

ASSIGNMENT 1

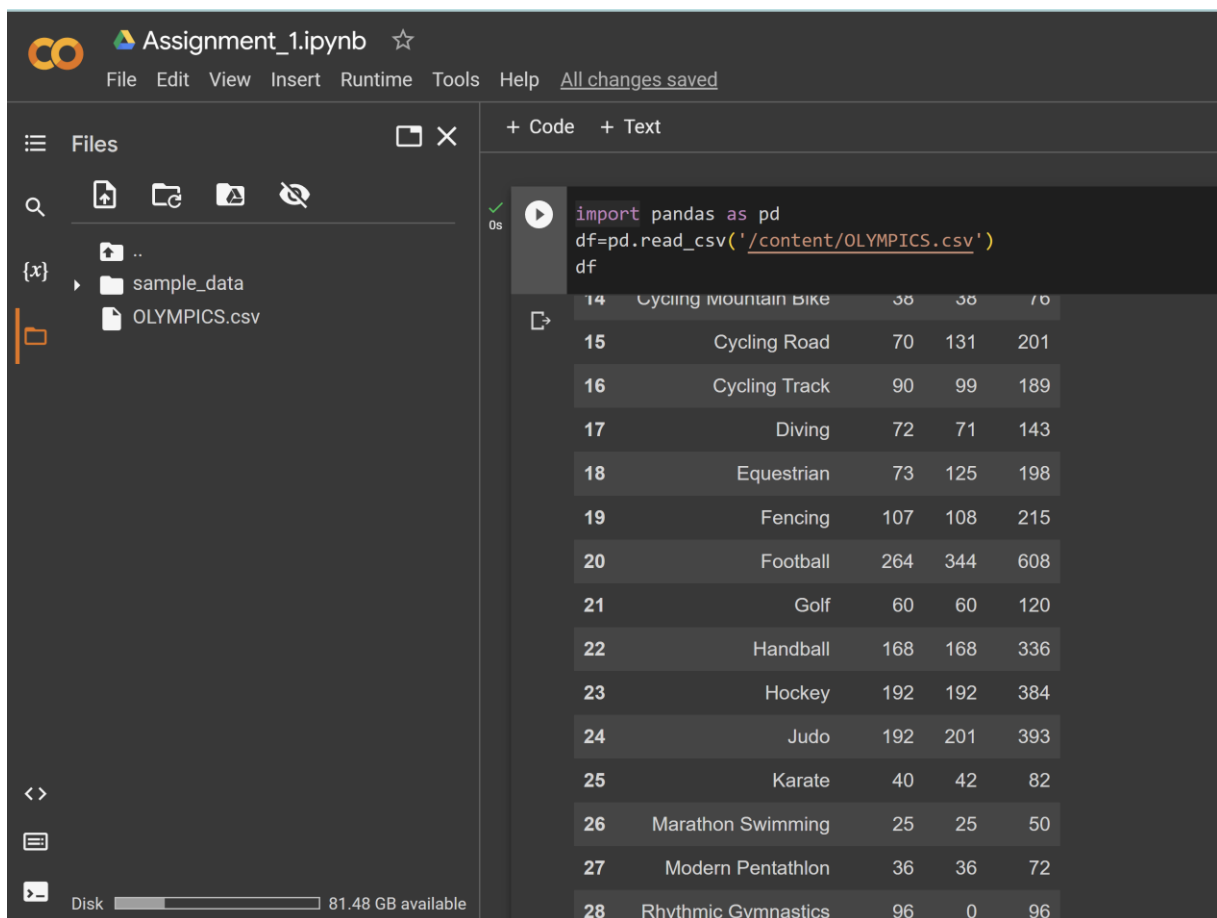
Siddharth Bhardwaj

21BCE2001

Git Link:

https://github.com/siDarth0803/AI_ML/blob/main/assignment_1.ipynb

Task - 1 Create a pandas dataframe (DataFrame name as 'df')




The screenshot shows a Jupyter Notebook titled "Assignment_1.ipynb". The left sidebar displays a file explorer with a folder named "sample_data" containing a file "OLYMPICS.csv". The main area shows a code cell with the following Python code:

```
import pandas as pd
df=pd.read_csv('/content/OLYMPICS.csv')
df
```

Below the code, the resulting DataFrame is displayed as a table with 28 rows and 5 columns. The table contains data for various Olympic sports, including Cycling, Diving, Equestrian, Fencing, Football, Golf, Handball, Hockey, Judo, Karate, Marathon Swimming, Modern Pentathlon, and Rhythmic Gymnastics.

14	Cycling Mountain Bike	36	36	76
15	Cycling Road	70	131	201
16	Cycling Track	90	99	189
17	Diving	72	71	143
18	Equestrian	73	125	198
19	Fencing	107	108	215
20	Football	264	344	608
21	Golf	60	60	120
22	Handball	168	168	336
23	Hockey	192	192	384
24	Judo	192	201	393
25	Karate	40	42	82
26	Marathon Swimming	25	25	50
27	Modern Pentathlon	36	36	72
28	Rhythmic Gymnastics	96	0	96

 `df.tail(10)`



	Play	Female	Male	Total
36	Swimming	361	418	779
37	Table Tennis	86	86	172
38	Taekwondo	65	65	130
39	Tennis	94	97	191
40	Trampoline Gymnastics	16	16	32
41	Triathlon	55	55	110
42	Volleyball	144	144	288
43	Water Polo	122	146	268
44	Weightlifting	98	99	197
45	Wrestling	96	193	289



[3] `df.shape`

(46, 4)

[5] `df.Female.mean()`

118.08695652173913

[6] `df.Male.mode()`

0 144
Name: Male, dtype: int64

Task- 2 Check the info of 'df'



```
df.info()
```



```
<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 46 entries, 0 to 45  
Data columns (total 4 columns):  
#   Column  Non-Null Count  Dtype  
---  -  
0   Play    46 non-null      object  
1   Female  46 non-null      int64  
2   Male    46 non-null      int64  
3   Total   46 non-null      int64  
dtypes: int64(3), object(1)  
memory usage: 1.6+ KB
```

Task 3- Check the descriptive statistics of 'df'



```
df.describe()
```



	Female	Male	Total
count	46.000000	46.000000	46.000000
mean	118.086957	127.913043	246.000000
std	147.169717	166.900746	312.46134
min	10.000000	0.000000	19.000000
25%	42.750000	40.250000	85.500000
50%	90.000000	97.500000	190.000000
75%	138.750000	149.750000	288.750000
max	969.000000	1072.000000	2041.000000



Task 4- check the 4th index observation with 'loc' slicing operator.

```
[9] df.loc[4]
```

Play	Athletics
Female	969
Male	1072
Total	2041

Name: 4, dtype: object

Task 5 - Check the null values in your 'df'

```
df.isnull().any()
```

Play	False
Female	False
Male	False
Total	False

dtype: bool