

Hello, welcome to another lesson on python and the data science libraries (Numpy, Pandas and Matplotlib). This is another dummy dataset that i worked on as i learned and grow on this career path. I want you to connect deeply with these dataset and feel encouraged about what you're doing. We all do it a little poorly until we get better. That's why i share my journey and my growth process.

So, Let's get started !!!

The first step is to import numpy and pandas libraries and load in the dataset. I also read the first 10 lines of the imported data as well as the last 10 using the .head() and .tail() functions respectively. This is for me to have an idea of wha my Dataset looks like.

```
In [1]: import numpy as np
import pandas as pd
filez = pd.read_excel("Excursion Portfolio.xls")
filez.head(10)
```

Out[1]:

	Unnamed: 0	Unnamed: 1	Unnamed: 2	Unnamed: 3	Unnamed: 4	Unnamed: 5	Unnamed: 6	Unnamed: 7	Unnamed: 8
0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
1	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
3	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
4	NaN	NaN	NaN	NaN	S/N	Name	Status	Amount Paid	T-shirt
5	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	Type
6	NaN	NaN	NaN	NaN	1	Victor T. NaAllah	Member	12000	Lo Sleeved
7	NaN	NaN	NaN	NaN	2	Olajide Mattew	Member	12000	Lo Sleeved
8	NaN	NaN	NaN	NaN	3	Abel Modu Timothy	Member	12000	Lo Sleeved
9	NaN	NaN	NaN	NaN	4	Egwim Jones Udojuaku	Member	12000	Lo Sleeved

```
In [2]: filez.tail(10)
```

Out[2]:

	Unnamed: 0	Unnamed: 1	Unnamed: 2	Unnamed: 3	Unnamed: 4	Unnamed: 5	Unnamed: 6	Unnamed: 7	Unnamed: 8
48	NaN	NaN	NaN	NaN	43	Shekinah Ajobola	Member	12000	12000
49	NaN	NaN	NaN	NaN	44	Adanu David	Member	9000	9000
50	NaN	NaN	NaN	NaN	45	Oche Muscle	Member	9500	9500
51	NaN	NaN	NaN	NaN	46	Ocheje Jeremiah	Member	12000	12000
52	NaN	NaN	NaN	NaN	47	Oguntowo Basit Ifedolapo	Member	12000	12000
53	NaN	NaN	NaN	NaN	48	Simeon Iganga	Member	12000	12000
54	NaN	NaN	NaN	NaN	49	Caleb Onuoja Aaron	Ex-Exco	8000	8000
55	NaN	NaN	NaN	NaN	TOTAL	NaN	NaN	537500	537500
56	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	NaN	NaN	Total Members Registered	NaN	NaN	49	49

I also checked other properties of the dataset which includes its shape size dimension and the data types respectively

```
In [3]: filez.shape
```

Out[3]: (58, 11)

```
In [4]: filez.size
```

Out[4]: 638

```
In [5]: filez.ndim
```

Out[5]: 2

```
In [6]: filez.dtypes
```

```
Out[6]: Unnamed: 0    float64  
        Unnamed: 1    float64  
        Unnamed: 2    float64  
        Unnamed: 3    float64  
        Unnamed: 4     object  
        Unnamed: 5     object  
        Unnamed: 6     object  
        Unnamed: 7     object  
        Unnamed: 8     object  
        Unnamed: 9     object  
        Unnamed: 10    object  
        dtype: object
```

Next, i renamed the columns of the dataset to make it more relatable. And after that i deleted the columns that contains irrelevant information (the columns that contained "nan" in all its rows)

```
In [7]: filez.columns = ['COL1', 'COL2', 'COL3', 'COL4', 'S/N', 'NAME', 'STATUS', 'AMOUNT_PAID', 'TSHIRT_TYPE', 'TSHIRT_AMOUNT', 'COMMENTS']
```

```
In [8]: filez.head(15)
```

Out[8]:

	COL1	COL2	COL3	COL4	S/N	NAME	STATUS	AMOUNT_PAID	TSHIRT_TYPE	TSHIRT_
0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
1	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
2	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
3	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
4	NaN	NaN	NaN	NaN	S/N	Name	Status	Amount Paid	T-shirt	
5	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	Type	
6	NaN	NaN	NaN	NaN	1	Victor T. Na'Allah	Member	12000	Long Sleeve	
7	NaN	NaN	NaN	NaN	2	Olajide Mattew	Member	12000	Long Sleeve	
8	NaN	NaN	NaN	NaN	3	Abel Modu Timothy	Member	12000	Long Sleeve	
9	NaN	NaN	NaN	NaN	4	Egwim Jones Udojuaku	Member	12000	Long Sleeve	
10	NaN	NaN	NaN	NaN	5	Nwachukwu Emmanuel Benedict	Member	12000	Long Sleeve	
11	NaN	NaN	NaN	NaN	6	Adole John A.	Member	12000	Short Sleeve	
12	NaN	NaN	NaN	NaN	7	Faleti Ayodeji Peter	Member	12000	Short Sleeve	
13	NaN	NaN	NaN	NaN	8	Ayantoye Ridwan Ayomide	Member	12000	NaN	
14	NaN	NaN	NaN	NaN	9	Marvellous T. Isaac	Member	12000	Short Sleeve	

```
In [9]: filez1 = filez.drop(['COL1', 'COL2', 'COL3', 'COL4', 'S/N', 'COMMENTS'], axis=1)
```

```
In [10]: filez1.head(15)
```

```
Out[10]:
```

	NAME	STATUS	AMOUNT_PAID	TSHIRT_TYPE	TSHIRT_AMOUNT
0	NaN	NaN	NaN	NaN	NaN
1	NaN	NaN	NaN	NaN	NaN
2	NaN	NaN	NaN	NaN	NaN
3	NaN	NaN	NaN	NaN	NaN
4	Name	Status	Amount Paid	T-shirt	NaN
5	NaN	NaN	NaN	Type	Amount
6	Victor T. Na'Allah	Member	12000	Long Sleeve	2500
7	Olajide Matthew	Member	12000	Long Sleeve	2500
8	Abel Modu Timothy	Member	12000	Long Sleeve	2500
9	Egwim Jones Udojuaku	Member	12000	Long Sleeve	2500
10	Nwachukwu Emmanuel Benedict	Member	12000	Long Sleeve	2500
11	Adole John A.	Member	12000	Short Sleeve	2100
12	Faleti Ayodeji Peter	Member	12000	Short Sleeve	2100
13	Ayantoye Ridwan Ayomide	Member	12000	NaN	NaN
14	Marvellous T. Isaac	Member	12000	Short Sleeve	2100

```
In [11]: filez1.tail(10)
```

```
Out[11]:
```

	NAME	STATUS	AMOUNT_PAID	TSHIRT_TYPE	TSHIRT_AMOUNT
48	Shekinah Ajibola	Member	12000	Long Sleeve	2500
49	Adanu David	Member	9000	NIL	NIL
50	Oche Muscle	Member	9500	NIL	NIL
51	Ocheje Jeremiah	Member	12000	Short Sleeve	2100
52	Oguntowo Basit Ifedolapo	Member	12000	NaN	NaN
53	Simeon Iganga	Member	12000	NaN	NaN
54	Caleb Onuoja Aaron	Ex-Exco	8000	NaN	NaN
55	NaN	NaN	537500	NaN	80700
56	NaN	NaN	NaN	NaN	NaN
57	NaN	NaN	49	NaN	NaN

Next , i dropped all the rows which contains "nan" all across the dataset or irrelevant data

```
In [12]: filez2 = filez1.drop([ 0,1,2,3,4,5,55,56,57])
```

```
In [13]: filez2.head(10)
```

```
Out[13]:
```

	NAME	STATUS	AMOUNT_PAID	TSHIRT_TYPE	TSHIRT_AMOUNT
6	Victor T. Na'Allah	Member	12000	Long Sleeve	2500
7	Olajide Mattew	Member	12000	Long Sleeve	2500
8	Abel Modu Timothy	Member	12000	Long Sleeve	2500
9	Egwim Jones Udojuaku	Member	12000	Long Sleeve	2500
10	Nwachukwu Emmanuel Benedict	Member	12000	Long Sleeve	2500
11	Adole John A.	Member	12000	Short Sleeve	2100
12	Faleti Ayodeji Peter	Member	12000	Short Sleeve	2100
13	Ayantoye Ridwan Ayomide	Member	12000	NaN	NaN
14	Marvellous T. Isaac	Member	12000	Short Sleeve	2100
15	Shenge Raphael Saarshatar	Member	12000	Short Sleeve	2100

```
In [14]: filez2.tail(10)
```

```
Out[14]:
```

	NAME	STATUS	AMOUNT_PAID	TSHIRT_TYPE	TSHIRT_AMOUNT
45	Daleng Elisha Nandi	Member	9000	NaN	NaN
46	Ukande Aondongu Cephas	Member	12000	Long Sleeve	2500
47	De-Gold David Tarki	Member	12000	NaN	NaN
48	Shekinah Ajibola	Member	12000	Long Sleeve	2500
49	Adanu David	Member	9000	NIL	NIL
50	Oche Muscle	Member	9500	NIL	NIL
51	Ocheje Jeremiah	Member	12000	Short Sleeve	2100
52	Oguntowo Basit Ifedolapo	Member	12000	NaN	NaN
53	Simeon Iganga	Member	12000	NaN	NaN
54	Caleb Onuoja Aaron	Ex-Exco	8000	NaN	NaN

Next, I made the "NAME" column the Index label. And i dropped the column after this step

Note that this step of setting a column as an index can also be achieved by using the "set_index()" method will you'll learn in later lessons

```
In [15]: filez2.index = filez2["NAME"]
```

```
In [16]: filez2.head()
```

Out[16]:

		NAME	STATUS	AMOUNT_PAID	TSHIRT_TYPE	TSHIRT_AMOUNT
		NAME				
	Victor T. Na'Allah	Victor T. Na'Allah	Member	12000	Long Sleeve	2500
	Olajide Mattew	Olajide Mattew	Member	12000	Long Sleeve	2500
	Abel Modu Timothy	Abel Modu Timothy	Member	12000	Long Sleeve	2500
	Egwim Jones Udojuaku	Egwim Jones Udojuaku	Member	12000	Long Sleeve	2500
	Nwachukwu Emmanuel Benedict	Nwachukwu Emmanuel Benedict	Member	12000	Long Sleeve	2500

```
In [17]: filez3 = filez2.drop("NAME", axis = "columns")
```

```
In [18]: filez3.head()
```

Out[18]:

		STATUS	AMOUNT_PAID	TSHIRT_TYPE	TSHIRT_AMOUNT
		NAME			
	Victor T. Na'Allah	Member	12000	Long Sleeve	2500
	Olajide Mattew	Member	12000	Long Sleeve	2500
	Abel Modu Timothy	Member	12000	Long Sleeve	2500
	Egwim Jones Udojuaku	Member	12000	Long Sleeve	2500
	Nwachukwu Emmanuel Benedict	Member	12000	Long Sleeve	2500

```
In [19]: filez3.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Index: 49 entries, Victor T. Na'Allah to Caleb Onuoja Aaron
Data columns (total 4 columns):
STATUS          49 non-null object
AMOUNT_PAID     48 non-null object
TSHIRT_TYPE     39 non-null object
TSHIRT_AMOUNT   39 non-null object
dtypes: object(4)
memory usage: 980.0+ bytes
```

Next i checked if there are null values on the Dataset and where excatly they are located

```
In [20]: filez3.isnull().sum()
```

```
Out[20]: STATUS          0
AMOUNT_PAID            1
TSHIRT_TYPE           10
TSHIRT_AMOUNT         10
dtype: int64
```

```
In [21]: filez3["AMOUNT_PAID"].head(10)
```

```
Out[21]: NAME
Victor T. Na'Allah          12000
Olajide Matthew             12000
Abel Modu Timothy           12000
Egwim Jones Udojuaku        12000
Nwachukwu Emmanuel Benedict 12000
Adole John A.               12000
Faleti Ayodeji Peter        12000
Ayantoye Ridwan Ayomide     12000
Marvellous T. Isaac         12000
Shenge Raphael Saarshatar   12000
Name: AMOUNT_PAID, dtype: object
```

I filled the null values in the "AMOUNT PAID" column with 0, and set inplace to be True. And i checked to confirm if the change has been effected

```
In [22]: filez3["AMOUNT_PAID"].fillna(0, inplace = True)
```

```
In [23]: filez3.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Index: 49 entries, Victor T. Na'Allah to Caleb Onuoja Aaron
Data columns (total 4 columns):
STATUS          49 non-null object
AMOUNT_PAID     49 non-null int64
TSHIRT_TYPE     39 non-null object
TSHIRT_AMOUNT   39 non-null object
dtypes: int64(1), object(3)
memory usage: 1.1+ KB
```

```
In [24]: filez3["TSHIRT_TYPE"].head()
```

```
Out[24]: NAME
Victor T. Na'Allah          Long Sleeve
Olajide Matthew             Long Sleeve
Abel Modu Timothy           Long Sleeve
Egwim Jones Udojuaku        Long Sleeve
Nwachukwu Emmanuel Benedict Long Sleeve
Name: TSHIRT_TYPE, dtype: object
```


If filled the null values contained in the "T-SHIRT TYPE" column with Long Sleeve and set inplace to True.

I replaced the null values contained in "T-SHIRT AMOUNT" accordingly

I also checked to confirm the change has been effected

```
In [25]: filez3["TSHIRT_TYPE"].fillna("Long Sleeve", inplace = True)
```

```
In [26]: filez3.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Index: 49 entries, Victor T. Na'Allah to Caleb Onuoja Aaron
Data columns (total 4 columns):
STATUS          49 non-null object
AMOUNT_PAID     49 non-null int64
TSHIRT_TYPE     49 non-null object
TSHIRT_AMOUNT   39 non-null object
dtypes: int64(1), object(3)
memory usage: 1.1+ KB
```

```
In [27]: filez3["TSHIRT_AMOUNT"].head()
```

```
Out[27]: NAME
Victor T. Na'Allah          2500
Olajide Mattew             2500
Abel Modu Timothy          2500
Egwim Jones Udojuaku       2500
Nwachukwu Emmanuel Benedict 2500
Name: TSHIRT_AMOUNT, dtype: object
```

```
In [28]: filez3["TSHIRT_AMOUNT"].fillna(2500, inplace = True)
```

```
In [29]: filez3.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Index: 49 entries, Victor T. Na'Allah to Caleb Onuoja Aaron
Data columns (total 4 columns):
STATUS          49 non-null object
AMOUNT_PAID     49 non-null int64
TSHIRT_TYPE     49 non-null object
TSHIRT_AMOUNT   49 non-null object
dtypes: int64(1), object(3)
memory usage: 1.1+ KB
```

In [30]: filez3

Out[30]:

	STATUS	AMOUNT_PAID	TSHIRT_TYPE	TSHIRT_AMOUNT
NAME				
Victor T. Na'Allah	Member	12000	Long Sleeve	2500
Olajide Matthew	Member	12000	Long Sleeve	2500
Abel Modu Timothy	Member	12000	Long Sleeve	2500
Egwim Jones Udojuaku	Member	12000	Long Sleeve	2500
Nwachukwu Emmanuel Benedict	Member	12000	Long Sleeve	2500
Adole John A.	Member	12000	Short Sleeve	2100
Faleti Ayodeji Peter	Member	12000	Short Sleeve	2100
Ayantoye Ridwan Ayomide	Member	12000	Long Sleeve	2500
Marvellous T. Isaac	Member	12000	Short Sleeve	2100
Shenge Raphael Saarshatar	Member	12000	Short Sleeve	2100
Reuben O.Enoch	Member	12000	Long Sleeve	2500
Mercy Ajayi	Member	12000	Long Sleeve	2500
Michael Kpoco	Ex-Exco	7000	Short Sleeve	2100
Saad	Member	12000	Long Sleeve	2500
Terzungwe Caleb	Ex-Exco	5000	Short Sleeve	2100
Faith	Member	12000	Long Sleeve	2500
Gonet Zion	Member	12000	Long Sleeve	2500
Hawwau Adeboyin Adeyemo\nFor => Jubrin Omeiza	Member	12000	Long Sleeve	2500
Nicholas Otonoku	Member	12000	Short Sleeve	2100
Timothy Ignitus Agbor	Member	12000	Long Sleeve	2500
Madumche Chidibere	Member	12000	Long Sleeve	2500
Nwokocha Ethelbert	Ex-Exco	8000	Short Sleeve	2100
Effiong Ubon Alasi	Member	12000	Long Sleeve	2500
Daniel Overcomer	Member	12000	Short Sleeve	2100
Aguwa Wisdom	Member	12000	Long Sleeve	2500
Ganiyu Mujeeb	Member	12000	Long Sleeve	2500
Hezekiel Joel	Member	12000	Long Sleeve	2500
Shitu Mustapha Ibrahim	Member	12000	Long Sleeve	2500
Orogu Francis Israel	Member	12000	Short Sleeve	2100
Agha Elizabeth	Member	12000	Short Sleeve	2100
Muhammed Zainab	Member	12000	Long Sleeve	2500
Isaac Priscilla	Member	12000	Short Sleeve	2100
Paul Elizabeth Ladi	Member	12000	Short Sleeve	2100
Omaji Samuel Owoicho	Ex-Exco	8000	Short Sleeve	2100

	STATUS	AMOUNT_PAID	TSHIRT_TYPE	TSHIRT_AMOUNT
NAME				
Esinome Abraham	Member	12000	Short Sleeve	2100
Umeh Audu Ayigba	Member	12000	Short Sleeve	2100
Sarah Kauna Edoja	Member	8000	Long Sleeve	2500
Ibrahim Hussein Chado	Member	10000	NIL	NIL
Fatima Ganiyu	Member	0	NIL	NIL
Daleng Elisha Nandi	Member	9000	Long Sleeve	2500
Ukande Aondongu Cephas	Member	12000	Long Sleeve	2500
De-Gold David Tarki	Member	12000	Long Sleeve	2500
Shekinah Ajibola	Member	12000	Long Sleeve	2500
Adanu David	Member	9000	NIL	NIL
Oche Muscle	Member	9500	NIL	NIL
Ocheje Jeremiah	Member	12000	Short Sleeve	2100
Oguntowo Basit Ifedolapo	Member	12000	Long Sleeve	2500
Simeon Iganga	Member	12000	Long Sleeve	2500
Caleb Onuoja Aaron	Ex-Exco	8000	Long Sleeve	2500

```
In [31]: filez3["TSHIRT_TYPE"].unique()
```

```
Out[31]: array(['Long Sleeve', 'Short Sleeve', 'NIL'], dtype=object)
```

```
In [32]: filez3["TSHIRT_AMOUNT"].unique()
```

```
Out[32]: array([2500, 2100, 'NIL'], dtype=object)
```

Checking the info of the Dataset, it's revealed that there are some values that are wrong. Like "NIL", and they are replaced accordingly using the replace function.

Note that this not a null value and requires a different treatment from "nan". If it were to be null values, "dropna" / "fillna" wolud have taken care of it

```
In [33]: filez3["TSHIRT_TYPE"].replace({"NIL" : "No Sleeve"}, inplace = True)
```

```
In [34]: filez3["TSHIRT_AMOUNT"].replace({"NIL" : 0}, inplace = True)
```

In [35]: filez3

Out[35]:

	STATUS	AMOUNT_PAID	TSHIRT_TYPE	TSHIRT_AMOUNT
NAME				
Victor T. Na'Allah	Member	12000	Long Sleeve	2500
Olajide Matthew	Member	12000	Long Sleeve	2500
Abel Modu Timothy	Member	12000	Long Sleeve	2500
Egwim Jones Udojuaku	Member	12000	Long Sleeve	2500
Nwachukwu Emmanuel Benedict	Member	12000	Long Sleeve	2500
Adole John A.	Member	12000	Short Sleeve	2100
Faleti Ayodeji Peter	Member	12000	Short Sleeve	2100
Ayantoye Ridwan Ayomide	Member	12000	Long Sleeve	2500
Marvellous T. Isaac	Member	12000	Short Sleeve	2100
Shenge Raphael Saarshatar	Member	12000	Short Sleeve	2100
Reuben O.Enoch	Member	12000	Long Sleeve	2500
Mercy Ajayi	Member	12000	Long Sleeve	2500
Michael Kpoco	Ex-Exco	7000	Short Sleeve	2100
Saad	Member	12000	Long Sleeve	2500
Terzungwe Caleb	Ex-Exco	5000	Short Sleeve	2100
Faith	Member	12000	Long Sleeve	2500
Gonet Zion	Member	12000	Long Sleeve	2500
Hawwau Adeboyin Adeyemo\nFor => Jubrin Omeiza	Member	12000	Long Sleeve	2500
Nicholas Otonoku	Member	12000	Short Sleeve	2100
Timothy Ignitus Agbor	Member	12000	Long Sleeve	2500
Madumche Chidibere	Member	12000	Long Sleeve	2500
Nwokocha Ethelbert	Ex-Exco	8000	Short Sleeve	2100
Effiong Ubon Alasi	Member	12000	Long Sleeve	2500
Daniel Overcomer	Member	12000	Short Sleeve	2100
Aguwa Wisdom	Member	12000	Long Sleeve	2500
Ganiyu Mujeeb	Member	12000	Long Sleeve	2500
Hezekiel Joel	Member	12000	Long Sleeve	2500
Shitu Mustapha Ibrahim	Member	12000	Long Sleeve	2500
Orogu Francis Israel	Member	12000	Short Sleeve	2100
Agha Elizabeth	Member	12000	Short Sleeve	2100
Muhammed Zainab	Member	12000	Long Sleeve	2500
Isaac Priscilla	Member	12000	Short Sleeve	2100
Paul Elizabeth Ladi	Member	12000	Short Sleeve	2100
Omaji Samuel Owoicho	Ex-Exco	8000	Short Sleeve	2100

	STATUS	AMOUNT_PAID	TSHIRT_TYPE	TSHIRT_AMOUNT
NAME				
Esinome Abraham	Member	12000	Short Sleeve	2100
Umeh Audu Ayigba	Member	12000	Short Sleeve	2100
Sarah Kauna Edoja	Member	8000	Long Sleeve	2500
Ibrahim Hussein Chado	Member	10000	No Sleeve	0
Fatima Ganiyu	Member	0	No Sleeve	0
Daleng Elisha Nandi	Member	9000	Long Sleeve	2500
Ukande Aondongu Cephas	Member	12000	Long Sleeve	2500
De-Gold David Tarki	Member	12000	Long Sleeve	2500
Shekinah Ajibola	Member	12000	Long Sleeve	2500
Adanu David	Member	9000	No Sleeve	0
Oche Muscle	Member	9500	No Sleeve	0
Ocheje Jeremiah	Member	12000	Short Sleeve	2100
Oguntowo Basit Ifedolapo	Member	12000	Long Sleeve	2500
Simeon Iganga	Member	12000	Long Sleeve	2500
Caleb Onuoja Aaron	Ex-Exco	8000	Long Sleeve	2500

The file is ready for analysis, visualization or any other process.

I checked a few of its properties to double check all I've done.

In [36]: `filez3.info()`

```
<class 'pandas.core.frame.DataFrame'>
Index: 49 entries, Victor T. Na'Allah to Caleb Onuoja Aaron
Data columns (total 4 columns):
STATUS          49 non-null object
AMOUNT_PAID     49 non-null int64
TSHIRT_TYPE     49 non-null object
TSHIRT_AMOUNT   49 non-null int64
dtypes: int64(2), object(2)
memory usage: 1.3+ KB
```

In [37]: filez3.describe()

Out[37]:

	AMOUNT_PAID	TSHIRT_AMOUNT
count	49.000000	49.000000
mean	10969.387755	2157.142857
std	2319.321388	676.387463
min	0.000000	0.000000
25%	12000.000000	2100.000000
50%	12000.000000	2500.000000
75%	12000.000000	2500.000000
max	12000.000000	2500.000000

In [38]: filez3.shape

Out[38]: (49, 4)

In [39]: filez3.size

Out[39]: 196

In [40]: filez3.dtypes

Out[40]: STATUS object
AMOUNT_PAID int64
TSHIRT_TYPE object
TSHIRT_AMOUNT int64
dtype: object

In [41]: filez3.ndim

Out[41]: 2

Now, you can write the file back into a document. In this case, i choose to write it back to an excel file

In [42]: filez3.to_excel("ExcurPort Document.xlsx")

In []:

This document has been successfully prepared for analysis. And I'm really excited 😄😄😄 about that. It's also been successfully written back to an excel file. And that's another giant feat achieved. I'll keep the document safe for visualization analysis when I'm done and ready with Matplotlib. I thank God for this that I've been able to achieve.

ThankfulHeart.

GratefulHeart

JourneyToGreatAndExtraordinaryFeats.

TheBecomingOfAWorldReveredDataScientist

TheMakingOfADataAnalyst

This brings this lesson to an end. The best way to learn is by doing, so get started and get practicing. It's only a matter of improved effort and you'd have developed your skills to take on giant projects.

Wishing you Goodspeed and Happy Learning

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []: