Problem Set 1

Chase Bookin

June 24, 2020

Question 1

```
sample_1 <- c(3, 4, 5, 6, 7)
sample_2 <- c(67, 68, 69, 70, 71)

mean_samp1 <- sum(sample_1) / length(sample_1)

sd_samp1 <- round(sqrt(((3-mean_samp1)^2 + (4-mean_samp1)^2 + (5-mean_samp1)^2
                  + (6-mean_samp1)^2 + (7-mean_samp1)^2) / (length(sample_1) - 1)), 3)

mean_samp2 <- sum(sample_2) / length(sample_2)

sd_samp2 <- round(sqrt(((67-mean_samp2)^2 + (68-mean_samp2)^2 + (69-mean_samp2)^2
                  + (70-mean_samp2)^2 + (71-mean_samp2)^2) / (length(sample_2) - 1)), 3)
```

Sample 1:
- Mean: 5
- Standard Deviation: 1.581

Sample 2:
-Mean: 69
-Standard Deviation: 1.581

The standard deviations of sample 1 and sample 2 are equal. This shows that although the samples have different means, they are spread around their center with the same variance.

Question 2

```
z_tokyo <- (380000 - 420000) / 20000

z_germany <- (3100 - 3200) / 57
```

Relative to their peers, the worker in Germany is earning more than the worker in Tokyo. This is demonstrated by the z-score of each worker's salary. The z-score of the workers' salaries tells the relative position of their salary to their peers using the mean and standard deviation. In this case, the German worker's z-score of roughly -1.75 is greater than the Tokyo worker's z-score of -2, demonstrating that the German worker is earning comparatively more than the Tokyo worker.

Question 3

a)

```
z_prob_keane <- 1-.192
z_keane <- 0.87
sd_keane <- (25000 - 21000) / 0.87
```

Standard Deviation: 4597.7

b)

```r
z_42nd <- -0.2
percentile_42 <- (sd_keane * z_42nd) + 21000
```

42nd Percentile: 20080.46

c)

```r
# -1.55 <= z <= 1.55
percentile_94 <- (sd_keane * 1.55) + 21000
percentile_06 <- (sd_keane * -1.55) + 21000
```

Middle 88% values: (13873.56, 28126.44)