Week 9 & 10 - Assignment 5.2

PYTHON

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

In [2]:
```
education_df = pd.read_csv("education.csv")
print(education_df.shape)
education_df.head(5)
```

```
(52, 7)
```

Out[2]:

	state	reading	math	writing	percent_graduates_sat	pupil_staff_ratio	dropout_rate
0	United States	501	515	493	46	7.9	4.4
1	Alabama	557	552	549	7	6.7	2.3
2	Alaska	520	516	492	46	7.9	7.3
3	Arizona	516	521	497	26	10.4	7.6
4	Arkansas	572	572	556	5	6.8	4.6

In [3]:
```
crime_rate_by_state_df = pd.read_csv("crimeratesbystate-formatted.csv")
print(crime_rate_by_state_df.shape)
crime_rate_by_state_df.head(5)
```

```
(52, 8)
```

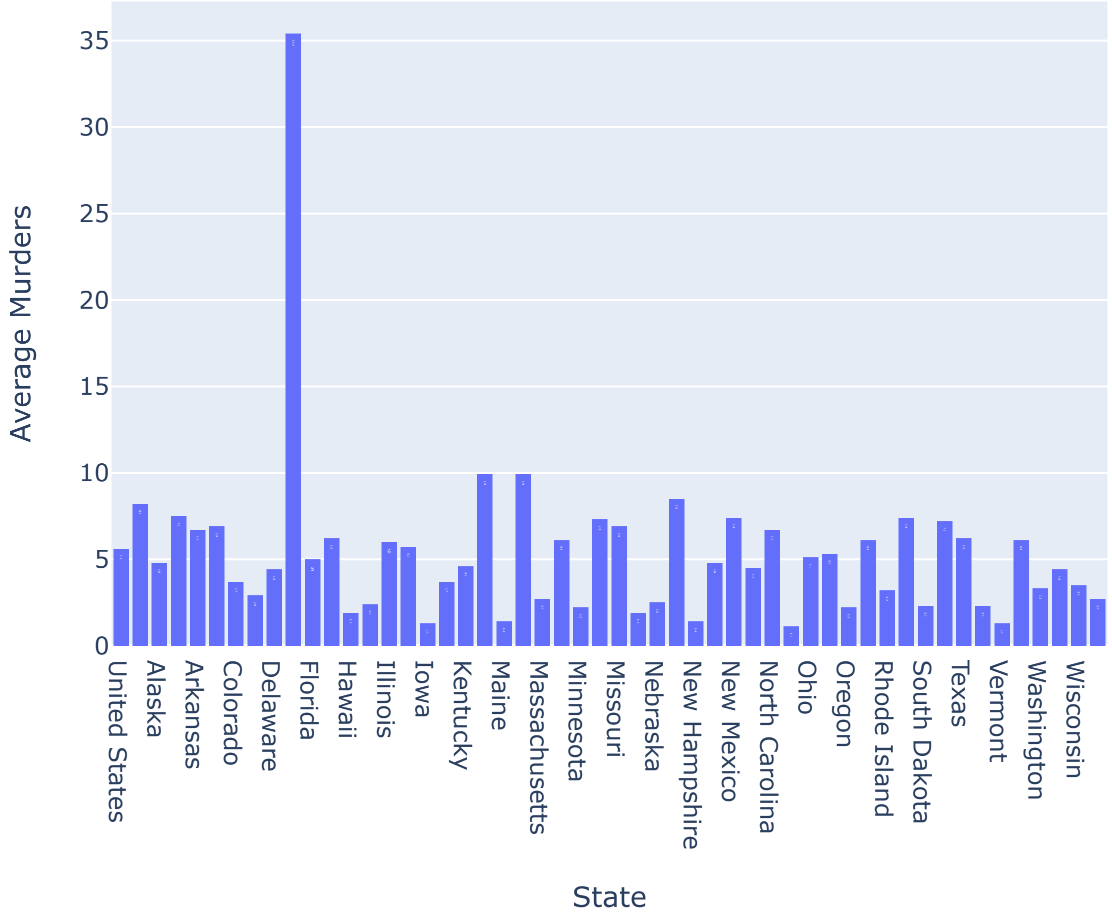
Out[3]:

	state	murder	forcible_rape	robbery	aggravated_assault	burglary	larceny_theft	motor_vehicle_theft
0	United States	5.6	31.7	140.7	291.1	726.7	2286.3	416.7
1	Alabama	8.2	34.3	141.4	247.8	953.8	2650.0	288.3
2	Alaska	4.8	81.1	80.9	465.1	622.5	2599.1	391.0
3	Arizona	7.5	33.8	144.4	327.4	948.4	2965.2	924.4
4	Arkansas	6.7	42.9	91.1	386.8	1084.6	2711.2	262.1

Python - Histogram

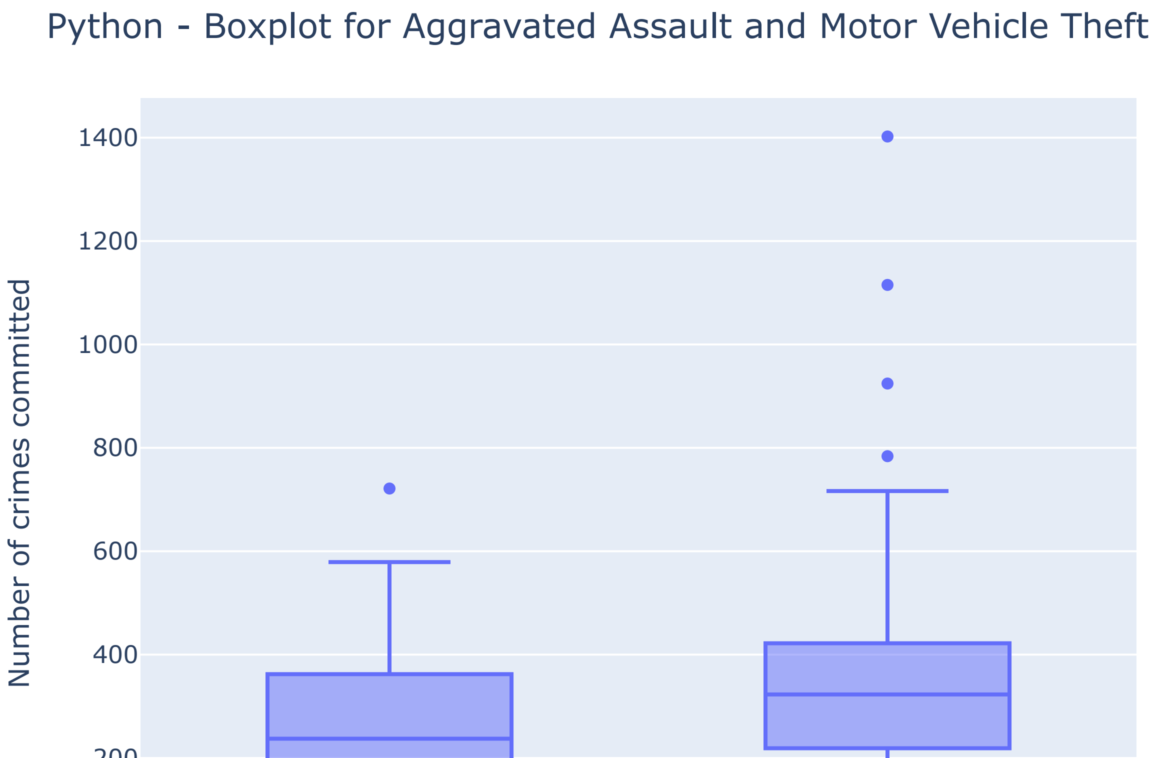
In [4]:
```
import plotly.express as px

fig = px.histogram(crime_rate_by_state_df, x="state", y="murder", histfunc="avg", nbins=
fig.update_layout(title = 'Python - Histogram for Murders per State',
                  xaxis_title="State",
                  yaxis_title="Average Murders")
fig.show("notebook")
```

Python - Histogram for Murders per State



Python - Box Plot

```
In [5]:  fig = px.box(crime_rate_by_state_df, y=["aggravated_assault","motor_vehicle_theft"])
         fig.update_layout(title = 'Python - Boxplot for Aggravated Assault and Motor Vehicle The
                           xaxis_title="Crime Category",
                           yaxis_title="Number of crimes committed")
         fig.show("notebook")
```



Python - Bullet chart

In [6]:

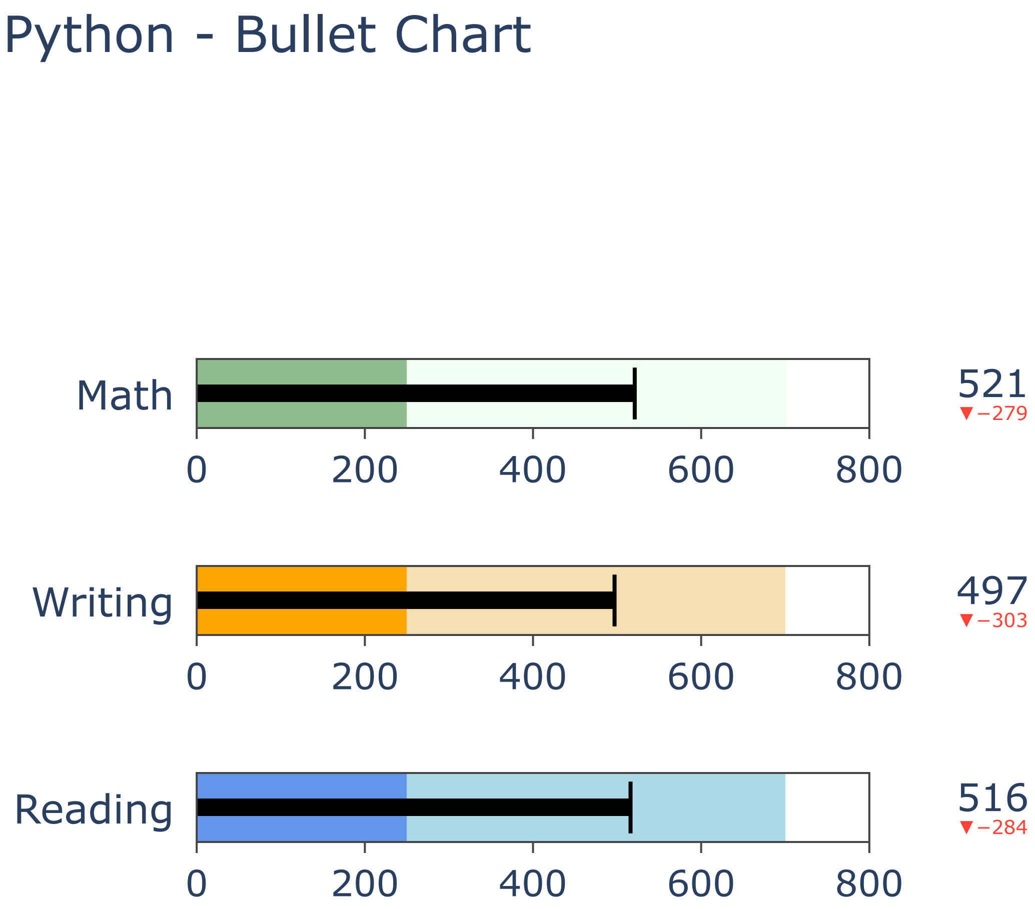
```
bullet_chart_df = education_df[(education_df.state.str.strip() == "Arizona")].reset_inde
bullet_chart_df
```

Out[6]:

	state	reading	math	writing	percent_graduates_sat	pupil_staff_ratio	dropout_rate
0	Arizona	516	521	497	26	10.4	7.6

In [7]:

```
fig = go.Figure()
fig.add_trace(go.Indicator(
    mode = "number+gauge+delta", value = bullet_chart_df.reading.iloc[0],
    delta = {'reference': 800},
    domain = {'x': [0.1, 1], 'y': [0, 0.1]},
    title = {'text': "Reading"},
    gauge = {
        'shape': "bullet",
        'axis': {'range': [None, 800]},
        'threshold': {
            'line': {'color': "black", 'width': 2},
            'thickness': 0.75,
            'value':bullet_chart_df.reading.iloc[0]},
        'steps': [
            {'range': [0,250], 'color': "cornflowerblue"},
            {'range': [250, 700], 'color': "lightblue"}],
        'bar': {'color': "black"}}))

fig.add_trace(go.Indicator(
    mode = "number+gauge+delta", value = bullet_chart_df.writing.iloc[0],
    delta = {'reference': 800},
    domain = {'x': [0.1, 1], 'y': [0.3, 0.4]},
    title = {'text': "Writing"},
    gauge = {
        'shape': "bullet",
        'axis': {'range': [None, 800]},
        'threshold': {
            'line': {'color': "black", 'width': 2},
            'thickness': 0.75,
            'value':bullet_chart_df.writing.iloc[0]},
        'steps': [
            {'range': [0,250], 'color': "orange"},
            {'range': [250, 700], 'color': "wheat"}],
        'bar': {'color': "black"}}))

fig.add_trace(go.Indicator(
    mode = "number+gauge+delta", value = bullet_chart_df.math.iloc[0],
    delta = {'reference': 800},
    domain = {'x': [0.1, 1], 'y': [0.6, 0.7]},
    title = {'text': "Math"},
    gauge = {
        'shape': "bullet",
        'axis': {'range': [None, 800]},
        'threshold': {
            'line': {'color': "black", 'width': 2},
            'thickness': 0.75,
            'value':bullet_chart_df.math.iloc[0]},
        'steps': [
            {'range': [0,250], 'color': "darkseagreen"},
            {'range': [250, 700], 'color': "honeydew"}],
        'bar': {'color': "black"}}))
fig.update_layout(title="Python - Bullet Chart", font=dict(size=18))
fig.show()
```



Following is a bullet chart with marker for practice

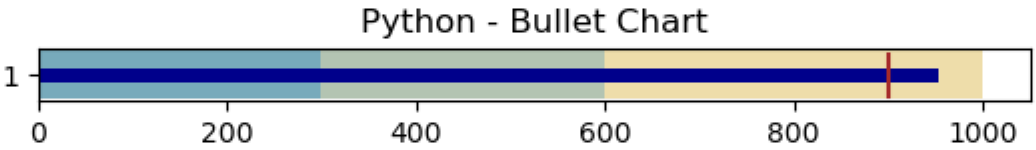
In [8]:

```
import seaborn as sns
crime_bullet = crime_rate_by_state_df[crime_rate_by_state_df["state"].str.strip()=="Alab
crime_bullet['target'] = 900
#print(crime_bullet)
crime_bullet_tuple = [tuple(x) for x in crime_bullet.values][0]
print('crime_bullet_tuple : ', crime_bullet_tuple)

limits = [300, 600, 1000]
palette = sns.color_palette("blend:#7AB,#EDA", len(limits))
fig, ax = plt.subplots()
ax.set_aspect('equal')
ax.set_yticks([1])
ax.set_title("Python - Bullet Chart")

prev_limit = 0
for idx, lim in enumerate(limits):
    #print(idx, lim)
    ax.barh([1], lim-prev_limit, left=prev_limit, height=50, color=palette[idx])
    prev_limit = lim

ax.barh([1], crime_bullet_tuple[1], color='darkblue', height=15)
ax.axvline(crime_bullet_tuple[2], color="brown", ymin=0.10, ymax=0.9)
```

```
crime_bullet_tuple :  ('Alabama ', 953.8, 900)
```

Out[8]:

```
<matplotlib.lines.Line2D at 0x17d428ec730>
```



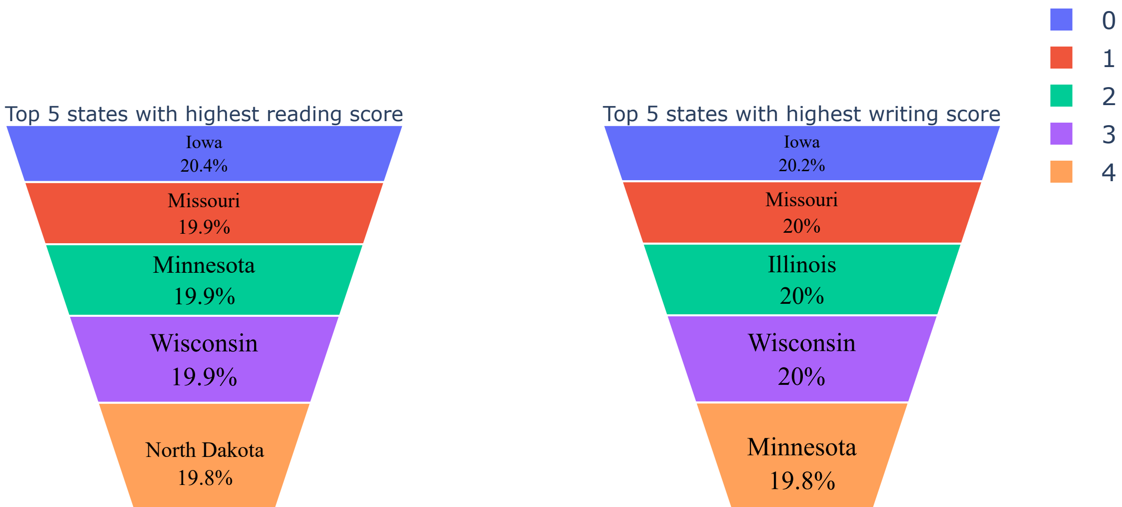
Python - Additional chart : Funnel Chart

In [9]:

```
#create separate dataframes for reading and writing with the state names. Create lists f
reading_df = education_df[["state","reading"]].sort_values("reading",ascending=False)  #S
read_val = list(x for x in reading_df.head(5)["reading"])
read_state_val = list(x for x in reading_df.head(5)["state"])  #get the top 5 states


writing_df = education_df[["state","writing"]].sort_values("writing",ascending=False)  #S
write_val = list(x for x in writing_df.head(5)["writing"])
write_state_val = list(x for x in writing_df.head(5)["state"])  #get the top 5 states
```

In [10]:

```
fig = go.Figure()

fig.add_trace(go.Funnelarea(
    title = {"position": "top center", "text": "Top 5 states with highest reading score"
    domain = {"x": [0, 0.4], "y": [0.12, 1]},
    values = read_val, text = read_state_val,
    textfont = {"family": "Old Standard TT, serif", "size": 13, "color": "black"}))

fig.add_trace(go.Funnelarea(
    title = {"position": "top left", "text": "Top 5 states with highest writing score","
    domain = {"x": [0.6,1], "y": [0.12, 1]},
    values = write_val, text = write_state_val,
    textfont = {"family": "Old Standard TT, serif", "size": 13, "color": "black"}))

fig.show()
```


Top 5 states with highest reading score
Iowa
Missouri
Minnesota
Wisconsin
North Dakota
Top 5 states with highest writing score
Iowa
Missouri
Illinois
Wisconsin
Minnesota