QANT 530 Homework #1

Phan Nguyen

07/12/2022

```
##   UtilityCosts MaintenceCosts
## 1     29.89512      17.818890
## 2     35.77786      14.547290
## 3     36.56817      18.462665
## 4     40.46540       8.226194
## 5     40.63734       7.087728
## 6     42.57490      18.452979
```

Figure 1: Dataset example

1.Descriptive Statistics

```
##   UtilityCosts   MaintenceCosts
##  Min.   :29.90   Min.   :-1.711
##  1st Qu.:48.42   1st Qu.: 7.582
##  Median :54.69   Median :10.828
##  Mean   :55.56   Mean   :10.427
##  3rd Qu.:61.43   3rd Qu.:14.255
##  Max.   :90.68   Max.   :20.801
```

Table 2: Descriptive Statistics table for two variables

```
##   UtilityCosts MaintenceCosts
##      0.4959448     -0.2601461
```

Table 3: Skewness for UtilityCosts and MaintenanceCosts

```
##   UtilityCosts MaintenceCosts
##       3.807389       2.486682
```

Table 4: Kurtosis for UtilityCosts and MaintenanceCosts

Measure of center

It looks like UtilityCosts has a normal range of values. The mean is slightly bigger than the median, suggesting that UtilityCosts' distribution is slight right skewed.
MaintenanceCosts has a minimum of -1.711, a negative number. This is quite strange, but since we don't know the context of this dataset, we won't know if this number makes sense. Mean and median for MaintenanceCosts are approximately equal, thus meaning the distribution for this variable is symmetrical.

Measure of variability

The interquartile range for UtilityCosts is between 48.42 and 61.43, meaning 50% of values are in this range. On the other hand, the interquartile range for MaintenanceCosts is between 7.58 and 14.25, meaning 50% of values for this variable are in this range.

Skewness and kurtosis

Skewness for UtilityCosts and MaintenanceCosts are very close to 0, suggesting that the distribution for the two variables are not skewed completely to left or right. The measures of center also suggested the same. UtilityCosts might be a bit more skewed to the right than MaintenanceCosts The distributions of the two variables are not normal, since the kurtosis is not equal to 3. The peak of UtilityCosts is higher than that of a normal distribution, since the kurtosis is larger than 3. MaintenanceCosts's distribution has heavier tails and lower bell, due to its kurtosis to be lower than 3.

2. Scatterplot analysis

Figure 5: Scatterplot for UtilityCosts and MaintenanceCosts