

## Azure Databricks Assignment - 2

### Mitushi Vishwakarma

#### Data Exploration and Visualization in Databricks

Week - 4

Azure Databricks  
Day - 2

Databricks has built-in support for charts and visualizations ~~on them~~

To Create a new visualization.

1. Read data from file or create dataframe.
2. display the dataframe : `display()`

Then click '+' and select visualization then editor appears.

select type of visualization from drop-down.  
pie,  
bar chart  
scatter.

Change colors accordingly

Create data-profile using '+' symbol

data profile shows summary statistics.

48	27	28	29	30
47	20	21	22	23
46	13	14	15	16
45	6	7	8	9
44	1	2	3	4
43				5

2023  
THURSDAY  
WK 48 • DAY 334-031  
NOVEMBER

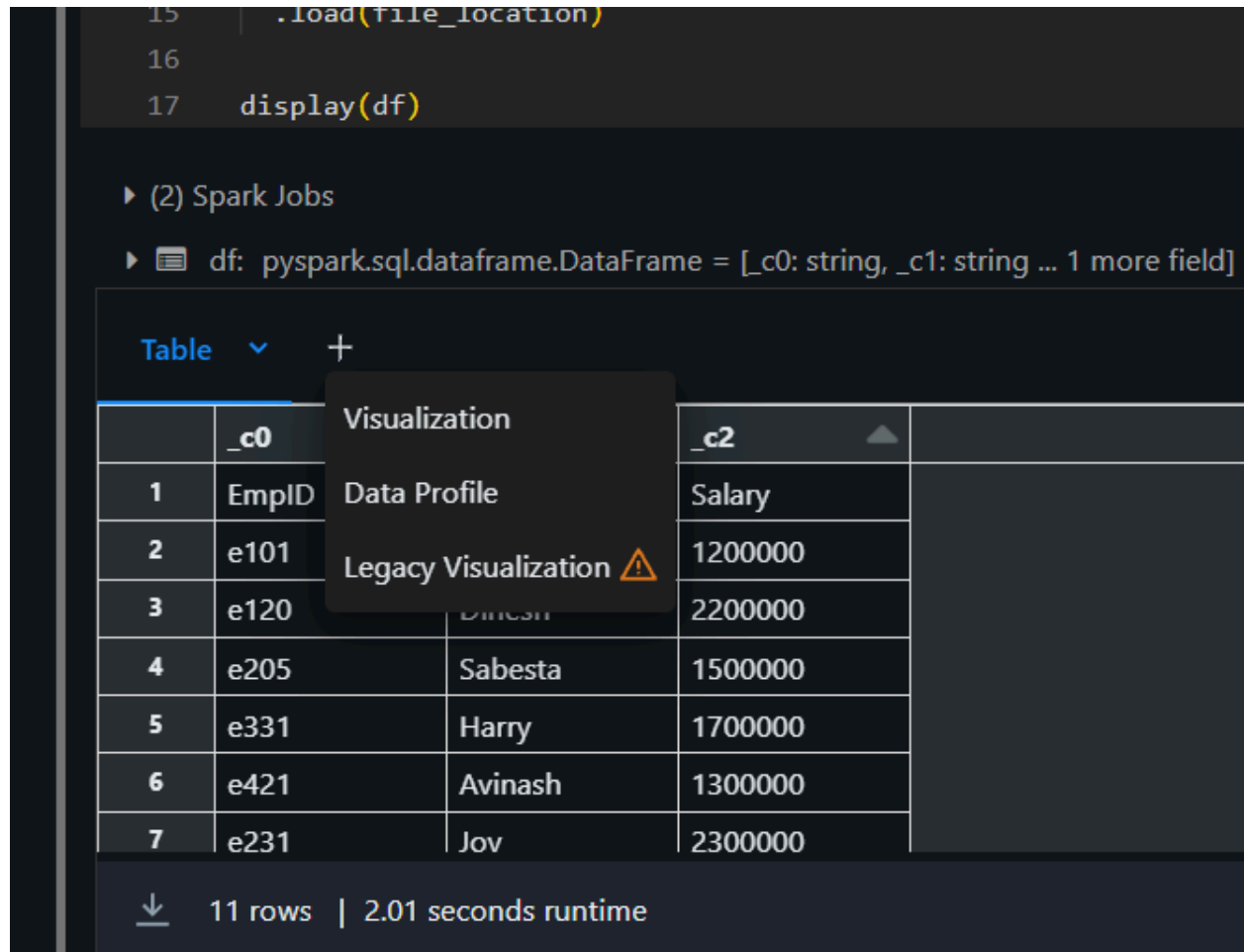
30  
30-11-2023

## Hands-on Exercise: Visualizing data in Databricks

### Creating new visualizations:

In order to create visualizations, we need to have data.  
Here I provided data and used the display() method.

- After creating a table
- Click on + symbol
- Click on visualization.



The screenshot shows the Databricks interface. At the top, a code editor displays the following code:

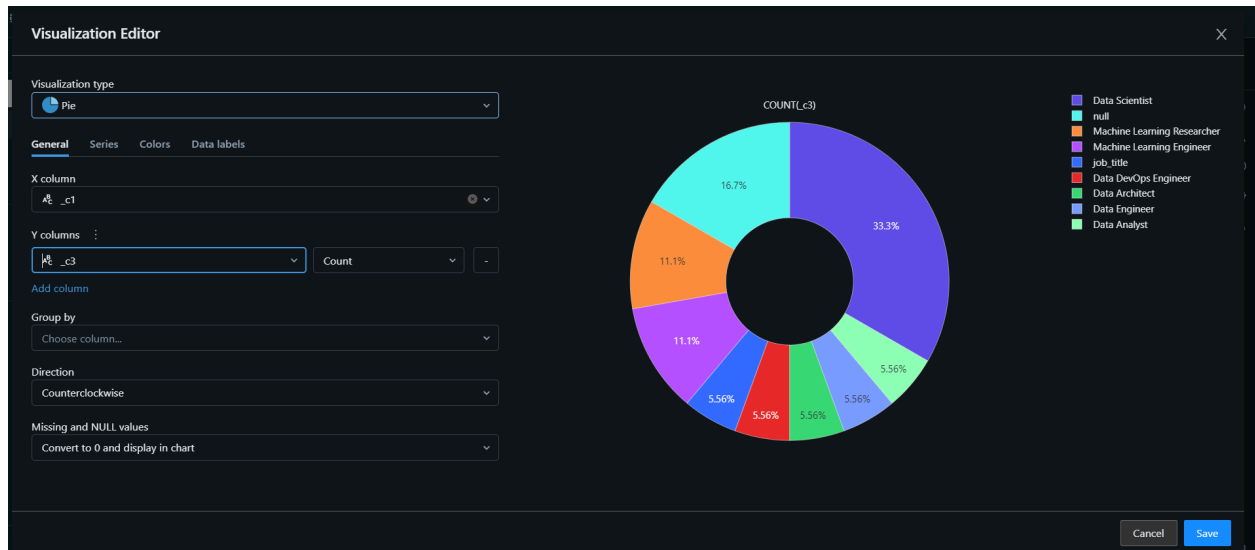
```
15 | .load(file_location)
16 |
17 | display(df)
```

Below the code, the Spark Jobs section shows a job named 'df' with the type 'pyspark.sql.dataframe.DataFrame' and a schema of '[\_c0: string, \_c1: string ... 1 more field]'. A table view of the data is displayed below the job information. The table has columns '\_c0', '\_c1', and '\_c2'. The data is as follows:

	_c0	_c1	_c2
1	EmplID		Salary
2	e101		1200000
3	e120		2200000
4	e205	Sabesta	1500000
5	e331	Harry	1700000
6	e421	Avinash	1300000
7	e231	Jov	2300000

A dropdown menu is open over the table, showing three options: 'Visualization', 'Data Profile', and 'Legacy Visualization' (marked with a warning icon). The 'Visualization' option is selected. At the bottom of the table, it shows '11 rows | 2.01 seconds runtime'.

In the visualization editor, We can select which type of chart we want in the drop down. There are various options then to groupBy, change colors,etc.



You can also change the color of the chart visualization by clicking “edit” on the bottom left corner.

You can edit , delete , view and download visualization charts.



After creating the visualization, we can also create data profile.

**Data profile :** Used to analyse the data based on some trends, you can also order and filter the data.

