

Question: Analyze a small dataset using Python, documenting the step-by-step process of the data analysis journey.

1. IMPORTING THE LIBRARY

```
1 # Importing a pandas library
import pandas as pd
```

UPLOADING THE DATASET

```
2 # uploading a dataset
Man = pd.read_excel('/content/sample_data/TEACHERS DATA.xls')
```

DISPLAYING THE SAMPLE OF THE DATASET

```
#printing the sample of the dataset
Man.head()
```

	Registration No	Category	Surname	Firstname	Middlename	Date of Birth	Sex	Phor Number
0	NaN	D	MUHAMMED	ATIKU	HUSSAINI	1986-11-12 00:00:00	M	NaN
1	NaN	D	SHAFI'U	ABDULLAHI	NaN	1979-01-24 00:00:00	M	NaN
2	NaN	D	MUSTAPHA	MUHAMMAD	DAUDA	1987-05-14 00:00:00	M	NaN
3	NaN	D	ABUBAKAR	HARUNA	NaN	1985-06-07 00:00:00	M	NaN
4	NaN	D	MUHAMMAD	USMAN	ABDULLAHI	1978-01-24 00:00:00	NaN	NaN

5 rows × 25 columns

FIRST THREE ROWS OF THE DATASET

```
# printing the 3 rows of the dataset
Man.head(3)
```

	Registration No	Category	Surname	Firstname	Middlename	Date of Birth	Sex	Phon Numbe
0	NaN	D	MUHAMMED	ATIKU	HUSSAINI	1986-11-12 00:00:00	M	Na
1	NaN	D	SHAFI'U	ABDULLAHI	NaN	1979-01-24 00:00:00	M	Na
2	NaN	D	MUSTAPHA	MUHAMMAD	DAUDA	1987-05-14 00:00:00	M	Na

3 rows \times 25 columns

```
# Display the last 5 rows of the dataset
Man.tail()
```

	Registration No	Category	Surname	Firstname	Middlename	Date of Birth	Sex	Phone Number
63230	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
63231	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
63232	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
63233	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
63234	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

5 rows x 25 columns

INFORMATION OF THE DATASET

```
#Display the information of the dataset
Man.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 63235 entries, 0 to 63234
Data columns (total 25 columns):
 #   Column              Non-Null Count  Dtype
  ...
```

```

---
0  Registration No      1 non-null    object
1  Category            121 non-null  object
2  Surname             121 non-null  object
3  Firstname           121 non-null  object
4  Middlename          40 non-null   object
5  Date of Birth       119 non-null  object
6  Sex                 119 non-null  object
7  Phone Number        0 non-null    float64
8  eMail               0 non-null    float64
9  Nationality         121 non-null  object
10 State of Origin      121 non-null  object
11 LGA of Origin       121 non-null  object
12 Current Employer    121 non-null  object
13 Office Address      121 non-null  object
14 Area of Specialization 121 non-null  object
15 Form No             0 non-null    float64
16 Institutions Attended 121 non-null  object
17 Work Experience      0 non-null    float64
18 Qualification wt Dates 121 non-null  object
19 Bank Name           121 non-null  object
20 Reg Amount Paid     121 non-null  float64
21 Unnamed: 21         121 non-null  object
22 Date Paid           121 non-null  datetime64[ns]
23 Registration Date   0 non-null    float64
24 License Amt Paid    1 non-null    object
dtypes: datetime64[ns](1), float64(6), object(18)
memory usage: 12.1+ MB

```

DESCRIPTION OF THE DATASET

```

# Display the description of the dataset
Man.describe()

```

	Phone Number	eMail	Form No	Work Experience	Reg Amount Paid	Registration Date
count	0.0	0.0	0.0	0.0	121.000000	0.0
mean	NaN	NaN	NaN	NaN	3495.867769	NaN
std	NaN	NaN	NaN	NaN	1118.957602	NaN
min	NaN	NaN	NaN	NaN	3000.000000	NaN
25%	NaN	NaN	NaN	NaN	3000.000000	NaN
50%	NaN	NaN	NaN	NaN	3000.000000	NaN
75%	NaN	NaN	NaN	NaN	3000.000000	NaN
max	NaN	NaN	NaN	NaN	6000.000000	NaN

