head(3)
```

```
print(three)
```

Output:

| | # | Date | Film | Gross | Notes |
|---|---|-----------|--------------------------------|---------------|-------|
| 0 | 1 | 7-Jan-18 | Star Wars: The Last Jedi | US\$6,557,062 | [1] |
| 1 | 2 | 14-Jan-18 | Jumanji: Welcome to the Jungle | US\$2,127,871 | [2] |
| 2 | 3 | 21-Jan-18 | Brillantissime | US\$2,006,033 | [3] |

c)

```
import numpy as np
```

```
import pandas as pd
```

```
frame= pd.read_table('Movies.csv',sep=',')
```

```
three=frame.dtypes
```

```
print(three)
```

Output:

```
#          int64
Date      object
Film      object
Gross     object
Notes     object
dtype: object
```

d)

To check a missing data, it would be easier to check by using the null function.

```
import numpy as np
```

```
import pandas as pd
```

```
frame= pd.read_table('Movies.csv',sep=',')
```

```
print(frame.isnull())
```

Output:

| # | Date | Film | Gross | Notes |
|----|-------|-------|-------|-------|
| 0 | False | False | False | False |
| 1 | False | False | False | False |
| 2 | False | False | False | False |
| 3 | False | False | False | False |
| 4 | False | False | False | False |
| 5 | False | False | False | False |
| 6 | False | False | False | False |
| 7 | False | False | False | False |
| 8 | False | False | False | False |
| 9 | False | False | False | False |
| 10 | False | False | False | False |
| 11 | False | False | False | False |
| 12 | False | False | False | False |
| 13 | False | False | False | False |
| 14 | False | False | False | False |
| 15 | False | False | False | False |
| 16 | False | False | False | False |
| 17 | False | False | False | False |
| 18 | False | False | False | False |
| 19 | False | False | False | False |
| 20 | False | False | False | False |
| 21 | False | False | False | False |
| 22 | False | False | False | False |
| 23 | False | False | False | False |
| 24 | False | False | False | False |
| 25 | False | False | False | False |
| 26 | False | False | False | False |
| 27 | False | False | False | False |
| 28 | False | False | False | False |
| 29 | False | False | False | False |
| 30 | False | False | False | False |
| 31 | False | False | False | False |
| 32 | False | False | False | False |
| 33 | False | False | False | False |
| 34 | False | False | False | False |
| 35 | False | False | False | False |
| 36 | False | False | False | False |
| 37 | False | False | False | False |
| 38 | False | False | False | False |
| 39 | False | False | False | False |
| 40 | False | False | False | False |
| 41 | False | False | False | False |
| 42 | False | False | False | False |
| 43 | False | False | False | False |
| 44 | False | False | False | False |
| 45 | False | False | False | False |
| 46 | False | False | False | False |

```

47 False False False False False
48 False False False False False
49 False False False False False

```

Question 2)

a)

```
import numpy as np
```

```
import pandas as pd
```

```
frame= pd.read_table('Movies.csv',sep=',')
```

```
frame['Gross']=frame['Gross'].str.replace(r'\D','')
```

```
print(frame)
```

Output:

| | # | Date | Film | Gross | Not |
|------|----|-----------|--------------------------------|----------|-----|
| es | | | | | |
| 0 | 1 | 7-Jan-18 | Star Wars: The Last Jedi | 6557062 | |
| [1] | | | | | |
| 1 | 2 | 14-Jan-18 | Jumanji: Welcome to the Jungle | 2127871 | |
| [2] | | | | | |
| 2 | 3 | 21-Jan-18 | Brillantissime | 2006033 | |
| [3] | | | | | |
| 3 | 4 | 28-Jan-18 | The Post | 2771269 | |
| [4] | | | | | |
| 4 | 5 | 4-Feb-18 | Les Tuche 3 | 16604101 | |
| [5] | | | | | |
| 5 | 6 | 11-Feb-18 | Les Tuche 3 | 8753427 | |
| [6] | | | | | |
| 6 | 7 | 18-Feb-18 | Black Panther | 7627163 | |
| [7] | | | | | |
| 7 | 8 | 25-Feb-18 | Black Panther | 5637555 | |
| [8] | | | | | |
| 8 | 9 | 4-Mar-18 | La ch'tite famille | 16739183 | |
| [9] | | | | | |
| 9 | 10 | 11-Mar-18 | La ch'tite famille | 9546133 | |
| [10] | | | | | |
| 10 | 11 | 18-Mar-18 | La ch'tite famille | 5278842 | |
| [11] | | | | | |

| | | | | |
|------|----|-----------|---------------------------------------|----------|
| 11 | 12 | 25-Mar-18 | Rolling to You | 3073866 |
| [12] | | | | |
| 12 | 13 | 1-Apr-18 | Ready Player One | 5484062 |
| [13] | | | | |
| 13 | 14 | 8-Apr-18 | Ready Player One | 3140585 |
| [14] | | | | |
| 14 | 15 | 15-Apr-18 | Taxi 5 | 10568544 |
| [15] | | | | |
| 15 | 16 | 22-Apr-18 | Taxi 5 | 4109407 |
| [16] | | | | |
| 16 | 17 | 29-Apr-18 | Avengers: Infinity War | 17645304 |
| [17] | | | | |
| 17 | 18 | 6-May-18 | Avengers: Infinity War | 6891789 |
| [18] | | | | |
| 18 | 19 | 13-May-18 | Avengers: Infinity War | 6273652 |
| [19] | | | | |
| 19 | 20 | 20-May-18 | Deadpool 2 | 8836807 |
| [20] | | | | |
| 20 | 21 | 27-May-18 | Solo: A Star Wars Story | 4015319 |
| [21] | | | | |
| 21 | 22 | 3-Jun-18 | Solo: A Star Wars Story | 2770078 |
| [22] | | | | |
| 22 | 23 | 10-Jun-18 | Jurassic World: Fallen Kingdom | 9579503 |
| [23] | | | | |
| 23 | 24 | 17-Jun-18 | Jurassic World: Fallen Kingdom | 4791805 |
| [24] | | | | |
| 24 | 25 | 24-Jun-18 | Jurassic World: Fallen Kingdom | 2520159 |
| [25] | | | | |
| 25 | 26 | 1-Jul-18 | Jurassic World: Fallen Kingdom | 1740016 |
| [26] | | | | |
| 26 | 27 | 8-Jul-18 | Incredibles 2 | 9454086 |
| [27] | | | | |
| 27 | 28 | 15-Jul-18 | Incredibles 2 | 5152960 |
| [28] | | | | |
| 28 | 29 | 22-Jul-18 | Incredibles 2 | 4859634 |
| [29] | | | | |
| 29 | 30 | 29-Jul-18 | Hotel Transylvania 3: Summer Vacation | 3929857 |
| [30] | | | | |
| 30 | 31 | 5-Aug-18 | Mission: Impossible - Fallout | 8193447 |
| [31] | | | | |
| 31 | 32 | 12-Aug-18 | Mission: Impossible - Fallout | 4350586 |
| [32] | | | | |
| 32 | 33 | 19-Aug-18 | Mission: Impossible - Fallout | 2873390 |
| [33] | | | | |
| 33 | 34 | 26-Aug-18 | The Meg | 4513632 |
| [34] | | | | |
| 34 | 35 | 2-Sep-18 | The Meg | 2513203 |
| [35] | | | | |
| 35 | 36 | 9-Sep-18 | Photo de Famille | 1414632 |
| [36] | | | | |
| 36 | 37 | 16-Sep-18 | Première année | 2222791 |
| [37] | | | | |
| 37 | 38 | 23-Sep-18 | The Nun | 4425722 |
| [38] | | | | |

```

38 39 30-Sep-18 The Nun 2254244
[39]
39 40 7-Oct-18 Alad'2 6228621
[40]
40 41 14-Oct-18 Venom 6646527
[41]
41 42 21-Oct-18 Venom 3019547
[42]
42 43 28-Oct-18 Sink or Swim 9051705
[43]
43 44 4-Nov-18 Sink or Swim 9676385
[44]
44 45 11-Nov-18 Bohemian Rhapsody 5154997
[45]
45 46 18-Nov-18 Fantastic Beasts: The Crimes of Grindelwald 11438486
[46]
46 47 25-Nov-18 Fantastic Beasts: The Crimes of Grindelwald 6378619
[47]
47 48 2-Dec-18 Fantastic Beasts: The Crimes of Grindelwald 3656409
[48]
48 49 9-Dec-18 Asterix: The Secret of the Magic Potion 6863693
[49]
49 50 16-Dec-18 Asterix: The Secret of the Magic Potion 4187527
[50]

```

b)

```
import numpy as np
```

```
import pandas as pd
```

```
frame= pd.read_table('Movies.csv',sep=',')
```

```
frame['Gross']=frame['Gross'].str.replace(r'\D',").astype(int)
```

```
print(frame)
```

Output:

```

#      Date      Film      Gross Note
s
0  1  7-Jan-18  Star Wars: The Last Jedi  6557062
[1]
1  2 14-Jan-18  Jumanji: Welcome to the Jungle  2127871
[2]

```

| | | | | |
|------|----|-----------|--------------------------------|----------|
| 2 | 3 | 21-Jan-18 | Brillantissime | 2006033 |
| [3] | | | | |
| 3 | 4 | 28-Jan-18 | The Post | 2771269 |
| [4] | | | | |
| 4 | 5 | 4-Feb-18 | Les Tuche 3 | 16604101 |
| [5] | | | | |
| 5 | 6 | 11-Feb-18 | Les Tuche 3 | 8753427 |
| [6] | | | | |
| 6 | 7 | 18-Feb-18 | Black Panther | 7627163 |
| [7] | | | | |
| 7 | 8 | 25-Feb-18 | Black Panther | 5637555 |
| [8] | | | | |
| 8 | 9 | 4-Mar-18 | La ch'tite famille | 16739183 |
| [9] | | | | |
| 9 | 10 | 11-Mar-18 | La ch'tite famille | 9546133 |
| [10] | | | | |
| 10 | 11 | 18-Mar-18 | La ch'tite famille | 5278842 |
| [11] | | | | |
| 11 | 12 | 25-Mar-18 | Rolling to You | 3073866 |
| [12] | | | | |
| 12 | 13 | 1-Apr-18 | Ready Player One | 5484062 |
| [13] | | | | |
| 13 | 14 | 8-Apr-18 | Ready Player One | 3140585 |
| [14] | | | | |
| 14 | 15 | 15-Apr-18 | Taxi 5 | 10568544 |
| [15] | | | | |
| 15 | 16 | 22-Apr-18 | Taxi 5 | 4109407 |
| [16] | | | | |
| 16 | 17 | 29-Apr-18 | Avengers: Infinity War | 17645304 |
| [17] | | | | |
| 17 | 18 | 6-May-18 | Avengers: Infinity War | 6891789 |
| [18] | | | | |
| 18 | 19 | 13-May-18 | Avengers: Infinity War | 6273652 |
| [19] | | | | |
| 19 | 20 | 20-May-18 | Deadpool 2 | 8836807 |
| [20] | | | | |
| 20 | 21 | 27-May-18 | Solo: A Star Wars Story | 4015319 |
| [21] | | | | |
| 21 | 22 | 3-Jun-18 | Solo: A Star Wars Story | 2770078 |
| [22] | | | | |
| 22 | 23 | 10-Jun-18 | Jurassic World: Fallen Kingdom | 9579503 |
| [23] | | | | |
| 23 | 24 | 17-Jun-18 | Jurassic World: Fallen Kingdom | 4791805 |
| [24] | | | | |
| 24 | 25 | 24-Jun-18 | Jurassic World: Fallen Kingdom | 2520159 |
| [25] | | | | |
| 25 | 26 | 1-Jul-18 | Jurassic World: Fallen Kingdom | 1740016 |
| [26] | | | | |
| 26 | 27 | 8-Jul-18 | Incredibles 2 | 9454086 |
| [27] | | | | |
| 27 | 28 | 15-Jul-18 | Incredibles 2 | 5152960 |
| [28] | | | | |
| 28 | 29 | 22-Jul-18 | Incredibles 2 | 4859634 |
| [29] | | | | |

| | | | | |
|------|----|-----------|---|----------|
| 29 | 30 | 29-Jul-18 | Hotel Transylvania 3: Summer Vacation | 3929857 |
| [30] | | | | |
| 30 | 31 | 5-Aug-18 | Mission: Impossible - Fallout | 8193447 |
| [31] | | | | |
| 31 | 32 | 12-Aug-18 | Mission: Impossible - Fallout | 4350586 |
| [32] | | | | |
| 32 | 33 | 19-Aug-18 | Mission: Impossible - Fallout | 2873390 |
| [33] | | | | |
| 33 | 34 | 26-Aug-18 | The Meg | 4513632 |
| [34] | | | | |
| 34 | 35 | 2-Sep-18 | The Meg | 2513203 |
| [35] | | | | |
| 35 | 36 | 9-Sep-18 | Photo de Famille | 1414632 |
| [36] | | | | |
| 36 | 37 | 16-Sep-18 | Première année | 2222791 |
| [37] | | | | |
| 37 | 38 | 23-Sep-18 | The Nun | 4425722 |
| [38] | | | | |
| 38 | 39 | 30-Sep-18 | The Nun | 2254244 |
| [39] | | | | |
| 39 | 40 | 7-Oct-18 | Alad'2 | 6228621 |
| [40] | | | | |
| 40 | 41 | 14-Oct-18 | Venom | 6646527 |
| [41] | | | | |
| 41 | 42 | 21-Oct-18 | Venom | 3019547 |
| [42] | | | | |
| 42 | 43 | 28-Oct-18 | Sink or Swim | 9051705 |
| [43] | | | | |
| 43 | 44 | 4-Nov-18 | Sink or Swim | 9676385 |
| [44] | | | | |
| 44 | 45 | 11-Nov-18 | Bohemian Rhapsody | 5154997 |
| [45] | | | | |
| 45 | 46 | 18-Nov-18 | Fantastic Beasts: The Crimes of Grindelwald | 11438486 |
| [46] | | | | |
| 46 | 47 | 25-Nov-18 | Fantastic Beasts: The Crimes of Grindelwald | 6378619 |
| [47] | | | | |
| 47 | 48 | 2-Dec-18 | Fantastic Beasts: The Crimes of Grindelwald | 3656409 |
| [48] | | | | |
| 48 | 49 | 9-Dec-18 | Asterix: The Secret of the Magic Potion | 6863693 |
| [49] | | | | |
| 49 | 50 | 16-Dec-18 | Asterix: The Secret of the Magic Potion | 4187527 |
| [50] | | | | |

c)

```
import numpy as np
```

```
import pandas as pd
```



```

frame= pd.read_table('Movies.csv',sep=',')

frame['Gross']=frame['Gross'].str.replace(r'\D',").astype(int)

frames=frame.drop(['Film','Date','Notes'],axis=1)

print(frames)

```

Output:

| | # | Gross |
|----|----|----------|
| 0 | 1 | 6557062 |
| 1 | 2 | 2127871 |
| 2 | 3 | 2006033 |
| 3 | 4 | 2771269 |
| 4 | 5 | 16604101 |
| 5 | 6 | 8753427 |
| 6 | 7 | 7627163 |
| 7 | 8 | 5637555 |
| 8 | 9 | 16739183 |
| 9 | 10 | 9546133 |
| 10 | 11 | 5278842 |
| 11 | 12 | 3073866 |
| 12 | 13 | 5484062 |
| 13 | 14 | 3140585 |
| 14 | 15 | 10568544 |
| 15 | 16 | 4109407 |
| 16 | 17 | 17645304 |
| 17 | 18 | 6891789 |
| 18 | 19 | 6273652 |
| 19 | 20 | 8836807 |
| 20 | 21 | 4015319 |
| 21 | 22 | 2770078 |
| 22 | 23 | 9579503 |
| 23 | 24 | 4791805 |
| 24 | 25 | 2520159 |
| 25 | 26 | 1740016 |
| 26 | 27 | 9454086 |
| 27 | 28 | 5152960 |
| 28 | 29 | 4859634 |
| 29 | 30 | 3929857 |
| 30 | 31 | 8193447 |
| 31 | 32 | 4350586 |
| 32 | 33 | 2873390 |
| 33 | 34 | 4513632 |
| 34 | 35 | 2513203 |

| | | |
|----|----|----------|
| 35 | 36 | 1414632 |
| 36 | 37 | 2222791 |
| 37 | 38 | 4425722 |
| 38 | 39 | 2254244 |
| 39 | 40 | 6228621 |
| 40 | 41 | 6646527 |
| 41 | 42 | 3019547 |
| 42 | 43 | 9051705 |
| 43 | 44 | 9676385 |
| 44 | 45 | 5154997 |
| 45 | 46 | 11438486 |
| 46 | 47 | 6378619 |
| 47 | 48 | 3656409 |
| 48 | 49 | 6863693 |
| 49 | 50 | 4187527 |

Question 3)

a)

The statistics shown on the graph show the difference cost and amount of films during the specific dates that are in ordered.

b)

```
import numpy as np
```

```
import pandas as pd
```

```
frame= pd.read_table('Movies.csv',sep=',')
```

```
frame.reset_index()
```

```
frame.set_index('Film',inplace=True)
```

```
frame['Gross']=frame['Gross'].str.replace(r'\D','').astype(int)
```

```
print(frame.loc[(frame['Gross']>=600000)])
```

Output:

Film

| | | | | |
|---|----|-----------|----------|------|
| Star Wars: The Last Jedi | 1 | 7-Jan-18 | 6557062 | [1] |
| Jumanji: Welcome to the Jungle | 2 | 14-Jan-18 | 2127871 | [2] |
| Brillantissime | 3 | 21-Jan-18 | 2006033 | [3] |
| The Post | 4 | 28-Jan-18 | 2771269 | [4] |
| Les Tuche 3 | 5 | 4-Feb-18 | 16604101 | [5] |
| Les Tuche 3 | 6 | 11-Feb-18 | 8753427 | [6] |
| Black Panther | 7 | 18-Feb-18 | 7627163 | [7] |
| Black Panther | 8 | 25-Feb-18 | 5637555 | [8] |
| La ch'tite famille | 9 | 4-Mar-18 | 16739183 | [9] |
| La ch'tite famille | 10 | 11-Mar-18 | 9546133 | [10] |
| La ch'tite famille | 11 | 18-Mar-18 | 5278842 | [11] |
| Rolling to You | 12 | 25-Mar-18 | 3073866 | [12] |
| Ready Player One | 13 | 1-Apr-18 | 5484062 | [13] |
| Ready Player One | 14 | 8-Apr-18 | 3140585 | [14] |
| Taxi 5 | 15 | 15-Apr-18 | 10568544 | [15] |
| Taxi 5 | 16 | 22-Apr-18 | 4109407 | [16] |
| Avengers: Infinity War | 17 | 29-Apr-18 | 17645304 | [17] |
| Avengers: Infinity War | 18 | 6-May-18 | 6891789 | [18] |
| Avengers: Infinity War | 19 | 13-May-18 | 6273652 | [19] |
| Deadpool 2 | 20 | 20-May-18 | 8836807 | [20] |
| Solo: A Star Wars Story | 21 | 27-May-18 | 4015319 | [21] |
| Solo: A Star Wars Story | 22 | 3-Jun-18 | 2770078 | [22] |
| Jurassic World: Fallen Kingdom | 23 | 10-Jun-18 | 9579503 | [23] |
| Jurassic World: Fallen Kingdom | 24 | 17-Jun-18 | 4791805 | [24] |
| Jurassic World: Fallen Kingdom | 25 | 24-Jun-18 | 2520159 | [25] |
| Jurassic World: Fallen Kingdom | 26 | 1-Jul-18 | 1740016 | [26] |
| Incredibles 2 | 27 | 8-Jul-18 | 9454086 | [27] |
| Incredibles 2 | 28 | 15-Jul-18 | 5152960 | [28] |
| Incredibles 2 | 29 | 22-Jul-18 | 4859634 | [29] |
| Hotel Transylvania 3: Summer Vacation | 30 | 29-Jul-18 | 3929857 | [30] |
| Mission: Impossible - Fallout | 31 | 5-Aug-18 | 8193447 | [31] |
| Mission: Impossible - Fallout | 32 | 12-Aug-18 | 4350586 | [32] |
| Mission: Impossible - Fallout | 33 | 19-Aug-18 | 2873390 | [33] |
| The Meg | 34 | 26-Aug-18 | 4513632 | [34] |
| The Meg | 35 | 2-Sep-18 | 2513203 | [35] |
| Photo de Famille | 36 | 9-Sep-18 | 1414632 | [36] |
| Première année | 37 | 16-Sep-18 | 2222791 | [37] |
| The Nun | 38 | 23-Sep-18 | 4425722 | [38] |
| The Nun | 39 | 30-Sep-18 | 2254244 | [39] |
| Alad'2 | 40 | 7-Oct-18 | 6228621 | [40] |
| Venom | 41 | 14-Oct-18 | 6646527 | [41] |
| Venom | 42 | 21-Oct-18 | 3019547 | [42] |
| Sink or Swim | 43 | 28-Oct-18 | 9051705 | [43] |
| Sink or Swim | 44 | 4-Nov-18 | 9676385 | [44] |
| Bohemian Rhapsody | 45 | 11-Nov-18 | 5154997 | [45] |
| Fantastic Beasts: The Crimes of Grindelwald | 46 | 18-Nov-18 | 11438486 | [46] |
| Fantastic Beasts: The Crimes of Grindelwald | 47 | 25-Nov-18 | 6378619 | [47] |
| Fantastic Beasts: The Crimes of Grindelwald | 48 | 2-Dec-18 | 3656409 | [48] |
| Asterix: The Secret of the Magic Potion | 49 | 9-Dec-18 | 6863693 | [49] |
| Asterix: The Secret of the Magic Potion | 50 | 16-Dec-18 | 4187527 | [50] |

c)

There seemed to be no outliers within the data and to show from a graph.

```
import numpy as np
```

```
import pandas as pd
```

```
import matplotlib.pyplot as plt
```

```
frame= pd.read_table('Movies.csv',sep=',')
```

```
y=frame['Date']
```

```
x=frame['Gross']
```

```
plt.axis([0,50,0,50])
```

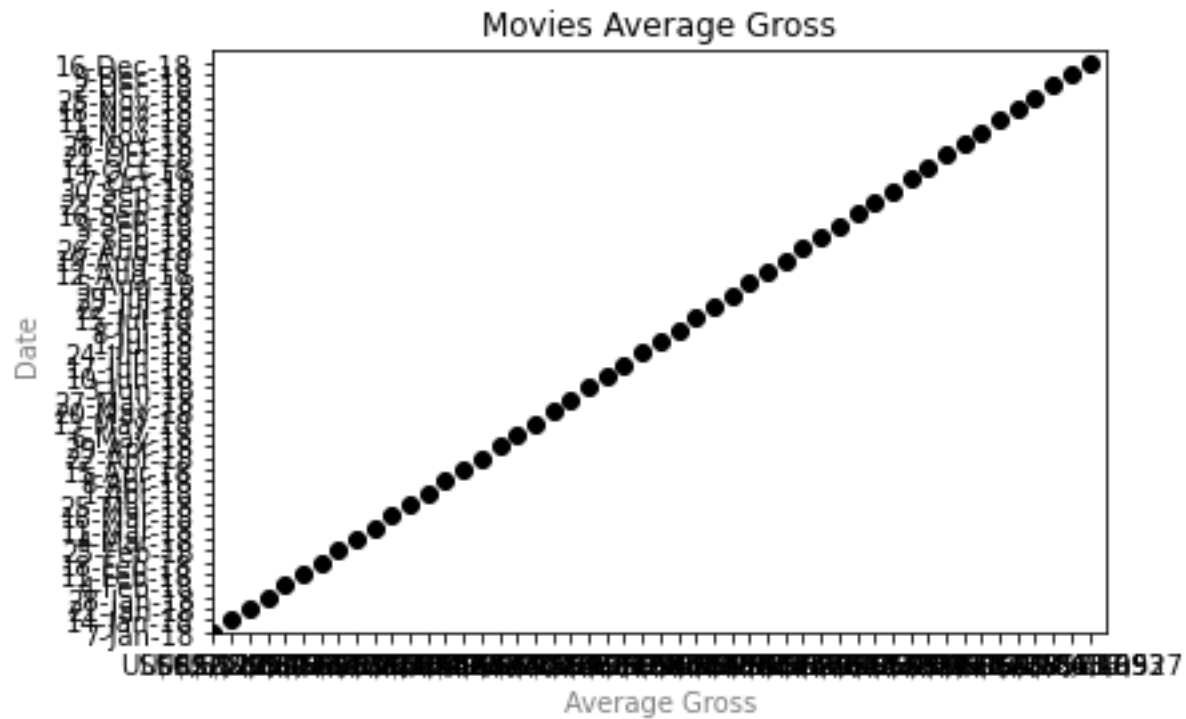
```
plt.title('Movies Average Gross')
```

```
plt.xlabel('Average Gross',color='gray')
```

```
plt.ylabel('Date',color='gray')
```

```
plt.plot(x,y,'ko')
```

Output:



d)

```
import numpy as np
```

```
import pandas as pd
```

```
frame= pd.read_table('Movies.csv',sep=',')
```

```
frame.reset_index()
```

```
frame.set_index('Film',inplace=True)
```

```
frame['Gross']=frame['Gross'].str.replace(r'\D','').astype(int)
```

```
print(frame['Gross'].nlargest(n=10))
```

Output:

```
Film
Avengers: Infinity War
```

```
17645304
```

| | |
|---|----------|
| La ch'tite famille | 16739183 |
| Les Tuche 3 | 16604101 |
| Fantastic Beasts: The Crimes of Grindelwald | 11438486 |
| Taxi 5 | 10568544 |
| Sink or Swim | 9676385 |
| Jurassic World: Fallen Kingdom | 9579503 |
| La ch'tite famille | 9546133 |
| Incredibles 2 | 9454086 |
| Sink or Swim | 9051705 |
| Name: Gross, dtype: int32 | |

e)

```
import numpy as np
```

```
import pandas as pd
```

```
frame= pd.read_table('Movies.csv',sep=',')
```

```
frame.reset_index()
```

```
frame.set_index('Film',inplace=True)
```

```
frame['Gross']=frame['Gross'].str.replace(r"\D',").astype(int)
```

```
print(frame['Gross'].nsmallest(n=10))
```

Output:

| | |
|--------------------------------|---------|
| Film | |
| Photo de Famille | 1414632 |
| Jurassic World: Fallen Kingdom | 1740016 |
| Brillantissime | 2006033 |
| Jumanji: Welcome to the Jungle | 2127871 |
| Première année | 2222791 |
| The Nun | 2254244 |
| The Meg | 2513203 |
| Jurassic World: Fallen Kingdom | 2520159 |
| Solo: A Star Wars Story | 2770078 |
| The Post | 2771269 |
| Name: Gross, dtype: int32 | |

