DSC 423: Data Analysis and Regression
Assignment 01: Review

Your submission must include your name and student ID. Your submission must include the honor statement: "I have completed this work independently. The solutions given are entirely my own work." Your submission must be submitted as a PDF.

1. Short Essay (10 pts.) Consider the following two scenarios. A) take a simple random sample of 100 graduate students at DePaul university and b) take a simple random sample of 100 graduate students studying Data Science. For each sample you record the amount spent on textbooks used for classes. Which sample do you expect to have the smaller standard deviation? Explain your answer.
2. Empirical rule (20 pts.) The 222 students enrolled in online-learning courses offered by a college ranged from 18 to 64 years of age. The mean age was 28 with standard deviation equal to 4. Use the 68-95-99.7 rule to answer the following questions:
    a. (10 pts.) Compute the percentage of students that are between 24 and 32 years old. Show your work.
    b. (10 pts.) Compute the percentage of students that are older than 36 years. Show your work.
3. Z-scores (10 pts.) Monthly sale figures for a particular e-retailer tend to be normally distributed with mean equal to 150 thousand dollars and a standard deviation of 35 thousand dollars. Use the normal distribution to determine the top 1% monthly sale figure (a.k.a. 99th percentile)? Show your work.
4. Hypothesis Testing (10 pts.) A network provider investigated the number of blocked intrusions to its network, and found that there were, on average, 45 blocked intrusions per day. After a change in firewall settings, the mean number of intrusions during the next 35 days was 42 with a standard deviation equal to 15.5. Perform a hypothesis test to determine if the change in firewall settings reduced the number of intrusions. Show your work.