Assignment – Day 17

-Sarthak Niranjan Kulkarni (Maverick)

- [sarthakkul2311@gmail.com](mailto:sarthakkul2311@gmail.com) - (+91) 93256 02791

**28/11/2024 (Thursday)**

**Practice on Visualization: -**

1. **Loading and Displaying Data from the 'loan' Table**

🡪

A screenshot of a computer

Description automatically generated

A screenshot of a graph

Description automatically generated

1. **Loading and Displaying Data from 'export' Table and Delta Location**

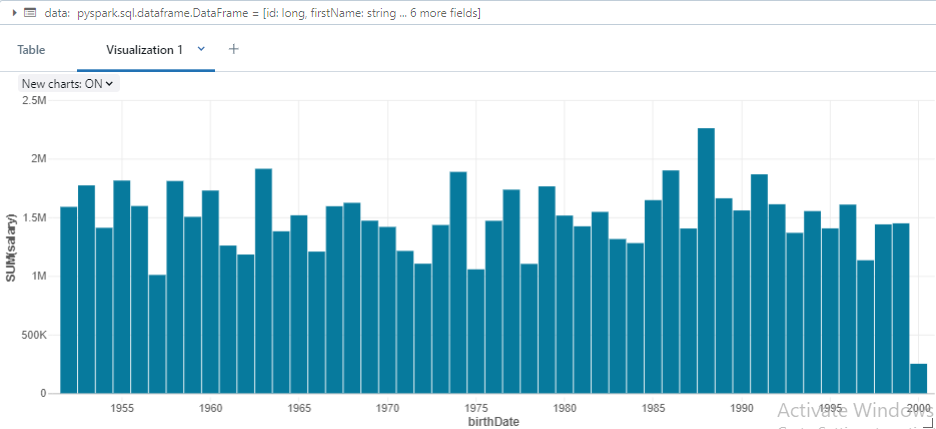
🡪 spark.table("default.export")

data = spark.read.format("delta").load("dbfs:/user/hive/warehouse/export")

data.display()

A screenshot of a computer

Description automatically generated



**Summary on Visualization: -**

In Azure Databricks, data visualization using PySpark can help you easily interpret and present data insights. PySpark allows you to work with large datasets and perform complex transformations before visualizing the results. The display() function in Databricks provides a powerful way to visualize DataFrames directly in the notebook interface. When you load data into PySpark, whether from a table or a Delta file, you can quickly visualize it using Databricks' built-in visualization tools.

Visualizations like bar charts, line graphs, and scatter plots can be created with just a few clicks, providing an intuitive way to explore data patterns. You can create custom visualizations to examine trends over time, compare categories, or understand distributions. Databricks also supports interactive visualization, which means you can drill down into the data, filter values, and adjust axes for better clarity.

By using PySpark for data processing and Databricks for visualization, you can enhance your data exploration experience without switching tools or environments. This integration makes it easier to share insights with others and perform interactive analysis in real-time. Visualizations are crucial for communicating data findings in an understandable and impactful way.