3rd March Assignment

March 14, 2023

1 Assignment 27

Que 1: Name any five plots that we can plot using the Seaborn library. Also, state the uses of each plot.

Ans - Seaborn is a Python data visualization library based on Matplotlib. It provides a high-level interface for creating informative and attractive statistical graphics. Here are five types of plots that can be created using Seaborn along with their uses:

1. Scatter plot: A scatter plot displays the relationship between two continuous variables. It can be used to identify patterns and trends in data.
2. Line plot: A line plot displays the relationship between two continuous variables using lines. It can be used to show changes in data over time or to display trends.
3. Bar plot: A bar plot displays categorical data using rectangular bars. It can be used to compare the values of different categories.
4. Histogram: A histogram displays the distribution of a continuous variable. It can be used to identify the shape of the distribution and to identify outliers.
5. Heatmap: A heatmap displays the relationship between two categorical variables using color. It can be used to identify patterns and relationships between different categories

Que 2: Load the "fmri" dataset using the load_dataset function of seaborn. Plot a line plot using x ="timepoint" and y = "signal" for different events and regions.

Ans.

[1]:
```
import seaborn as sns
```

[5]:
```
d=sns.load_dataset('fmri')
```

[6]:
```
d
```

[6]:

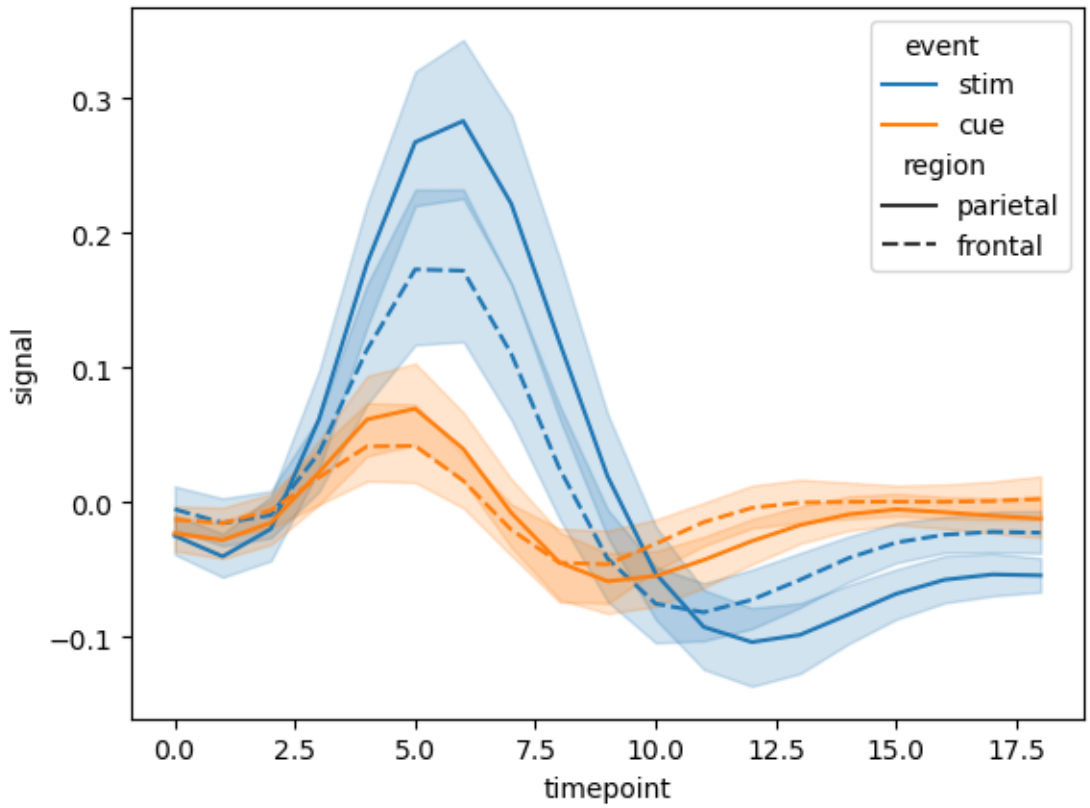
	subject	timepoint	event	region	signal
0	s13	18	stim	parietal	-0.017552
1	s5	14	stim	parietal	-0.080883
2	s12	18	stim	parietal	-0.081033
3	s11	18	stim	parietal	-0.046134
4	s10	18	stim	parietal	-0.037970
…	…	…	…	…	…
1059	s0	8	cue	frontal	0.018165

```
1060     s13          7   cue   frontal -0.029130
1061     s12          7   cue   frontal -0.004939
1062     s11          7   cue   frontal -0.025367
1063      s0          0   cue  parietal -0.006899

[1064 rows x 5 columns]
```

[13]:
```
sns.lineplot(data=d,x=d.timepoint,y=d.signal,hue='event',style='region')
```

[13]:
```
<AxesSubplot: xlabel='timepoint', ylabel='signal'>
```



Que 3: Load the "titanic" dataset using the load_dataset function of seaborn. Plot two box plots using x ='pclass', y = 'age' and y = 'fare'.

Ans.

[14]:
```
t=sns.load_dataset('titanic')
```

[15]:
```
t
```

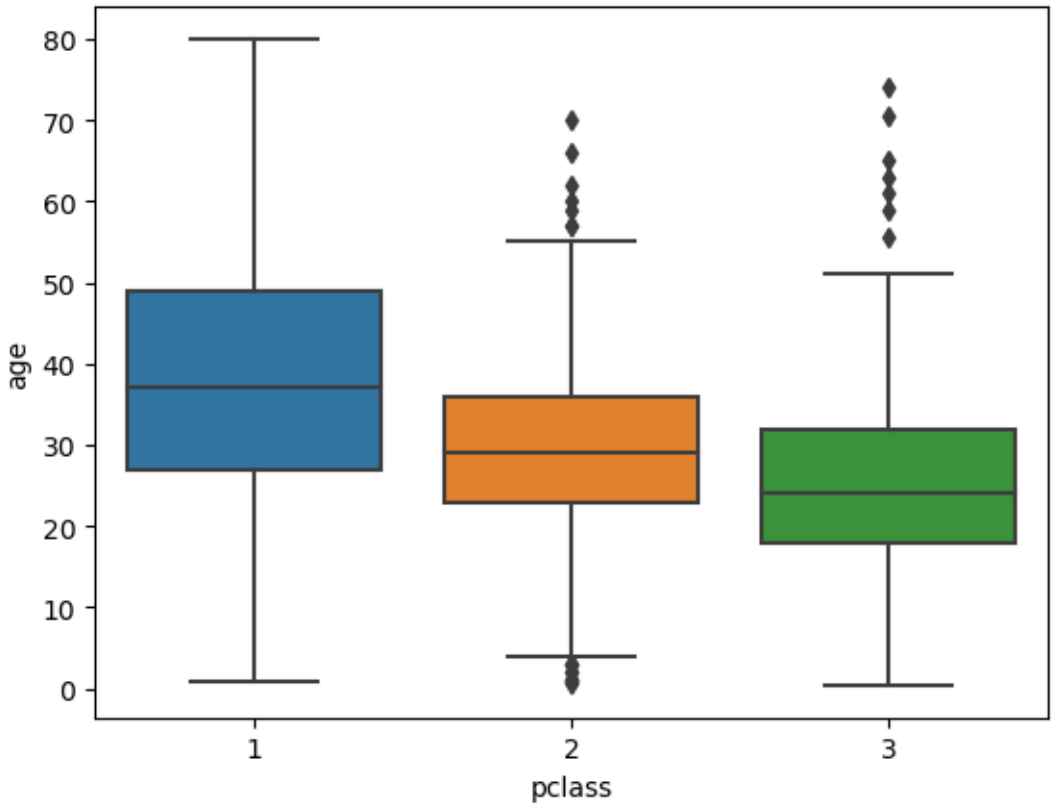
[15]:
```
     survived  pclass     sex   age  sibsp  parch     fare embarked   class  \
0           0       3    male  22.0      1      0   7.2500        S   Third
1           1       1  female  38.0      1      0  71.2833        C   First
2           1       3  female  26.0      0      0   7.9250        S   Third
3           1       1  female  35.0      1      0  53.1000        S   First
4           0       3    male  35.0      0      0   8.0500        S   Third
..        …     …      …   …    …    …       …      …     …
886         0       2    male  27.0      0      0  13.0000        S  Second
887         1       1  female  19.0      0      0  30.0000        S   First
888         0       3  female   NaN      1      2  23.4500        S   Third
889         1       1    male  26.0      0      0  30.0000        C   First
890         0       3    male  32.0      0      0   7.7500        Q   Third

       who  adult_male deck  embark_town alive  alone
0      man        True  NaN  Southampton    no  False
1    woman       False    C    Cherbourg   yes  False
2    woman       False  NaN  Southampton   yes   True
3    woman       False    C  Southampton   yes  False
4      man        True  NaN  Southampton    no   True
..     …          …  …          …   …    …
886    man        True  NaN  Southampton    no   True
887  woman       False    B  Southampton   yes   True
888  woman       False  NaN  Southampton    no  False
889    man        True    C    Cherbourg   yes   True
890    man        True  NaN   Queenstown    no   True

[891 rows x 15 columns]
```

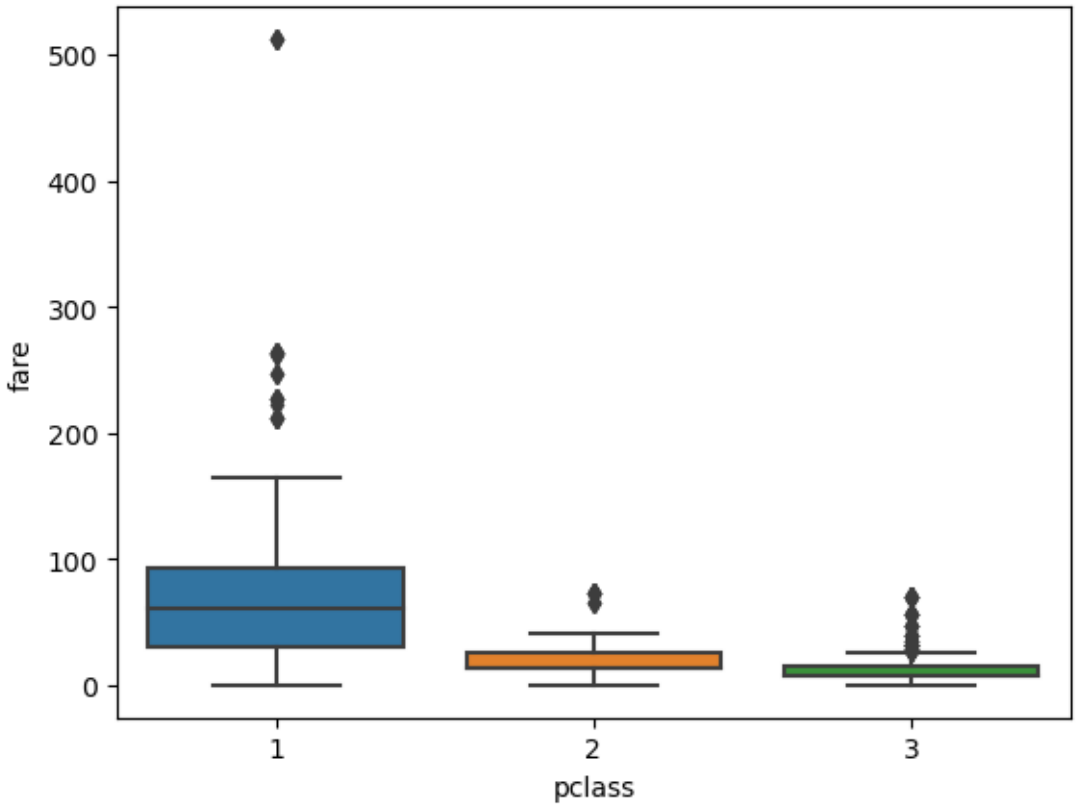
[20]:
```
sns.boxplot(data=t,x=t.pclass,y=t.age)
```

[20]:
```
<AxesSubplot: xlabel='pclass', ylabel='age'>
```



[19]:
```
sns.boxplot(data=t,x=t.pclass,y=t.fare)
```

[19]:
```
<AxesSubplot: xlabel='pclass', ylabel='fare'>
```



Que 4: Use the "diamonds" dataset from seaborn to plot a histogram for the 'price' column. Use the hue parameter for the 'cut' column of the diamonds dataset.

Ans.

[21]:
```
diamond=sns.load_dataset('diamonds')
```

[22]:
```
diamond
```

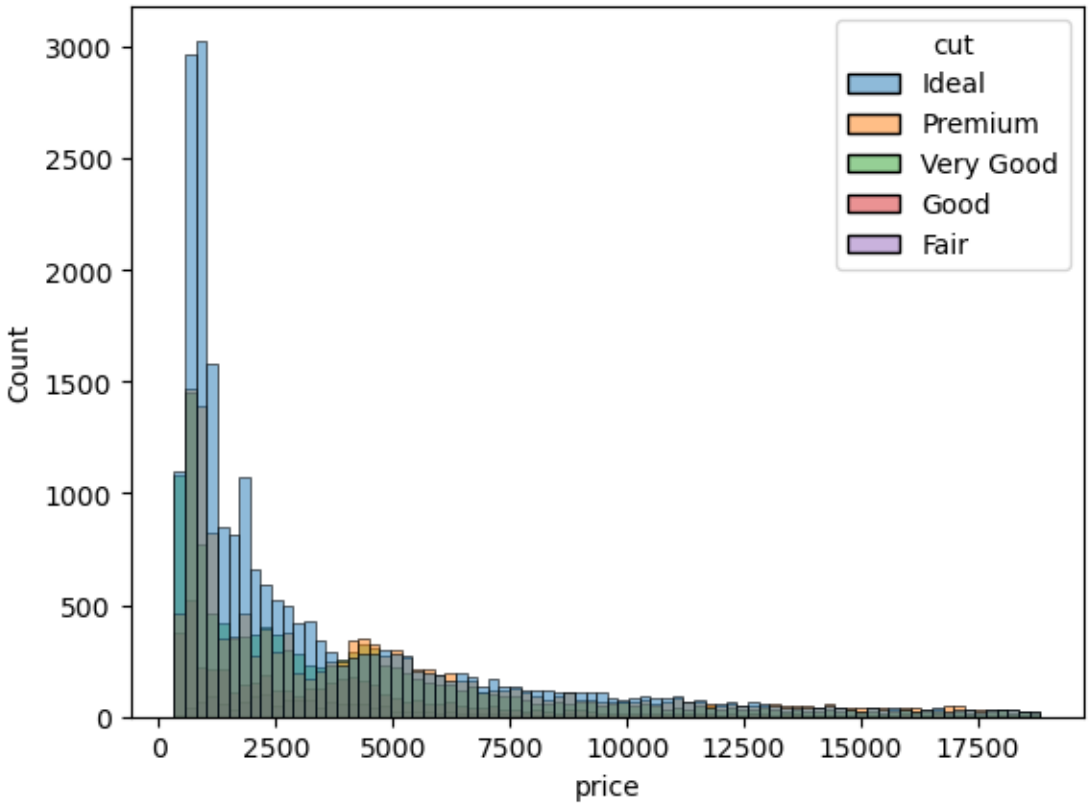
[22]:

	carat	cut	color	clarity	depth	table	price	x	y	z
0	0.23	Ideal	E	SI2	61.5	55.0	326	3.95	3.98	2.43
1	0.21	Premium	E	SI1	59.8	61.0	326	3.89	3.84	2.31
2	0.23	Good	E	VS1	56.9	65.0	327	4.05	4.07	2.31
3	0.29	Premium	I	VS2	62.4	58.0	334	4.20	4.23	2.63
4	0.31	Good	J	SI2	63.3	58.0	335	4.34	4.35	2.75
…	…	…	…	…	…	…	…	…	…	…
53935	0.72	Ideal	D	SI1	60.8	57.0	2757	5.75	5.76	3.50
53936	0.72	Good	D	SI1	63.1	55.0	2757	5.69	5.75	3.61
53937	0.70	Very Good	D	SI1	62.8	60.0	2757	5.66	5.68	3.56
53938	0.86	Premium	H	SI2	61.0	58.0	2757	6.15	6.12	3.74
53939	0.75	Ideal	D	SI2	62.2	55.0	2757	5.83	5.87	3.64

[53940 rows x 10 columns]

[23]:
```
sns.histplot(data=diamond,x=diamond.price,hue=diamond.cut)
```

[23]: <AxesSubplot: xlabel='price', ylabel='Count'>



Que 5: Use the "iris" dataset from seaborn to plot a pair plot. Use the hue parameter for the "species" column of the iris dataset.

Ans.

[24]:
```
iris=sns.load_dataset('iris')
```

[25]:
```
iris
```

[25]:

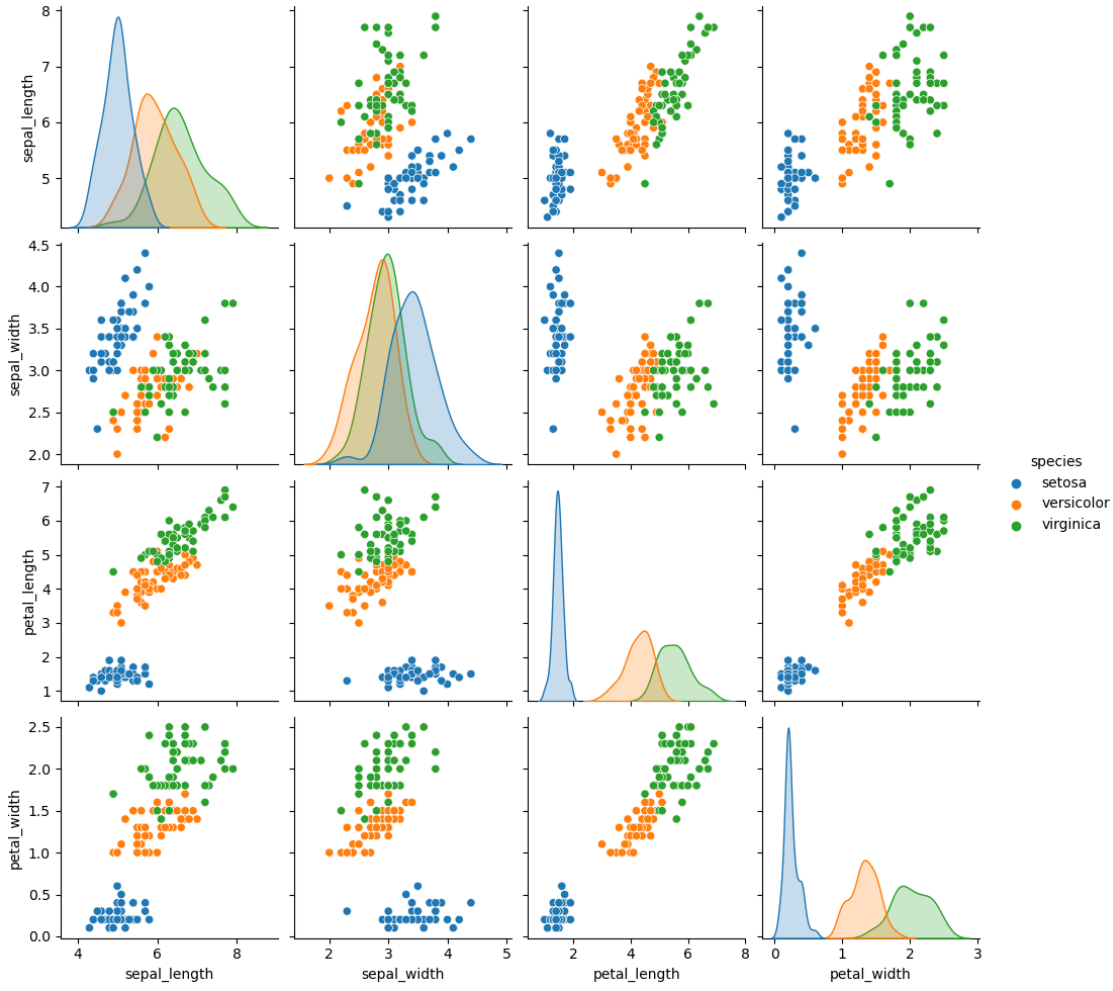
	sepal_length	sepal_width	petal_length	petal_width	species
0	5.1	3.5	1.4	0.2	setosa
1	4.9	3.0	1.4	0.2	setosa
2	4.7	3.2	1.3	0.2	setosa
3	4.6	3.1	1.5	0.2	setosa

```
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 x 5 columns]
```

[26]:
```
sns.pairplot(data=iris,hue='species')
```

[26]:
```
<seaborn.axisgrid.PairGrid at 0x7f3c1cd02f80>
```



Que 6: Use the "flights" dataset from seaborn to plot a heatmap.

Ans.

[28]:
```
f=sns.load_dataset('flights')
```

[31]:
```
f
```

[31]:
```
     year month  passengers
0    1949   Jan         112
1    1949   Feb         118
2    1949   Mar         132
3    1949   Apr         129
4    1949   May         121
..    …    …          …
139  1960   Aug         606
140  1960   Sep         508
141  1960   Oct         461
142  1960   Nov         390
143  1960   Dec         432

[144 rows x 3 columns]
```

[32]:
```
month_dict = {
    'Jan': 1,
    'Feb': 2,
    'Mar': 3,
    'Apr': 4,
    'May': 5,
    'Jun': 6,
    'Jul': 7,
    'Aug': 8,
    'Sep': 9,
    'Oct': 10,
    'Nov': 11,
    'Dec': 12
}
```

[33]:
```
f['month']=f['month'].map(month_dict)
```

[34]:
```
f
```

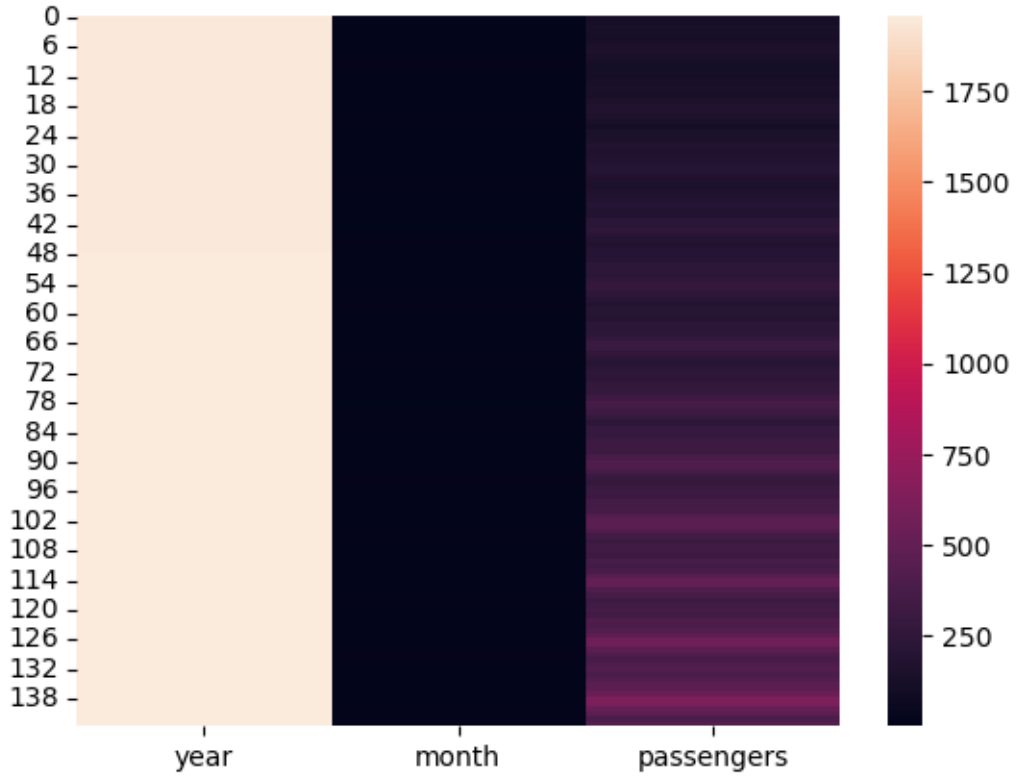
[34]:
```
     year month  passengers
0    1949     1         112
1    1949     2         118
2    1949     3         132
3    1949     4         129
4    1949     5         121
..    …    …          …
139  1960     8         606
```

```
140  1960     9          508
141  1960    10          461
142  1960    11          390
143  1960    12          432

[144 rows x 3 columns]
```

[35]:
```
sns.heatmap(data=f)
```

[35]:
```
<AxesSubplot: >
```



[ ]: