

```
In [1]: import pandas as pd
import matplotlib.pyplot as plt
```

```
In [2]: data=pd.read_csv("/Users/bikashadhikari/Desktop/GMU Assignment/HAP 725/1/Assignment 3/hos_revised_flatfiles_ar
```

```
In [3]: data.head()
```

Out[3]:

	Facility ID	Facility Name	Address	City	State	ZIP Code	County Name	Phone Number	Measure ID	Measure Name	Compared to National	Denomina
0	010001	SOUTHEAST ALABAMA MEDICAL CENTER	1108 ROSS CLARK CIRCLE	DOTHAN	AL	36301	HOUSTON	(334) 793-8701	COMP_HIP_KNEE	Rate of complications for hip/knee replacement...	No Different Than the National Rate	:
1	010001	SOUTHEAST ALABAMA MEDICAL CENTER	1108 ROSS CLARK CIRCLE	DOTHAN	AL	36301	HOUSTON	(334) 793-8701	MORT_30_AMI	Death rate for heart attack patients	No Different Than the National Rate	{
2	010001	SOUTHEAST ALABAMA MEDICAL CENTER	1108 ROSS CLARK CIRCLE	DOTHAN	AL	36301	HOUSTON	(334) 793-8701	MORT_30_CABG	Death rate for CABG surgery patients	No Different Than the National Rate	:
3	010001	SOUTHEAST ALABAMA MEDICAL CENTER	1108 ROSS CLARK CIRCLE	DOTHAN	AL	36301	HOUSTON	(334) 793-8701	MORT_30_COPD	Death rate for COPD patients	No Different Than the National Rate	4
4	010001	SOUTHEAST ALABAMA MEDICAL CENTER	1108 ROSS CLARK CIRCLE	DOTHAN	AL	36301	HOUSTON	(334) 793-8701	MORT_30_HF	Death rate for heart failure patients	No Different Than the National Rate	{

```
In [4]: data2= data[data['Measure ID']=='COMP_HIP_KNEE']
```

```
In [5]: data2
```

Out[5]:

	Facility ID	Facility Name	Address	City	State	ZIP Code	County Name	Phone Number	Measure ID	Measure Name	Co
0	010001	SOUTHEAST ALABAMA MEDICAL CENTER	1108 ROSS CLARK CIRCLE	DOTHAN	AL	36301	HOUSTON	(334) 793-8701	COMP_HIP_KNEE	Rate of complications for hip/knee replacement...	[
19	010005	MARSHALL MEDICAL CENTERS	2505 U S HIGHWAY 431 NORTH	BOAZ	AL	35957	MARSHALL	(256) 593-8310	COMP_HIP_KNEE	Rate of complications for hip/knee replacement...	[
38	010006	NORTH ALABAMA MEDICAL CENTER	1701 VETERANS DRIVE	FLORENCE	AL	35630	LAUDERDALE	(256) 768-8400	COMP_HIP_KNEE	Rate of complications for hip/knee replacement...	[

	Facility ID	Facility Name	Address	City	State	ZIP Code	County Name	Phone Number	Measure ID	Measure Name	Comparison
57	010007	MIZELL MEMORIAL HOSPITAL	702 N MAIN ST	OPP	AL	36467	COVINGTON	(334) 493-3541	COMP_HIP_KNEE	Rate of complications for hip/knee replacement...	CT
76	010008	CRENSHAW COMMUNITY HOSPITAL	101 HOSPITAL CIRCLE	LUVERNE	AL	36049	CRENSHAW	(334) 335-3374	COMP_HIP_KNEE	Rate of complications for hip/knee replacement...	A
...
91285	670125	TEXAS CENTER FOR INFECTIOUS DISEASE	2303 SE MILITARY DR P O BOX 23340	SAN ANTONIO	TX	78223	BEXAR	(210) 534-8857	COMP_HIP_KNEE	Rate of complications for hip/knee replacement...	A
91304	670127	EL PASO LTAC HOSPITAL	1221 N COTTON ST 3RD FLOOR	EL PASO	TX	79902	EL PASO	(915) 496-9687	COMP_HIP_KNEE	Rate of complications for hip/knee replacement...	A
91323	670128	BAYLOR SCOTT & WHITE MEDICAL CENTER PFLUGERVILLE	2600 EAST PFLUGERVILLE PARKWAY	PFLUGERVILLE	TX	78660	TRAVIS	(512) 654-6100	COMP_HIP_KNEE	Rate of complications for hip/knee replacement...	A
91342	670129	THE HEIGHTS HOSPITAL	1917 ASHLAND STREET, SUITE 300	HOUSTON	TX	77008	HARRIS	(713) 574-6814	COMP_HIP_KNEE	Rate of complications for hip/knee replacement...	A
91361	670130	SOUTHCROSS HOSPITAL	4243 EAST SOUTHCROSS BLVD SUITE 100	SAN ANTONIO	TX	78222	BEXAR	(210) 507-4392	COMP_HIP_KNEE	Rate of complications for hip/knee replacement...	A

4896 rows × 18 columns

In [6]: data2.dtypes

```
Out[6]: Facility ID      object
Facility Name    object
Address          object
City             object
State           object
ZIP Code        int64
County Name     object
Phone Number    object
Measure ID      object
Measure Name    object
Compared to National  object
Denominator     object
Score           object
Lower Estimate   object
Higher Estimate  object
Footnote        float64
Start Date      object
End Date        object
dtype: object
```

In [7]: data2['Score']=pd.to_numeric(data['Score'],errors='coerce')

/var/folders/cy/wngl16rn4hg_0y7s1s7v1f740000gn/T/ipykernel_12283/3319984331.py:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.

Try using `.loc[row_indexer,col_indexer] = value` instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#re-creating-a-view-versus-a-copy
`data2['Score']=pd.to_numeric(data['Score'],errors='coerce')`

```
In [8]: data2['Denominator']=pd.to_numeric(data['Denominator'],errors='coerce')
```

`/var/folders/cy/wngl16rn4hg_0y7s1s7v1f740000gn/T/ipykernel_12283/259439869.py:1: SettingWithCopyWarning:`
A value is trying to be set on a copy of a slice from a DataFrame.
Try using `.loc[row_indexer,col_indexer] = value` instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#re-creating-a-view-versus-a-copy
`data2['Denominator']=pd.to_numeric(data['Denominator'],errors='coerce')`

```
In [9]: data2
```

Out[9]:

	Facility ID	Facility Name	Address	City	State	ZIP Code	County Name	Phone Number	Measure ID	Measure Name	Co
											N
0	010001	SOUTHEAST ALABAMA MEDICAL CENTER	1108 ROSS CLARK CIRCLE	DOTHAN	AL	36301	HOUSTON	(334) 793-8701	COMP_HIP_KNEE	Rate of complications for hip/knee replacement...	[T]
19	010005	MARSHALL MEDICAL CENTERS	2505 U S HIGHWAY 431 NORTH	BOAZ	AL	35957	MARSHALL	(256) 593-8310	COMP_HIP_KNEE	Rate of complications for hip/knee replacement...	[T]
38	010006	NORTH ALABAMA MEDICAL CENTER	1701 VETERANS DRIVE	FLORENCE	AL	35630	LAUDERDALE	(256) 768-8400	COMP_HIP_KNEE	Rate of complications for hip/knee replacement...	[T]
57	010007	MIZELL MEMORIAL HOSPITAL	702 N MAIN ST	OPP	AL	36467	COVINGTON	(334) 493-3541	COMP_HIP_KNEE	Rate of complications for hip/knee replacement...	[T]
76	010008	CRENSHAW COMMUNITY HOSPITAL	101 HOSPITAL CIRCLE	LUVERNE	AL	36049	CRENSHAW	(334) 335-3374	COMP_HIP_KNEE	Rate of complications for hip/knee replacement...	A
...	
91285	670125	TEXAS CENTER FOR INFECTIOUS DISEASE	2303 SE MILITARY DR P O BOX 23340	SAN ANTONIO	TX	78223	BEXAR	(210) 534-8857	COMP_HIP_KNEE	Rate of complications for hip/knee replacement...	A
91304	670127	EL PASO LTAC HOSPITAL	1221 N COTTON ST 3RD FLOOR	EL PASO	TX	79902	EL PASO	(915) 496-9687	COMP_HIP_KNEE	Rate of complications for hip/knee replacement...	A
91323	670128	BAYLOR SCOTT & WHITE MEDICAL CENTER PFLUGERVILLE	2600 EAST PFLUGERVILLE PARKWAY	PFLUGERVILLE	TX	78660	TRAVIS	(512) 654-6100	COMP_HIP_KNEE	Rate of complications for hip/knee replacement...	A
91342	670129	THE HEIGHTS HOSPITAL	1917 ASHLAND STREET, SUITE 300	HOUSTON	TX	77008	HARRIS	(713) 574-6814	COMP_HIP_KNEE	Rate of complications for hip/knee replacement...	A

	Facility ID	Facility Name	Address	City	State	ZIP Code	County Name	Phone Number	Measure ID	Measure Name	Complication Rate
91361	670130	SOUTHCROSS HOSPITAL	4243 EAST SOUTHCROSS BLVD SUITE 100	SAN ANTONIO	TX	78222	BEXAR	(210) 507-4392	COMP_HIP_KNEE	Rate of complications for hip/knee replacement...	A

4896 rows × 18 columns

In [10]: data2.dtypes

Out[10]: Facility ID object
Facility Name object
Address object
City object
State object
ZIP Code int64
County Name object
Phone Number object
Measure ID object
Measure Name object
Compared to National object
Denominator float64
Score float64
Lower Estimate object
Higher Estimate object
Footnote float64
Start Date object
End Date object
dtype: object

In [11]: finaldata=data2[['Facility Name','Score','Denominator']]

In [12]: finaldata

Out[12]:

	Facility Name	Score	Denominator
0	SOUTHEAST ALABAMA MEDICAL CENTER	3.2	292.0
19	MARSHALL MEDICAL CENTERS	2.8	257.0
38	NORTH ALABAMA MEDICAL CENTER	2.6	389.0
57	MIZELL MEMORIAL HOSPITAL	2.8	31.0
76	CRENSHAW COMMUNITY HOSPITAL	NaN	NaN
...
91285	TEXAS CENTER FOR INFECTIOUS DISEASE	NaN	NaN
91304	EL PASO LTAC HOSPITAL	NaN	NaN
91323	BAYLOR SCOTT & WHITE MEDICAL CENTER PFLUGERVILLE	NaN	NaN
91342	THE HEIGHTS HOSPITAL	NaN	NaN
91361	SOUTHCROSS HOSPITAL	NaN	NaN

4896 rows × 3 columns

In [13]: finaldata['Complication Rate']= (finaldata['Score']/ finaldata['Denominator'])

/var/folders/cy/wngl16rn4hg_0y7s1s7v1f740000gn/T/ipykernel_12283/1365446010.py:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#re-turning-a-view-versus-a-copy

```
finaldata['Complication Rate'] = (finaldata['Score'] / finaldata['Denominator'])
```

In [14]:

```
finaldata
```

Out[14]:

	Facility Name	Score	Denominator	Complication Rate
0	SOUTHEAST ALABAMA MEDICAL CENTER	3.2	292.0	0.010959
19	MARSHALL MEDICAL CENTERS	2.8	257.0	0.010895
38	NORTH ALABAMA MEDICAL CENTER	2.6	389.0	0.006684
57	MIZELL MEMORIAL HOSPITAL	2.8	31.0	0.090323
76	CRENSHAW COMMUNITY HOSPITAL	NaN	NaN	NaN
...
91285	TEXAS CENTER FOR INFECTIOUS DISEASE	NaN	NaN	NaN
91304	EL PASO LTAC HOSPITAL	NaN	NaN	NaN
91323	BAYLOR SCOTT & WHITE MEDICAL CENTER PFLUGERVILLE	NaN	NaN	NaN
91342	THE HEIGHTS HOSPITAL	NaN	NaN	NaN
91361	SOUTHCROSS HOSPITAL	NaN	NaN	NaN

4896 rows × 4 columns

In [15]:

```
finaldata=finaldata.dropna(how="any")
```

In [16]:

```
finaldata
```

Out[16]:

	Facility Name	Score	Denominator	Complication Rate
0	SOUTHEAST ALABAMA MEDICAL CENTER	3.2	292.0	0.010959
19	MARSHALL MEDICAL CENTERS	2.8	257.0	0.010895
38	NORTH ALABAMA MEDICAL CENTER	2.6	389.0	0.006684
57	MIZELL MEMORIAL HOSPITAL	2.8	31.0	0.090323
95	ST. VINCENT'S EAST	2.9	107.0	0.027103
...
91038	BAYLOR SCOTT & WHITE MEDICAL CENTER - MARBLE F...	2.2	61.0	0.036066
91095	LEGENT ORTHOPEDIC + SPINE	3.4	81.0	0.041975
91133	WISE HEALTH SYSTEM	2.8	35.0	0.080000
91190	PROVIDENCE HOSPITAL OF NORTH HOUSTON LLC	2.6	34.0	0.076471
91247	HOUSTON METHODIST THE WOODLANDS HOSPITAL	2.5	82.0	0.030488

2727 rows × 4 columns

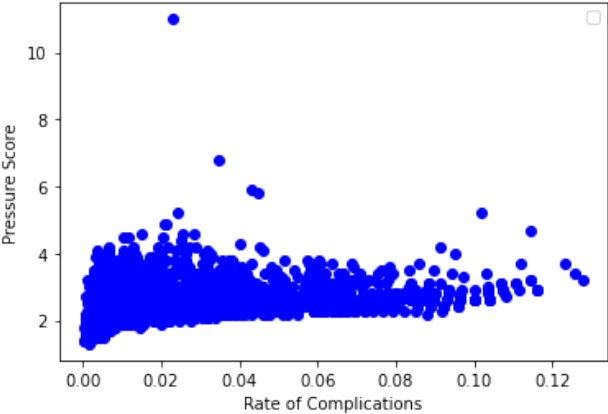
In [17]:

```
plt.scatter(finaldata['Complication Rate'],finaldata['Score'],color='b',)
plt.title('Relationship between rate of Complications for hip/knee replacements patients at different hospital')
plt.xlabel('Rate of Complications')
plt.ylabel('Pressure Score')
plt.legend()
```

No handles with labels found to put in legend.

Out[17]: <matplotlib.legend.Legend at 0x7f8f983f1550>

Realtionship between rate of Complications for hip/knee replacements patients at different hospital



In []:

In []:

In []: