

## **Task 1 – Report (Day 1)**

### **Task Name – Manual Audit and Logic Check**

**Day – Tuesday**

**Date – 10<sup>th</sup> Feb 2026**

**Time – 7:00 AM – 10:00 AM**

**As this task was assigned yesterday, but I missed the session due to my mistake. I apologize for that and I am submitting the completed report today.**

**Dataset – Café\_Sales**

**Dataset Source –** The dataset was downloaded from Kaggle.

#### **Problems Faced -**

- As the instructions mentioned in task sheet that any dirty dataset with 5000 rows, but I downloaded the available dataset with approximately 10000 rows, since the dataset not found as per requirements.
- But it taken a long time to found the dataset as per the requirements.
- Also, I faced issue with my laptop's Excel Application. so, I completed the task using a desktop PC.

## **Dataset Overview –**

The dataset contains sales transaction from café. It includes information about orders, customer details, items, quantities, prices etc.

## **Dataset Structure –**

<b>Columns</b>	<b>Rows</b>
<b>08</b>	<b>10000</b>

- As I have went through first 5000 rows to understand the dataset, I observed that there are missing values are present, also has the mismatched data formats.
- Also, I observed that some item names are written differently as some names are written in All CAPS format, where some in small. That means it has inconsistency format.
- After that, I have gone through the column names, so the column names were so unique and simple to understand its meaning but it also has space in between the names.

**Columns are –** Transaction ID, Item, Quantity, Price Per Unit, Total Spent, Payment Method, Location, Transaction Date

- Today I went through the dataset and understand the dataset structure, from tomorrow I will start to clean the dataset.