In [15]:
```
import numpy as np
import pandas as pd
from PIL import Image
```

In [16]:
```
data = pd.read_csv('https://cocl.us/datascience_survey_data', index_col=
0)
```

In [17]:
```
data.head()
```

Out[17]:

	Very interested	Somewhat interested	Not interested
Big Data (Spark / Hadoop)	1332	729	127
Data Analysis / Statistics	1688	444	60
Data Journalism	429	1081	610
Data Visualization	1340	734	102
Deep Learning	1263	770	136

In [18]:
```
data1 = data.sort_values(by = ['Very interested'], ascending = False)
```

In [19]:
```
data2 = round((data1/2233)*100, 2)
```

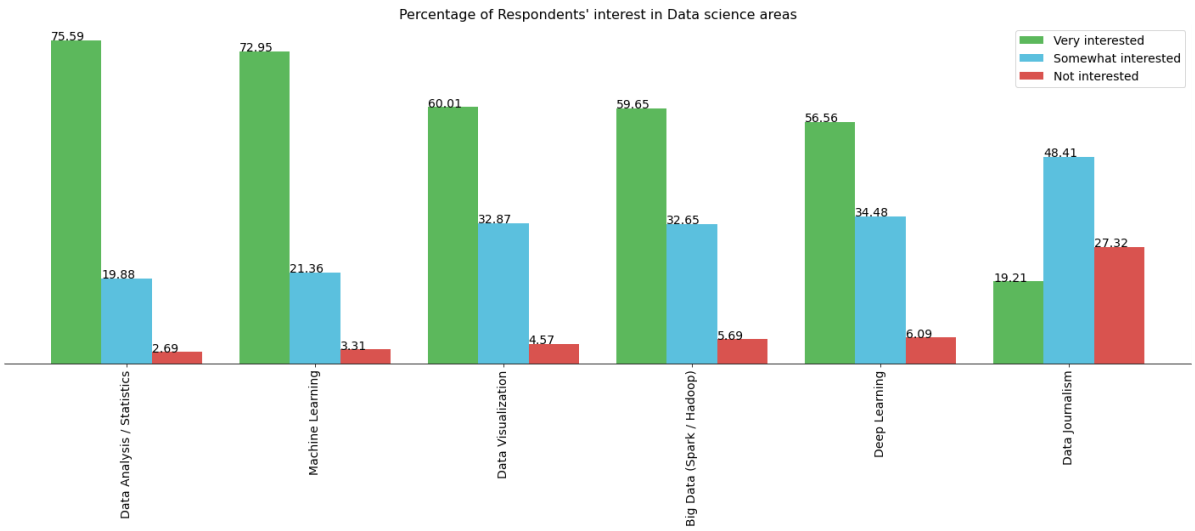
In [20]:
```
data2
```

Out[20]:

	Very interested	Somewhat interested	Not interested
Data Analysis / Statistics	75.59	19.88	2.69
Machine Learning	72.95	21.36	3.31
Data Visualization	60.01	32.87	4.57
Big Data (Spark / Hadoop)	59.65	32.65	5.69
Deep Learning	56.56	34.48	6.09
Data Journalism	19.21	48.41	27.32

In [21]:
```
import matplotlib as mpl
import matplotlib.pyplot as plt
ax = data2.plot(kind = 'bar', figsize=(20, 8), width = 0.8, color = ['#5
cb85c','#5bc0de', '#d9534f'])
plt.title("Percentage of Respondents' interest in Data science areas", s
ize=16)
from decimal import Decimal
for p in ax.patches:
    ax.annotate('{:.2f}'.format(Decimal(str(p.get_height()))), (p.get_x
(), p.get_height()), fontsize=14)
plt.tight_layout()
plt.gca().spines['right'].set_color('none')
plt.gca().spines['top'].set_color('none')
plt.gca().spines['left'].set_color('none')
plt.yticks([])
plt.legend(fontsize=14)
plt.xticks(fontsize=14)
```

Out[21]:
```
(array([0, 1, 2, 3, 4, 5]),
 [Text(0, 0, 'Data Analysis / Statistics'),
  Text(1, 0, 'Machine Learning'),
  Text(2, 0, 'Data Visualization'),
  Text(3, 0, 'Big Data (Spark / Hadoop)'),
  Text(4, 0, 'Deep Learning'),
  Text(5, 0, 'Data Journalism')])
```



In [22]:
```
center = pd.read_csv('https://cocl.us/sanfran_crime_dataset')
```

In [23]:
```
center.head()
```

Out[23]:

	IncidntNum	Category	Descript	DayOfWeek	Date	Time	PdDistrict	Resolutic
0	120058272	WEAPON LAWS	POSS OF PROHIBITED WEAPON	Friday	01/29/2016 12:00:00 AM	11:00	SOUTHERN	ARRES BOOKE
1	120058272	WEAPON LAWS	FIREARM, LOADED, IN VEHICLE, POSSESSION OR USE	Friday	01/29/2016 12:00:00 AM	11:00	SOUTHERN	ARRES BOOKE
2	141059263	WARRANTS	WARRANT ARREST	Monday	04/25/2016 12:00:00 AM	14:59	BAYVIEW	ARRES BOOKE
3	160013662	NON-CRIMINAL	LOST PROPERTY	Tuesday	01/05/2016 12:00:00 AM	23:50	TENDERLOIN	NON
4	160002740	NON-CRIMINAL	LOST PROPERTY	Friday	01/01/2016 12:00:00 AM	00:30	MISSION	NON

In [24]:
```
center1 = center.groupby(["PdDistrict"]).count().reset_index()
center1.drop(center1.columns.difference(['PdDistrict','IncidntNum']), 1,
inplace=True)
center1.rename(columns={'PdDistrict':'Neighborhood','IncidntNum':'Count'
}, inplace=True)
center1
```

Out[24]:

	Neighborhood	Count
0	BAYVIEW	14303
1	CENTRAL	17666
2	INGLESIDE	11594
3	MISSION	19503
4	NORTHERN	20100
5	PARK	8699
6	RICHMOND	8922
7	SOUTHERN	28445
8	TARAVAL	11325
9	TENDERLOIN	9942

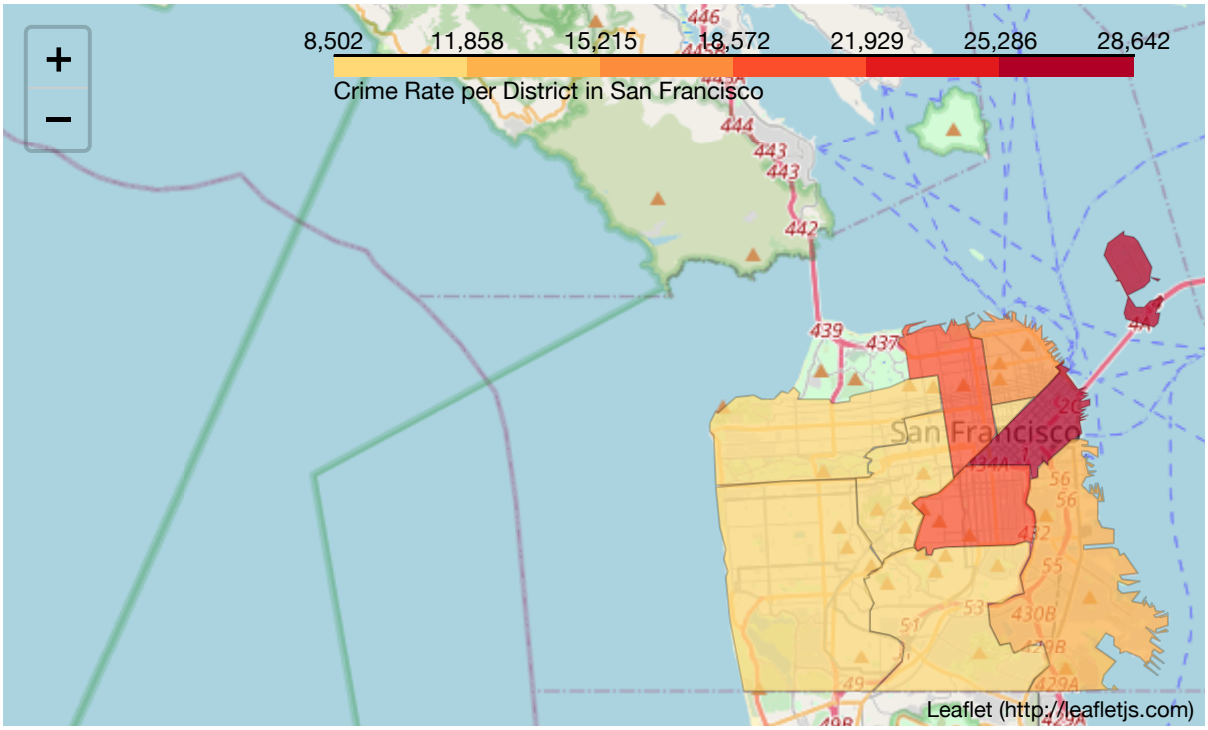
In [25]:
```
import folium
```

In [26]:
```
latitude = 37.77
longitude = -122.42
sfmap = folium.Map(location=[latitude,longitude], zoom_start=12)
```

In [27]:
```
sfmap.choropleth(
    geo_data='https://cocl.us/sanfran_geojson',
    data=center1,
    columns=['Neighborhood', 'Count'],
    key_on='feature.properties.DISTRICT',
    fill_color='YlOrRd',
    fill_opacity=0.7,
    line_opacity=0.2,
    legend_name='Crime Rate per District in San Francisco')
```

In [28]:
```
sfmap
```

Out[28]:



In [ ]:

In [ ]: