



Varun Sagar
Theegala

Date	PERMNO	RETURN	monthlyD	w
8/1/1930	10096	0.015112	1930-08	0.022184
8/1/1930	10114	0.015114	1930-08	0.02255
8/1/1930	10030	0.023216	1930-08	0.023727
8/1/1930	10115	0.008732	1930-08	0.009838
8/1/1930	10173	0.07617	1930-08	0.027199
8/1/1930	10081	-4.00E-05	1930-08	0.021026
8/1/1930	10102	0.02134	1930-08	0.017168
8/1/1930	10110	0.02134	1930-08	0.012924
8/1/1930	10137	-0.02692	1930-08	0.022955
8/1/1930	10145	-4.00E-05	1930-08	0.005208
8/1/1930	10153	0.013353	1930-08	0.004051
8/1/1930	10161	0.019	1930-08	0.002508
8/1/1930	10196	0.019	1930-08	0.026042
8/1/1930	10209	-4.00E-05	1930-08	0.026813
8/1/1930	10217	-0.01216	1930-08	0.003279
8/1/1930	10225	0.032747	1930-08	0.01794
8/1/1930	10233	-4.00E-05	1930-08	0.01331
8/1/1930	10241	0.01918	1930-08	0.003188

MAKING EDA SIMPLE!

STEP 1

Install the library Dtale.

If using Jupyter Notebook, you can run the following command:

```
In [1]: %pip install dtale
```

```
Collecting dtale
  Obtaining dependency information for dtale from https://pypi.org/project/dtale/3.10.0
  Downloading dtale-3.10.0-py2.py3-none-any.whl.metadata (1.5 kB)
Collecting dash-colorscales (from dtale)
  Downloading dash_colorscales-0.0.4.tar.gz (62 kB)
  Preparing metadata (setup.py) ... done
Collecting dash-daq (from dtale)
  Downloading dash_daq-0.5.0.tar.gz (642 kB)
  Preparing metadata (setup.py) ... done
Collecting Flask-Compress (from dtale)
  Obtaining dependency information for Flask-Compress from https://pypi.org/project/Flask-Compress/1.14.0
  Downloading Flask_Compress-1.14.0-py3-none-any.whl.metadata (1.5 kB)
Requirement already satisfied: future>=0.14.0 in /Users/varun/.venv/dtale/lib/python3.10/site-packages (from Flask-Compress)
Requirement already satisfied: kaleido in /Users/varun/.venv/dtale/lib/python3.10/site-packages (from dtale)
```

Restart the kernel once completed



STEP 2

Import the library Seaborn.

Use function “load_dataset” to import a dataset from the library

```
In [4]: import seaborn as sns  
df = sns.load_dataset("iris")  
display(df)
```

	sepal_length	sepal_width	petal_length	petal_width	species
0	5.1	3.5	1.4	0.2	setosa
1	4.9	3.0	1.4	0.2	setosa
2	4.7	3.2	1.3	0.2	setosa
3	4.6	3.1	1.5	0.2	setosa
4	5.0	3.6	1.4	0.2	setosa
...
145	6.7	3.0	5.2	2.3	virginica
146	6.3	2.5	5.0	1.9	virginica
147	6.5	3.0	5.2	2.0	virginica
148	6.2	3.4	5.4	2.3	virginica
149	5.9	3.0	5.1	1.8	virginica

150 rows x 5 columns

NOTE - You also can use your own data, if available




STEP 3

Import the library Dtale.

Use the function “.show()” and insert your dataset as argument.

In [6]: `import dtale`
`dtale.show(df)`



	5	sepal_length :	sepal_width :	petal_length :	petal_width :	species :	
0		5.10	3.50	1.40	0.20	setosa	
1		4.90	3.00	1.40	0.20	setosa	
2		4.70	3.20	1.30	0.20	setosa	
3		4.60	3.10	1.50	0.20	setosa	
4		5.00	3.60	1.40	0.20	setosa	
5		5.40	3.90	1.70	0.40	setosa	
6		4.60	3.40	1.40	0.30	setosa	
7		5.00	3.40	1.50	0.20	setosa	
8		4.40	2.90	1.40	0.20	setosa	
9		4.90	3.10	1.50	0.10	setosa	
10		5.40	3.70	1.50	0.20	setosa	
11		4.80	3.40	1.60	0.20	setosa	
12		4.80	3.00	1.40	0.10	setosa	
13		4.30	3.00	1.10	0.10	setosa	
14		5.80	4.00	1.20	0.20	setosa	
15		5.70	4.40	1.50	0.40	setosa	
16		5.40	3.90	1.30	0.40	setosa	

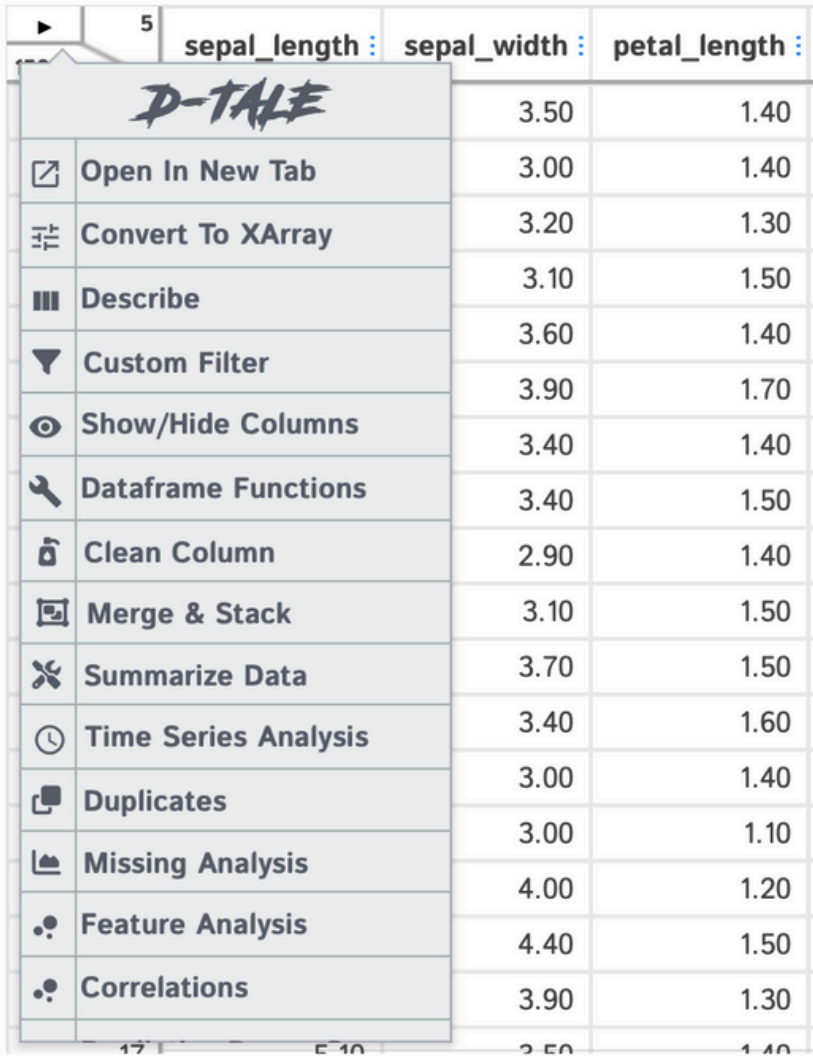
To get drop down - click the arrow at upper left corner in the output



THAT'S IT!

NOW YOU CAN ACCESS THE WHOLE SUITE OF EDA FUNCTIONS TO PERFORM ON YOUR DATA -

In [6]: `import dtale`
`dtale.show(df)`



The screenshot shows the dtale web interface. A data table is displayed with columns: `sepal_length`, `sepal_width`, and `petal_length`. A context menu is open over the first row, listing various EDA functions. The menu items are:

- Open In New Tab
- Convert To XArray
- Describe
- Custom Filter
- Show/Hide Columns
- Dataframe Functions
- Clean Column
- Merge & Stack
- Summarize Data
- Time Series Analysis
- Duplicates
- Missing Analysis
- Feature Analysis
- Correlations

	<code>sepal_length</code>	<code>sepal_width</code>	<code>petal_length</code>
1	3.50	1.40	
2	3.00	1.40	
3	3.20	1.30	
4	3.10	1.50	
5	3.60	1.40	
6	3.90	1.70	
7	3.40	1.40	
8	3.40	1.50	
9	2.90	1.40	
10	3.10	1.50	
11	3.70	1.50	
12	3.40	1.60	
13	3.00	1.40	
14	3.00	1.10	
15	4.00	1.20	
16	4.40	1.50	
17	3.90	1.30	
18	3.50	1.40	

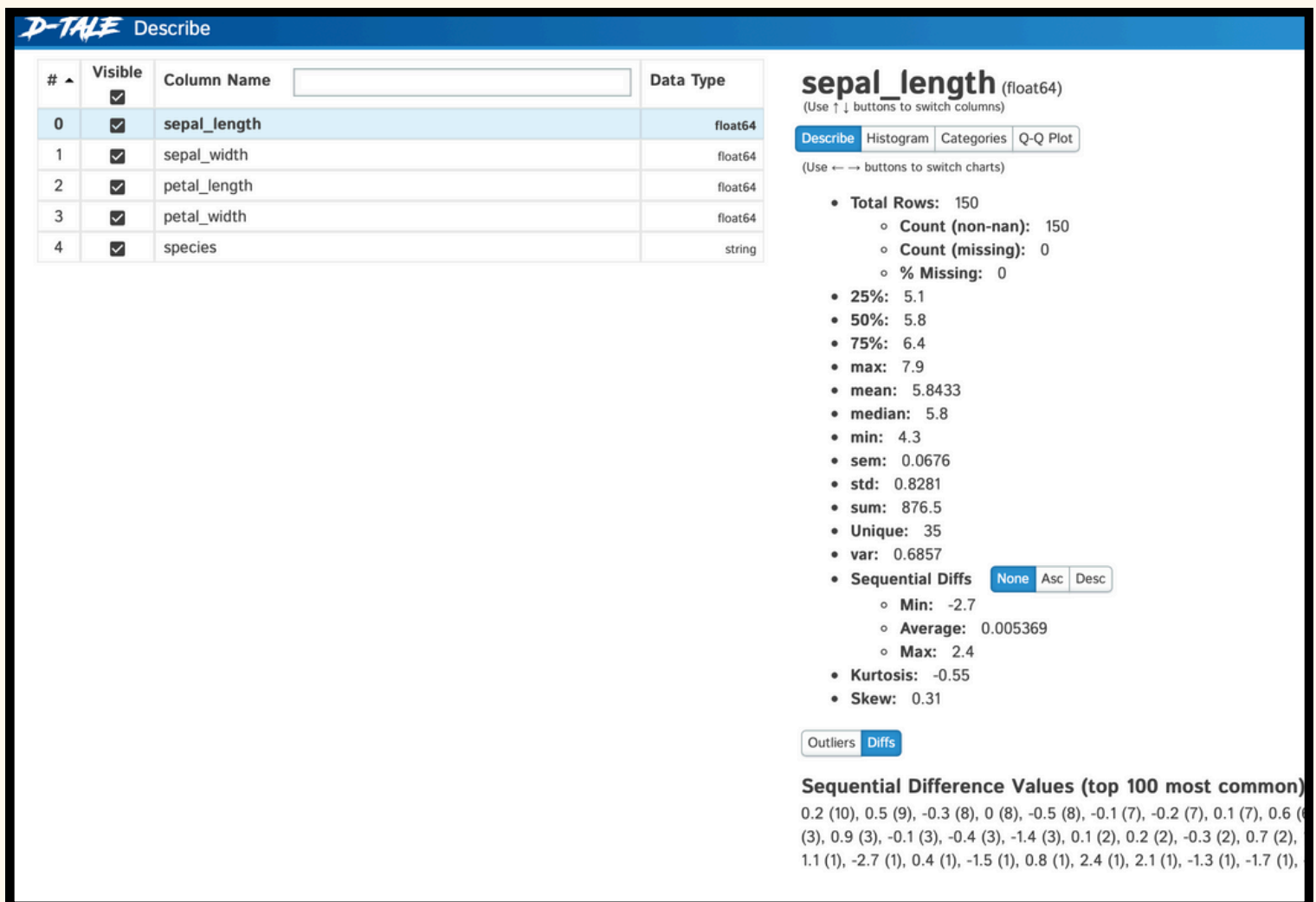


SOME OF THE CRAZY STUFF YOU CAN DO



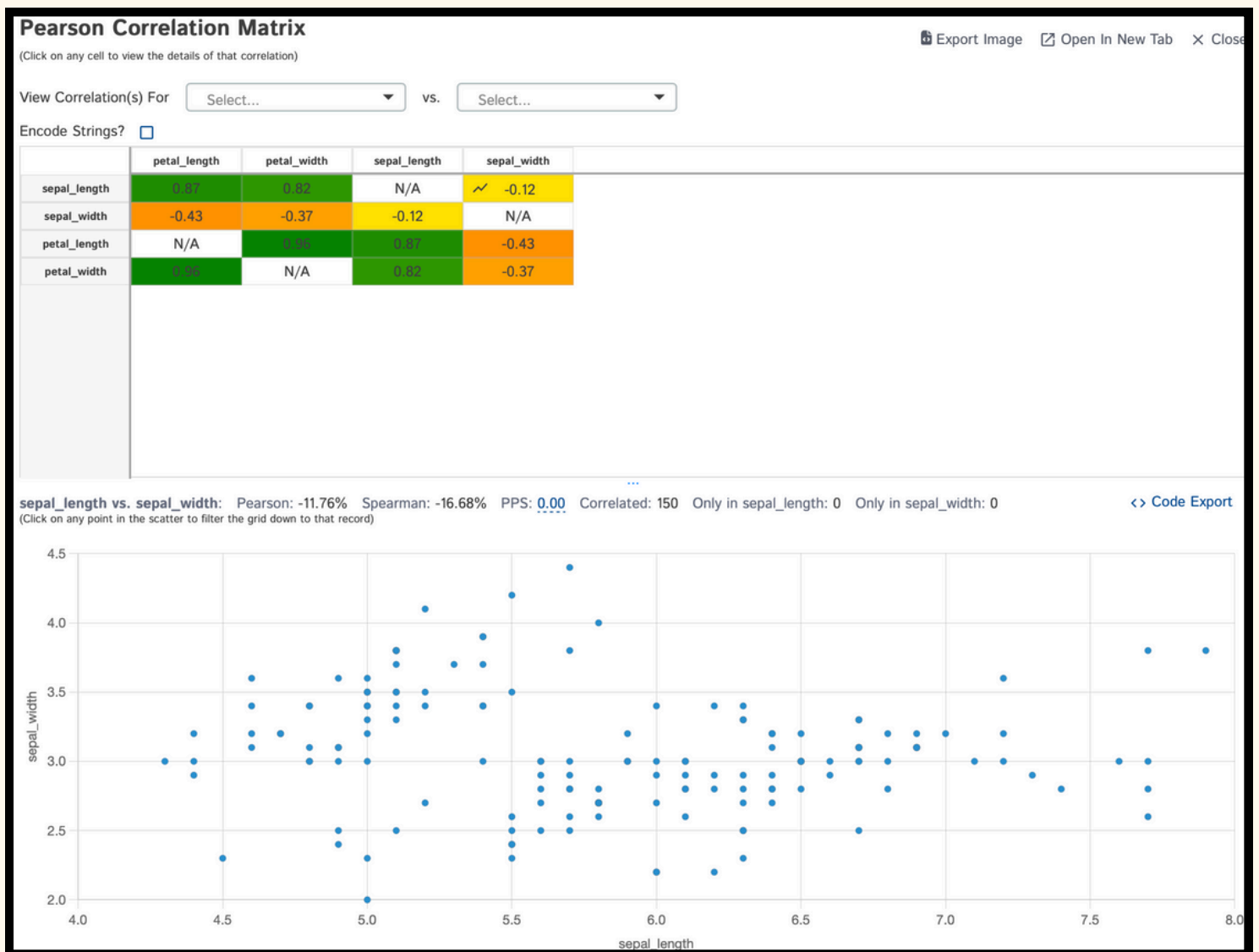
DESCRIPTIVE STATISTICS

Get the statistical summary for the
columns of your choosing!



CORRELATION ANALYSIS

Understand the relationship between variables in your data



DUPLICATE SUMMARY

Identify the presence & count of duplicate values in select fields

Duplicates

Operation

Remove Duplicate Columns

Remove Duplicate Column Names

Remove Duplicate Rows

Show Duplicates

Show all duplicates data or duplicate data for a specific value.

Column(s)

petal_width

X

View Duplicates

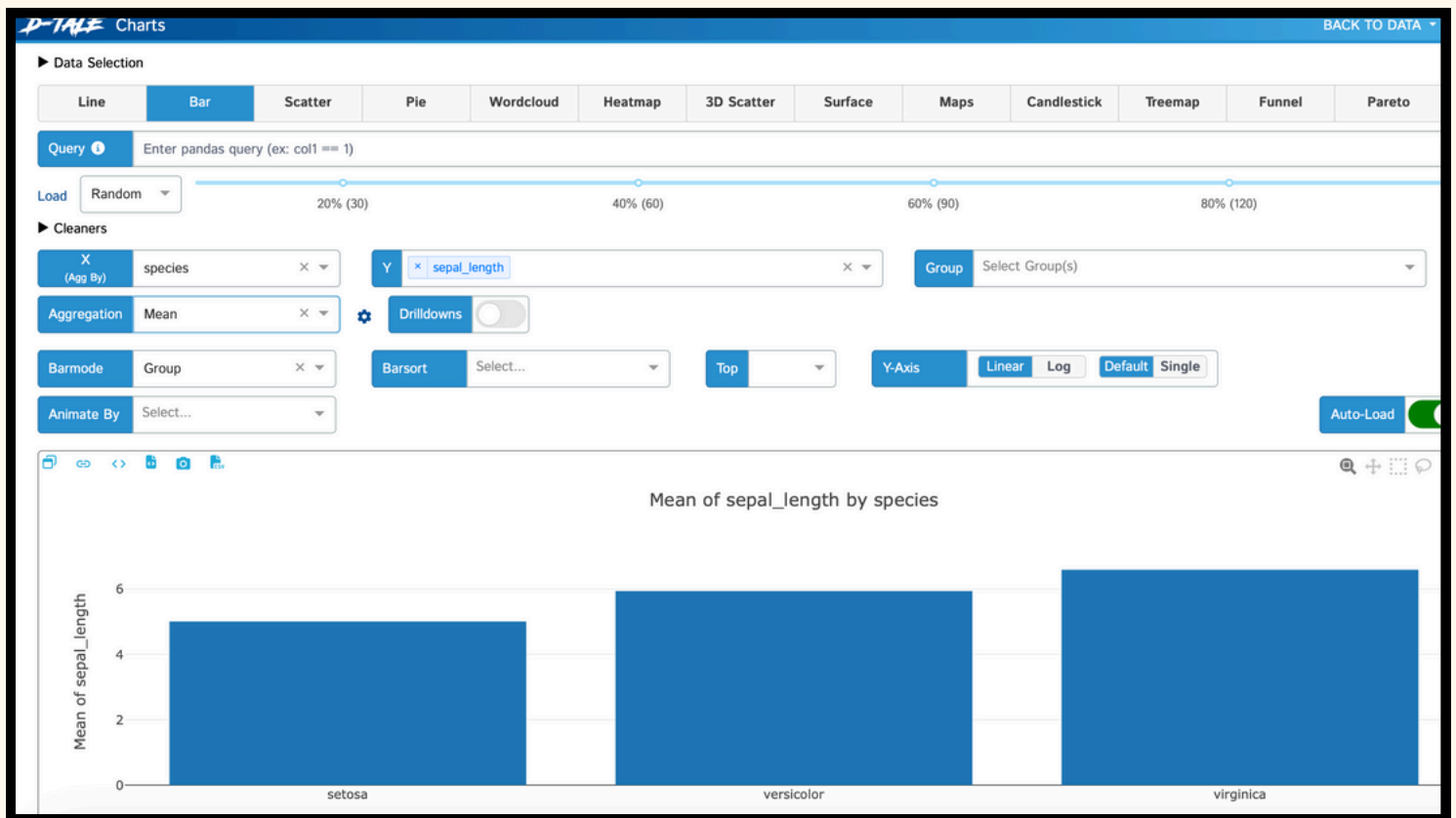
Duplicates exist for the following (petal_width) groups:
Total Duplicates: 148

- ☐ 1: 7
- ☐ 2: 6
- ☐ 0.1: 5
- ☐ 0.2: 29
- ☐ 0.3: 7
- ☐ 0.4: 7
- ☐ 1.1: 3
- ☐ 1.2: 5
- ☐ 1.3: 13
- ☐ 1.4: 8
- ☐ 1.5: 12
- ☐ 1.6: 4
- ☐ 1.7: 2
- ☐ 1.8: 12
- ☐ 1.9: 5
- - - -



VISUALISE YOUR DATA USING CHARTS

Build summaries & visuals based on required fields, aggregation, etc



IDENTIFY RELEVANT ROW USING RANGE HIGHLIGHT

Highlight the rows you're looking for based on applying criteria(s)

length	sepal_width	petal_length	petal_width	species
5.10	3.50	1.40	0.20	setosa
4.90	3.00	1.40	0.20	setosa
4.70	3.20	1.30	0.20	setosa
4.60	3.10	1.50	0.20	setosa
5.00	3.60	1.40	0.20	setosa
4.40	3.90	1.70	0.40	setosa
4.60	3.40	1.40	0.30	setosa
5.00	3.40	1.50	0.20	setosa
4.40	2.90	1.40	0.20	setosa
4.90	3.10	1.50	0.10	setosa
4.40	3.70	1.50	0.20	setosa
4.80	3.40	1.60	0.20	setosa
4.80	3.00	1.40	0.10	setosa
4.30	3.00	1.10	0.10	setosa
4.80	4.00	1.20	0.20	setosa
4.70	4.40	1.50	0.40	setosa
4.40	3.90	1.30	0.40	setosa
5.10	3.50	1.40	0.30	setosa
4.70	3.80	1.70	0.30	setosa
5.10	3.80	1.50	0.30	setosa
4.40	3.40	1.70	0.20	setosa
5.10	3.70	1.50	0.40	setosa
4.60	3.60	1.00	0.20	setosa
5.10	3.30	1.70	0.50	setosa
4.80	3.40	1.90	0.20	setosa
5.00	3.00	1.60	0.20	setosa
5.00	3.40	1.60	0.40	setosa

Range Highlights
Col: petal_length
☒ Equals 1.5
☒ Greater Than 1.5
☐ Less Than
Apply
☒ petal_length: Equals 1.5 or Greater Than 1.5



**WE BARELY
SCRATCHED THE
SURFACE!**

**THERE'S A LOT
MORE TO
EXPLORE.**

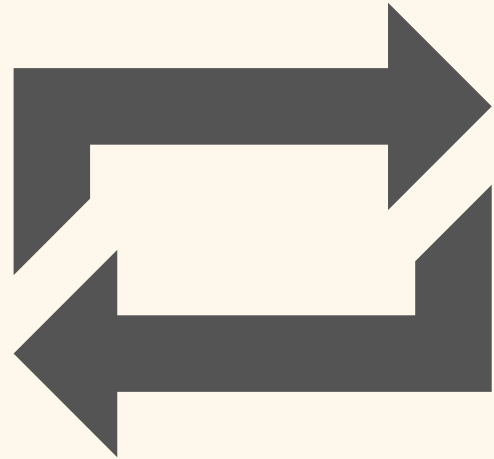




Varun Sagar
Theegala

SHARE THIS

If you think
your network
would
find this
valuable



FOLLOW ME

I help you
**GROW &
SUSTAIN** as a
Data Analyst

