

```
In [258]: print("Name :")
print("We will learn about bubble plot today and find some interesting facts about production of movies their revenue")
```

Name :

We will learn about bubble plot today and find some interesting facts about production of movies their revenue

## Activity 1 - Which movies has collected more than Four hundred million dollars in box office

```
In [265]: import pandas as pd
import matplotlib.pyplot as plt
import numpy as np

disney_movies = pd.read_csv('disney_movies.csv')
disney_movies
```

```
Out[265]:
```

	movie_title	release_date	genre	mpaa_rating	total_gross	inflation_adjusted_gross
0	Snow White and the Seven Dwarfs	1937-12-21	Musical	G	184925485	5228953251
1	Pinocchio	1940-02-09	Adventure	G	84300000	2188229052
2	Fantasia	1940-11-13	Musical	G	83320000	2187090808
3	Song of the South	1946-11-12	Adventure	G	65000000	1078510579
4	Cinderella	1950-02-15	Drama	G	85000000	920608730
...	...	...	...	...	...	...
574	The Light Between Oceans	2016-09-02	Drama	PG-13	12545979	12545979
575	Queen of Katwe	2016-09-23	Drama	PG	8874389	8874389
576	Doctor Strange	2016-11-04	Adventure	PG-13	232532923	232532923
577	Moana	2016-11-23	Adventure	PG	246082029	246082029
578	Rogue One: A Star Wars Story	2016-12-16	Adventure	PG-13	529483936	529483936

579 rows × 6 columns

```
In [267]: #Find list of movies who has collected more than Four hundred million dollars
movies_collection = disney_movies.loc[disney_movies['total_gross'] > 400000]
movies_collection
```

```
Out[267]:
```

	movie_title	release_date	genre	mpaa_rating	total_gross	inflation_adjusted_gross
179	The Lion King	1994-06-15	Adventure	G	422780140	761640898
441	Pirates of the Caribbean: Dead Man's...	2006-07-07	Adventure	PG-13	423315812	544817142
499	Toy Story 3	2010-06-18	Adventure	G	415004880	443408255
524	The Avengers	2012-05-04	Action	PG-13	623279547	660081224
532	Iron Man 3	2013-05-03	Action	PG-13	408992272	424084233
539	Frozen	2013-11-22	Adventure	PG	400738009	414997174
558	Avengers: Age of Ultron	2015-05-01	Action	PG-13	459005868	459005868
564	Star Wars Ep. VII: The Force Awakens	2015-12-18	Adventure	PG-13	936662225	936662225
569	Captain America: Civil War	2016-05-06	Action	PG-13	408084349	408084349
571	Finding Dory	2016-06-17	Adventure	PG	486295561	486295561
578	Rogue One: A Star Wars Story	2016-12-16	Adventure	PG-13	529483936	529483936

```
In [268]: #Plot a bubble graph to show which movies who has collected more then Four
plt.figure(figsize=(12, 8))
plt.xticks(rotation='vertical')

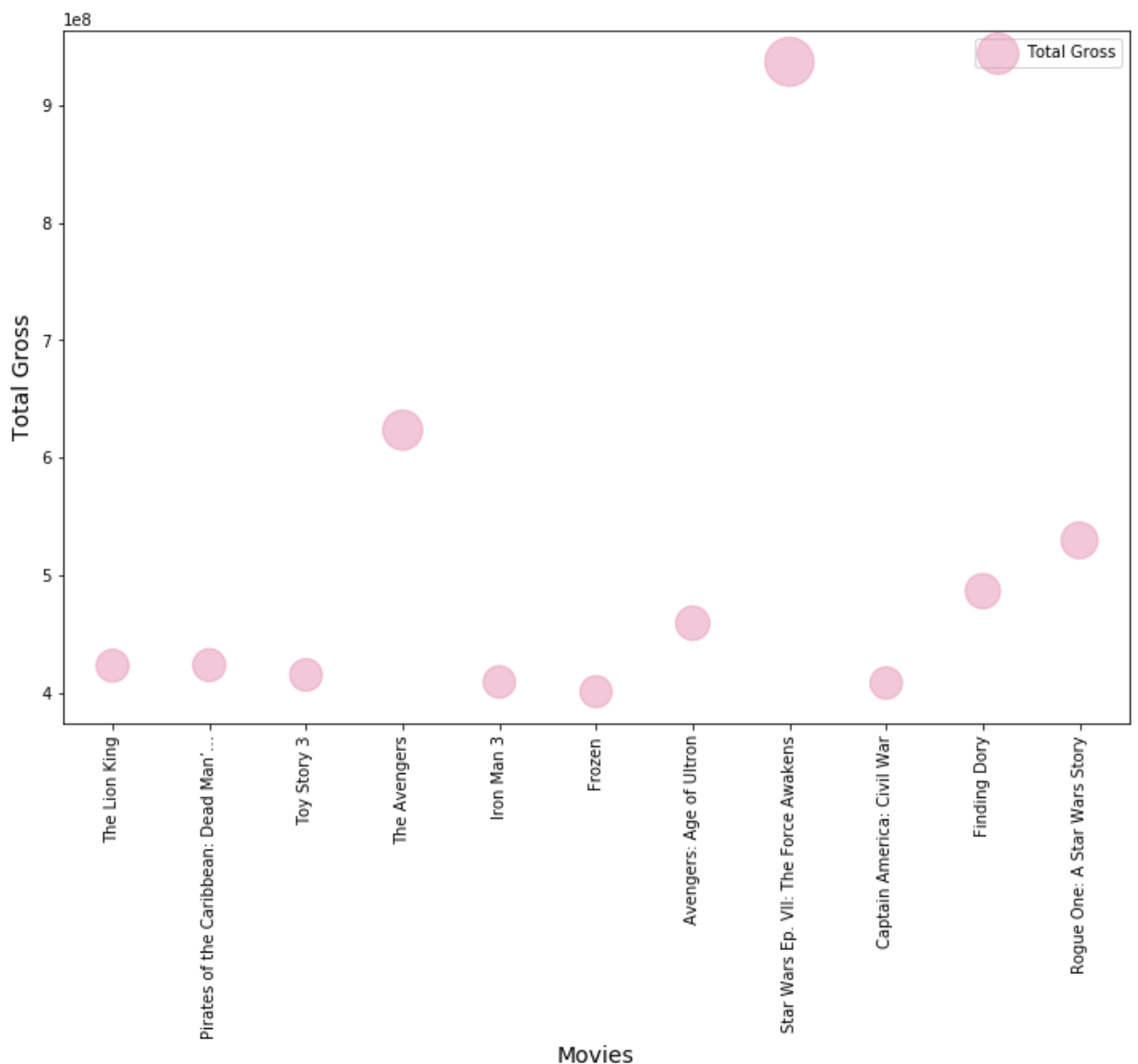
rgb = np.random.rand(3)
print(rgb)

plt.scatter(movies_collection['movie_title'], movies_collection['total_gross'],
            c=[rgb],label = 'Total Gross',
            alpha=0.5,
            s = movies_collection['total_gross']/1000000)

plt.legend()
plt.xlabel("Movies", size=14)
plt.ylabel("Total Gross", size=14)

[0.90222446 0.56304913 0.69719519]
```

Out[268]: Text(0, 0.5, 'Total Gross')



Conclusion - Star wars Ep 7: The Force Awakers and The Avengers are 2 movies which made the most in box office till now

## Activity 2 - Find the total number of movies per year and plot a bubble chart, to find out in which year the maximum movies were produced

```
In [1]: #Import libraries and read csv 'movie_update.csv'

import pandas as pd
import matplotlib.pyplot as plt
import numpy as np

disney_movies = pd.read_csv('disney_movies.csv')
disney_movies
```

```
In [2]: #Group by year column and find total number of movies released each year

group_year = movie_update.groupby('Year')['Title'].count().reset_index()
group_year
```

```
In [2]: Plot a bubble graph to show total number of movies released each year

import matplotlib.pyplot as plt

plt.figure(figsize=(20,80))

plt.scatter(group_year['Year'], group_year['Title'], color= 'green', label='Number of Movies')

plt.xticks(rotation='vertical')
plt.legend()
plt.xlabel("Years", size=14)
plt.ylabel('Number of Movies', size=14)
```

```
-----
--
ModuleNotFoundError                                Traceback (most recent call last)
t)
/var/folders/ld/dx5rjvf9lqq7f432s9k1r3780000gp/T/ipykernel_45827/38703988
56.py in <module>
      1 #Plot a bubble graph to show total number of movies released each
      2 year
----> 3 import matplotlib.pyplot as plt
      4
      5 plt.figure(figsize=(20,80))

ModuleNotFoundError: No module named 'matplotlib'
```

Conclusion -

## Activity 3 - Find which category of movie has the maximum income as per you.

```
In [4]: #Read 'disney_movies.csv'

#Put a condition as per you to find out which category of movie has the max
```

```
In [5]: #Group by genre and sum the total income of the movies
```

```
In [ ]: #Plot a pie chart as per the genre and fetch which category of movie to mak
```

Conclusion :

```
In [ ]:
```

In [ ]:

In [ ]:

In [ ]:

In [ ]: