Assignment overview

Estimated Reading Time: 5 minutes

Data wrangling, also known as data munging, involves cleaning and preparing datasets to make them ready for analysis. In this assignment, you will perform a series of data wrangling tasks using the Stack Overflow survey data:

A. Identify and remove duplicate rows

- Locate any duplicate rows within the dataset.
- Remove these duplicates to maintain data integrity.

B. Analyze missing values

- Calculate the number of missing values in each column to assess data completeness.

C. Examine employment status

- Use the Employment column to identify value counts and analyze employment distribution among respondents.

D. Normalize data for comparative analysis

- Choose two relevant columns, such as CompTotal (total compensation) and ConvertedCompYearly (annual compensation), and normalize their values to enable comparison or further analysis.