 campusx-official Add files via upload 🕒 History

👤 1 contributor

1378 lines (1378 sloc) | 298 KB ⋮

```
In [36]: import pandas as pd
import seaborn as sns
```

```
In [37]: tips = sns.load_dataset('tips')
```

```
In [39]: tips.head()
```

```
Out[39]:
```

| | total_bill | tip | sex | smoker | day | time | size |
|---|------------|------|--------|--------|-----|--------|------|
| 0 | 16.99 | 1.01 | Female | No | Sun | Dinner | 2 |
| 1 | 10.34 | 1.66 | Male | No | Sun | Dinner | 3 |
| 2 | 21.01 | 3.50 | Male | No | Sun | Dinner | 3 |
| 3 | 23.68 | 3.31 | Male | No | Sun | Dinner | 2 |
| 4 | 24.59 | 3.61 | Female | No | Sun | Dinner | 4 |

```
In [18]: titanic = pd.read_csv('train.csv')
```

```
In [21]: flights = sns.load_dataset('flights')
```

```
In [40]: flights.head()
```

```
Out[40]:
```

| | year | month | passengers |
|---|------|----------|------------|
| 0 | 1949 | January | 112 |
| 1 | 1949 | February | 118 |
| 2 | 1949 | March | 132 |
| 3 | 1949 | April | 129 |
| 4 | 1949 | May | 121 |

```
In [22]: iris = sns.load_dataset('iris')
```

```
In [41]: iris
```

```
Out[41]:
```

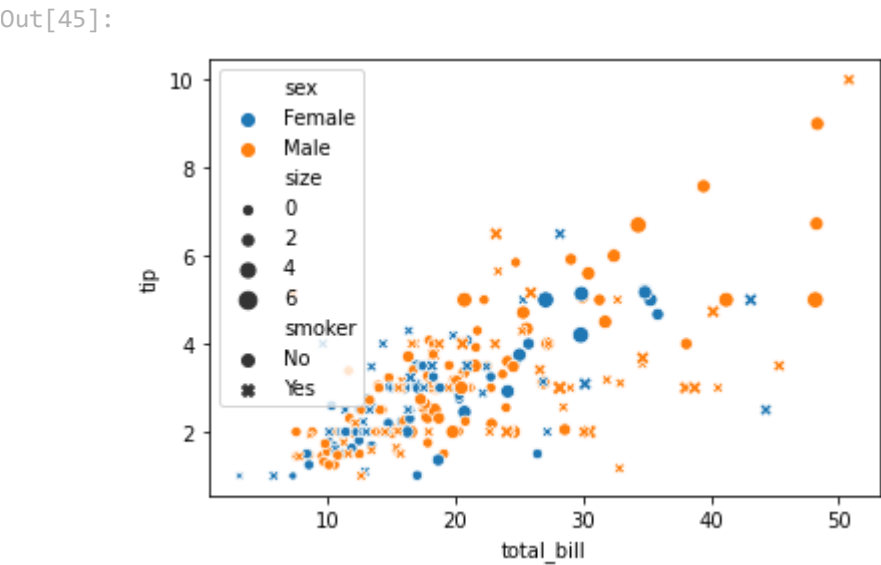
| | sepal_length | sepal_width | petal_length | petal_width | species |
|---|--------------|-------------|--------------|-------------|---------|
| 0 | 5.1 | 3.5 | 1.4 | 0.2 | setosa |
| 1 | 4.9 | 3.0 | 1.4 | 0.2 | setosa |
| 2 | 4.7 | 3.2 | 1.3 | 0.2 | setosa |
| 3 | 4.6 | 3.1 | 1.5 | 0.2 | setosa |

| | | | | | |
|-----|-----|-----|-----|-----|-----------|
| 4 | 5.0 | 3.6 | 1.4 | 0.2 | setosa |
| ... | ... | ... | ... | ... | ... |
| 145 | 6.7 | 3.0 | 5.2 | 2.3 | virginica |
| 146 | 6.3 | 2.5 | 5.0 | 1.9 | virginica |
| 147 | 6.5 | 3.0 | 5.2 | 2.0 | virginica |
| 148 | 6.2 | 3.4 | 5.4 | 2.3 | virginica |
| 149 | 5.9 | 3.0 | 5.1 | 1.8 | virginica |

150 rows × 5 columns

1. Scatterplot (Numerical - Numerical)

```
In [45]: sns.scatterplot(tips['total_bill'],tips['tip'],hue=df['sex'],style=df['smoker'])
```



2. Bar Plot (Numerical - Categorical)

```
In [47]: titanic.head()
```

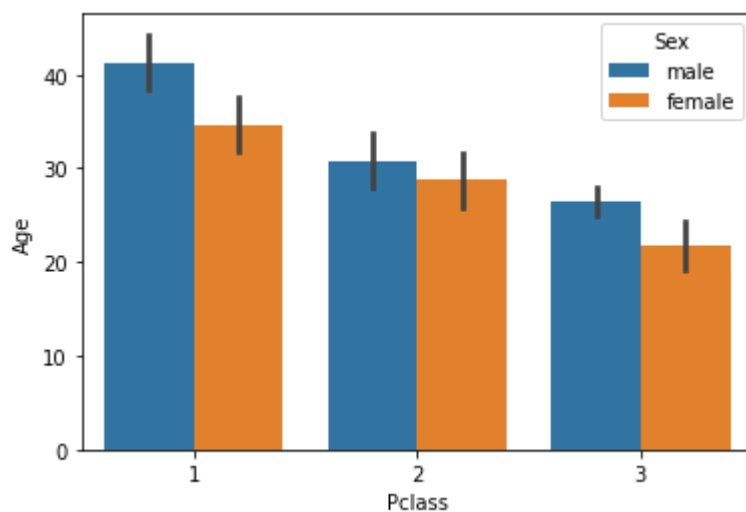
Out[47]:

| | PassengerId | Survived | Pclass | Name | Sex | Age | SibSp | Parch | Ticket | Fare | Ca |
|---|-------------|----------|--------|---|--------|------|-------|-------|-----------|---------|----|
| 0 | 1 | 0 | 3 | Braund, Mr. Owen Harris | male | 22.0 | 1 | 0 | A/5 21171 | 7.2500 | ↑ |
| 1 | 2 | 1 | 1 | Cumings, Mrs. John Bradley (Florence Briggs Th... | female | 38.0 | 1 | 0 | PC 17599 | 71.2833 | |

| | | | | | | | | | | | |
|---|---|---|---|--|--------|------|---|---|------------------|---------|---|
| 2 | 3 | 1 | 3 | Heikkinen, Miss. Laina | female | 26.0 | 0 | 0 | STON/O2. 3101282 | 7.9250 | ↑ |
| 3 | 4 | 1 | 1 | Futrelle, Mrs. Jacques Heath (Lily May Peel) | female | 35.0 | 1 | 0 | 113803 | 53.1000 | C |
| 4 | 5 | 0 | 3 | Allen, Mr. William Henry | male | 35.0 | 0 | 0 | 373450 | 8.0500 | ↑ |

In [52]: `sns.barplot(titanic['Pclass'],titanic['Age'],hue=titanic['Sex'])`

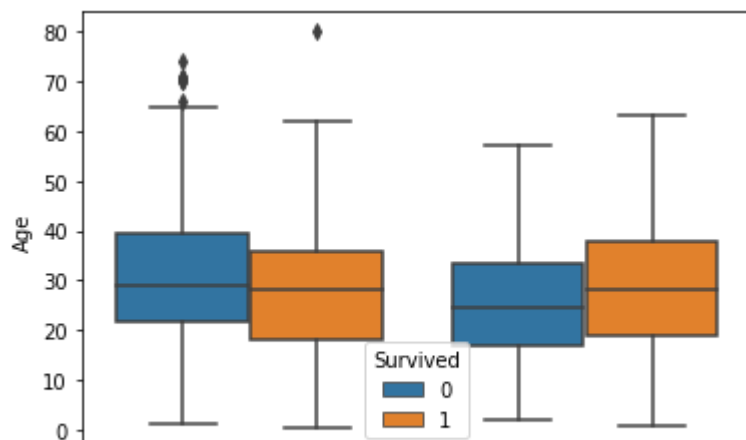
Out[52]:



3. Box Plot (Numerical - Categorical)

In [54]: `sns.boxplot(titanic['Sex'],titanic['Age'],hue=titanic['Survived'])`

Out[54]:



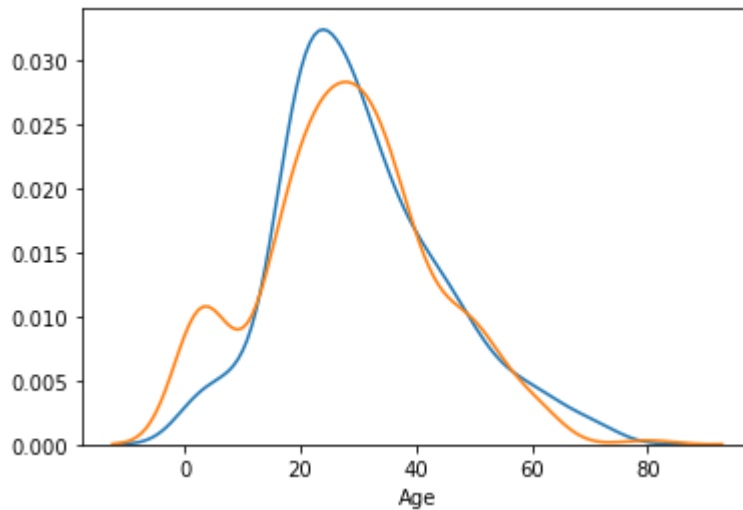
male female

Sex

4. Distplot (Numerical - Categorical)

```
In [61]: sns.distplot(titanic[titanic['Survived']==0]['Age'],hist=False)
sns.distplot(titanic[titanic['Survived']==1]['Age'],hist=False)
```

Out[61]:



5. HeatMap (Categorical - Categorical)

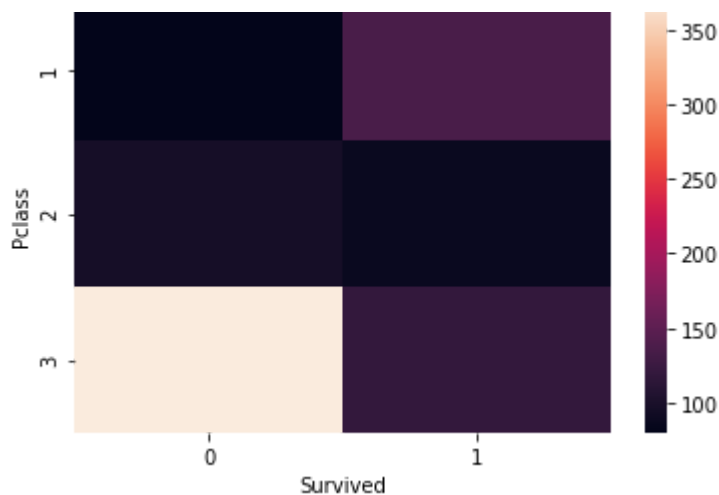
```
In [62]: titanic.head(3)
```

Out[62]:

| | PassengerId | Survived | Pclass | Name | Sex | Age | SibSp | Parch | Ticket | Fare | Cabin |
|---|-------------|----------|--------|---|--------|------|-------|-------|------------------|---------|-------|
| 0 | 1 | 0 | 3 | Braund, Mr. Owen Harris | male | 22.0 | 1 | 0 | A/5 21171 | 7.2500 | nan |
| 1 | 2 | 1 | 1 | Cumings, Mrs. John Bradley (Florence Briggs Th... | female | 38.0 | 1 | 0 | PC 17599 | 71.2833 | nan |
| 2 | 3 | 1 | 3 | Heikkinen, Miss. Laina | female | 26.0 | 0 | 0 | STON/O2. 3101282 | 7.9250 | nan |

```
In [64]: sns.heatmap(pd.crosstab(titanic['Pclass'],titanic['Survived']))
```

Out[64]:



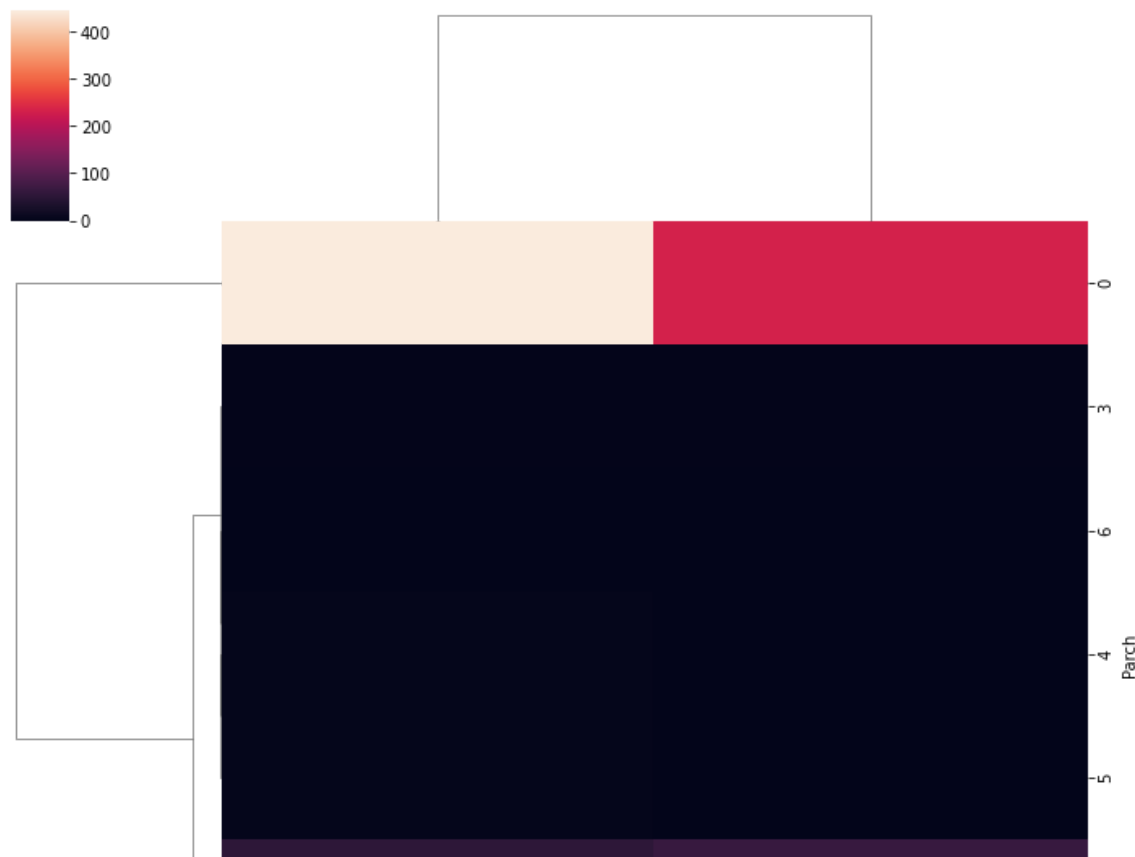
```
In [72]: (titanic.groupby('Embarked').mean()['Survived']*100)
```

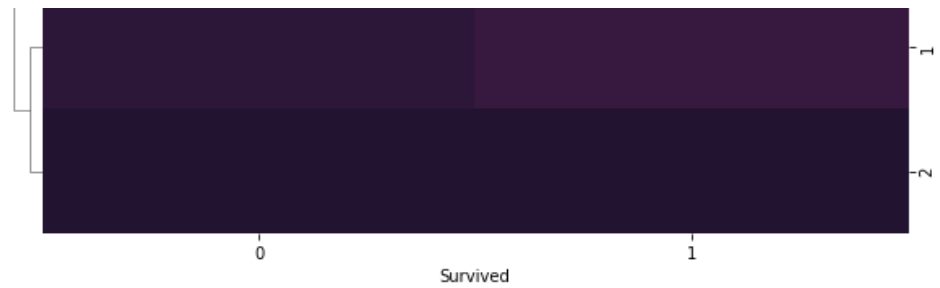
```
Out[72]: Embarked  
C      55.357143  
Q      38.961039  
S      33.695652  
Name: Survived, dtype: float64
```

6. ClusterMap (Categorical - Categorical)

```
In [75]: sns.clustermap(pd.crosstab(titanic['Parch'],titanic['Survived']))
```

```
Out[75]:
```





7. Pairplot

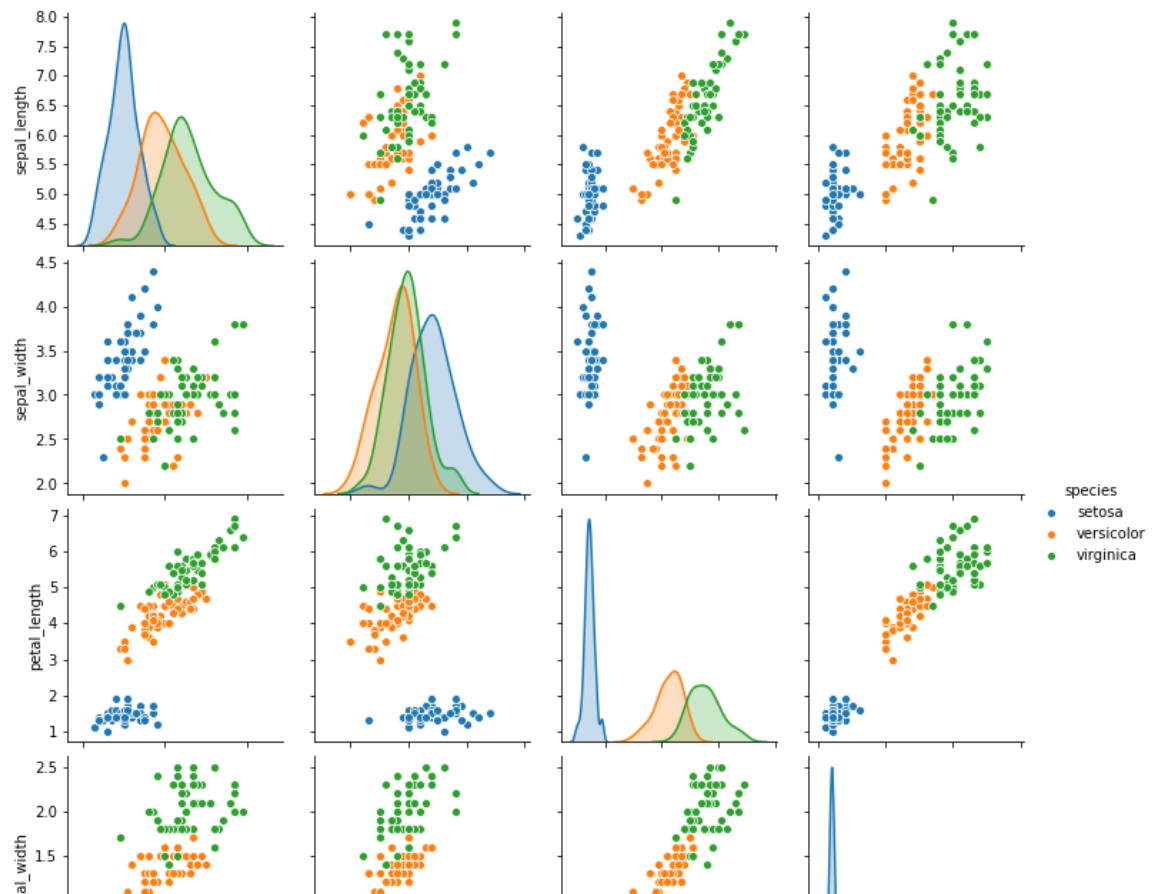
In [76]: `iris.head()`

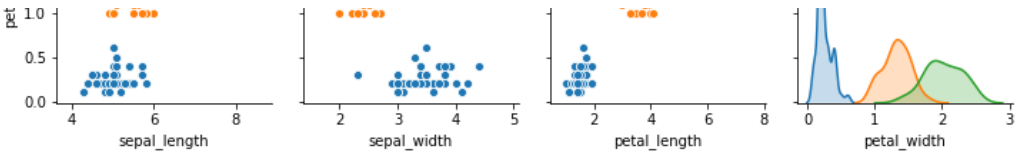
Out[76]:

| | sepal_length | sepal_width | petal_length | petal_width | species |
|---|--------------|-------------|--------------|-------------|---------|
| 0 | 5.1 | 3.5 | 1.4 | 0.2 | setosa |
| 1 | 4.9 | 3.0 | 1.4 | 0.2 | setosa |
| 2 | 4.7 | 3.2 | 1.3 | 0.2 | setosa |
| 3 | 4.6 | 3.1 | 1.5 | 0.2 | setosa |
| 4 | 5.0 | 3.6 | 1.4 | 0.2 | setosa |

In [78]: `sns.pairplot(iris,hue='species')`

Out[78]:





8. Lineplot (Numerical - Numerical)

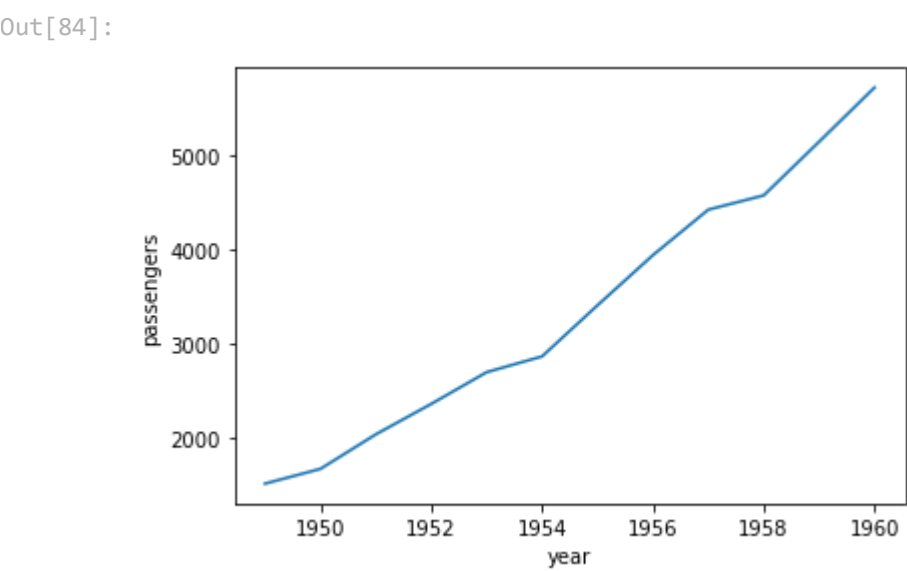
```
In [79]: flights.head()
```

Out[79]:

| | year | month | passengers |
|---|------|----------|------------|
| 0 | 1949 | January | 112 |
| 1 | 1949 | February | 118 |
| 2 | 1949 | March | 132 |
| 3 | 1949 | April | 129 |
| 4 | 1949 | May | 121 |

```
In [83]: new = flights.groupby('year').sum().reset_index()
```

```
In [84]: sns.lineplot(new['year'],new['passengers'])
```



```
In [85]: flights
```

Out[85]:

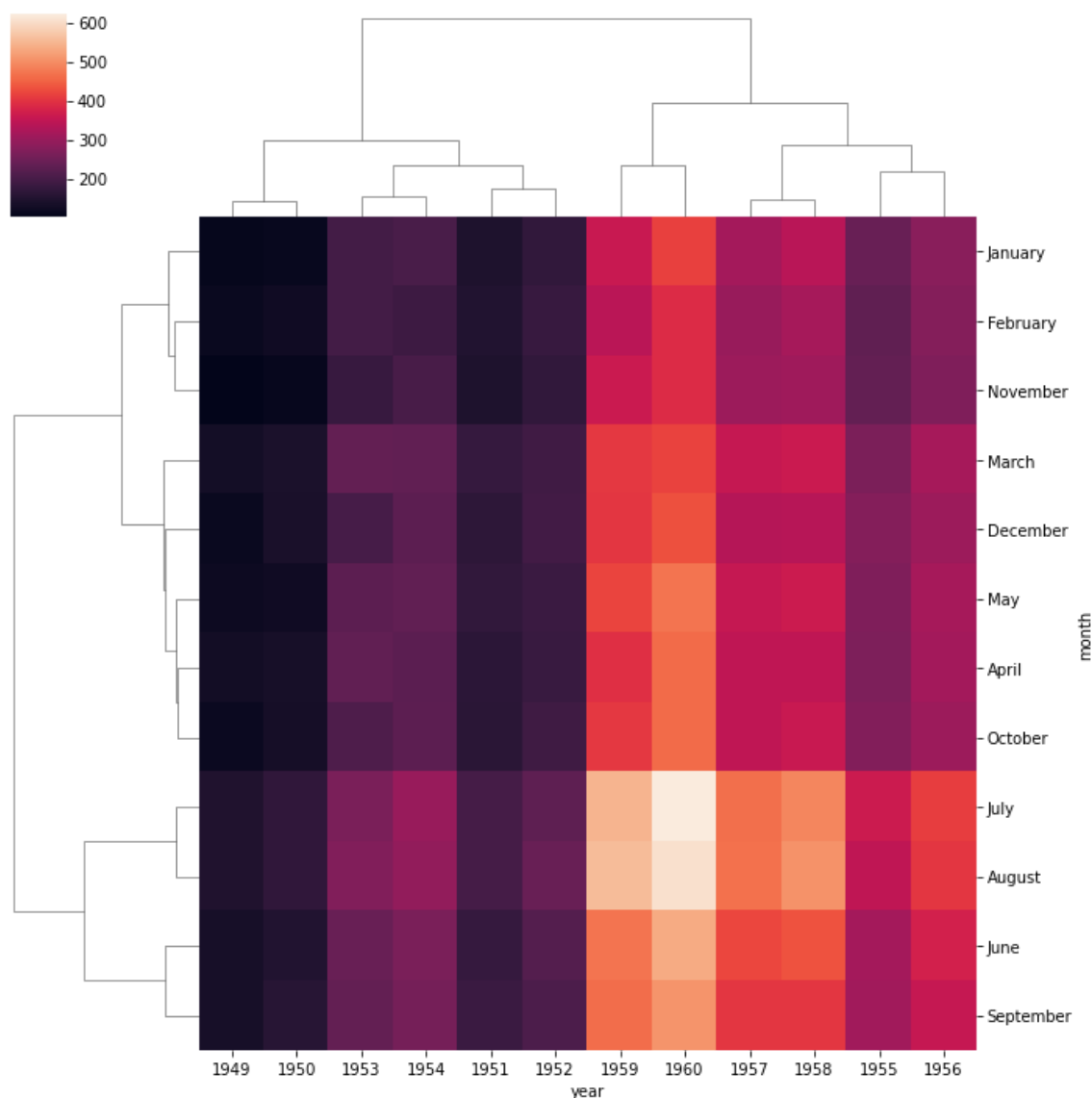
| | year | month | passengers |
|---|------|----------|------------|
| 0 | 1949 | January | 112 |
| 1 | 1949 | February | 118 |
| 2 | 1949 | March | 132 |
| 3 | 1949 | April | 129 |

| | | | |
|------------|------|-----------|-----|
| 4 | 1949 | May | 121 |
| ... | ... | ... | ... |
| 139 | 1960 | August | 606 |
| 140 | 1960 | September | 508 |
| 141 | 1960 | October | 461 |
| 142 | 1960 | November | 390 |
| 143 | 1960 | December | 432 |

144 rows × 3 columns

In [88]: `sns.clustermap(flights.pivot_table(values='passengers',index='month',columns='y`

Out[88]:



In []:

