Assignment2\_Jo

Josue Iradukunda

2024-11-11

1. Import the CSV (comma-separated value) using the read.csv() command. Name it survey\_data.

# Data cleaning using grepl

## Field Count Percentage  
## 1 Applied Statistics 26 55.32  
## 6 Industrial Art and Design 4 8.51  
## 2 Civil Engineering 3 6.38  
## 4 Data Science 3 6.38  
## 10 Travel and Tourism Management 3 6.38  
## 3 Data science 2 4.26  
## 7 Information System 2 4.26  
## 9 Quantity Surveying 2 4.26  
## 5 Demography 1 2.13  
## 8 Information Technology 1 2.13

The table illustrates the distribution of respondents across nine fields of study, with **Applied Statistics** leading with 27 respondents, accounting for 52.94% of the total. **Data Science** comes next with 5 respondents (9.80%), while both **Industrial Art and Design** and **Information System** each have 4 respondents (7.84%). Fields like **Civil Engineering**, **Quantity Surveying**, and **Travel and Tourism Management** each have 3 respondents (5.88%). The least represented fields, **Demography** and **Information Technology**, each have only 1 respondent, making up 1.96%. In summary, the data is predominantly skewed toward Applied Statistics, with most other fields showing limited participation.

