Joseph Annand

DSE5004

Assignment 01: Introduction

Greetings! My name is Joey, I am from Central Massachusetts, and I have been in Merrimack’s M.S in Data Science program since Summer 2023. This is my sixth course in the program. Prior to enrolling in this program, I graduated from Stonehill College with a B.A. in Chemistry, and I completed two courses, Introduction to Python Programming and Algorithms and Discrete Structures, as part of the M.S. in Computer Science program at Merrimack. I have a decent amount of experience with Python and R programming from both Merrimack coursework and self-teaching.

Currently, I work as a Scientist at a medical device manufacturer; my role is focused in the development of a new instrument for blood-clotting analysis. A major responsibility in my job is to present data to various multi-disciplinary, cross-functional audiences. The challenge of this responsibility is crafting engaging, informative presentations to discuss data from the experiments my colleagues and I perform in the laboratory. Something I have consistently struggled with since I started this job two years ago is translating numbers and statistics into comprehensible visualizations. I want to have the knowledge to choose the appropriate plot or chart that is either easy for my audience to quickly view and understand or easy for me to briefly explain to them during my presentations. Often in my data science projects thus far, I have learned to work with large data sets; however, at work, I am often analyzing data of small sample sizes, making it difficult for me to know how to translate what I have learned in my courses to my problems at work. My hope for this course is that I will better understand how to represent data at work through visuals and be better able to tell the story of my experiments using them.

The Customer Dataset consists of 5000 rows, or observations, and 60 columns, or variables.