Bar Chart

Bar Chart is a diagram in which numerical values are represented by the lengths of rectangles of equal width.

http://matplotlib.org/api/pyplot_api.html#matplotlib.pyplot.bar
(http://matplotlib.org/api/pyplot_api.html#matplotlib.pyplot.bar)

Horizontal bar chart: https://matplotlib.org/gallery/lines_bars_and_markers/barh.html
(https://matplotlib.org/gallery/lines_bars_and_markers/barh.html)

In [1]:

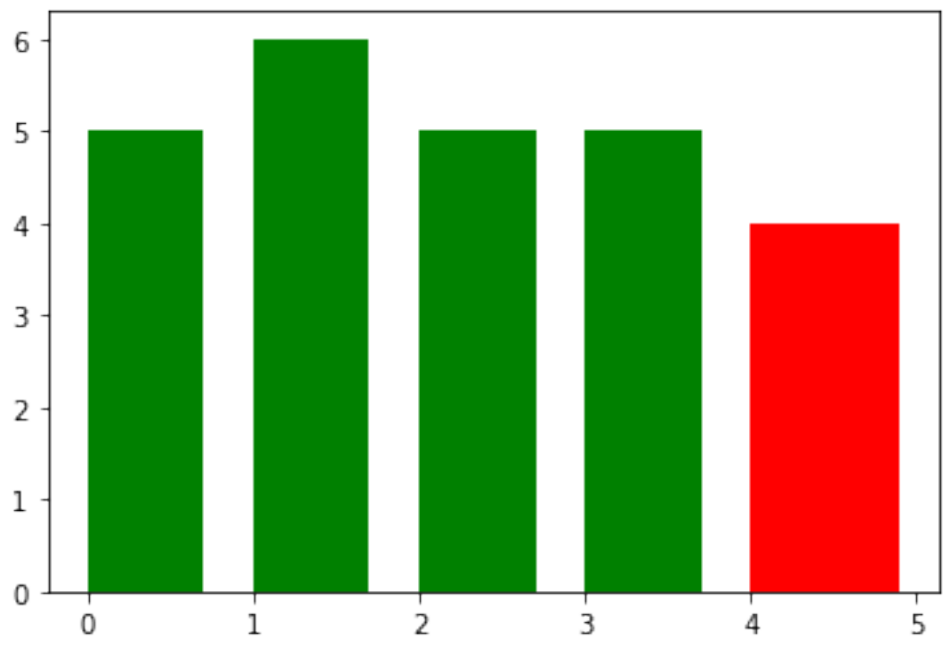
```
import matplotlib.pyplot as plt
import numpy as np
import pandas as pd
```

In [6]:

```
values = [5, 6, 5, 5, 4]
widths = [0.7, 0.7, 0.7, 0.7, 0.9]
colors = ['g', 'g', 'g', 'g', 'r']
left = range(0, 5) # the x coordinates of the left sides of the bars

# align is either 'center' or 'edge'
#, hatch="o"
plt.bar(left, values, width=widths, color=colors, align='edge')

plt.show()
```

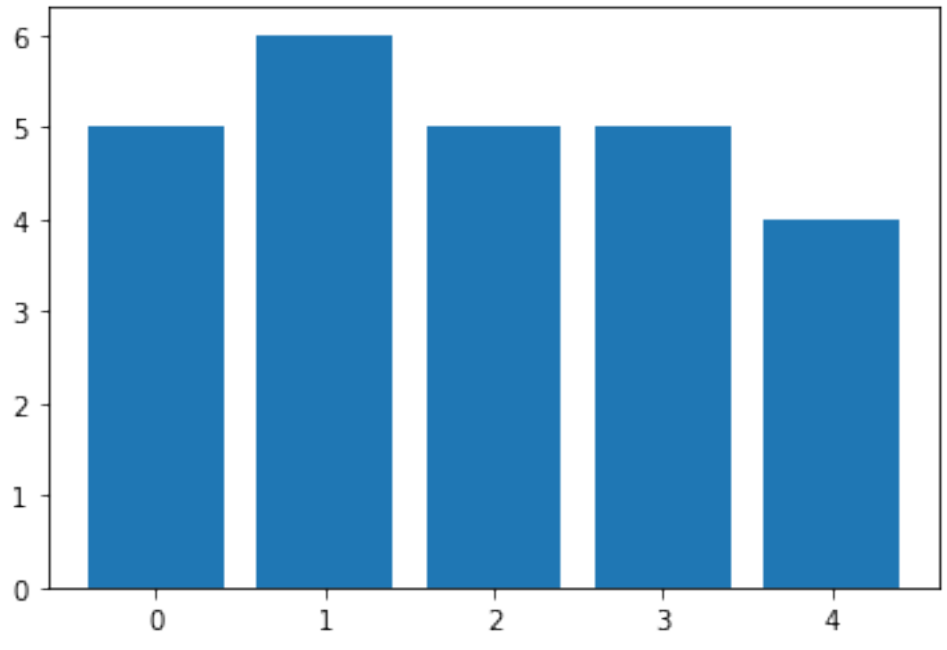


In [9]:

```
plt.bar(left, values)
```

Out[9]:

```
<BarContainer object of 5 artists>
```



Bar Plots from a DataFrame

In [10]:

```
np.random.seed(12348)
```

In [11]:

```
my_array = np.random.rand(6, 4)
my_array
```

Out[11]:

```
array([[0.37067042, 0.60279178, 0.22915897, 0.48674433],
       [0.42008234, 0.57165324, 0.0490237 , 0.88059193],
       [0.8145682 , 0.2771602 , 0.88031588, 0.43132587],
       [0.37401997, 0.89942028, 0.46030363, 0.10084331],
       [0.43326982, 0.125107  , 0.49467511, 0.96182537],
       [0.60164789, 0.4785759 , 0.20569042, 0.56054658]])
```

In [12]:

```
df = pd.DataFrame(my_array,
                  index=['one', 'two', 'three', 'four', 'five', 'six'],
                  columns=pd.Index(['A', 'B', 'C', 'D'], name='Genus')) #name refers
df
```

Out[12]:

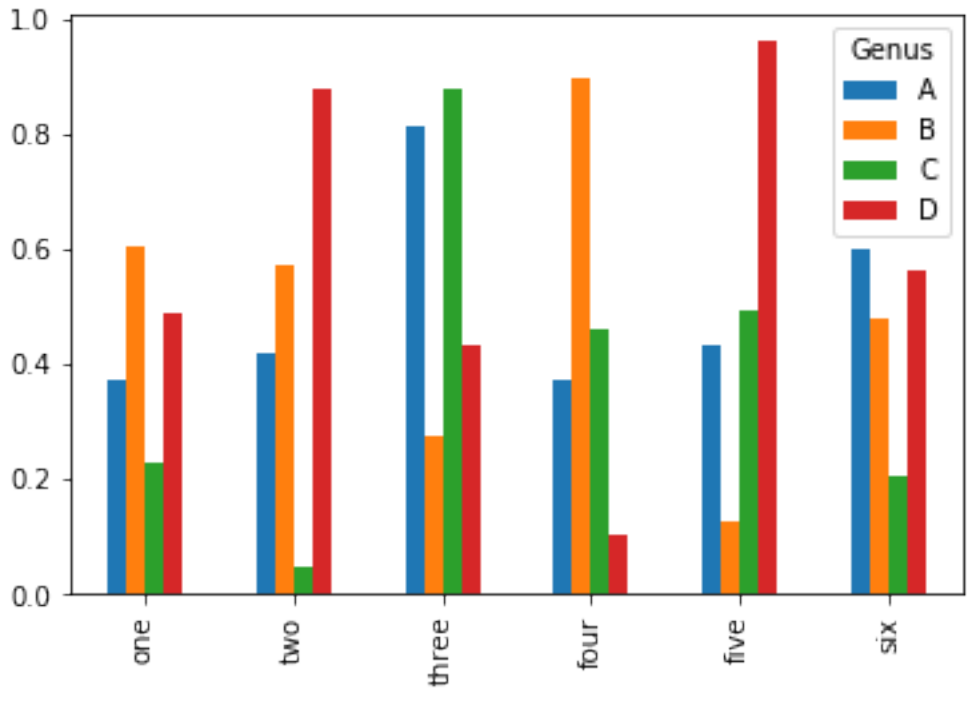
Genus	A	B	C	D
one	0.370670	0.602792	0.229159	0.486744
two	0.420082	0.571653	0.049024	0.880592
three	0.814568	0.277160	0.880316	0.431326
four	0.374020	0.899420	0.460304	0.100843
five	0.433270	0.125107	0.494675	0.961825
six	0.601648	0.478576	0.205690	0.560547

In [13]:

```
#df.plot.<tab> - what else can you plot?
df.plot.bar()
```

Out[13]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x11c6b5b38>
```

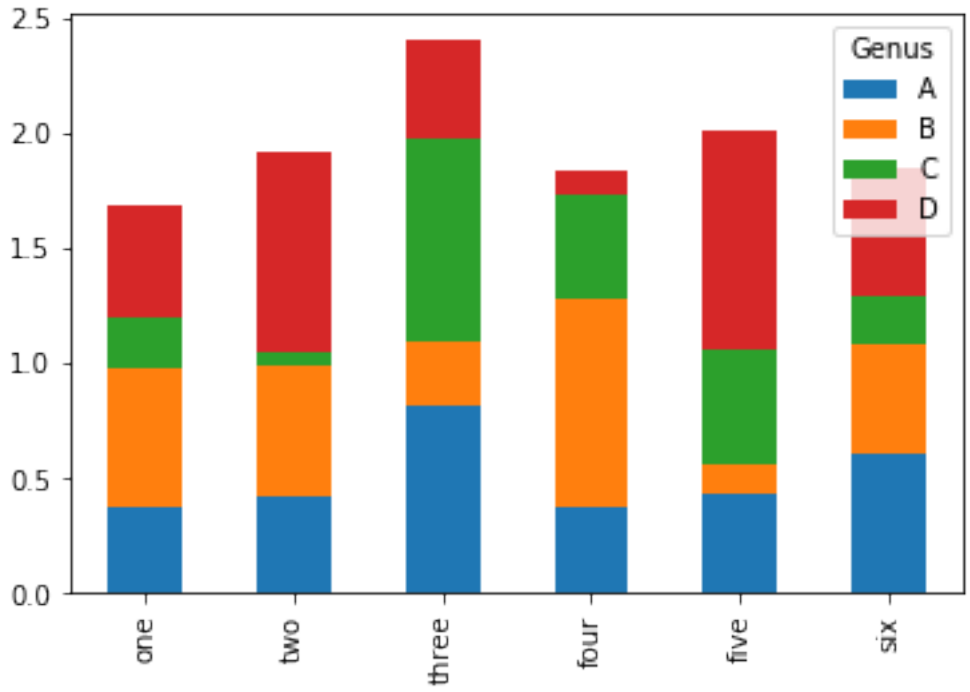


In [16]:

```
df.plot.bar(stacked=True)
```

Out[16]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x11ca01c88>
```

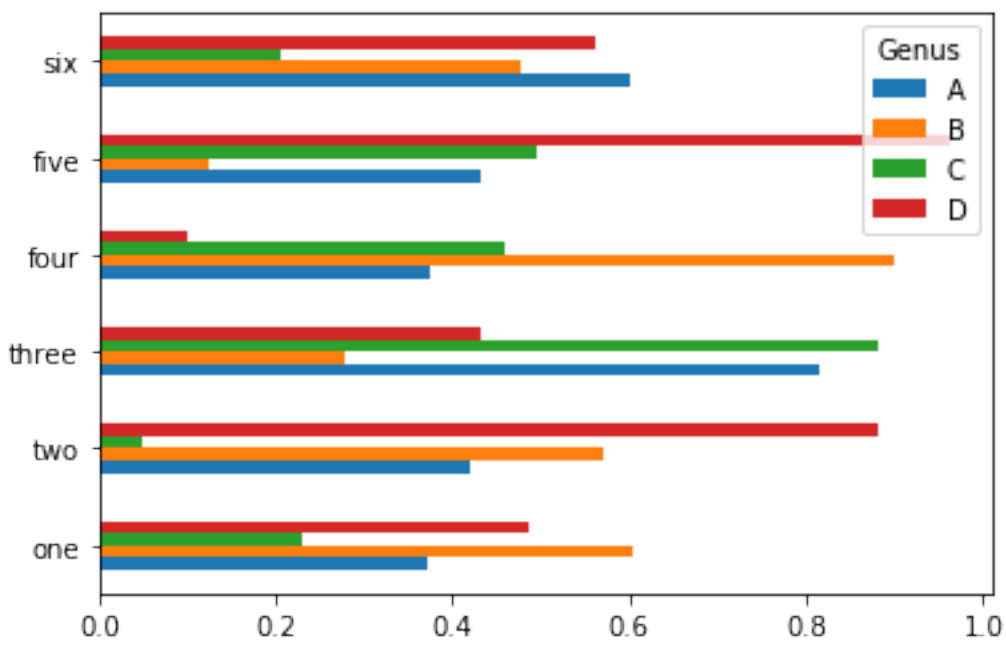


In [17]:

```
df.plot.barh()
```

Out[17]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x11ce53c50>
```

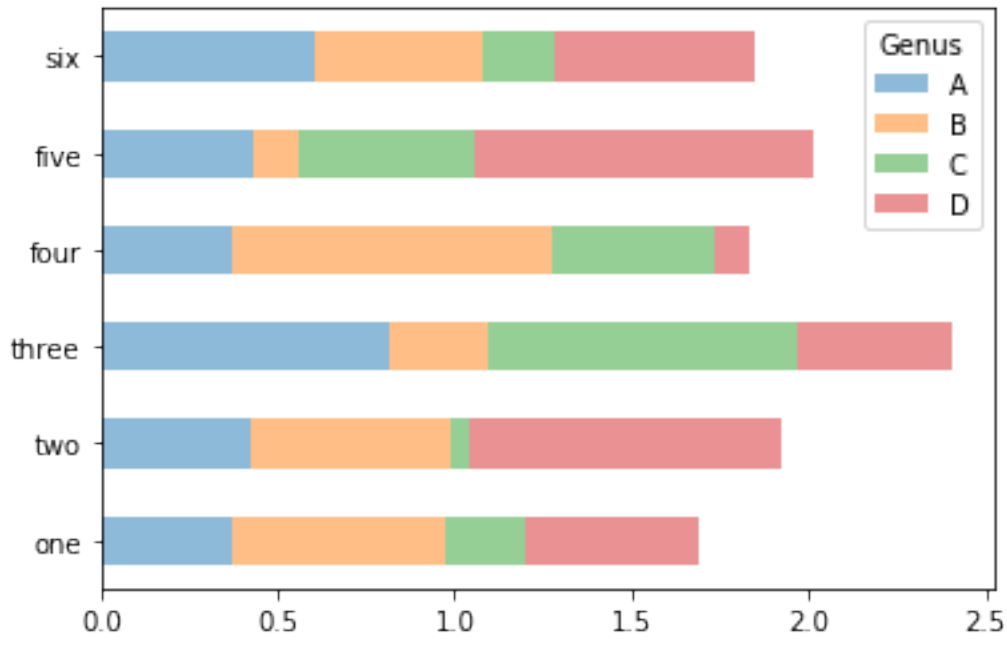


In [18]:

```
df.plot.barh(stacked=True, alpha=0.5)
```

Out[18]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x11ce2a2b0>
```



Tipping DataSet

In [19]:

```
tips = pd.read_csv('data/tips.csv')
tips.head(5)
```

Out[19]:

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

Crosstab (also known as contingency table or cross tabulation) is a table showing *frequency distribution* of one variable in rows and another on columns.

In [20]:

```
party_counts = pd.crosstab(tips['day'], tips['size'])
party_counts
```

Out[20]:

size	1	2	3	4	5	6
day						
Fri	1	16	1	1	0	0
Sat	2	53	18	13	1	0
Sun	0	39	15	18	3	1
Thur	1	48	4	5	1	3

In [21]:

```
party_counts = party_counts.reindex(index=['Thur', 'Fri', 'Sat', 'Sun'])
party_counts
```

Out[21]:

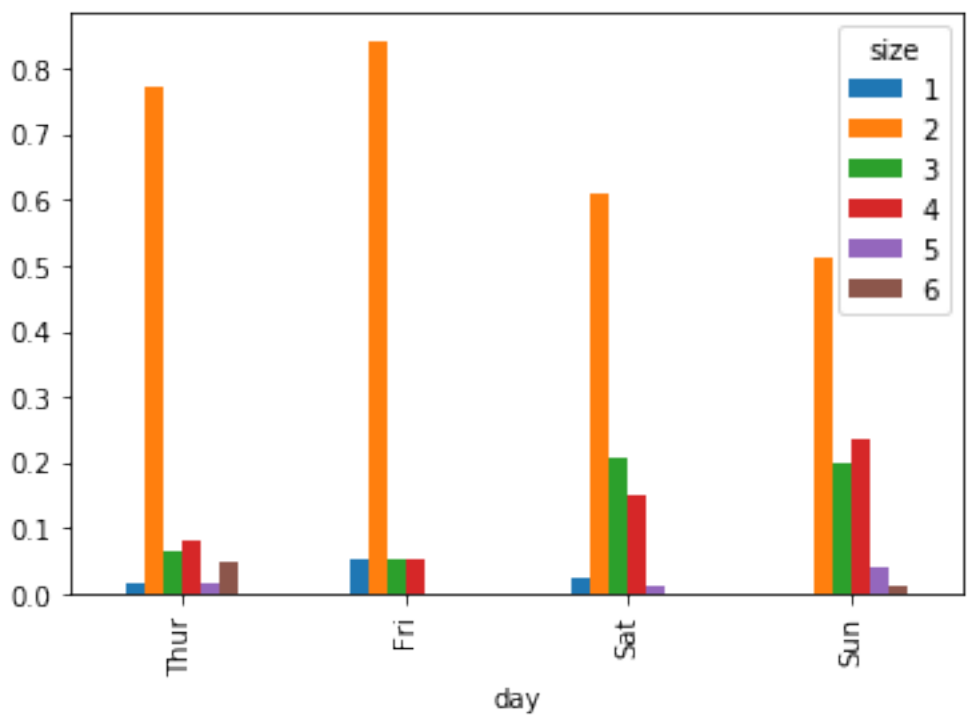
size	1	2	3	4	5	6
day						
Thur	1	48	4	5	1	3
Fri	1	16	1	1	0	0
Sat	2	53	18	13	1	0
Sun	0	39	15	18	3	1

In [22]:

```
party_pcts = party_counts.div(party_counts.sum(1), axis=0)
party_pcts.plot.bar()
```

Out[22]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x11ccadba8>
```



Party size appear to increase on the weekend in this dataset.

Bar Chart

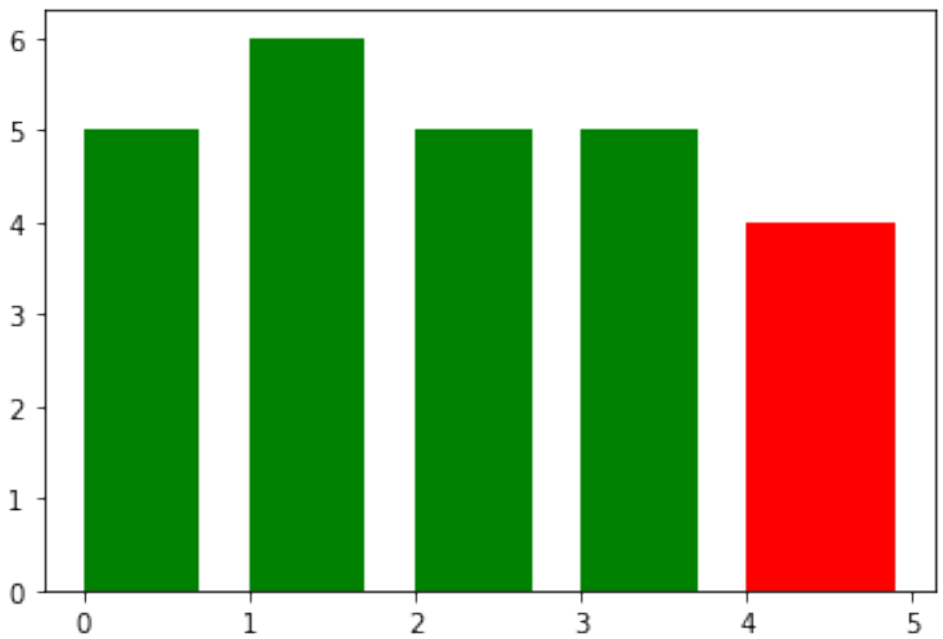
Bar Chart is a diagram in which numerical values are represented by the lengths of rectangles of equal width.

http://matplotlib.org/api/pyplot_api.html#matplotlib.pyplot.bar (http://matplotlib.org/api/pyplot_api.html#matplotlib.pyplot.bar)

Horizontal bar chart: https://matplotlib.org/gallery/lines_bars_and_markers/barh.html (https://matplotlib.org/gallery/lines_bars_and_markers/barh.html)

In [1]:

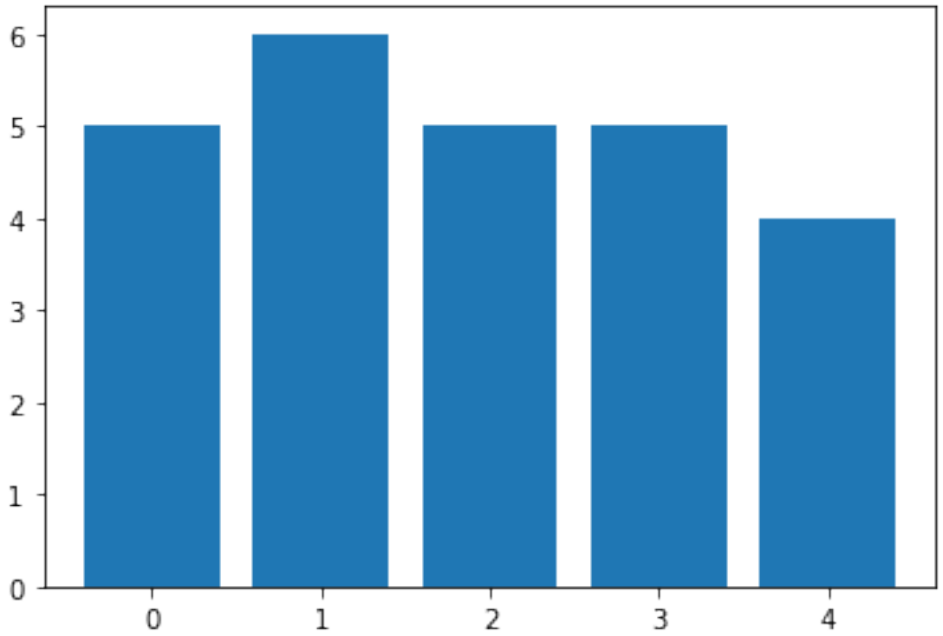
In [6]:



In [9]:

Out[9]:

```
<BarContainer object of 5 artists>
```



Bar Plots from a DataFrame

In [10]:

In [11]:

Out[11]:

```
array([[0.37067042, 0.60279178, 0.22915897, 0.48674433],
       [0.42008234, 0.57165324, 0.0490237 , 0.88059193],
       [0.8145682 , 0.2771602 , 0.88031588, 0.43132587],
       [0.37401997, 0.89942028, 0.46030363, 0.10084331],
       [0.43326982, 0.125107  , 0.49467511, 0.96182537],
       [0.60164789, 0.4785759 , 0.20569042, 0.56054658]])
```

In [12]:

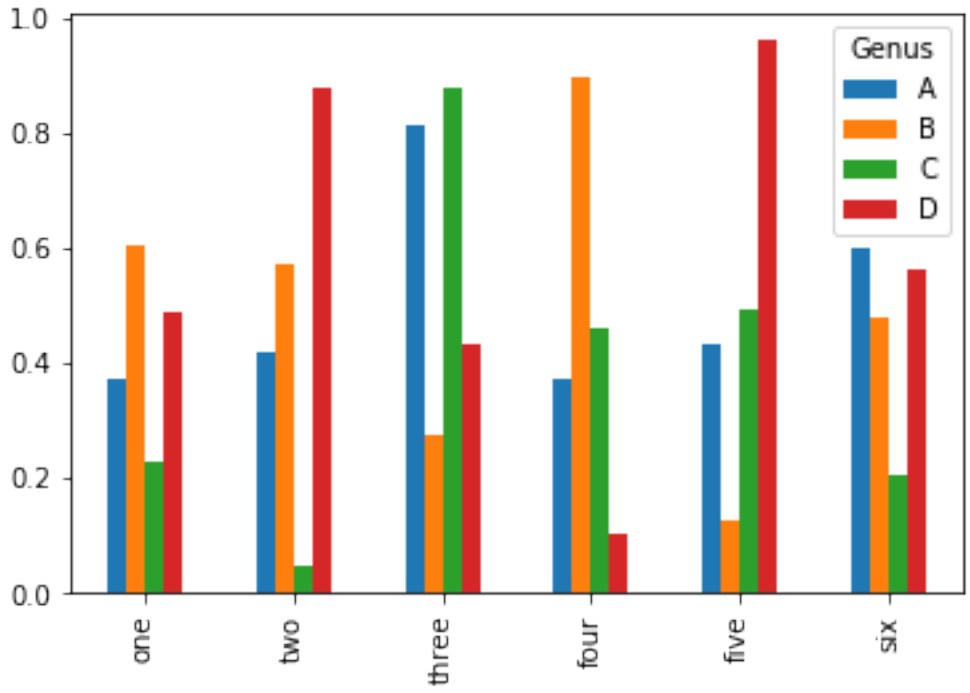
Out[12]:

Genus	A	B	C	D
one	0.370670	0.602792	0.229159	0.486744
two	0.420082	0.571653	0.049024	0.880592
three	0.814568	0.277160	0.880316	0.431326
four	0.374020	0.899420	0.460304	0.100843
five	0.433270	0.125107	0.494675	0.961825
six	0.601648	0.478576	0.205690	0.560547

In [13]:

Out[13]:

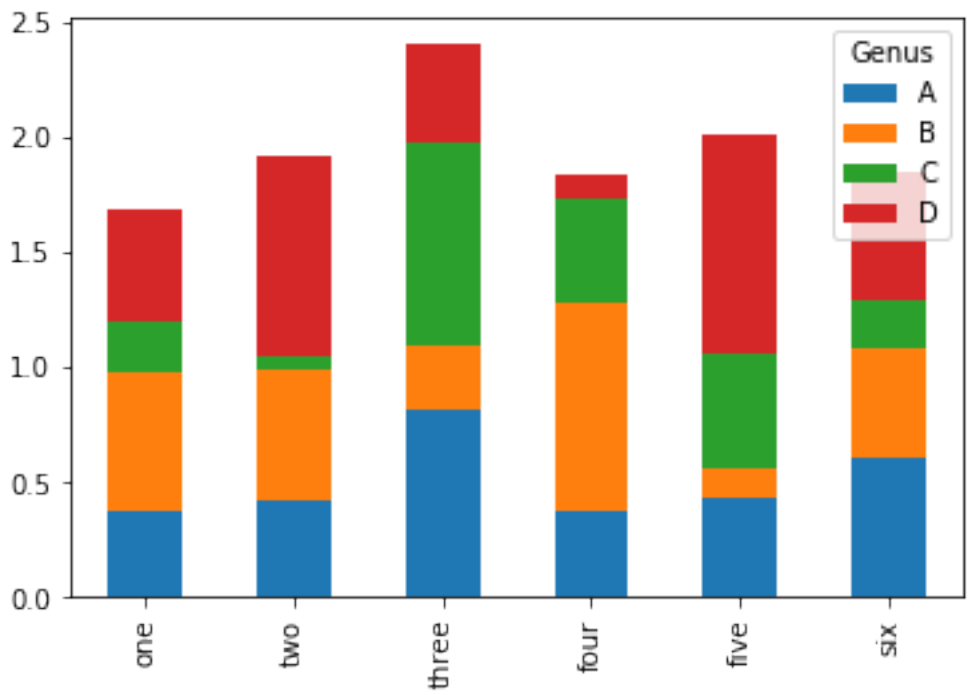
```
<matplotlib.axes._subplots.AxesSubplot at 0x11c6b5b38>
```



In [16]:

Out[16]:

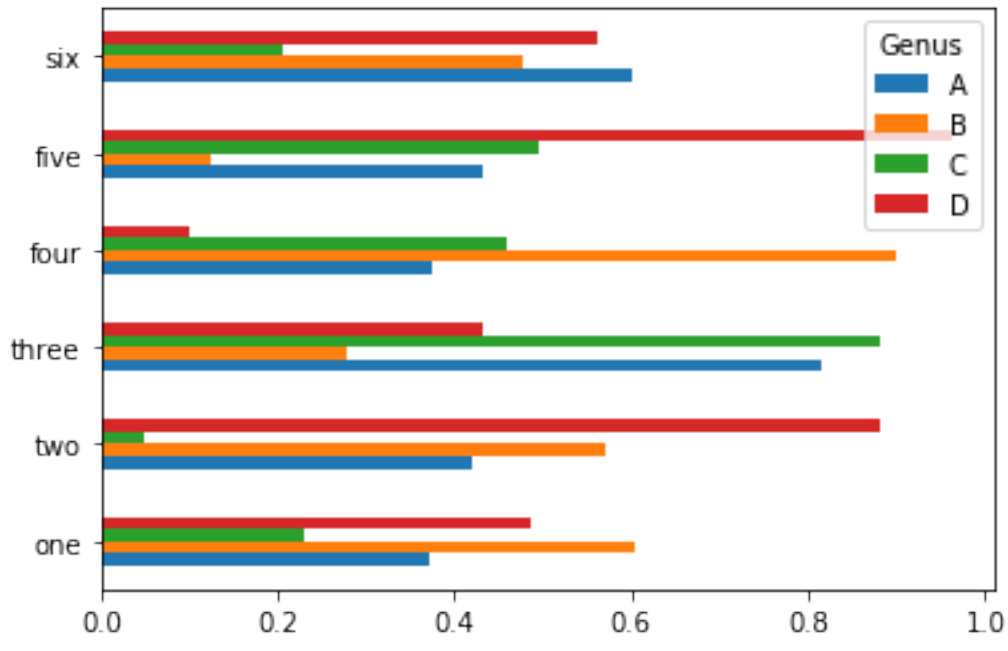
```
<matplotlib.axes._subplots.AxesSubplot at 0x11ca01c88>
```



In [17]:

Out[17]:

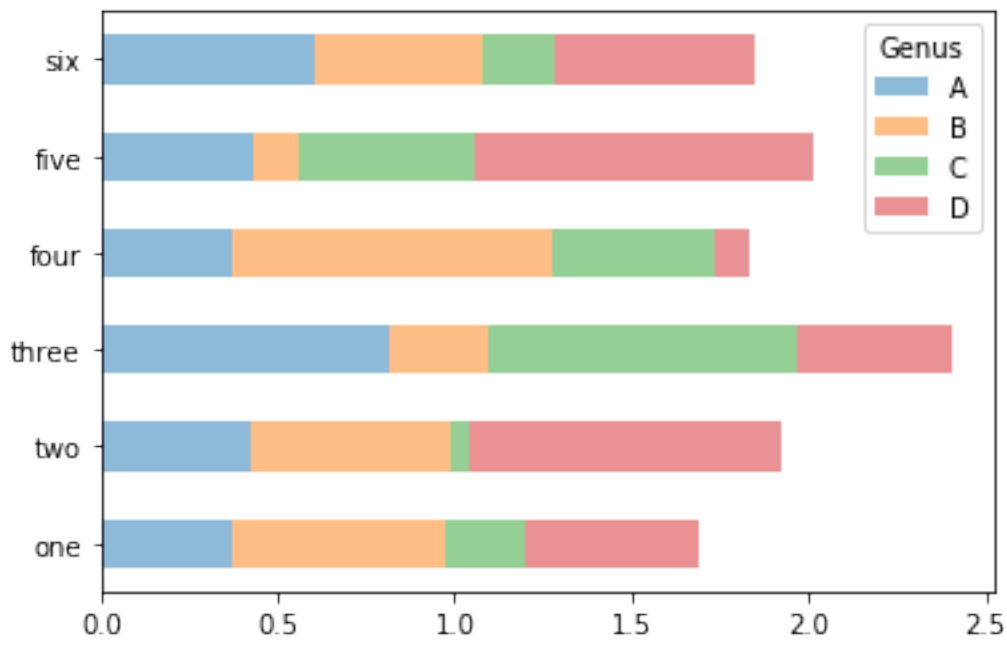
```
<matplotlib.axes._subplots.AxesSubplot at 0x11ce53c50>
```



In [18]:

Out[18]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x11ce2a2b0>
```



Tipping DataSet

In [ ]:

Crosstab (also known as contingency table or cross tabulation) is a table showing *frequency distribution* of one variable in rows and another on columns.

In [ ]:

```

```

In [ ]:

```

```

In [ ]:

```

```

Party size appear to increase on the weekend in this dataset.