Unit 2.4 Graded Assignment: Data visualization – Scatter Matrix

Instructions:

Download the Breast Cancer Wisconsin dataset from https://www.kaggle.com/datasets/uciml/breast-cancer-wisconsin-data.

After downloading, read about scatter matrix and implement it using plotly.

Limit it to only few (5-6) features of your choice. Try to make it as readable as possible (eg. use colors to represent target class).

Submitted By:

1. **Syed Muhammad Hammad Irshad (2303.KHI.DEG.032)**
2. **Arshad Shiwani (2303.KHI.DEG.026)**

Solution:

In this assignment we use plotly and pandas to import csv file and present data visualization on given dataset

first we use pandas to import csv file

```
import pandas as pd
```
✓ 0.4s Python

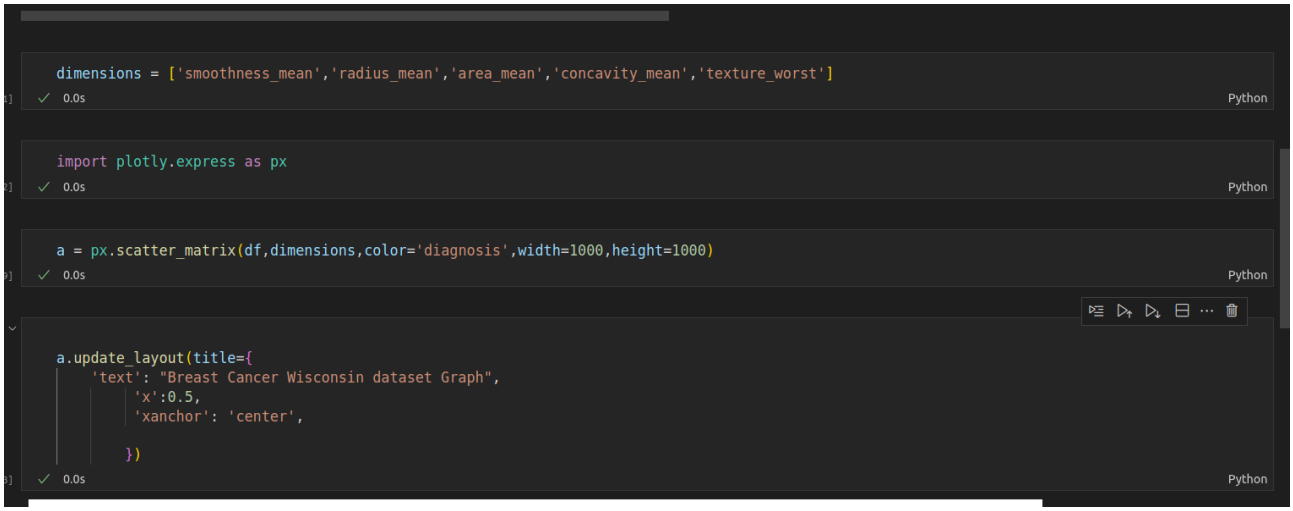
```
df = pd.read_csv('data.csv')
```
✓ 0.0s Python

```
df.head(10)
```
✓ 0.0s Python

	id	diagnosis	radius_mean	texture_mean	perimeter_mean	area_mean	smoothness_mean	compactness_mean	concavity_mean	concave points_mean	...	texture_wo
0	842302	M	17.99	10.38	122.80	1001.0	0.11840	0.27760	0.30010	0.14710	...	17.
1	842517	M	20.57	17.77	132.90	1326.0	0.08474	0.07864	0.08690	0.07017	...	23.
2	84300903	M	19.69	21.25	130.00	1203.0	0.10960	0.15990	0.19740	0.12790	...	25.
3	84348301	M	11.42	20.38	77.58	386.1	0.14250	0.28390	0.24140	0.10520	...	26.
4	84358402	M	20.29	14.34	135.10	1297.0	0.10030	0.13280	0.19800	0.10430	...	16.
5	843786	M	12.45	15.70	82.57	477.1	0.12780	0.17000	0.15780	0.08089	...	23.
6	844359	M	18.25	19.98	119.60	1040.0	0.09463	0.10900	0.11270	0.07400	...	27.
7	84458202	M	13.71	20.83	90.20	577.9	0.11890	0.16450	0.09366	0.05985	...	28.
8	844981	M	13.00	21.82	87.50	519.8	0.12730	0.19320	0.18590	0.09353	...	30.
9	84501001	M	12.46	24.04	83.97	475.9	0.11860	0.23960	0.22730	0.08543	...	40.

10 rows × 33 columns

second we use plotly to represent data in the form of graph



```
dimensions = ['smoothness_mean','radius_mean','area_mean','concavity_mean','texture_worst']
```

0.0s Python

```
import plotly.express as px
```

0.0s Python

```
a = px.scatter_matrix(df,dimensions,color='diagnosis',width=1000,height=1000)
```

0.0s Python

```
a.update_layout(title={
    'text': "Breast Cancer Wisconsin dataset Graph",
        'x':0.5,
        'xanchor': 'center',

       })
```

0.0s Python

after running above commands we get this graph