

<https://www.youtube.com/watch?v=DWVLRhnuGqI&t=1531s>
(<https://www.youtube.com/watch?v=DWVLRhnuGqI&t=1531s>)

https://colab.research.google.com/drive/1_Mk2NWYBzNxICokEtJFxf6haq_1YIDA?usp=sharing#scrollTo=LTIOlNe4sar7
(https://colab.research.google.com/drive/1_Mk2NWYBzNxICokEtJFxf6haq_1YIDA?usp=sharing#scrollTo=LTIOlNe4sar7)

```
In [1]: import seaborn as sns
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import plotly.express as px
```

1. Relational plot

scatter and line plot

```
In [2]: tips=sns.load_dataset('tips')
```

In [38]: tips

Out[38]:

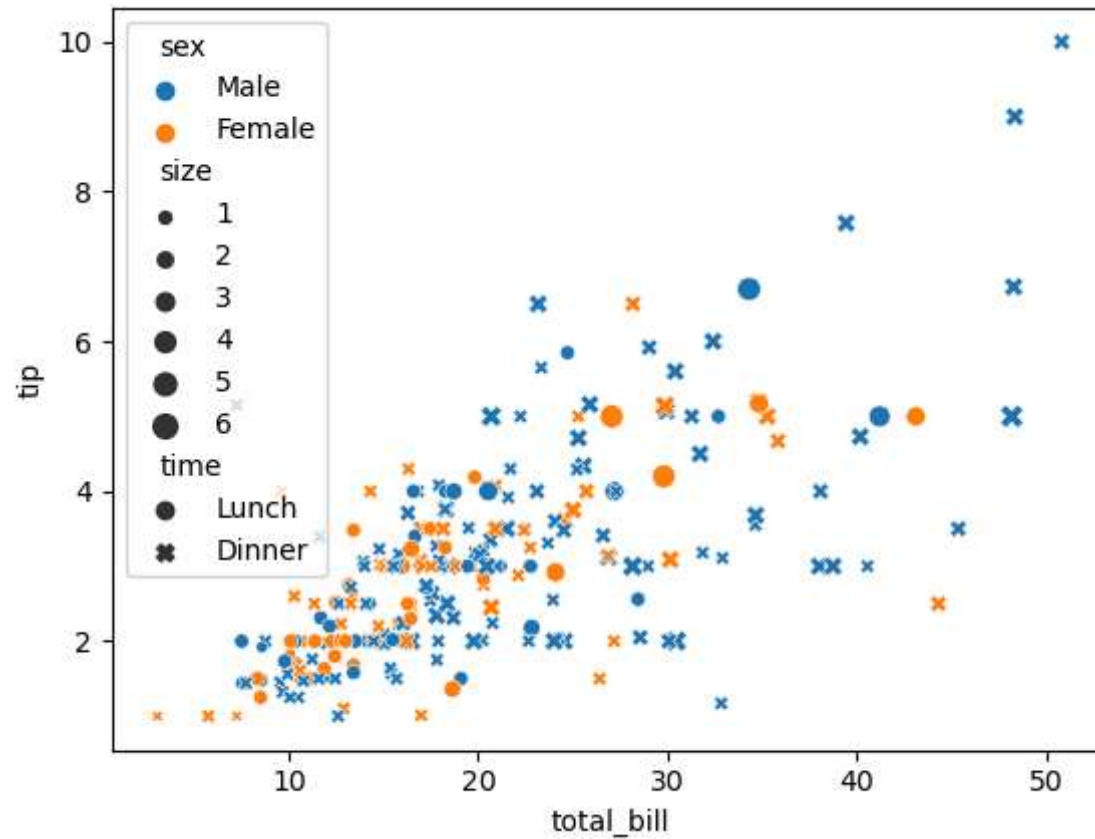
	total_bill	tip	sex	smoker	day	time	size
0	16.99	1.01	Female	No	Sun	Dinner	2
1	10.34	1.66	Male	No	Sun	Dinner	3
2	21.01	3.50	Male	No	Sun	Dinner	3
3	23.68	3.31	Male	No	Sun	Dinner	2
4	24.59	3.61	Female	No	Sun	Dinner	4
...
239	29.03	5.92	Male	No	Sat	Dinner	3
240	27.18	2.00	Female	Yes	Sat	Dinner	2
241	22.67	2.00	Male	Yes	Sat	Dinner	2
242	17.82	1.75	Male	No	Sat	Dinner	2
243	18.78	3.00	Female	No	Thur	Dinner	2

244 rows × 7 columns

1.1 scatter plot -> axis level function

```
In [39]: sns.scatterplot(data=tips,x='total_bill',y='tip',hue='sex',style='time',size='size')
```

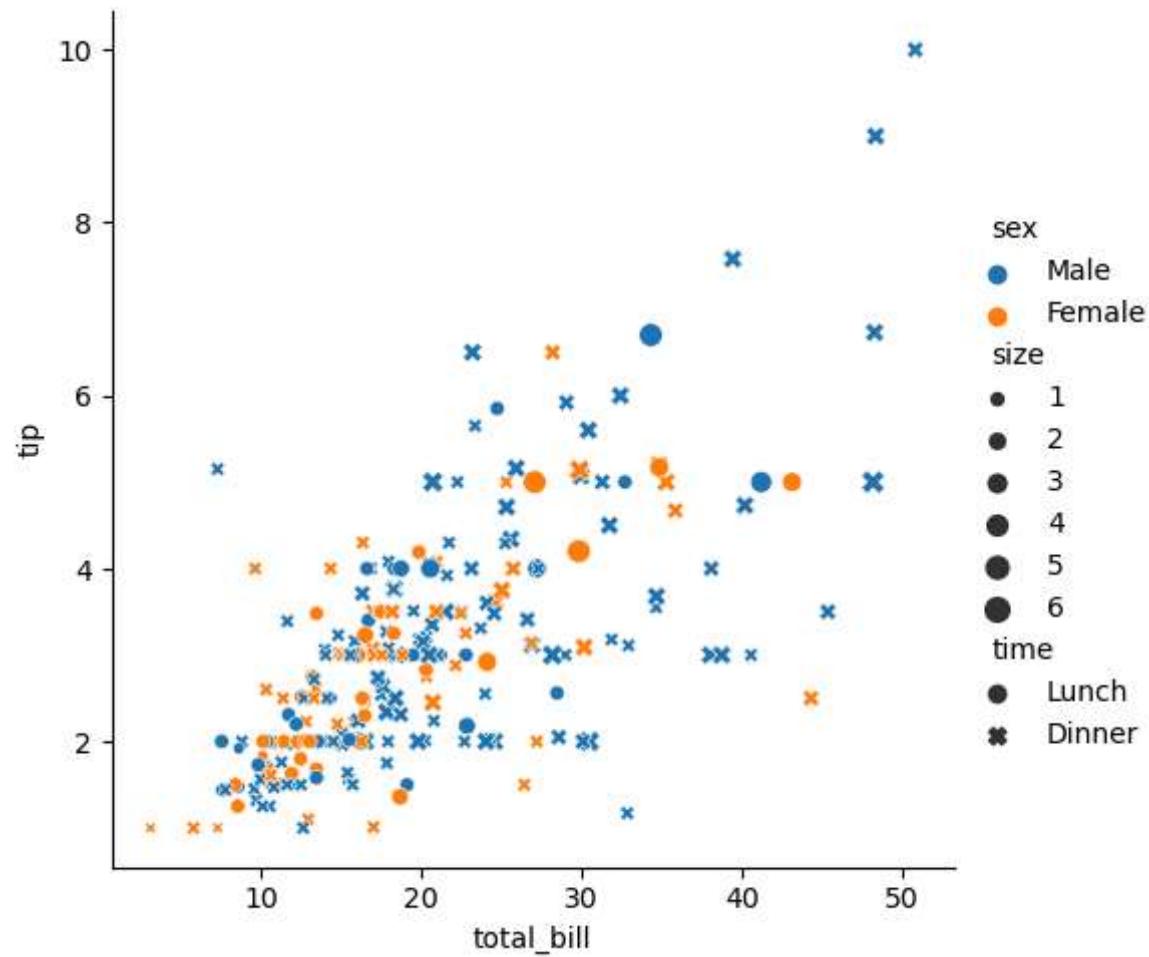
```
Out[39]: <Axes: xlabel='total_bill', ylabel='tip'>
```



relplot -> figure level

```
In [40]: sns.relplot(data=tips,x='total_bill',y='tip',kind='scatter',hue='sex',style='time',size='size')
```

```
Out[40]: <seaborn.axisgrid.FacetGrid at 0x22a03f62810>
```



1.2 line plot (mostly for timeseries)

```
In [3]: gap= px.data.gapminder()
```

```
In [42]: gap
```

```
Out[42]:
```

	country	continent	year	lifeExp	pop	gdpPercap	iso_alpha	iso_num
0	Afghanistan	Asia	1952	28.801	8425333	779.445314	AFG	4
1	Afghanistan	Asia	1957	30.332	9240934	820.853030	AFG	4
2	Afghanistan	Asia	1962	31.997	10267083	853.100710	AFG	4
3	Afghanistan	Asia	1967	34.020	11537966	836.197138	AFG	4
4	Afghanistan	Asia	1972	36.088	13079460	739.981106	AFG	4
...
1699	Zimbabwe	Africa	1987	62.351	9216418	706.157306	ZWE	716
1700	Zimbabwe	Africa	1992	60.377	10704340	693.420786	ZWE	716
1701	Zimbabwe	Africa	1997	46.809	11404948	792.449960	ZWE	716
1702	Zimbabwe	Africa	2002	39.989	11926563	672.038623	ZWE	716
1703	Zimbabwe	Africa	2007	43.487	12311143	469.709298	ZWE	716

1704 rows × 8 columns

```
In [43]: tempdf=gap[gap['country']=='India']
```

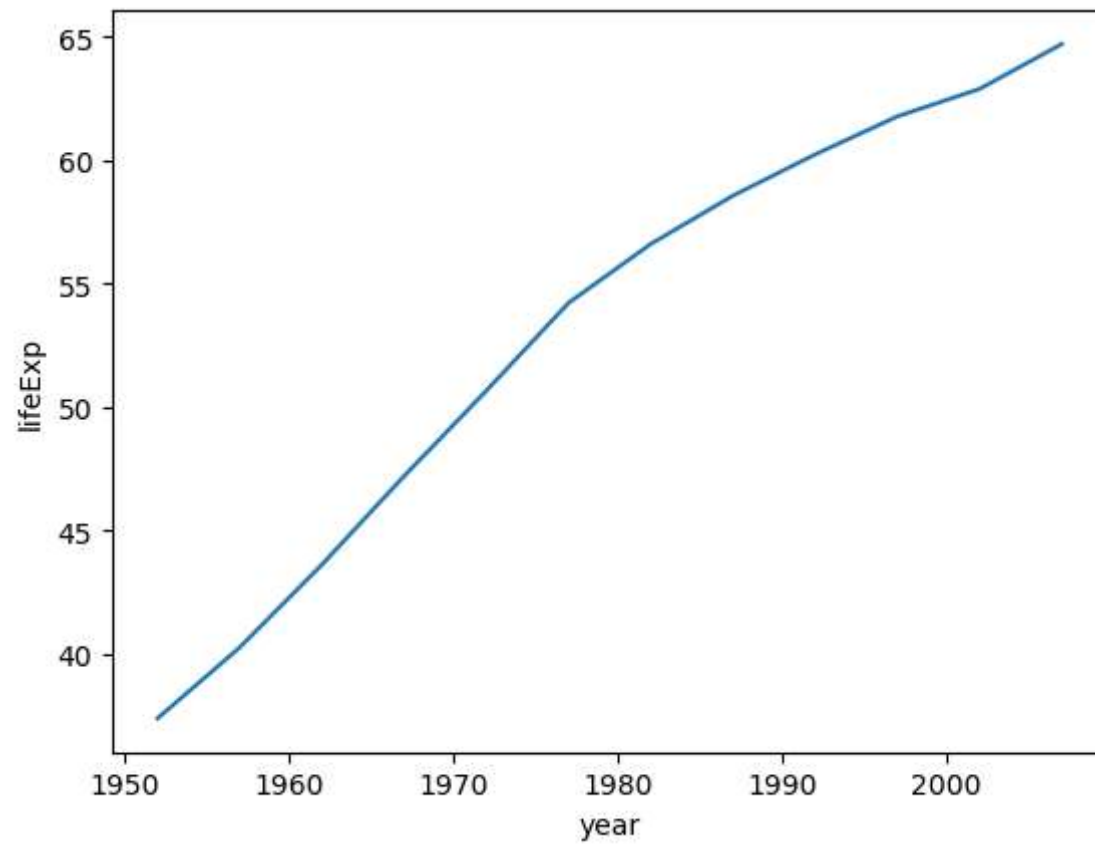
In [44]: tempdf

Out[44]:

	country	continent	year	lifeExp	pop	gdpPercap	iso_alpha	iso_num
696	India	Asia	1952	37.373	372000000	546.565749	IND	356
697	India	Asia	1957	40.249	409000000	590.061996	IND	356
698	India	Asia	1962	43.605	454000000	658.347151	IND	356
699	India	Asia	1967	47.193	506000000	700.770611	IND	356
700	India	Asia	1972	50.651	567000000	724.032527	IND	356
701	India	Asia	1977	54.208	634000000	813.337323	IND	356
702	India	Asia	1982	56.596	708000000	855.723538	IND	356
703	India	Asia	1987	58.553	788000000	976.512676	IND	356
704	India	Asia	1992	60.223	872000000	1164.406809	IND	356
705	India	Asia	1997	61.765	959000000	1458.817442	IND	356
706	India	Asia	2002	62.879	1034172547	1746.769454	IND	356
707	India	Asia	2007	64.698	1110396331	2452.210407	IND	356

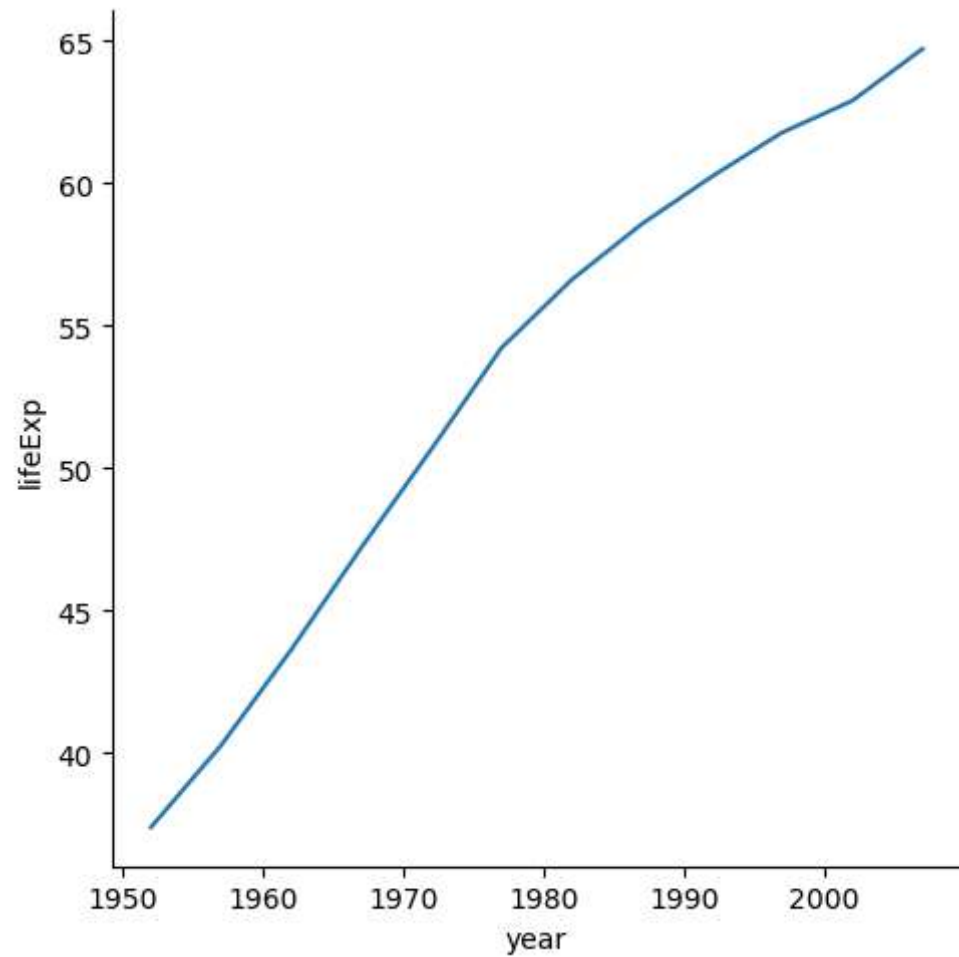
```
In [45]: sns.lineplot(data=tempdf,x='year',y='lifeExp',)
```

```
Out[45]: <Axes: xlabel='year', ylabel='lifeExp'>
```



```
In [46]: sns.relplot(data=tempdf,x='year',y='lifeExp',kind='line')
```

```
Out[46]: <seaborn.axisgrid.FacetGrid at 0x22a03ff3910>
```



```
In [47]: tempdf2=gap[gap['country'].isin(['India','Brazil','Germany'])]
```


In [48]: tempdf2

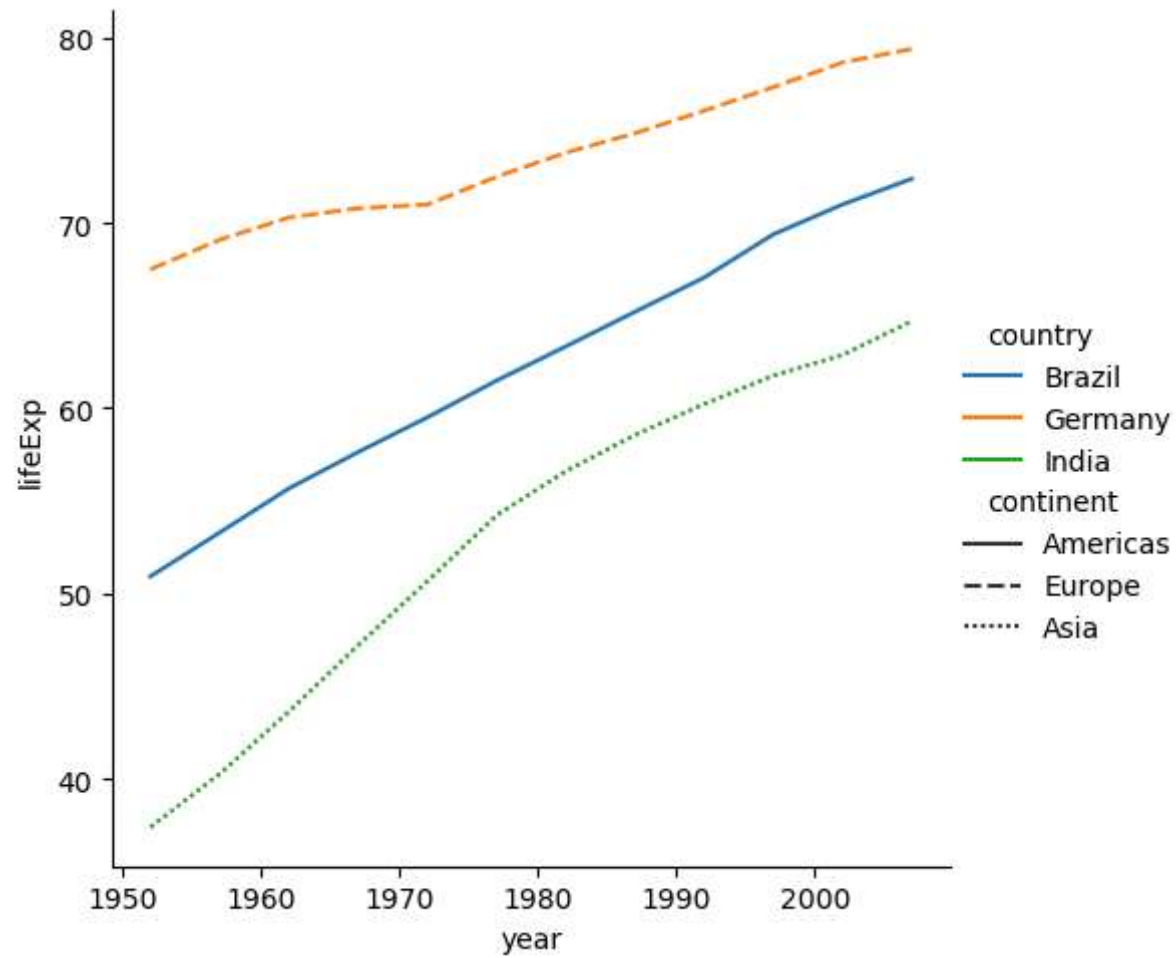
Out[48]:

	country	continent	year	lifeExp	pop	gdpPercap	iso_alpha	iso_num
168	Brazil	Americas	1952	50.917	56602560	2108.944355	BRA	76
169	Brazil	Americas	1957	53.285	65551171	2487.365989	BRA	76
170	Brazil	Americas	1962	55.665	76039390	3336.585802	BRA	76
171	Brazil	Americas	1967	57.632	88049823	3429.864357	BRA	76
172	Brazil	Americas	1972	59.504	100840058	4985.711467	BRA	76
173	Brazil	Americas	1977	61.489	114313951	6660.118654	BRA	76
174	Brazil	Americas	1982	63.336	128962939	7030.835878	BRA	76
175	Brazil	Americas	1987	65.205	142938076	7807.095818	BRA	76
176	Brazil	Americas	1992	67.057	155975974	6950.283021	BRA	76
177	Brazil	Americas	1997	69.388	168546719	7957.980824	BRA	76
178	Brazil	Americas	2002	71.006	179914212	8131.212843	BRA	76
179	Brazil	Americas	2007	72.390	190010647	9065.800825	BRA	76
564	Germany	Europe	1952	67.500	69145952	7144.114393	DEU	276
565	Germany	Europe	1957	69.100	71019069	10187.826650	DEU	276
566	Germany	Europe	1962	70.300	73739117	12902.462910	DEU	276
567	Germany	Europe	1967	70.800	76368453	14745.625610	DEU	276
568	Germany	Europe	1972	71.000	78717088	18016.180270	DEU	276
569	Germany	Europe	1977	72.500	78160773	20512.921230	DEU	276
570	Germany	Europe	1982	73.800	78335266	22031.532740	DEU	276
571	Germany	Europe	1987	74.847	77718298	24639.185660	DEU	276
572	Germany	Europe	1992	76.070	80597764	26505.303170	DEU	276
573	Germany	Europe	1997	77.340	82011073	27788.884160	DEU	276
574	Germany	Europe	2002	78.670	82350671	30035.801980	DEU	276
575	Germany	Europe	2007	79.406	82400996	32170.374420	DEU	276
696	India	Asia	1952	37.373	372000000	546.565749	IND	356
697	India	Asia	1957	40.249	409000000	590.061996	IND	356

	country	continent	year	lifeExp	pop	gdpPercap	iso_alpha	iso_num
698	India	Asia	1962	43.605	454000000	658.347151	IND	356
699	India	Asia	1967	47.193	506000000	700.770611	IND	356
700	India	Asia	1972	50.651	567000000	724.032527	IND	356
701	India	Asia	1977	54.208	634000000	813.337323	IND	356
702	India	Asia	1982	56.596	708000000	855.723538	IND	356
703	India	Asia	1987	58.553	788000000	976.512676	IND	356
704	India	Asia	1992	60.223	872000000	1164.406809	IND	356
705	India	Asia	1997	61.765	959000000	1458.817442	IND	356
706	India	Asia	2002	62.879	1034172547	1746.769454	IND	356
707	India	Asia	2007	64.698	1110396331	2452.210407	IND	356

```
In [49]: sns.relplot(data=tempdf2,x='year',y='lifeExp',kind='line',hue='country',style='continent')
```

```
Out[49]: <seaborn.axisgrid.FacetGrid at 0x22a049225d0>
```



Facet plot

```
In [50]: tips
```

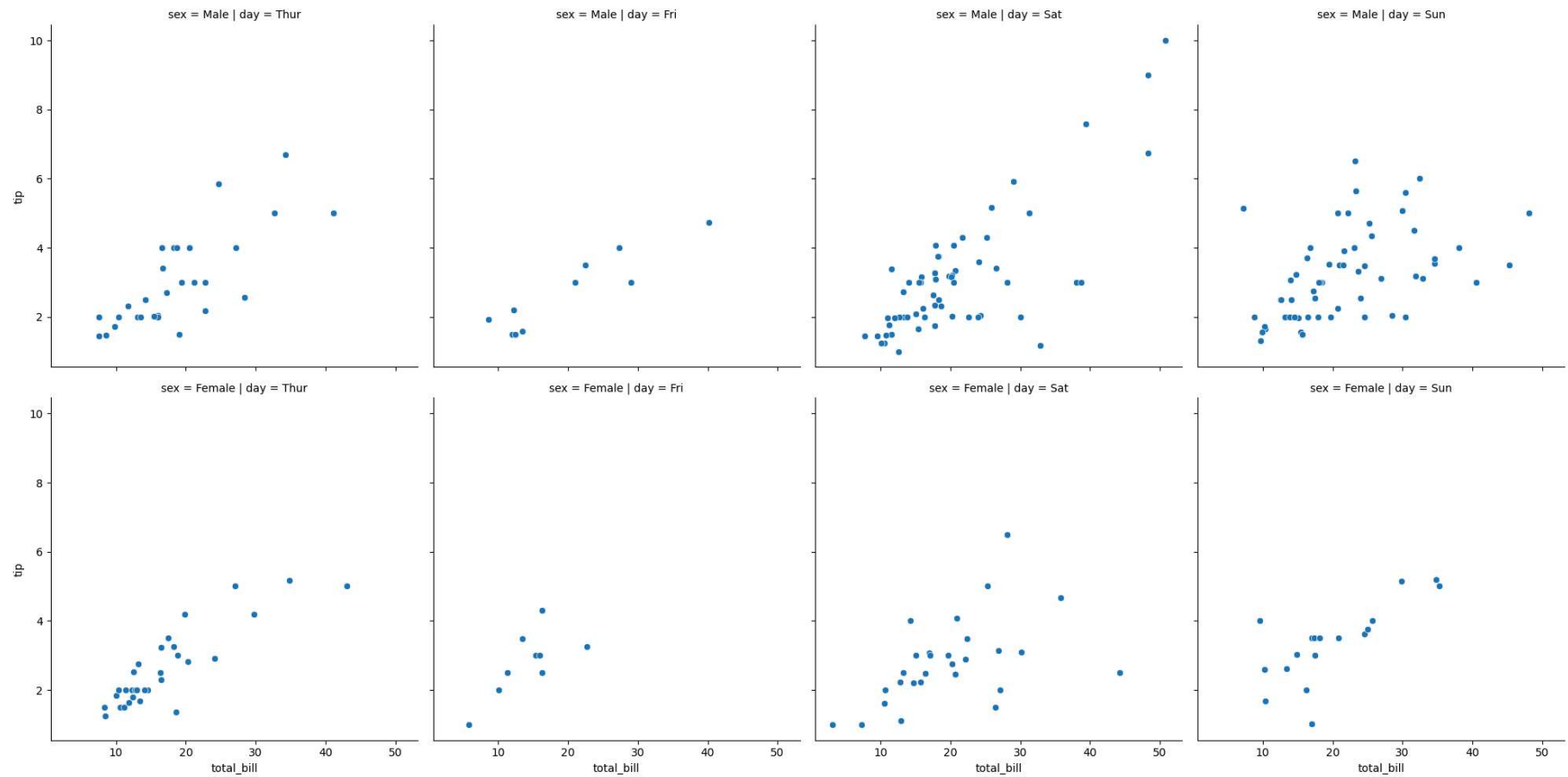
Out[50]:

	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
...
239	29.03	5.92	Male	No	Sat	Dinner	3
240	27.18	2.00	Female	Yes	Sat	Dinner	2
241	22.67	2.00	Male	Yes	Sat	Dinner	2
242	17.82	1.75	Male	No	Sat	Dinner	2
243	18.78	3.00	Female	No	Thur	Dinner	2

244 rows × 7 columns

```
In [51]: #facet plot is figure level function so it can only work with rel plot
sns.relplot(data=tips,x='total_bill',y='tip',kind='scatter',row='sex',col='day')
```

```
Out[51]: <seaborn.axisgrid.FacetGrid at 0x22a05bada90>
```



In [52]: gap

Out[52]:

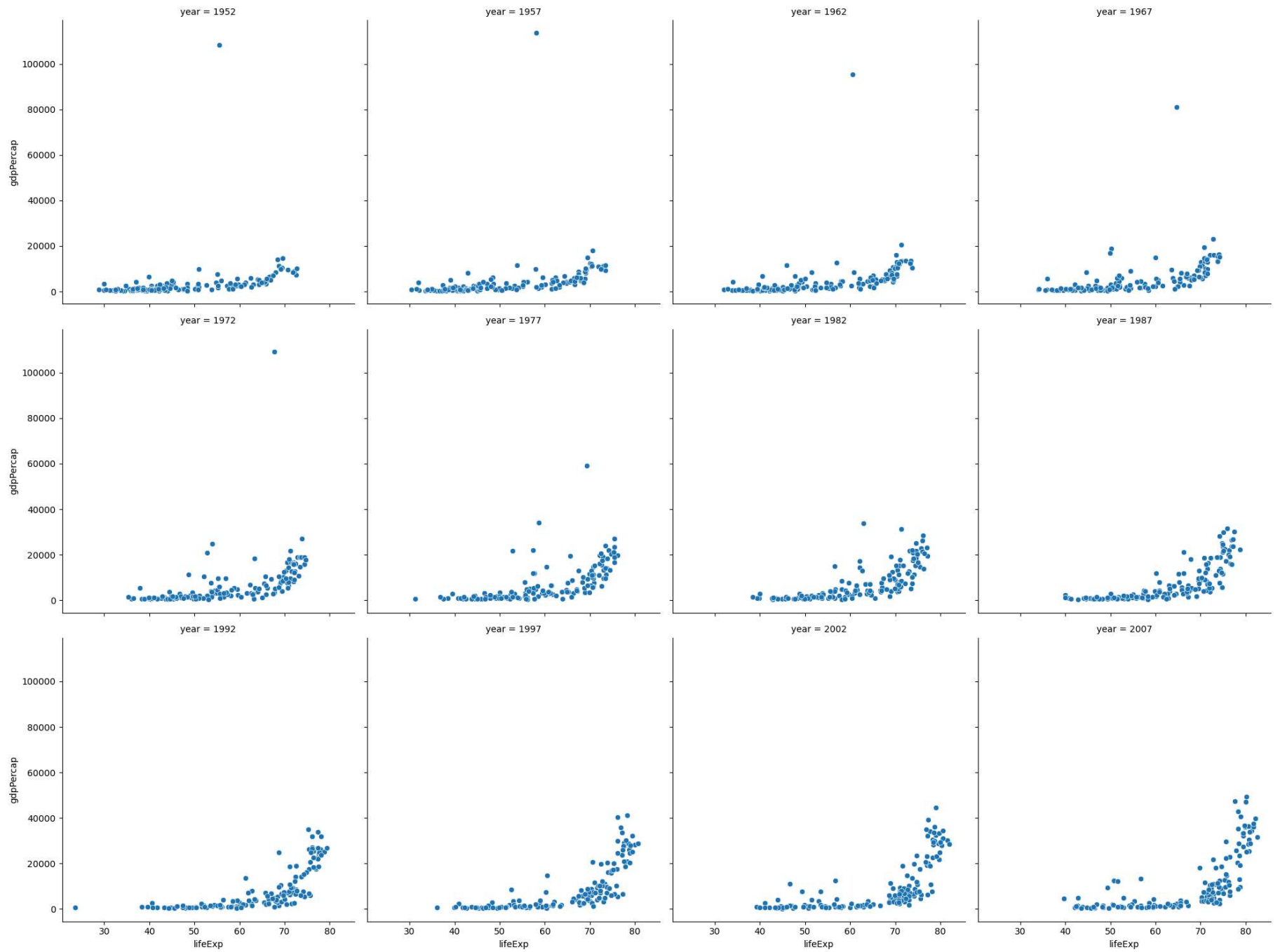
	country	continent	year	lifeExp	pop	gdpPercap	iso_alpha	iso_num
0	Afghanistan	Asia	1952	28.801	8425333	779.445314	AFG	4
1	Afghanistan	Asia	1957	30.332	9240934	820.853030	AFG	4
2	Afghanistan	Asia	1962	31.997	10267083	853.100710	AFG	4
3	Afghanistan	Asia	1967	34.020	11537966	836.197138	AFG	4
4	Afghanistan	Asia	1972	36.088	13079460	739.981106	AFG	4
...
1699	Zimbabwe	Africa	1987	62.351	9216418	706.157306	ZWE	716
1700	Zimbabwe	Africa	1992	60.377	10704340	693.420786	ZWE	716
1701	Zimbabwe	Africa	1997	46.809	11404948	792.449960	ZWE	716
1702	Zimbabwe	Africa	2002	39.989	11926563	672.038623	ZWE	716
1703	Zimbabwe	Africa	2007	43.487	12311143	469.709298	ZWE	716

1704 rows × 8 columns

column wrap

```
In [54]: sns.relplot(data=gap,x='lifeExp',y='gdpPercap',kind='scatter',col='year',col_wrap=4)
```

```
Out[54]: <seaborn.axisgrid.FacetGrid at 0x22a068a7510>
```

2. Distribution plots (histplot,kdeplot,rugplot)

```
In [55]: # used for univariate
# range of observation
# central tendencies

#figure level function: displot
#axis level function:histplot,kdeplot,rugplot
```

hist plot

```
In [56]: tips
```

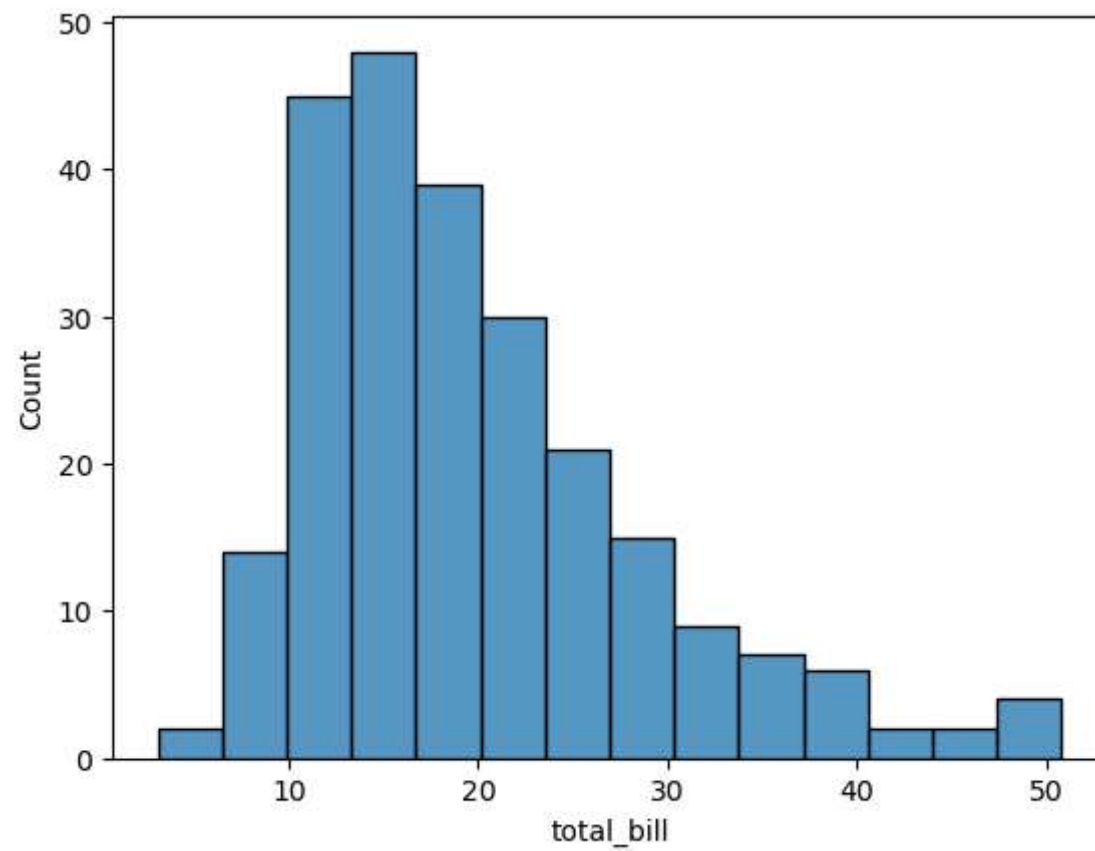
```
Out[56]:
```

	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
...
239	29.03	5.92	Male	No	Sat	Dinner	3
240	27.18	2.00	Female	Yes	Sat	Dinner	2
241	22.67	2.00	Male	Yes	Sat	Dinner	2
242	17.82	1.75	Male	No	Sat	Dinner	2
243	18.78	3.00	Female	No	Thur	Dinner	2

244 rows × 7 columns

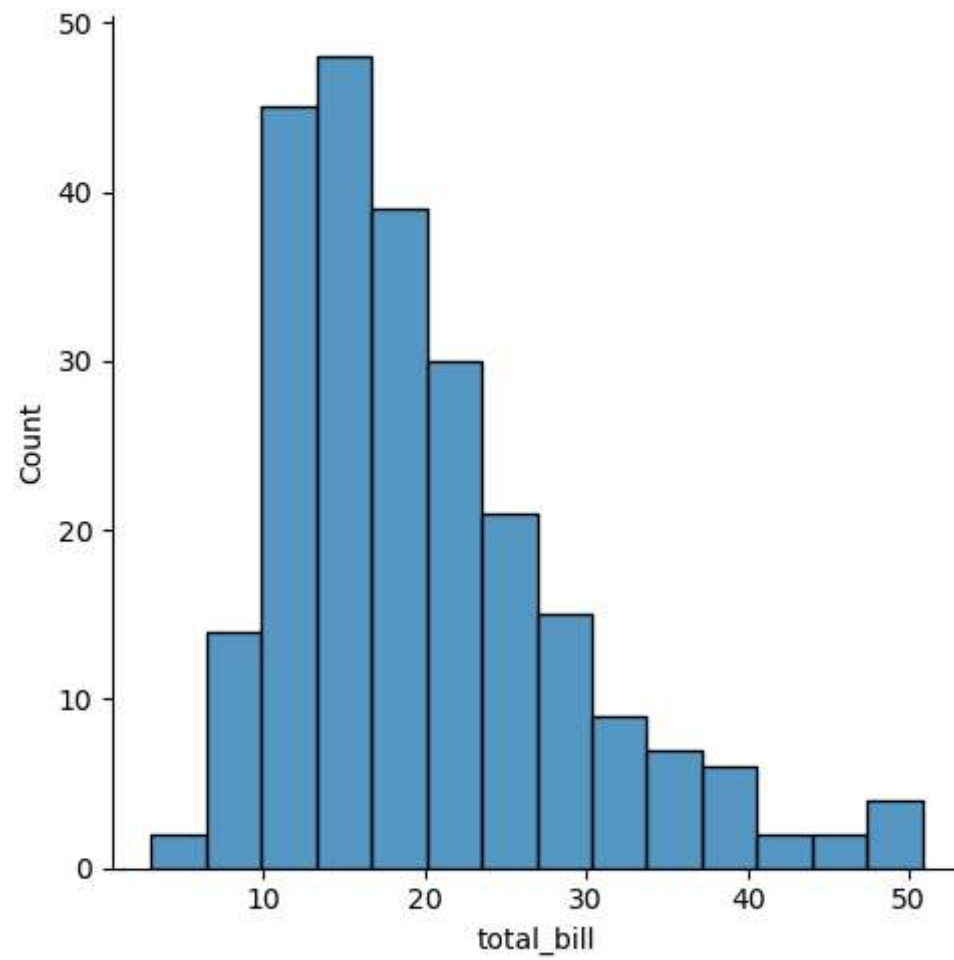
```
In [59]: sns.histplot(data=tips,x='total_bill')
```

```
Out[59]: <Axes: xlabel='total_bill', ylabel='Count'>
```



```
In [60]: sns.displot(data=tips,x='total_bill',kind='hist')
```

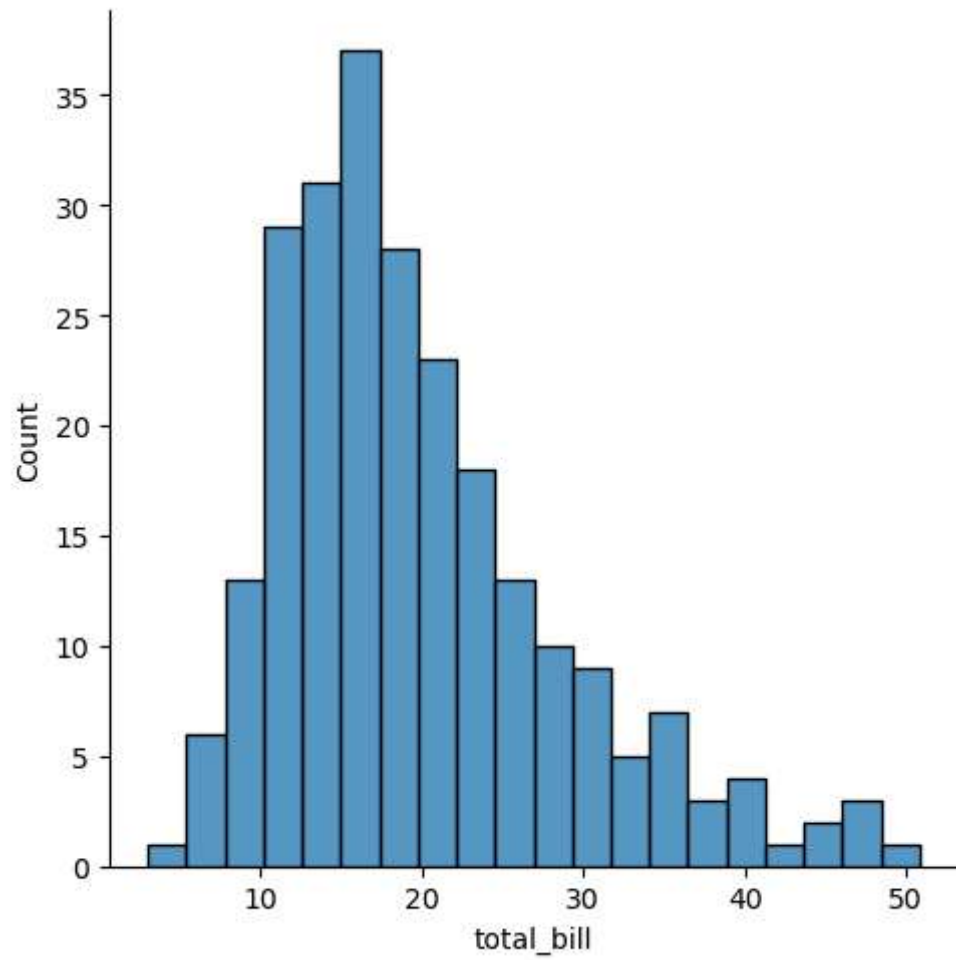
```
Out[60]: <seaborn.axisgrid.FacetGrid at 0x22a08c4a9d0>
```



```
In [61]: #bins parameter
```

```
In [62]: sns.displot(data=tips,x='total_bill',kind='hist',bins=20)
```

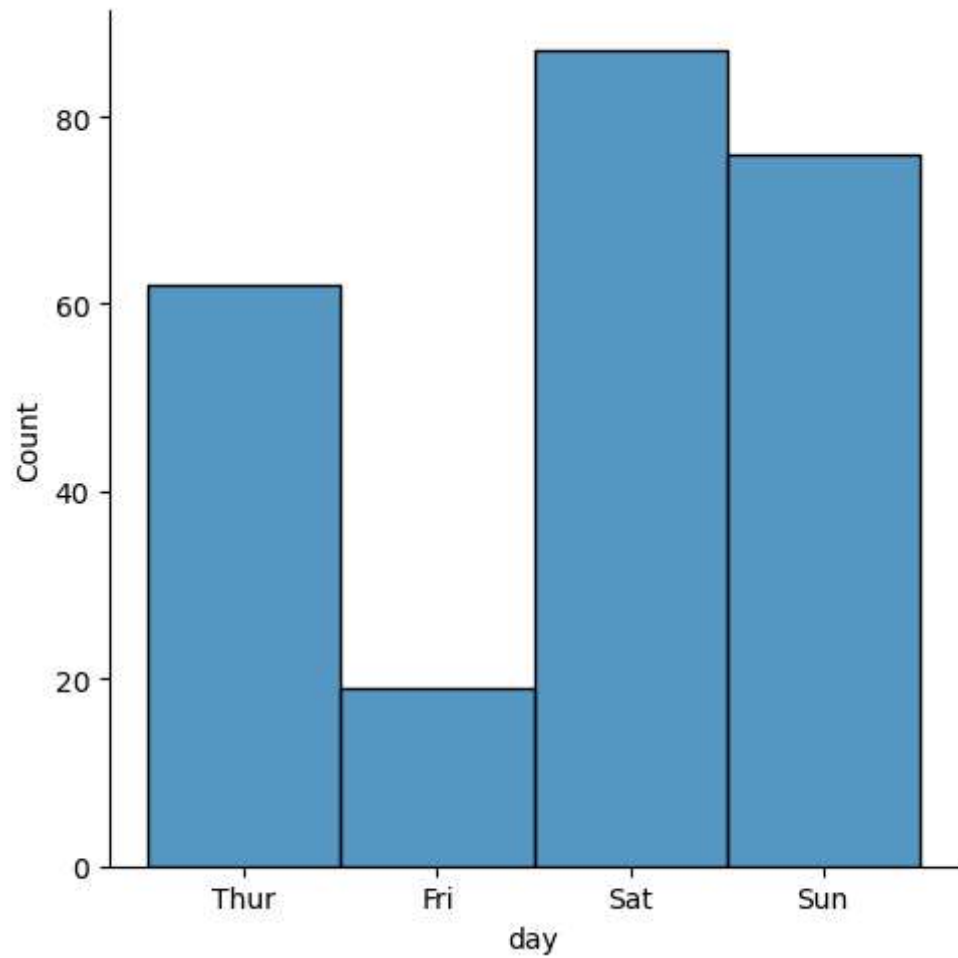
```
Out[62]: <seaborn.axisgrid.FacetGrid at 0x22a08c60610>
```



```
In [65]: #count plot
```

```
In [67]: sns.displot(data=tips,x='day',kind='hist')
```

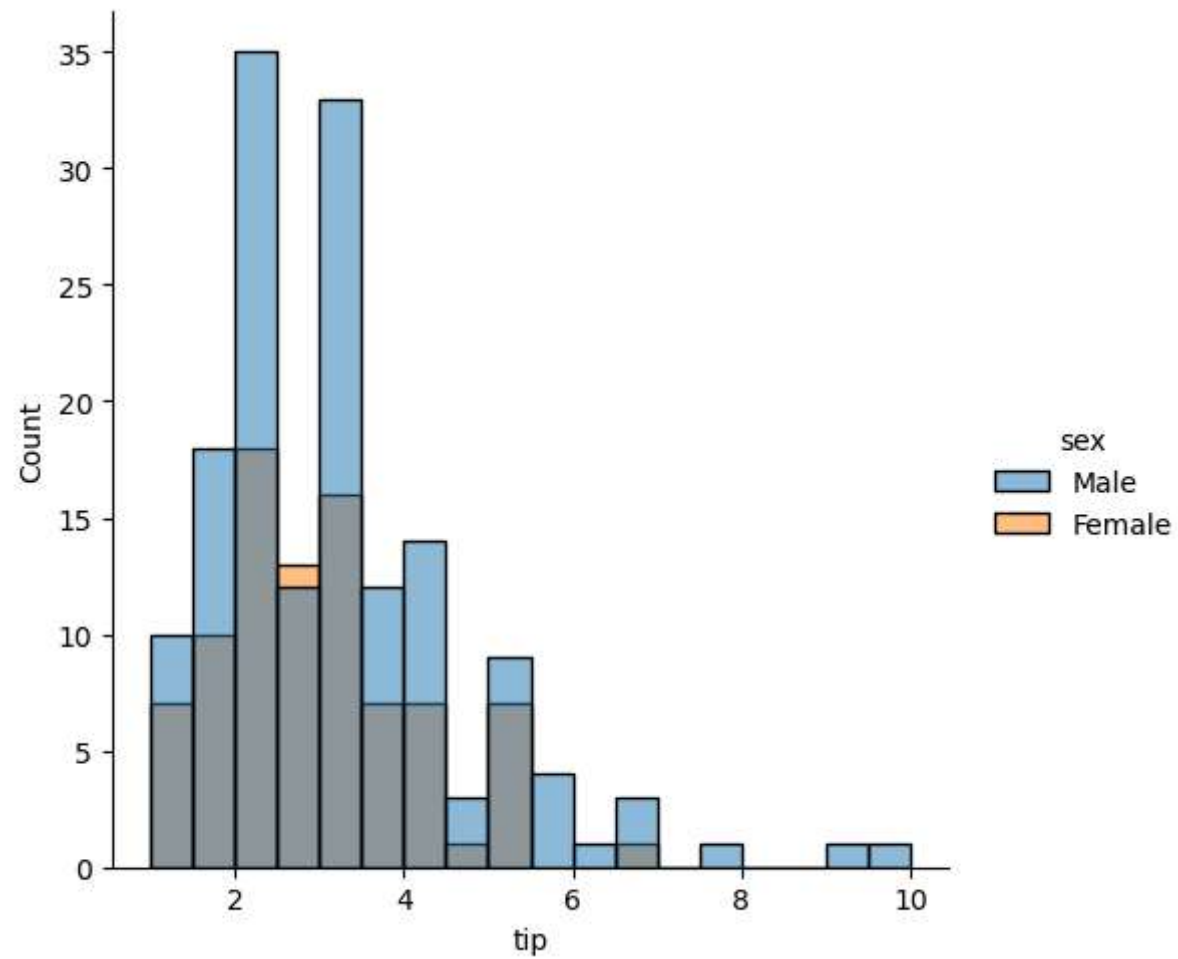
```
Out[67]: <seaborn.axisgrid.FacetGrid at 0x22a08ddb950>
```



```
In [63]: #hue parameter
```

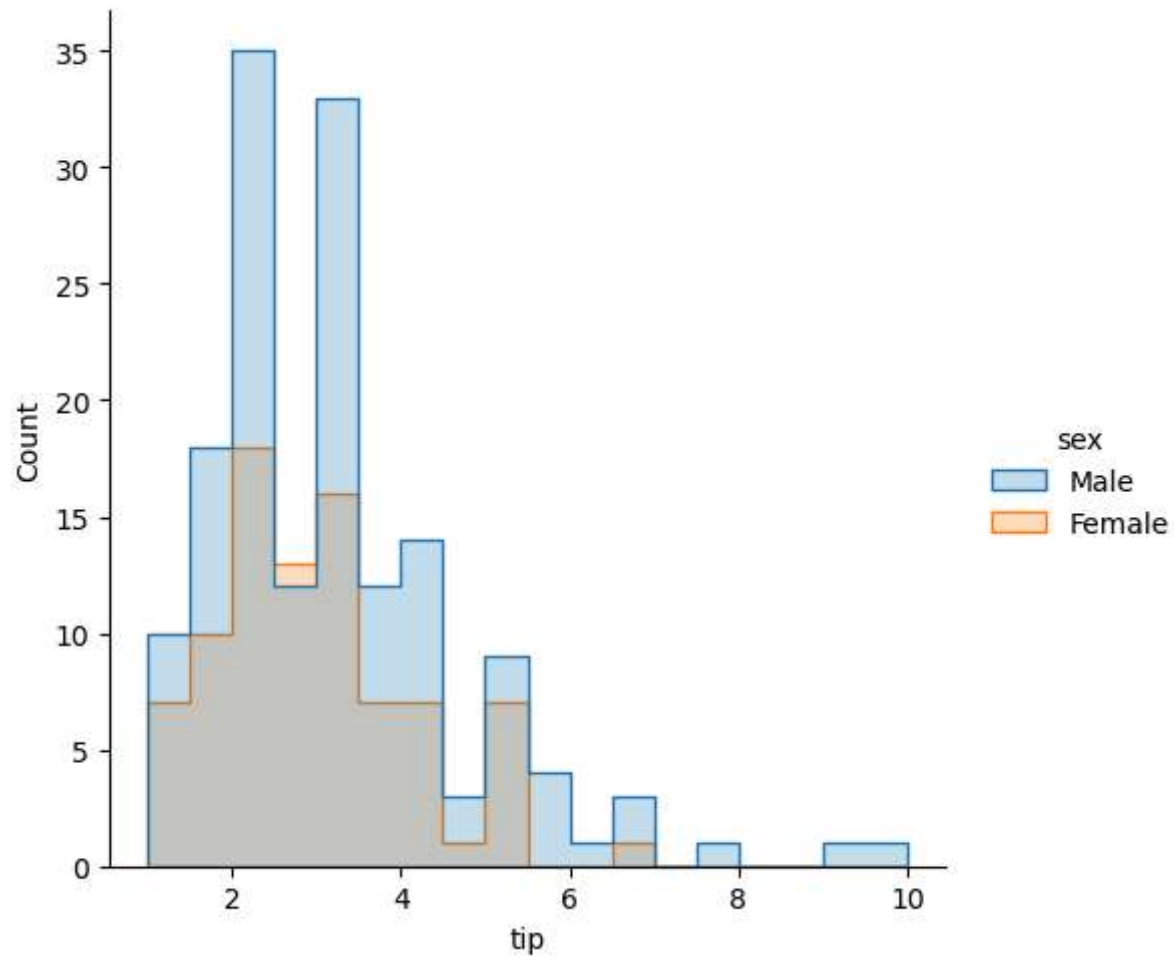
```
In [68]: sns.displot(data=tips,x='tip',kind='hist',hue='sex')
```

```
Out[68]: <seaborn.axisgrid.FacetGrid at 0x22a07fa0a10>
```



```
In [69]: sns.displot(data=tips,x='tip',kind='hist',hue='sex',element='step')
```

```
Out[69]: <seaborn.axisgrid.FacetGrid at 0x22a09ecb810>
```



```
In [70]: titanic=sns.load_dataset('titanic')
```



```
In [71]: titanic
```

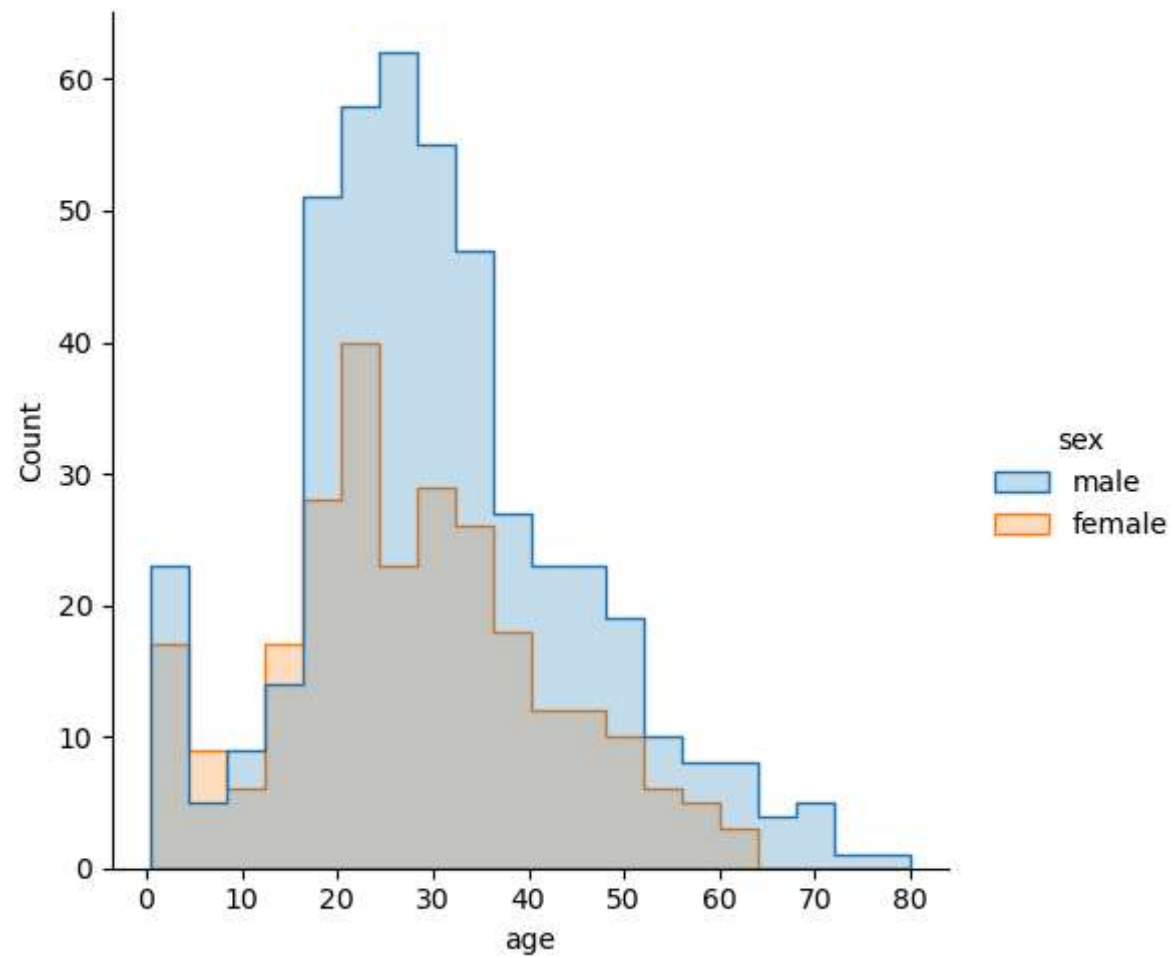
```
Out[71]:
```

	survived	pclass	sex	age	sibsp	parch	fare	embarked	class	who	adult_male	deck	embark_town	alive	alone
0	0	3	male	22.0	1	0	7.2500	S	Third	man	True	NaN	Southampton	no	False
1	1	1	female	38.0	1	0	71.2833	C	First	woman	False	C	Cherbourg	yes	False
2	1	3	female	26.0	0	0	7.9250	S	Third	woman	False	NaN	Southampton	yes	True
3	1	1	female	35.0	1	0	53.1000	S	First	woman	False	C	Southampton	yes	False
4	0	3	male	35.0	0	0	8.0500	S	Third	man	True	NaN	Southampton	no	True
...
886	0	2	male	27.0	0	0	13.0000	S	Second	man	True	NaN	Southampton	no	True
887	1	1	female	19.0	0	0	30.0000	S	First	woman	False	B	Southampton	yes	True
888	0	3	female	NaN	1	2	23.4500	S	Third	woman	False	NaN	Southampton	no	False
889	1	1	male	26.0	0	0	30.0000	C	First	man	True	C	Cherbourg	yes	True
890	0	3	male	32.0	0	0	7.7500	Q	Third	man	True	NaN	Queenstown	no	True

891 rows × 15 columns

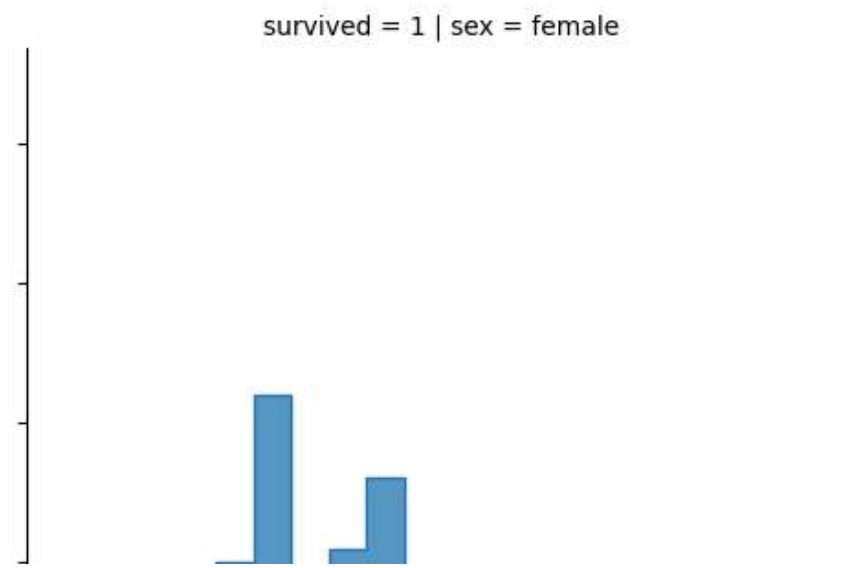
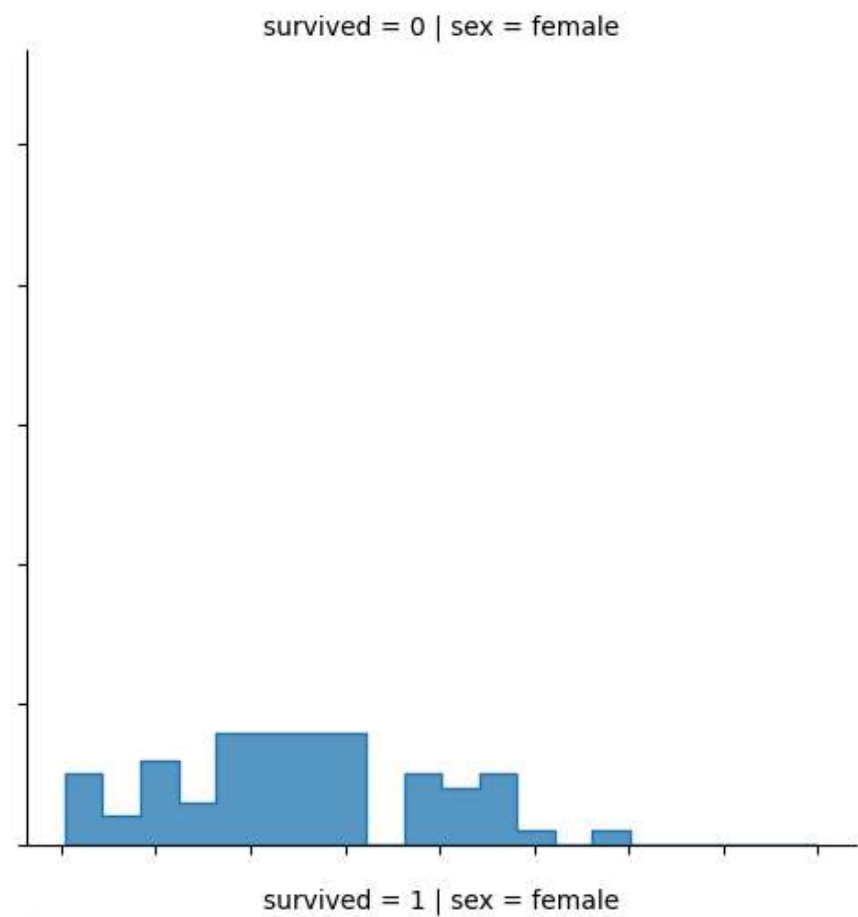
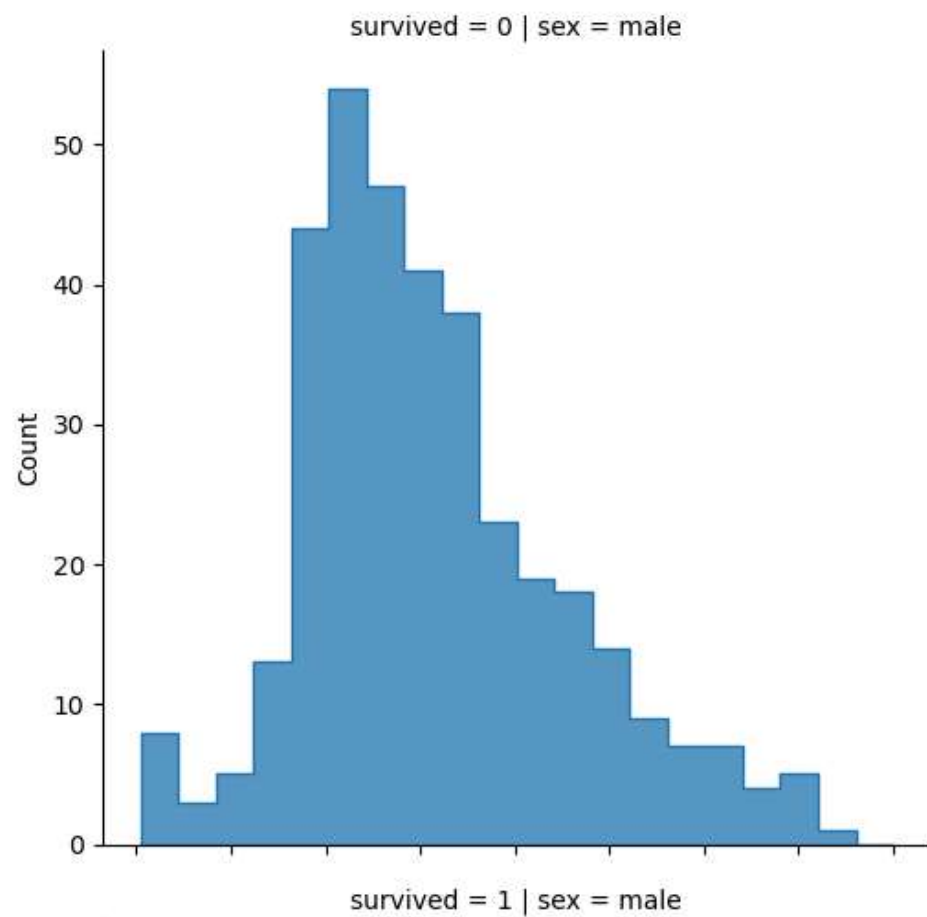
```
In [80]: sns.displot(data=titanic,x='age',element='step',hue='sex',kind='hist')
```

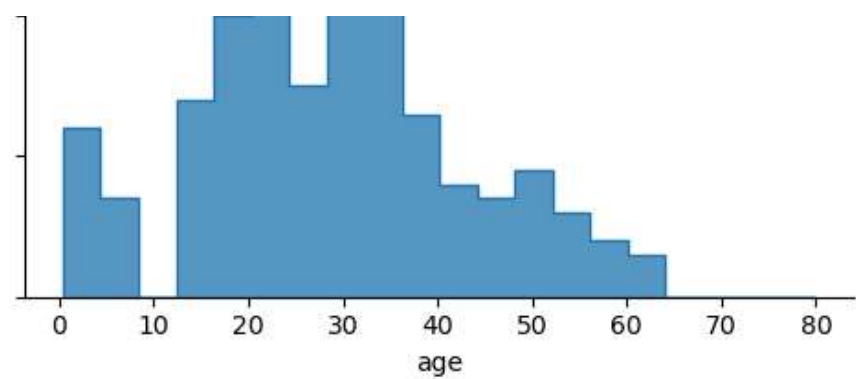
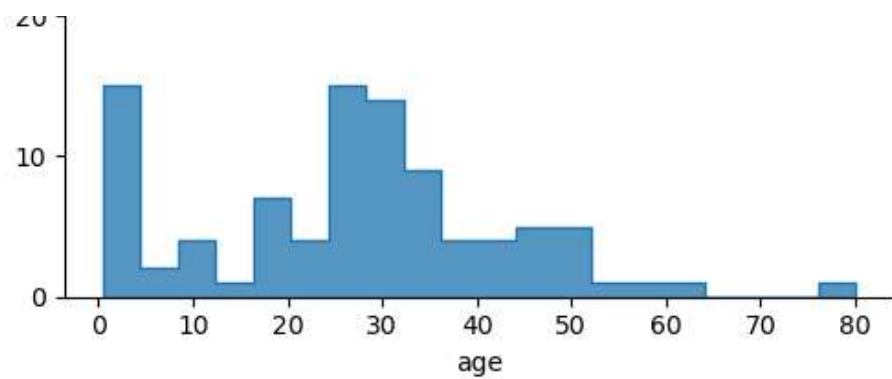
```
Out[80]: <seaborn.axisgrid.FacetGrid at 0x22a0c9108d0>
```



```
In [79]: sns.displot(data=titanic,x='age',element='step',col='sex',row='survived',kind='hist')
```

```
Out[79]: <seaborn.axisgrid.FacetGrid at 0x22a0bf83310>
```

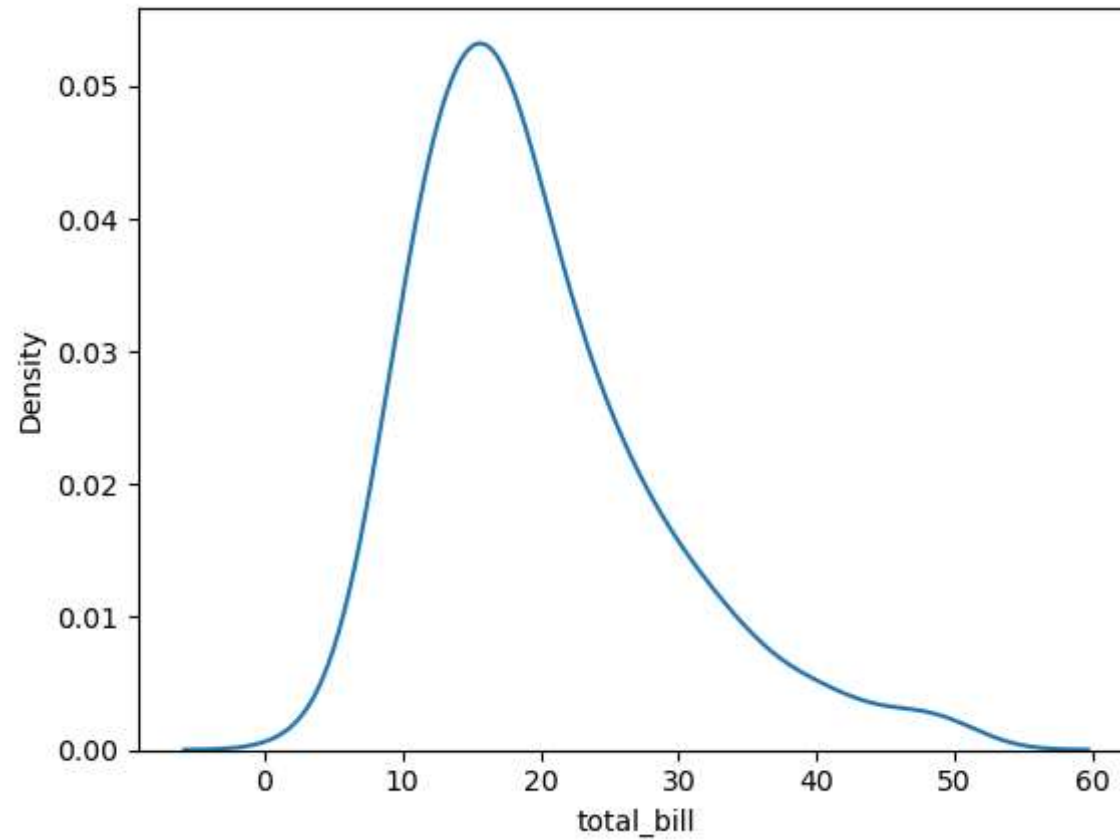





KDE plot

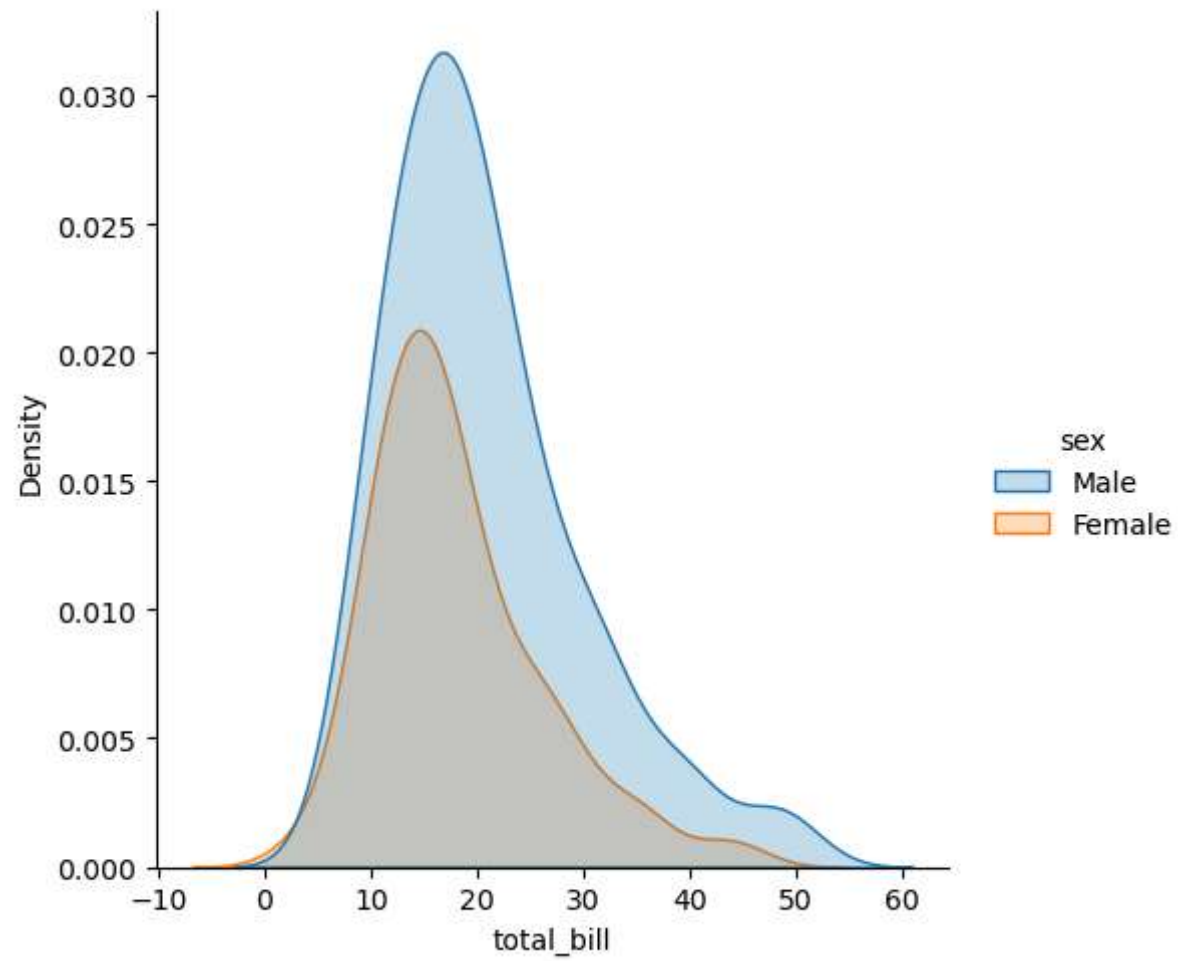
```
In [81]: sns.kdeplot(data=tips,x='total_bill')
```

```
Out[81]: <Axes: xlabel='total_bill', ylabel='Density'>
```



```
In [84]: sns.displot(data=tips,x='total_bill',kind='kde',hue='sex',fill=True)
```

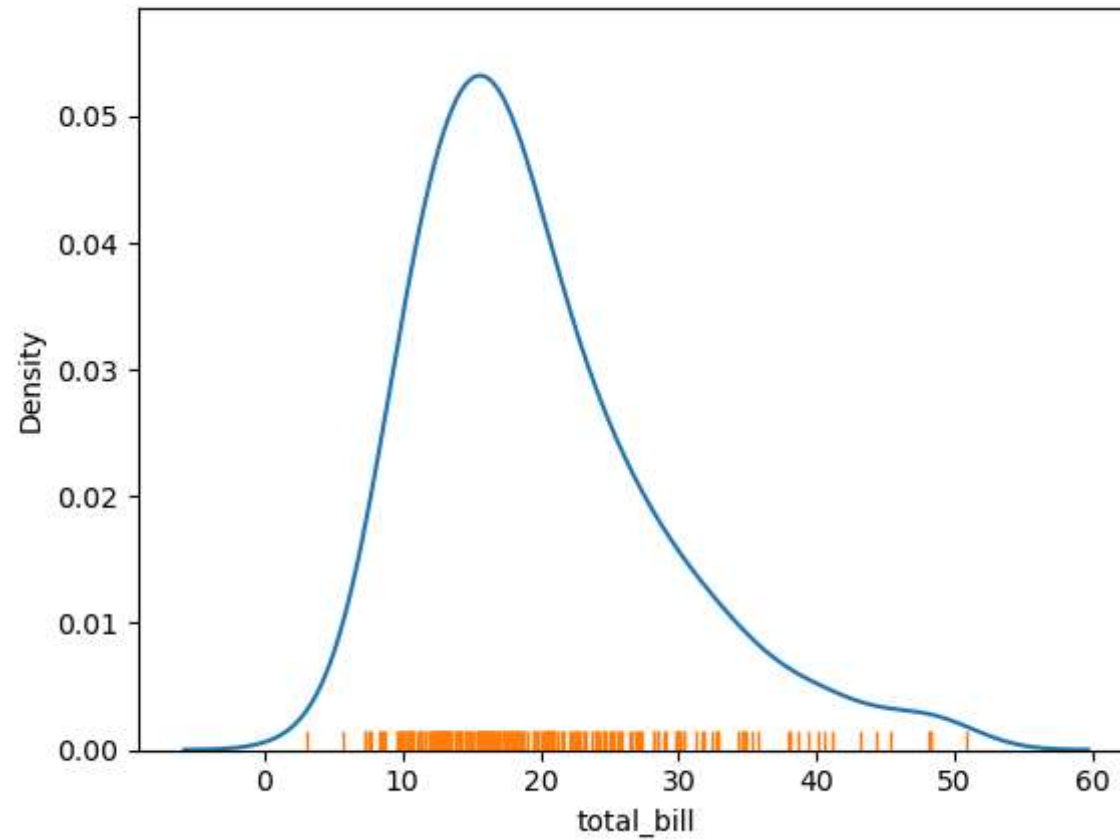
```
Out[84]: <seaborn.axisgrid.FacetGrid at 0x22a0cfa0110>
```



Rugplot

```
In [85]: sns.kdeplot(data=tips,x='total_bill')  
sns.rugplot(data=tips,x='total_bill')
```

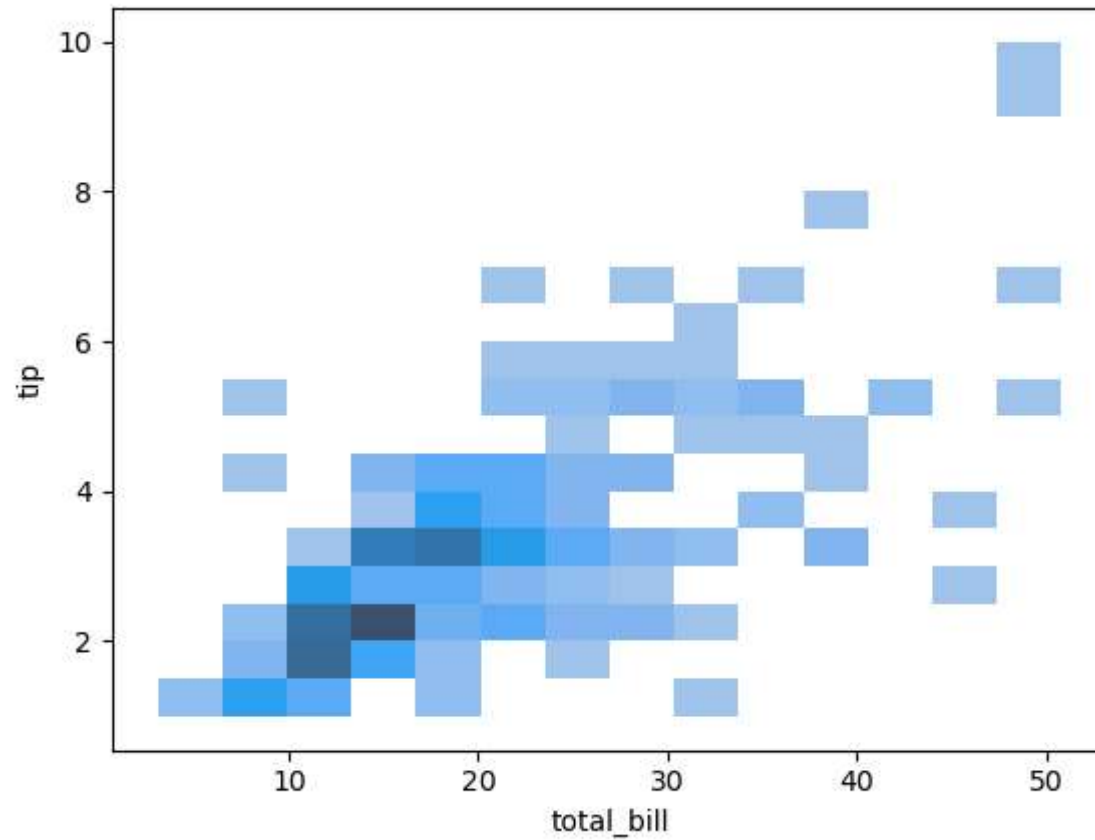
```
Out[85]: <Axes: xlabel='total_bill', ylabel='Density'>
```



bivariate histogram

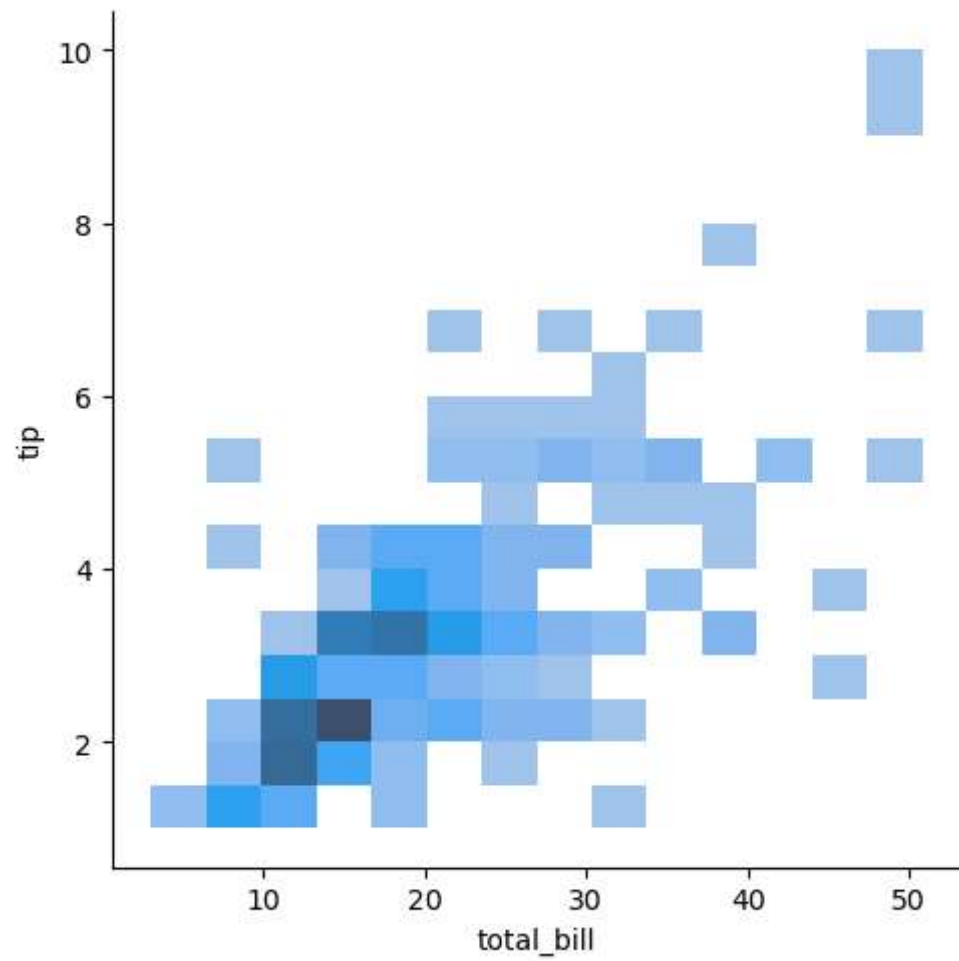
```
In [86]: sns.histplot(data=tips,x='total_bill',y='tip')
```

```
Out[86]: <Axes: xlabel='total_bill', ylabel='tip'>
```



```
In [87]: sns.displot(data=tips,x='total_bill',y='tip',kind='hist')
```

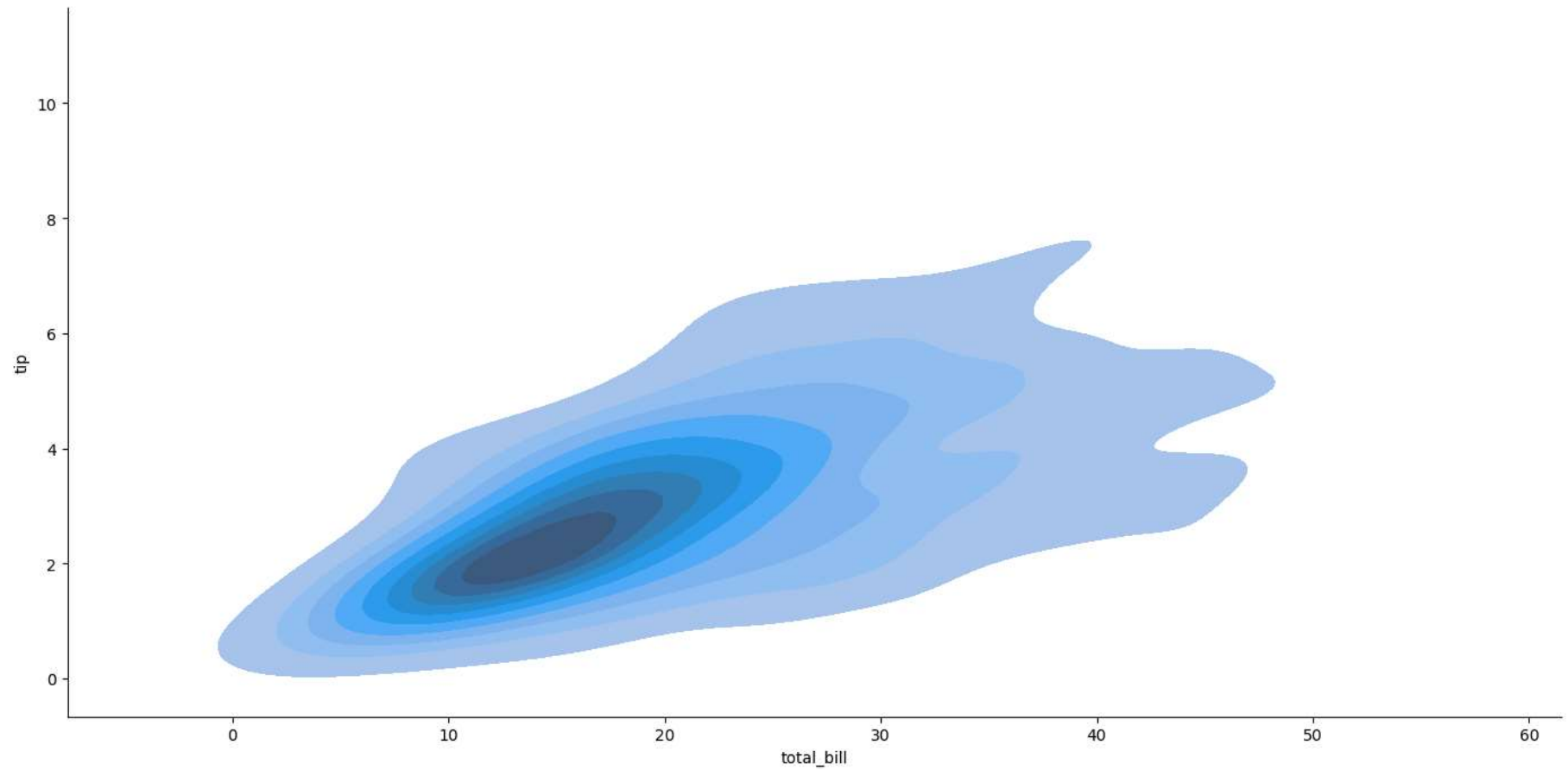
```
Out[87]: <seaborn.axisgrid.FacetGrid at 0x22a0d7dd210>
```



Bivariate kde

```
In [28]: sns.displot(data=tips,x='total_bill',y='tip',kind='kde',fill=True,height=7,aspect=2)    #important to increase size in j
```

```
Out[28]: <seaborn.axisgrid.FacetGrid at 0x222f5c0d850>
```



3. Matrix plot

3.1 Heatmap (axis level function)

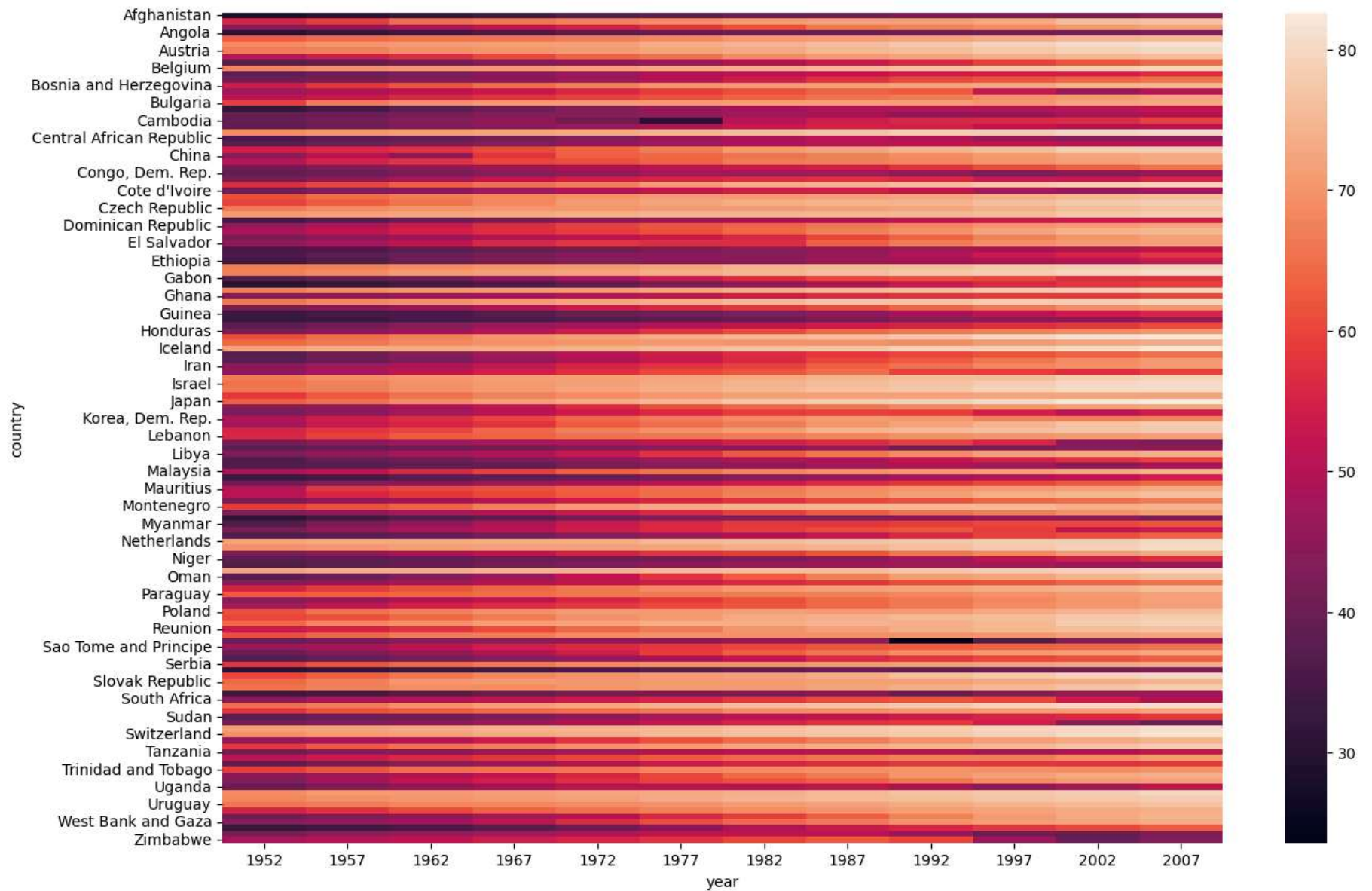
In [4]: *#rectangular data as color encoded matrix*

In [5]: *#what is wide and Long format ?*

In [13]: `tempdf3=gap.pivot(index='country',columns='year',values='lifeExp')`

```
In [14]: plt.figure(figsize=(15,10))
sns.heatmap(tempdf3)
```

```
Out[14]: <Axes: xlabel='year', ylabel='country'>
```



```
In [17]: tempdf4=gap[gap['continent']=='Europe'].pivot(index='country',columns='year',values='lifeExp')
```

In [16]: tempdf4

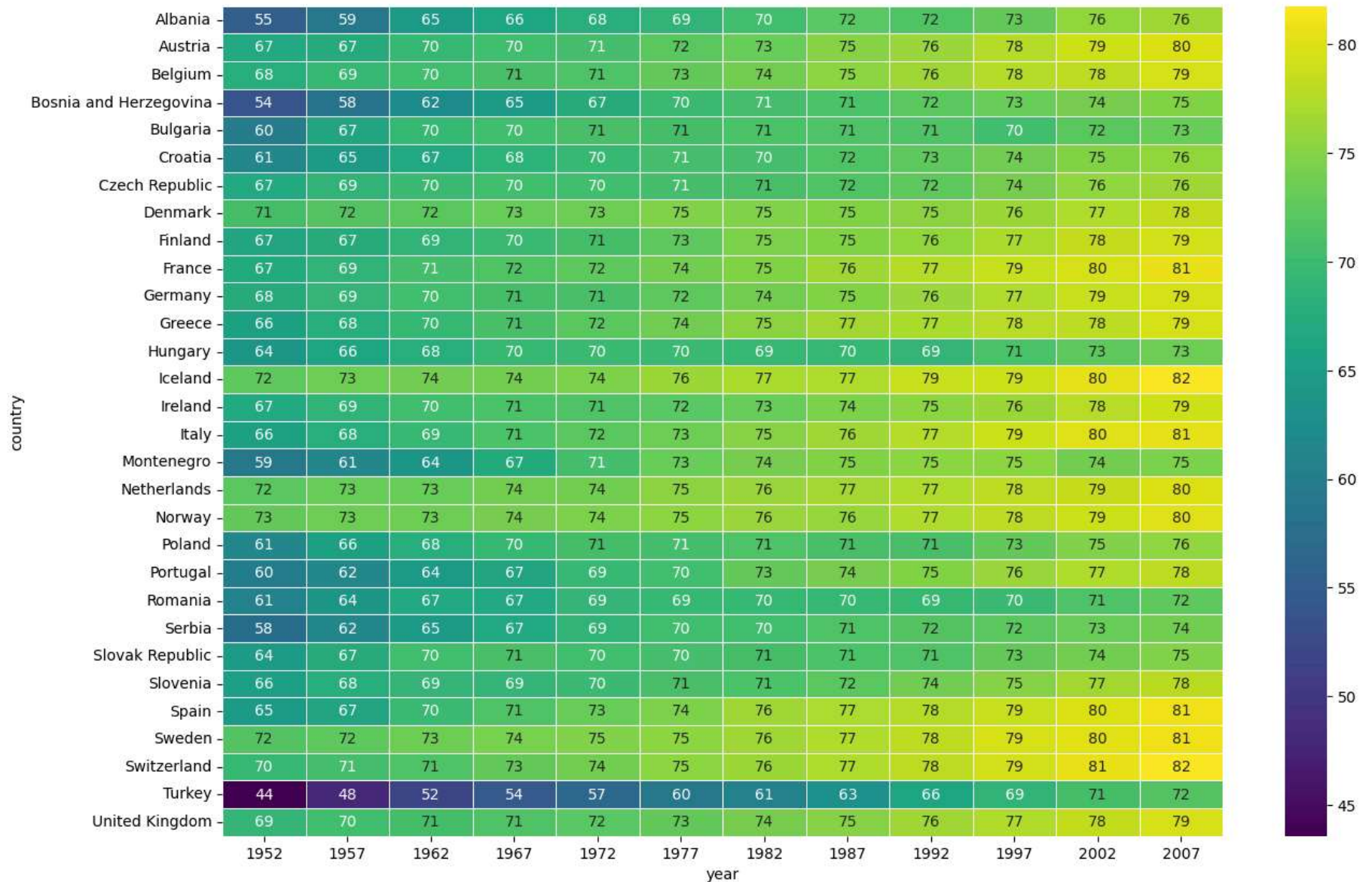
Out[16]:

	year	1952	1957	1962	1967	1972	1977	1982	1987	1992	1997	2002	2007
country													
Albania		55.230	59.280	64.820	66.220	67.690	68.930	70.420	72.000	71.581	72.950	75.651	76.423
Austria		66.800	67.480	69.540	70.140	70.630	72.170	73.180	74.940	76.040	77.510	78.980	79.829
Belgium		68.000	69.240	70.250	70.940	71.440	72.800	73.930	75.350	76.460	77.530	78.320	79.441
Bosnia and Herzegovina		53.820	58.450	61.930	64.790	67.450	69.860	70.690	71.140	72.178	73.244	74.090	74.852
Bulgaria		59.600	66.610	69.510	70.420	70.900	70.810	71.080	71.340	71.190	70.320	72.140	73.005
Croatia		61.210	64.770	67.130	68.500	69.610	70.640	70.460	71.520	72.527	73.680	74.876	75.748
Czech Republic		66.870	69.030	69.900	70.380	70.290	70.710	70.960	71.580	72.400	74.010	75.510	76.486
Denmark		70.780	71.810	72.350	72.960	73.470	74.690	74.630	74.800	75.330	76.110	77.180	78.332
Finland		66.550	67.490	68.750	69.830	70.870	72.520	74.550	74.830	75.700	77.130	78.370	79.313
France		67.410	68.930	70.510	71.550	72.380	73.830	74.890	76.340	77.460	78.640	79.590	80.657
Germany		67.500	69.100	70.300	70.800	71.000	72.500	73.800	74.847	76.070	77.340	78.670	79.406
Greece		65.860	67.860	69.510	71.000	72.340	73.680	75.240	76.670	77.030	77.869	78.256	79.483
Hungary		64.030	66.410	67.960	69.500	69.760	69.950	69.390	69.580	69.170	71.040	72.590	73.338
Iceland		72.490	73.470	73.680	73.730	74.460	76.110	76.990	77.230	78.770	78.950	80.500	81.757
Ireland		66.910	68.900	70.290	71.080	71.280	72.030	73.100	74.360	75.467	76.122	77.783	78.885
Italy		65.940	67.810	69.240	71.060	72.190	73.480	74.980	76.420	77.440	78.820	80.240	80.546
Montenegro		59.164	61.448	63.728	67.178	70.636	73.066	74.101	74.865	75.435	75.445	73.981	74.543
Netherlands		72.130	72.990	73.230	73.820	73.750	75.240	76.050	76.830	77.420	78.030	78.530	79.762
Norway		72.670	73.440	73.470	74.080	74.340	75.370	75.970	75.890	77.320	78.320	79.050	80.196
Poland		61.310	65.770	67.640	69.610	70.850	70.670	71.320	70.980	70.990	72.750	74.670	75.563
Portugal		59.820	61.510	64.390	66.600	69.260	70.410	72.770	74.060	74.860	75.970	77.290	78.098
Romania		61.050	64.100	66.800	66.800	69.210	69.460	69.660	69.530	69.360	69.720	71.322	72.476
Serbia		57.996	61.685	64.531	66.914	68.700	70.300	70.162	71.218	71.659	72.232	73.213	74.002
Slovak Republic		64.360	67.450	70.330	70.980	70.350	70.450	70.800	71.080	71.380	72.710	73.800	74.663
Slovenia		65.570	67.850	69.150	69.180	69.820	70.970	71.063	72.250	73.640	75.130	76.660	77.926

	year	1952	1957	1962	1967	1972	1977	1982	1987	1992	1997	2002	2007
	country												
	Spain	64.940	66.660	69.690	71.440	73.060	74.390	76.300	76.900	77.570	78.770	79.780	80.941
	Sweden	71.860	72.490	73.370	74.160	74.720	75.440	76.420	77.190	78.160	79.390	80.040	80.884
	Switzerland	69.620	70.560	71.320	72.770	73.780	75.390	76.210	77.410	78.030	79.370	80.620	81.701
	Turkey	43.585	48.079	52.098	54.336	57.005	59.507	61.036	63.108	66.146	68.835	70.845	71.777
	United Kingdom	69.180	70.420	70.760	71.360	72.010	72.760	74.040	75.007	76.420	77.218	78.471	79.425

```
In [21]: plt.figure(figsize=(15,10))
sns.heatmap(tempdf4,annot=True,linewidth=0.5,cmap='viridis')
```

```
Out[21]: <Axes: xlabel='year', ylabel='country'>
```



cluster map

```
In [23]: iris=px.data.iris()
```

```
In [24]: iris
```

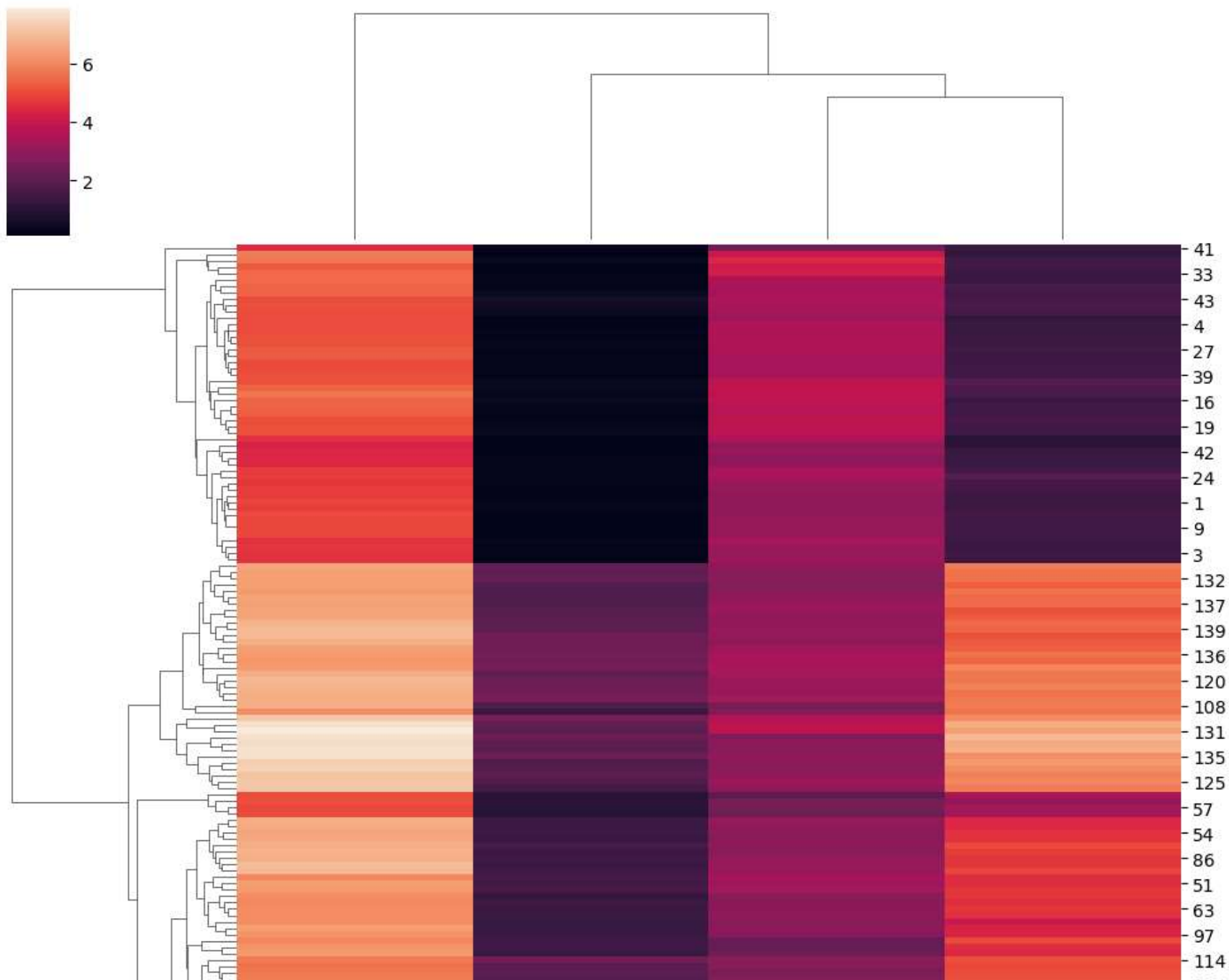
```
Out[24]:
```

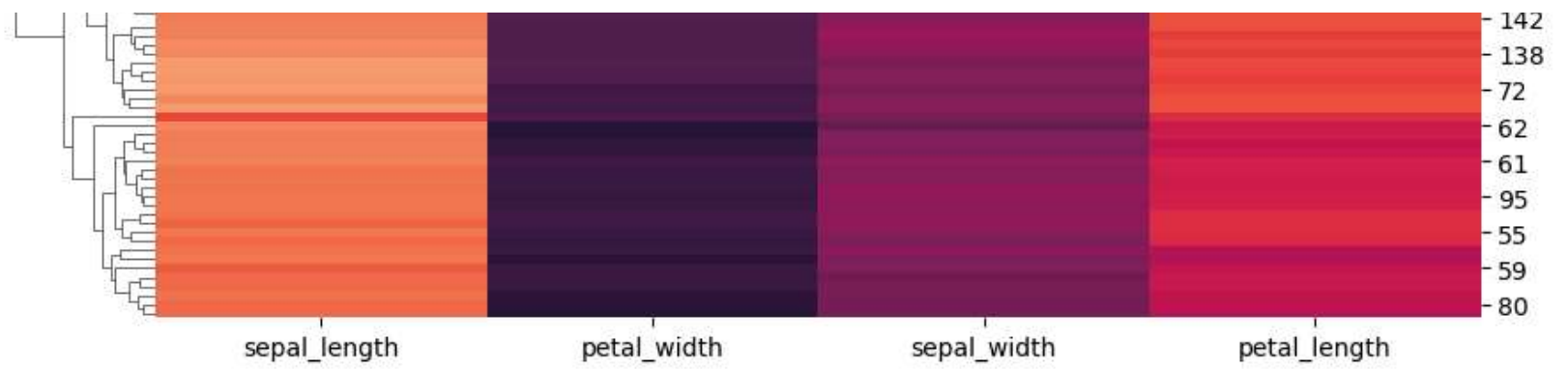
	sepal_length	sepal_width	petal_length	petal_width	species	species_id
0	5.1	3.5	1.4	0.2	setosa	1
1	4.9	3.0	1.4	0.2	setosa	1
2	4.7	3.2	1.3	0.2	setosa	1
3	4.6	3.1	1.5	0.2	setosa	1
4	5.0	3.6	1.4	0.2	setosa	1
...
145	6.7	3.0	5.2	2.3	virginica	3
146	6.3	2.5	5.0	1.9	virginica	3
147	6.5	3.0	5.2	2.0	virginica	3
148	6.2	3.4	5.4	2.3	virginica	3
149	5.9	3.0	5.1	1.8	virginica	3

150 rows × 6 columns

```
In [25]: sns.clustermap(iris.iloc[:,[0,1,2,3]])
```

```
Out[25]: <seaborn.matrix.ClusterGrid at 0x222f5c24110>
```



In []: