Honor statement: I have completed this work independently. The solutions given are entirely my own work.

1. Short Essay (10 pts.) Consider the following two scenarios.

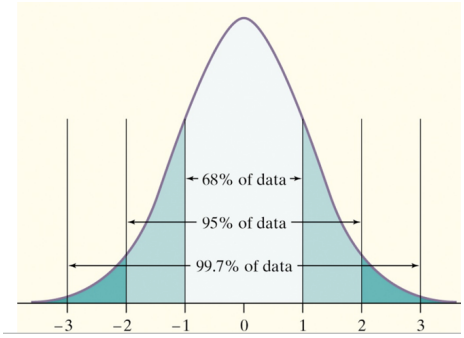
A) take a simple random sample of 100 graduate students at DePaul university and b) take a simple random sample of 100 graduate students studying Data Science. For each sample you record the amount spent on textbooks used for classes. Which sample do you expect to have the smaller standard deviation? Explain your answer.

I would say the B sample has smaller standard deviation.
Smaller standard deviation means that the values in a statistical data set is closer to the mean. The difference between the amount that each observations of graduate students at DePaul university spent and the average amount of graduate students at DePaul university spent could be larger than the graduate students at Data Science spent on textbook , since the graduate students at Data Science might need to buy similar textbook for the classes, which means that standard deviation of the amount that graduate students at Data Science could be smaller.

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.

Ans : Approximately 68% of students are between 24 and 32 years old .
68%fall within one standard deviation of mean.
Mean - sd = 28-4 =24
Mean + sd = 28+4 =32



b. (10 pts.) Compute the percentage of students that are older than 36 years. Show your work.

Ans : Approximately 68% of students are older than 36 years .

Mean + 2*sd = 28+8 =36

(1-95%)/2 = 2.5%

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.

Mean = 150,000
Standard deviation = 35,000
Z-score = X - mean/ standard deviation
> qnorm(0.99)
[1] 2.326348
2.326348 = (X-150,000)/ 35,000
X = 2.326348*35,000+150,000
X = 231422.2

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

H0 = 45
H1< 45
N= 35

T = x - mean / S/35**½ = 42-45 / 15.5/35**½ = -1.145048
t=-1.145048
df = 35-1 =34
lower.tail = TRUE

P- Value
> pt(-1.145048, 34 , lower.tail = TRUE)
[1] 0.1300929

Note: if p-value less than significance level α, reject $H0$

Assume α = 0.05
P-value > α Accept H0