

Question: Analyze a Small Dataset Using python, Documenting the Step by step procss of the Data analysis journey.

1. Importing the Library

```
# importing a pandas library
import pandas as pd
```

2. Uploading the Dataset

```
# Uploading a Dataset
user = pd.read_csv("/content/sample_data/UserData.csv")
```

3. Diplaying the Sample of Dataset

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

	PreferredSponsors	Gender	Country	Degree	Sign Up Date	city	zip
0	["GlobalShala","Grant Thornton China","Saint L...	Male	Nigeria	Undergraduate Student	2023-07-23T08:05:58.602Z	Owerri	460103
1	["GlobalShala","Grant Thornton China","Saint L...	Male	India	Undergraduate Student	2023-04-24T09:57:07.405Z	kottayam	686501
2	["GlobalShala","Illinois Institute of Technology"]	Male	United States	High School Student	2022-10-22T14:06:56.114Z	Austin	78701

```
# Printing thw 3 rows of the dataset
user.head(3)
```

	PreferredSponsors	Gender	Country	Degree	Sign Up Date	city	zip
0	["GlobalShala","Grant Thornton China","Saint L...	Male	Nigeria	Undergraduate Student	2023-07-23T08:05:58.602Z	Owerri	460103
1	["GlobalShala","Grant Thornton China","Saint L...	Male	India	Undergraduate Student	2023-04-24T09:57:07.405Z	kottayam	686501
2	["GlobalShala","Illinois Institute of Technology"]	Male	United States	High School Student	2022-10-22T14:06:56.114Z	Austin	78701


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

	PreferredSponsors	Gender	Country	Degree	Sign Up Date	city	zip
27557	["GlobalShala","Grant Thornton China","Saint L...	Female	Botswana	Undergraduate Student	2023-04-08T05:30:44.705Z	Gaborone	2400
27558	["GlobalShala","Saint Louis University","Illin...	Male	United States	Undergraduate Student	2023-02-01T20:46:32.637Z	Coppell	75019
27559	["GlobalShala","Illinois Institute of Technolo...	Male	United States	High School Student	2022-09-22T14:06:56.114Z	Austin	78701
27560	["GlobalShala","Grant Thornton China","Saint L...	Male	Pakistan	NaN	2023-06-16T04:18:38.811Z	Daraban kalan	75019

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

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 27562 entries, 0 to 27561
Data columns (total 8 columns):
#   Column                Non-Null Count  Dtype
---  -
0   PreferredSponsors    27562 non-null  object
1   Gender                18027 non-null  object
2   Country               27500 non-null  object
3   Degree                16750 non-null  object
4   Sign Up Date          27562 non-null  object
5   city                  18029 non-null  object
6   zip                   18028 non-null  object
7   isFromSocialMedia     27553 non-null  object
dtypes: object(8)
memory usage: 1.7+ MB
```

```
# Diplaying the description of the dataset
user.describe()
```



	PreferredSponsors	Gender	Country	Degree	Sign Up Date	city	zip	isFromSocialMedia
count	27562	18027	27500	16750	27562	18029	18028	27553
unique	94	4	169	4	27561	4728	7454	2
top	["GlobalShala","Grant Thornton China","Saint L...	Male	India	Undergraduate Student	2022-10-30T17:25:54.072Z	Hyderabad	63108	True
freq	22011	11027	11893	6527	2	743	629	13811

