Name: Shakib Ibna Shameem Polock

ID: 2307673

**What have you learned today?**

Ans: This exercise has exciting topics on PySpark and its architecture. I learned about Spark architecture, Spark sessions, and how Spark processes vast amounts of data in parallel using clusters of computers. Moreover, in the hands-on session, I applied various functions on Spark DataFrame for data manipulation and processing and created some visualizations. So far, it has been very interesting.

**What was difficult? At which step have you encountered problems?**

Ans: So far in this exercise all the instructions seemed clear and doable.

**Provide some figures by using matplotlib and comment what is visualized there.**

Ans: Here are the few visualizatoions of the dataset.

Here is the time series plot for the air\_temperature over the time. We can see some sesonality here.

A blue and white sound wave

Description automatically generated

Histograms for various columns.

A graph of a number of data

Description automatically generated with medium confidence