Analyse a small dataset using python,documenting the step by step process of the data anlyses journey.

importing a pandas library

```
import pandas as pd
```

uplaoding the dataset

```
user = pd.read_excel("/content/sample_data/sample data.xlsx")
```

printing the sample of the dataset

```
user.head()
```

	SERIAL/NO	NAME	GENDER	OCUPATION	COUNTRY
0	1	ABUBAKAR	MALE	LAWYER	NIGERIA
1	2	AMINA	FEMALE	TEACHER	NIGER
2	3	HASANA	FEMALE	NURSE	ALGERIA
3	4	BELLO	MALE	TAILOR	SUDAN
4	5	MUSA	MALE	DOCTOR	NIGERIA

printing thw 3 rows of the dataset

```
user.head(3)
```

	SERIAL/NO	NAME	GENDER	OCUPATION	COUNTRY
0	1	ABUBAKAR	MALE	LAWYER	NIGERIA
1	2	AMINA	FEMALE	TEACHER	NIGER
2	3	HASANA	FEMALE	NURSE	ALGERIA

display the last 5 rows of the dataset

```
user.tail()
```

	SERIAL/NO	NAME	GENDER	OCUPATION	COUNTRY
4	5	MUSA	MALE	DOCTOR	NIGERIA
5	6	SALISU	MALE	ENGINEER	GHANA
6	7	IBRAHIM	MALE	PILOT	NIGERIA
7	8	NAFISA	FEMALE	JOURNALIST	NIGERIA
8	9	YAHAYA	MALE	DOCTOR	NIGER

display the information of the dataset

```
user.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 9 entries, 0 to 8
Data columns (total 5 columns):
```

```
 #   Column     Non-Null Count  Dtype
---  ------     --------------  -----
 0   SERIAL/NO  9 non-null      int64
 1   NAME       9 non-null      object
 2   GENDER     9 non-null      object
 3   OCUPATION  9 non-null      object
 4   COUNTRY    9 non-null      object
dtypes: int64(1), object(4)
memory usage: 488.0+ bytes
```

displaying the discription of the dataset

```
user.describe()
```

	SERIAL/NO
count	9.000000
mean	5.000000
std	2.738613
min	1.000000
25%	3.000000
50%	5.000000
75%	7.000000
max	9.000000

Double-click (or enter) to edit