

## AI ML Assignment 1

Task - 1	Create a pandas dataframe (DataFrame name as 'df') with numpy random values (4 features and 4 observation)
Task - 2	Rename the task - 1 'df' dataframe column names to 'Random value 1', 'Random value 2', 'Random value 3' & 'Random value 4'
Task - 3	Find the descriptive statistics of the 'df' dataframe.
Task - 4	Check for the null values in 'df' and find the data type of the columns.
Task - 5	Display the 'Random value 2' & 'Random value 3' columns with location method and index location method.

### CODE:

```
#This assignment is made by:Soumya singh 21bce3014
import pandas as pd

# Creating a dictionary containing data
data = {
    'Name': ['soumya', 'abhishek', 'nishit', 'isha', 'ashu'],
    'Age': [25, 30, 22, 28, 24],
    'Country': ['USA', 'Canada', 'UK', 'India', 'Australia'],
    'gender': ['f', 'm', 'f', 'm', 'f'],
    'year_of_birth': [2000, None, None, 2002, 2003]
}

# Creating a DataFrame
df = pd.DataFrame(data)

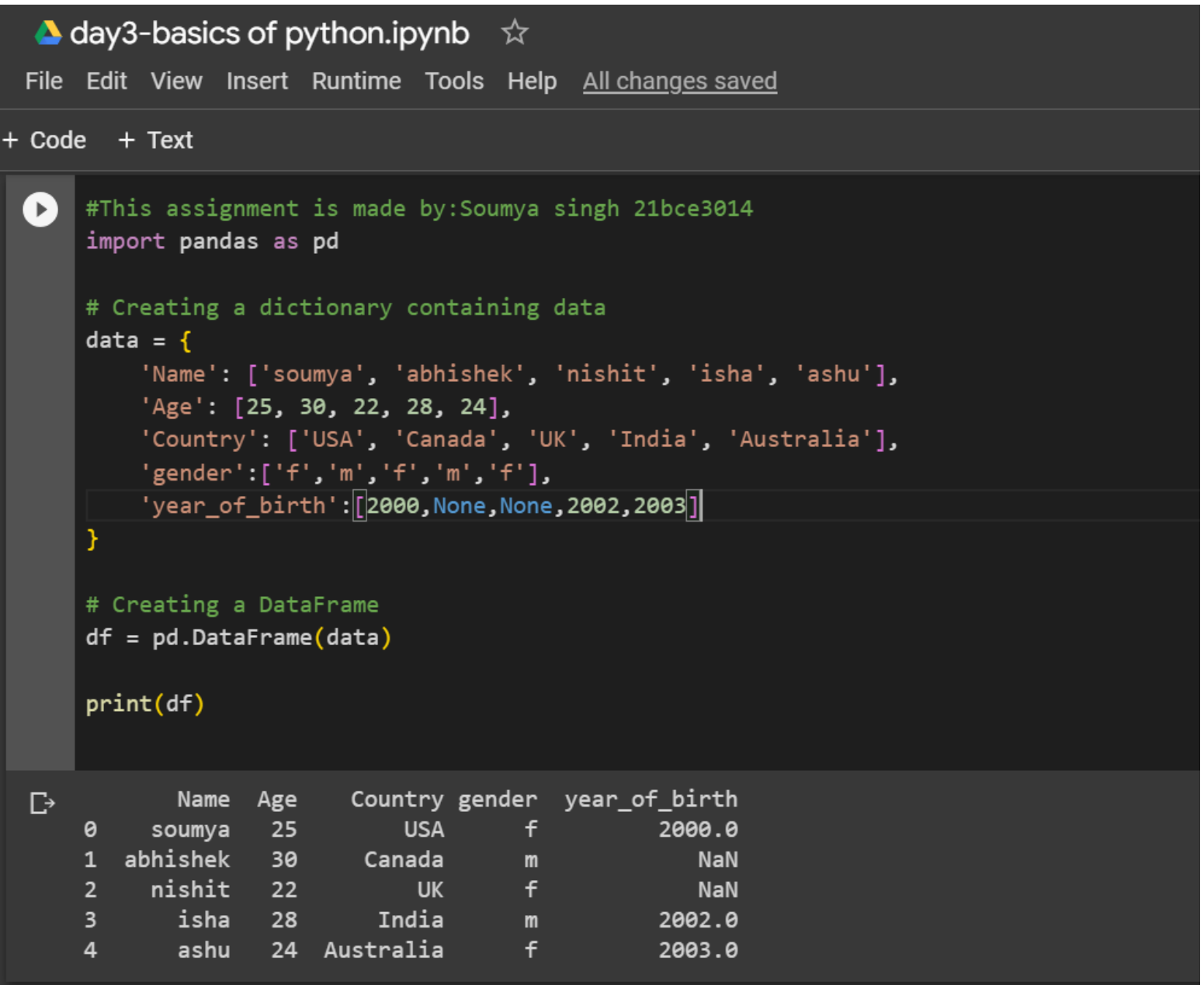
print(df)
df.info()
descriptive_stats = df.describe()

print(descriptive_stats)
observation_4 = df.loc[4]

print(observation_4)
null_counts = df.isnull().sum()

print(null_counts)
```

## OUTPUT SCREENSHOTS



The image shows a Jupyter Notebook interface with a dark theme. The top bar displays the file name 'day3-basics of python.ipynb' and a star icon. Below the top bar is a menu bar with 'File', 'Edit', 'View', 'Insert', 'Runtime', 'Tools', 'Help', and 'All changes saved'. The main area is divided into two sections: a code editor and an output area. The code editor contains the following Python code:

```
#This assignment is made by:Soumya singh 21bce3014
import pandas as pd

# Creating a dictionary containing data
data = {
    'Name': ['soumya', 'abhishek', 'nishit', 'isha', 'ashu'],
    'Age': [25, 30, 22, 28, 24],
    'Country': ['USA', 'Canada', 'UK', 'India', 'Australia'],
    'gender': ['f', 'm', 'f', 'm', 'f'],
    'year_of_birth': [2000, None, None, 2002, 2003]
}

# Creating a DataFrame
df = pd.DataFrame(data)

print(df)
```

The output area shows the resulting DataFrame as a table with 5 rows and 6 columns. The columns are 'Name', 'Age', 'Country', 'gender', and 'year\_of\_birth'. The rows are indexed from 0 to 4. The 'year\_of\_birth' column contains the values 2000.0, NaN, NaN, 2002.0, and 2003.0.

	Name	Age	Country	gender	year_of_birth
0	soumya	25	USA	f	2000.0
1	abhishek	30	Canada	m	NaN
2	nishit	22	UK	f	NaN
3	isha	28	India	m	2002.0
4	ashu	24	Australia	f	2003.0



+ Code + Text



df.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 5 entries, 0 to 4
Data columns (total 5 columns):
#   Column          Non-Null Count  Dtype
---  -
0   Name             5 non-null     object
1   Age              5 non-null     int64
2   Country          5 non-null     object
3   gender           5 non-null     object
4   year_of_birth    3 non-null     float64
dtypes: float64(1), int64(1), object(3)
memory usage: 328.0+ bytes
```



+ Code + Text



descriptive\_stats = df.describe()

print(descriptive\_stats)

```
count      Age  year_of_birth
mean    25.800000  2001.666667
std       3.193744    1.527525
min      22.000000  2000.000000
25%      24.000000  2001.000000
50%      25.000000  2002.000000
75%      28.000000  2002.500000
max      30.000000  2003.000000
```

+ Code + Text

✓  
0s

```
▶ observation_4 = df.loc[4]  
  
print(observation_4)
```

```
↳ Name          ashu  
   Age           24  
   Country      Australia  
   gender        f  
   year_of_birth 2003.0  
   Name: 4, dtype: object
```

✓  
0s

```
[17] null_counts = df.isnull().sum()  
  
print(null_counts)
```

```
Name          0  
Age           0  
Country       0  
gender        0  
year_of_birth 2  
dtype: int64
```