Final Report on Dashboard

Data Preparation and Cleaning Process

Step 1: Data Extraction and Initial Review

- The dataset was pulled from the metadata source and imported into Excel.

- A manual inspection was performed to verify the structure and integrity of data fields.

- The dataset includes the following key fields:

    o **Learner_name**

    o **Program_name**

    o **Codin_assignment**

    o **max_score**

    o **obtained_score**

    o **submitted_at**

Step 2: Data Cleaning

Step 2.1: Initial Data Cleaning

- The **submitted_at** column was removed as it was deemed unnecessary for further analysis.

- Any null values in the **obtained_score** column were replaced with **0** to ensure consistency in calculations.

- A new column, **Final**, was created to flag records where **obtained_score = max_score**.

    o Only records meeting this condition were retained.

    o Post-filtering, the dataset retained the following record counts:

        ▪ **July**: 17,003 records

        ▪ **September**: 9,624 records

        ▪ **October**: 15,193 records

        ▪ **November**: 18,234 records

Step 2.2: Handling Duplicate Entries

- Duplicate entries were identified based on **Learner_name**, **Program_name**, and **Codin_assignment**.

- Using Excel's **Remove Duplicates** feature (drop_duplicates logic), only the last occurrence of each duplicate combination was retained.

- After deduplication, the dataset had the following counts:

  - **July**: 12,386 records

  - **September**: 8,380 records

  - **October**: 12,772 records

  - **November**: 15,941 records

- A new column named **Month** was created by extracting the month information from the **Program_name** field.

Step 3: Data Consolidation and Reporting

- Utilizing Excel's **Power Query**, all cleaned data was consolidated into a master Excel file for comprehensive analysis.

- A **Pivot Table** was created with the following configurations:

  - **Learner_name** added to the **Values** area as a **Count of Learner_name**.

  - **Month** placed in the **Legend** field.

  - **Codin_assignment** positioned in the **Rows** area.

- A **Pivot Chart** was inserted to visually represent the summarized data.

- The final report was generated, providing a clear and structured analysis ready for stakeholder review.