

▼ importing pandas libraries

```
# importing pandas
import pandas as pd
```

▼ importing dataset

```
#uploading dataset
bb = pd.read_excel("/content/bb week5.xlsx")
```

▼ displaying dataset

```
#viewing dataset
bb
```

	Name	Address	Age	Gender	
0	Hussein Hakeem	Number 22 Fioye Crescent Surulere Lagos	17	Male	
1	Arojoye Samuel	11 Omolade Close Omole Estate Lagos	16	Male	
2	Alex Ezurum @	1 Adamu Lane, Abuja	14	Male	
3	Susan Nwaimo/	Number 58 Yaba Street, Kaduna State	16	Female	
4	Ajao Opeyemi	No12 Olubunmi Street, Abeokuta	18	Female	
5	Banjoko Adebisola	34 Ngige Street, Ugheli, Delta	14	Female	
6	/Muhammed Olabisi	13, ICAN road, Enugu	12	Female	
7	Oluwagbemi Mojisola	ACCA Lane, Onitsha	13	Female	
8	Hussein Hakeem	Number 22 Fioye Crescent Surulere Lagos	17	Male	
9	Arojoye Samuel	11 Omolade Close Omole Estate Lagos	16	Male	

Next steps:

[Generate code with bb](#)[View recommended plots](#)

▼ remove duplicates

```
#removing duplicates
bb = bb.drop_duplicates()
```

```
#displaying dataset to confirm changes
bb
```

	Name	Address	Age	Gender	
0	Hussein Hakeem	Number 22 Fioye Crescent Surulere Lagos	17	Male	
1	Arojoye Samuel	11 Omolade Close Omole Estate Lagos	16	Male	
2	Alex Ezurum @	1 Adamu Lane, Abuja	14	Male	
3	Susan Nwaimo/	Number 58 Yaba Street, Kaduna State	16	Female	
4	Ajao Opeyemi	No12 Olubunmi Street, Abeokuta	18	Female	
5	Banjoko Adebisola	34 Ngige Street, Ugheli, Delta	14	Female	
6	/Muhammed Olabisi	13, ICAN road, Enugu	12	Female	
7	Oluwagbemi Mojisola	ACCA Lane, Onitsha	13	Female	

Next steps:

[Generate code with bb](#)[View recommended plots](#)

removing the unwanted character

```
#remove unwanted character
```

```
bb["Name"] = bb["Name"].str.strip("/")
```

```
<ipython-input-19-d5fcf562fdcf>:2: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-c

```
bb["Name"] = bb["Name"].str.strip("/")
```

bb

	Name	Address	Age	Gender	
0	Hussein Hakeem	Number 22 Fioye Crescent Surulere Lagos	17	Male	
1	Arojoye Samuel	11 Omolade Close Omole Estate Lagos	16	Male	
2	Alex Ezurum	1 Adamu Lane, Abuja	14	Male	
3	Susan Nwaimo	Number 58 Yaba Street, Kaduna State	16	Female	
4	Ajao Opeyemi	No12 Olubunmi Street, Abeokuta	18	Female	
5	Banjoko Adebisola	34 Ngige Street, Ugheli, Delta	14	Female	
6	Muhammed Olabisi	13, ICAN road, Enugu	12	Female	
7	Oluwagbemi Mojisola	ACCA Lane, Onitsha	13	Female	

Next steps:

[Generate code with bb](#)
[View recommended plots](#)

split of address

```
#address splitting
```

```
bb[["Street_Address", "State"]]=bb["Address"].str.split(",",1, expand =True)
```

```
<ipython-input-38-7f62bc2985de>:2: FutureWarning: In a future version of pandas all arguments of StringMethods.split except for the arg
bb[["Street_Address", "State"]] =bb["Address"].str.split(",",1, expand =True)
<ipython-input-38-7f62bc2985de>:2: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-c

```
bb[["Street_Address", "State"]]=bb["Address"].str.split(",",1, expand =True)
```

bb

	Name	Address	Age	Gender	Street_Address	State	
0	Hussein Hakeem	Number 22 Fioye Crescent Surulere Lagos	17	Male	Number 22 Fioye Crescent Surulere Lagos	None	
1	Arojoye Samuel	11 Omolade Close Omole Estate Lagos	16	Male	11 Omolade Close Omole Estate Lagos	None	
2	Alex Ezurum	1 Adamu Lane, Abuja	14	Male	1 Adamu Lane	Abuja	
3	Susan Nwaimo	Number 58 Yaba Street, Kaduna State	16	Female	Number 58 Yaba Street	Kaduna State	
4	Ajao Opeyemi	No12 Olubunmi Street, Abeokuta	18	Female	No12 Olubunmi Street	Abeokuta	
5	Banjoko Adebisola	34 Ngige Street, Ugheli, Delta	14	Female	34 Ngige Street	Ugheli, Delta	
6	Muhammed Olabisi	13, ICAN road, Enugu	12	Female	13 ICAN road, Enugu		
7	Oluwagbemi Mojisola	ACCA Lane, Onitsha	13	Female	ACCA Lane	Onitsha	




Next steps:

[Generate code with bb](#)
[View recommended plots](#)

dropping in column

```
#dropping address columns
bb=bb.drop(columns = "Address")
```

bb

	Name	Age	Gender	Street_Address	State	
0	Hussein Hakeem	17	Male	Number 22 Fioye Crescent Surulere Lagos	None	
1	Arojoye Samuel	16	Male	11 Omolade Close Omole Estate Lagos	None	
2	Alex Ezurum	14	Male	1 Adamu Lane	Abuja	
3	Susan Nwaimo	16	Female	Number 58 Yaba Street	Kaduna State	
4	Ajao Opeyemi	18	Female	No12 Olubunmi Street	Abeokuta	
5	Banjoko Adebisola	14	Female	34 Ngige Street	Ugheli, Delta	
6	Muhammed Olabisi	12	Female	13 ICAN road, Enugu		
7	Oluwagbemi Mojisola	13	Female	ACCA Lane	Onitsha	

Next steps:




[Generate code with bb](#)

 [View recommended plots](#)

looping through columns


```
# loop columns
for b in bb.index:
    if bb.loc[b, "Street_Address"] == "Number 22 Fioye Crescent Surulere Lagos":
        bb.drop(b, inplace=True)
```

bb

	Name	Age	Gender	Street_Address	State	
1	Arojoye Samuel	16	Male	11 Omolade Close Omole Estate Lagos	None	
2	Alex Ezurum	14	Male	1 Adamu Lane	Abuja	
3	Susan Nwaimo	16	Female	Number 58 Yaba Street	Kaduna State	
4	Ajao Opeyemi	18	Female	No12 Olubunmi Street	Abeokuta	
5	Banjoko Adebisola	14	Female	34 Ngige Street	Ugheli, Delta	
6	Muhammed Olabisi	12	Female	13 ICAN road, Enugu		
7	Oluwagbemi Mojisola	13	Female	ACCA Lane	Onitsha	




Next steps:

[Generate code with bb](#)

 [View recommended plots](#)

```
# loop columns
for b in bb.index:
    if bb.loc[b, "Street_Address"] == "11 Omolade Close Omole Estate Lagos":
        bb.drop(b, inplace=True)
```

bb

	Name	Age	Gender	Street_Address	State	
2	Alex Ezurum	14	Male	1 Adamu Lane	Abuja	
3	Susan Nwaimo	16	Female	Number 58 Yaba Street	Kaduna State	
4	Ajao Opeyemi	18	Female	No12 Olubunmi Street	Abeokuta	
5	Banjoko Adebisola	14	Female	34 Ngige Street	Ugheli, Delta	
6	Muhammed Olabisi	12	Female	13 ICAN road, Enugu		
7	Oluwagbemi Mojisola	13	Female	ACCA Lane	Onitsha	

Next steps:

[Generate code with bb](#)

 [View recommended plots](#)

replacing element in a columnns

```
#replce element
```

```
bb["Gender"] = bb["Gender"].str.replace("Male", "M")
```

```
bb
```

	Name	Age	Gender	Street_Address	State
2	Alex Ezurum	14	M	1 Adamu Lane	Abuja
3	Susan Nwaimo	16	Female	Number 58 Yaba Street	Kaduna State
4	Ajao Opeyemi	18	Female	No12 Olubunmi Street	Abeokuta
5	Banjoko Adebisola	14	Female	34 Ngige Street	Ugheli, Delta
6	Muhammed Olabisi	12	Female	13 ICAN road,	Enugu
7	Oluwagbemi Mojisola	13	Female	ACCA Lane	Onitsha

Next steps:

[Generate code with bb](#)
[View recommended plots](#)

```
bb["Gender"] = bb["Gender"].str.replace("Female", "F")
```

```
bb
```

	Name	Age	Gender	Street_Address	State
2	Alex Ezurum	14	M	1 Adamu Lane	Abuja
3	Susan Nwaimo	16	F	Number 58 Yaba Street	Kaduna State
4	Ajao Opeyemi	18	F	No12 Olubunmi Street	Abeokuta
5	Banjoko Adebisola	14	F	34 Ngige Street	Ugheli, Delta
6	Muhammed Olabisi	12	F	13 ICAN road,	Enugu
7	Oluwagbemi Mojisola	13	F	ACCA Lane	Onitsha

Next steps:

[Generate code with bb](#)
[View recommended plots](#)

reseting index

```
#index reseting
```

```
bb = bb.reset_index()
```

```
bb
```

index	Name	Age	Gender	Street_Address	State
0	2 Alex Ezurum	14	M	1 Adamu Lane	Abuja
1	3 Susan Nwaimo	16	F	Number 58 Yaba Street	Kaduna State
2	4 Ajao Opeyemi	18	F	No12 Olubunmi Street	Abeokuta
3	5 Banjoko Adebisola	14	F	34 Ngige Street	Ugheli, Delta
4	6 Muhammed Olabisi	12	F	13 ICAN road,	Enugu
5	7 Oluwagbemi Mojisola	13	F	ACCA Lane	Onitsha

Next steps:

[Generate code with bb](#)
[View recommended plots](#)

removing index columnns

```
#remove index
```

```
bb = bb.drop(columns = "index")
```

final result

#displaying clean dataset
bb

	Name	Age	Gender	Street_Address	State	
0	Alex Ezurum	14	M	1 Adamu Lane	Abuja	
1	Susan Nwaimo	16	F	Number 58 Yaba Street	Kaduna State	
2	Ajao Opeyemi	18	F	No12 Olubunmi Street	Abeokuta	
3	Banjoko Adebisola	14	F	34 Ngige Street	Ugheli, Delta	
4	Muhammed Olabisi	12	F	13 ICAN road,	Enugu	
5	Oluwagbemi Mojisola	13	F	ACCA Lane	Onitsha	

Next steps: [Generate code with bb](#)  [View recommended plots](#)