**Assignment-2**

**Data cleaning using Tableau**

Subject: Web Analytics and Business Intelligence Tools

PROG8461 - Fall 2024

Professor: Qasim Ail

Student Name : Haripriya Naga Neelima Yalamkayala

Student ID : 8939330

Date: 26th November 2024

1. Load the Dataset into Tableau

A screenshot of a computer

Description automatically generated

1. Remove Unnecessary Columns: Columns like Query10, Response10, Query11, Response11, Query12, Response12, Query13, Response13 have more than 50% of data missing in them, so I’m removing those since they’re not much helpful for any analysis.

A screenshot of a computer

Description automatically generated

1. Handle Missing Data by Filtering and Replacing

Added filters for columns with high missing data.

A screenshot of a computer

Description automatically generated

Replaced missing data from worker ID column with median.

A screenshot of a computer

Description automatically generated

1. Detect and Remove Outliers in duration column.

Outliers in the Duration column were identified and flagged using the **IQR (Interquartile Range) method**. Rows marked as "Outliers" were excluded by applying a filter, leaving only "Normal" rows for further analysis.

A screenshot of a computer

Description automatically generated

1. Created a new column to convert Time column to Date Time instead of string datatype.

A screenshot of a computer

Description automatically generated

1. Create additional columns to extract Year, Month, Day and Date only.

A screenshot of a computer

Description automatically generated

Story on the cleaned data:

A screenshot of a computer

Description automatically generated