```
import pandas as pd
import matplotlib.pyplot as plt
import plotly.express as px
```

```
[ ] from google.colab import drive
    drive.mount('/content/drive')
```

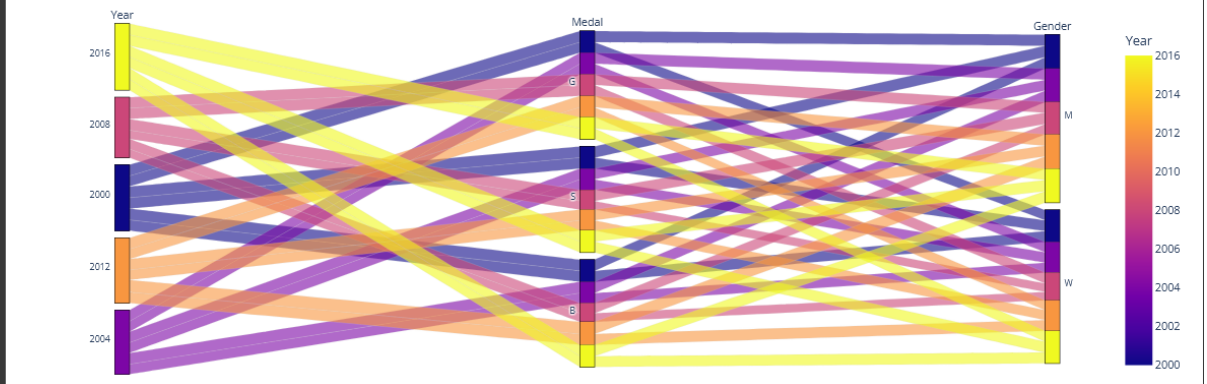
```
Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content/drive", force_remount=True).
```

Q1 Please use Pandas to read olympic_medals.csv and use parallel_categories function from plotly.express to visualize proportions of medal type for each gender from since year 2000. Please see the example in the Python notebook we walked through in the class.

```
[ ] #please use this cell to read and select your data

    df = pd.read_csv('/content/drive/MyDrive/DATA/olympic_medals.csv')
    df.head()
```

	Gender	Event	Location	Year	Medal	Name	Nationality	Result
0	M	10000M Men	Rio	2016	G	Mohamed FARAH	GBR	25:05.17
1	M	10000M Men	Rio	2016	S	Paul Kipngetich TANUI	KEN	27:05.64
2	M	10000M Men	Rio	2016	B	Tamirat TOLA	ETH	27:06.26
3	M	10000M Men	Beijing	2008	G	Kenenisa BEKELE	ETH	27:01.17
4	M	10000M Men	Beijing	2008	S	Sileshi SIHINE	ETH	27:02.77

```
[ ] #Please use this cell to create your your figure. Please use Year column to color your graph.

    df = df[df['Year'] >= 2000]

    px.parallel_categories(df, dimensions=['Year', 'Medal', 'Gender'], color="Year")
```

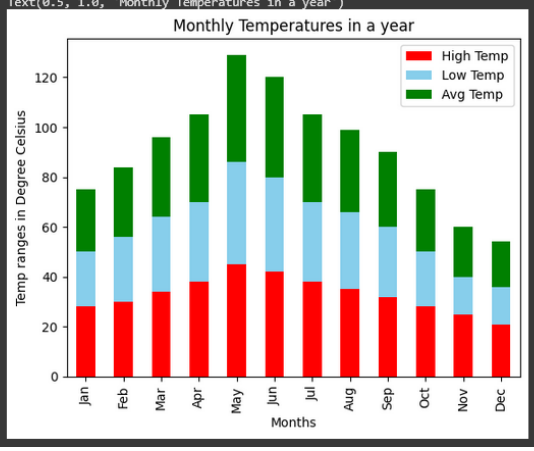


Q2 Please inspect the code below and observe how values are plotted by running it. Then, read the 2016elections.csv from the DATA folder and select rows for AR, MI, CA, and WI. Then, utilize stacked bar plot, to stack vote percentages for Trump, Clinton, Johnson, and Others. Please see 'pct_clinton', 'pct_trump', 'pct_johnson', 'pct_other' columns. Make sure that your x tick labels are those four states above.

+ Code + Text

```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt

# create DataFrame
df = pd.DataFrame({'High Temp': [28, 30, 34, 38, 45, 42,
                                 38, 35, 32, 28, 25, 21],
                   'Low Temp': [22, 26, 30, 32, 41, 38,
                                32, 31, 28, 22, 15, 15],
                   'Avg Temp': [25, 28, 32, 35, 43, 40,
                                35, 33, 30, 25, 20, 18]},
                  index=['Jan', 'Feb', 'Mar', 'Apr', 'May',
                         'Jun', 'Jul', 'Aug', 'Sep', 'Oct',
                         'Nov', 'Dec'])


# create stacked bar chart for monthly temperatures
df.plot(kind='bar', stacked=True, color=['red', 'skyblue', 'green'])

# labels for x & y axis
plt.xlabel('Months')
plt.ylabel('Temp ranges in Degree Celsius')

# title of plot
plt.title('Monthly Temperatures in a year')
```

```
Text(0.5, 1.0, 'Monthly Temperatures in a year')
```



```python
#You can use this cell to write your code. It is doable at most 4 lines of code.

df = pd.read_csv('/content/drive/MyDrive/DATA/2016elections.csv')

# technically one line from these two combined :)
# df = df[df['state'].isin(['Arizona', 'Michigan', 'California', 'Wisconsin'])]
# df = df.set_index('state')[['pct_clinton', 'pct_trump', 'pct_johnson', 'pct_other']]

df = (df[df['state'].isin(['Arizona', 'Michigan', 'California', 'Wisconsin'])]
        .set_index('state')[['pct_clinton', 'pct_trump', 'pct_johnson', 'pct_other']]
)

df.plot(kind='bar', stacked=True, color=['blue', 'red', 'yellow', 'gray'])

# another "one liner"
plt.gca().set(xlabel='States', ylabel='Vote Percentages', title='2016 Election Vote Percentages'); plt.show()
```