

Working With python DataFrames

Getting Started...

```
databank = {  
    "Name": ["Ben", "Jake", "Sully", "Bob", "Joy", "Matt", "Ace", "Hilary"],  
    "Age": [20, 18, 25, 35, 12, 20, 40, 38],  
    "Year": [2015, 2016, 2017, 2018, 2019, 2020, 2021, 2022],  
    "State": ["Lagos", "Kano", "Kogi", "Kano", "Gombe", "FCT", "Oyo", "Kano"],  
    "Language": ["English", "German", "Chinese", "French", "Latin", "Spanish", "Dutch", "Mandarin"],  
    "Population": [20000, 37000, 15000, 45000, 19000, 53000, 62000, 93000]  
}
```

Turning our data into a Frame

data =

pd. DataFrame(databank)

```
In [252]: data
```

```
Out[252]:
```

	Name	Age	Year	State	Language	Population
0	Ben	20	2015	Lagos	English	20000
1	Jake	18	2016	Borno	German	37000
2	Sully	25	2017	Zamfara	Chinese	15000
3	Bob	35	2018	Kano	French	45000
4	Joy	12	2019	Gombe	Latin	19000
5	Matt	20	2020	FCT	Cantonese	53000
6	Ace	40	2021	Oyo	Dutch	62000
7	Hilary	38	2022	Yola	Mandarin	93000

Slice out a column...

`data["Age"]`

```
In [255]: data["Age"]
```

```
Out[255]: 0      20
```

```
          1      18
```

```
          2      25
```

```
          3      35
```

```
          4      12
```

```
          5      20
```

```
          6      40
```

```
          7      38
```

```
          Name: Age, dtype: object
```

Slice out MULTIPLE columns...

```
data[["Age", "Language"]]
```

```
In [256]: data[["Age", "Language"]]
```

```
Out[256]:
```

	Age	Language
0	20	English
1	18	German
2	25	Chinese
3	35	French
4	12	Latin
5	20	Cantonese
6	40	Dutch
7	38	Mandarin

Call out a row...

`data.loc[2]`

```
In [257]: data.loc[2]
```

```
Out[257]: Name          Sully  
          Age           25  
          Year          2017  
          State         Zamfara  
          Language      Chinese  
          Population     15000  
          Name: 2, dtype: object
```

Call out multiple rows...

`data.loc[[2,3]]`

```
In [259]: data.loc[[2,3]]
```

```
Out[259]:
```

	Name	Age	Year	State	Language	Population
2	Sully	25	2017	Zamfara	Chinese	15000
3	Bob	35	2018	Kano	French	45000

Insert a column

```
data.insert(2, "Gender", ["M", "F", "M", "M", "F", "M", "F", "M"])
```

Out[276]:

	Name	Age	Gender	Year	State	Language	Population
0	Ben	20	M	2015	Lagos	English	20000
1	Jake	18	F	2016	Borno	German	37000
2	Sully	25	M	2017	Zamfara	Chinese	15000
3	Bob	35	M	2018	Kano	French	45000
4	Joy	12	F	2019	Gombe	Latin	19000
5	Matt	20	M	2020	FCT	Cantonese	53000
6	Ace	40	F	2021	Oyo	Dutch	62000
7	Hilary	38	M	2022	Yola	Mandarin	93000

Delete a column

Out[268]:

```
data.pop("Language")
```

	Name	Age	Gender	Year	State	Population
0	Ben	20	M	2015	Lagos	20000
1	Jake	18	F	2016	Borno	37000
2	Sully	25	M	2017	Zamfara	15000
3	Bob	35	M	2018	Kano	45000
4	Joy	12	F	2019	Gombe	19000
5	Matt	20	M	2020	FCT	53000
6	Ace	40	F	2021	Oyo	62000
7	Hilary	38	M	2022	Yola	93000

Dropping a column

`data.drop(columns = "Population")`

```
In [252]: data
```

```
Out[252]:
```

	Name	Age	Year	State	Language	Population
0	Ben	20	2015	Lagos	English	20000
1	Jake	18	2016	Borno	German	37000
2	Sully	25	2017	Zamfara	Chinese	15000
3	Bob	35	2018	Kano	French	45000
4	Joy	12	2019	Gombe	Latin	19000
5	Matt	20	2020	FCT	Cantonese	53000
6	Ace	40	2021	Oyo	Dutch	62000
7	Hilary	38	2022	Yola	Mandarin	93000

Dropping a Row

`data.drop(index = 8)`

```
In [252]: data
```

```
Out[252]:
```

	Name	Age	Year	State	Language	Population
0	Ben	20	2015	Lagos	English	20000
1	Jake	18	2016	Borno	German	37000
2	Sully	25	2017	Zamfara	Chinese	15000
3	Bob	35	2018	Kano	French	45000
4	Joy	12	2019	Gombe	Latin	19000
5	Matt	20	2020	FCT	Cantonese	53000
6	Ace	40	2021	Oyo	Dutch	62000
7	Hilary	38	2022	Yola	Mandarin	93000

Calling An Entry in a Frame

`data.loc[3,"State"]`

`data.loc[5,"Age"]`

	Name	Age	Gender	Year	State	Language	Population
0	Ben	20	M	2015	Lagos	English	20000
1	Jake	18	F	2016	Borno	German	37000
2	Sully	25	M	2017	Zamfara	Chinese	15000
3	Bob	35	M	2018	Kano	French	45000
4	Joy	12	F	2019	Gombe	Latin	19000
5	Matt	20	M	2020	FCT	Cantonese	53000
6	Ace	40	F	2021	Oyo	Dutch	62000
7	Hilary	38	M	2022	Yola	Mandarin	93000

Changing An Entry in a Frame

```
data.loc[3,"State"] = "Abia"
```

```
data.loc[5,"Age"] = 99
```

	Name	Age	Gender	Year	State	Language	Population
0	Ben	20	M	2015	Lagos	English	20000
1	Jake	18	F	2016	Borno	German	37000
2	Sully	25	M	2017	Zamfara	Chinese	15000
3	Bob	35	M	2018	Abia	French	45000
4	Joy	12	F	2019	Gombe	Latin	19000
5	Matt	99	M	2020	FCT	Cantonese	53000
6	Ace	40	F	2021	Oyo	Dutch	62000
7	Hilary	38	M	2022	Yola	Mandarin	93000

Styling Your DataFrame

db.style.

highlight_max(color='lightblue').

highlight_min(color='lightgreen')

	Name	Age	Year	State	Language	Population
0	Ben	20	2015	Lagos	English	20000
1	Jake	18	2016	Kano	German	37000
2	Sully	25	2017	Kogi	Chinese	15000
3	Bob	35	2018	Kano	French	45000
4	Joy	12	2019	Gombe	Latin	19000
5	Matt	20	2020	FCT	Spanish	53000
6	Ace	40	2021	Oyo	Dutch	62000
7	Hilary	38	2022	Kano	Mandarin	93000

Styling Your DataFrame

db.style.

**highlight_between(subset='Age' ,
left=20,right=40)**

	Name	Age	Year	State	Language	Population
0	Ben	20	2015	Lagos	English	20000
1	Jake	18	2016	Kano	German	37000
2	Sully	25	2017	Kogi	Chinese	15000
3	Bob	35	2018	Kano	French	45000
4	Joy	12	2019	Gombe	Latin	19000
5	Matt	20	2020	FCT	Spanish	53000
6	Ace	40	2021	Oyo	Dutch	62000
7	Hilary	38	2022	Kano	Mandarin	93000

Styling Your DataFrame

```
db[['Age','Population']].  
style.bar(color='lightgreen')
```

	Age	Population
0	20	20000
1	18	37000
2	25	15000
3	35	45000
4	12	19000
5	20	53000
6	40	62000
7	38	93000