Talha Khan (2303.009.KHI.DEG)

Muhammad Moiz Khan (2303.022.KHI.DEG)

ASSIGNMNET 2.4

```
[2]: import pandas as pd
     import plotly.express as px

[3]: data = pd.read_csv("./Downloads/scaler/data.csv")
     data
```

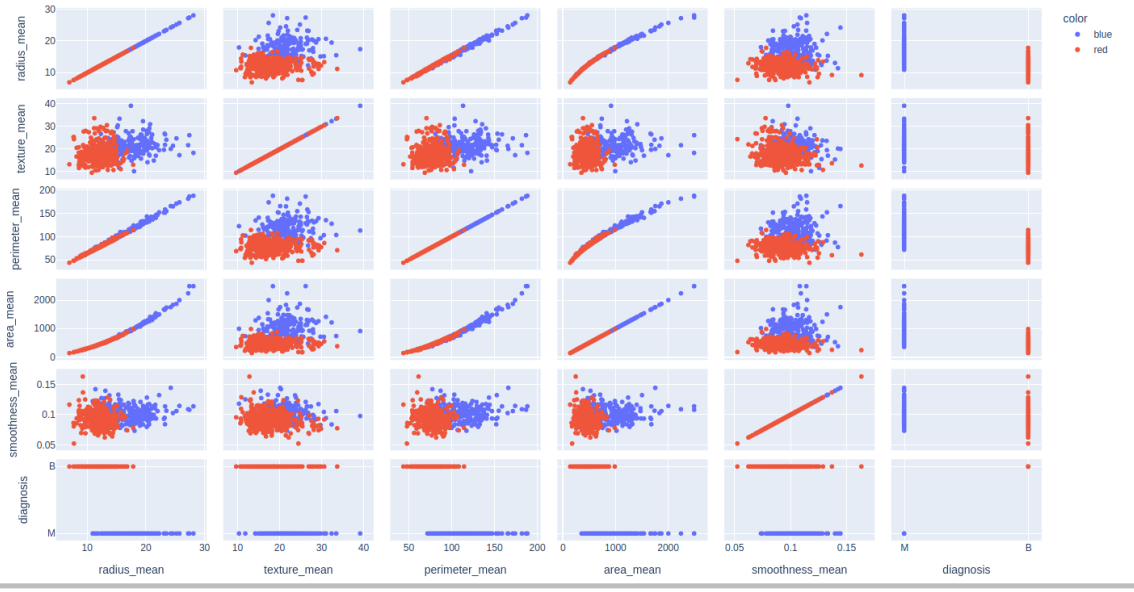
[3]:

	id	diagnosis	radius_mean	texture_mean	perimeter_mean	area_mean	smoothness_mean	compactness_mean	concavity_mean	concave points_mean	...	texture_worst	perimeter_worst	area_worst	smooth
0	842302	M	17.99	10.38	122.80	1001.0	0.11840	0.27760	0.30010	0.14710	...	17.33	184.60	2019.0	
1	842517	M	20.57	17.77	132.90	1326.0	0.08474	0.07864	0.08690	0.07017	...	23.41	158.80	1956.0	
2	84300903	M	19.69	21.25	130.00	1203.0	0.10960	0.15990	0.19740	0.12790	...	25.53	152.50	1709.0	
3	84348301	M	11.42	20.38	77.58	386.1	0.14250	0.28390	0.24140	0.10520	...	26.50	98.87	567.7	
4	84358402	M	20.29	14.34	135.10	1297.0	0.10030	0.13280	0.19800	0.10430	...	16.67	152.20	1575.0	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
564	926424	M	21.56	22.39	142.00	1479.0	0.11100	0.11590	0.24390	0.13890	...	26.40	166.10	2027.0	
565	926682	M	20.13	28.25	131.20	1261.0	0.09780	0.10340	0.14400	0.09791	...	38.25	155.00	1731.0	
566	926954	M	16.60	28.08	108.30	858.1	0.08455	0.10230	0.09251	0.05302	...	34.12	126.70	1124.0	
567	927241	M	20.60	29.33	140.10	1265.0	0.11780	0.27700	0.35140	0.15200	...	39.42	184.60	1821.0	
568	92751	B	7.76	24.54	47.92	181.0	0.05263	0.04362	0.00000	0.00000	...	30.37	59.16	268.6	

569 rows × 33 columns

- This code reads a CSV file located in the "Downloads/scaler/" directory and creates a scatter matrix plot using the px.scatter_matrix() function from the Plotly Express library.

```
[4]: fig = px.scatter_matrix(data,
         dimensions=["radius_mean", "texture_mean", "perimeter_mean", "area_mean", "smoothness_mean", "diagnosis"],color=data['diagnosis'].map({'B': 'red', 'M': 'blue'}))
     fig.update_layout(height=800,showlegend=True)
     fig.show()
```



- The plot will have six dimensions, namely radius_mean, texture_mean, perimeter_mean, area_mean, smoothness_mean, and diagnosis.
- The color parameter is used to color-code the points based on the diagnosis column, with 'B' (benign) shown in red and 'M' (malignant) shown in blue.
- Finally, the fig.update_layout() method is called to set the height of the plot and show the legend, and fig.show() is used to display the plot in the Jupyter notebook or in a web browser if you're running the code outside of Jupyter.