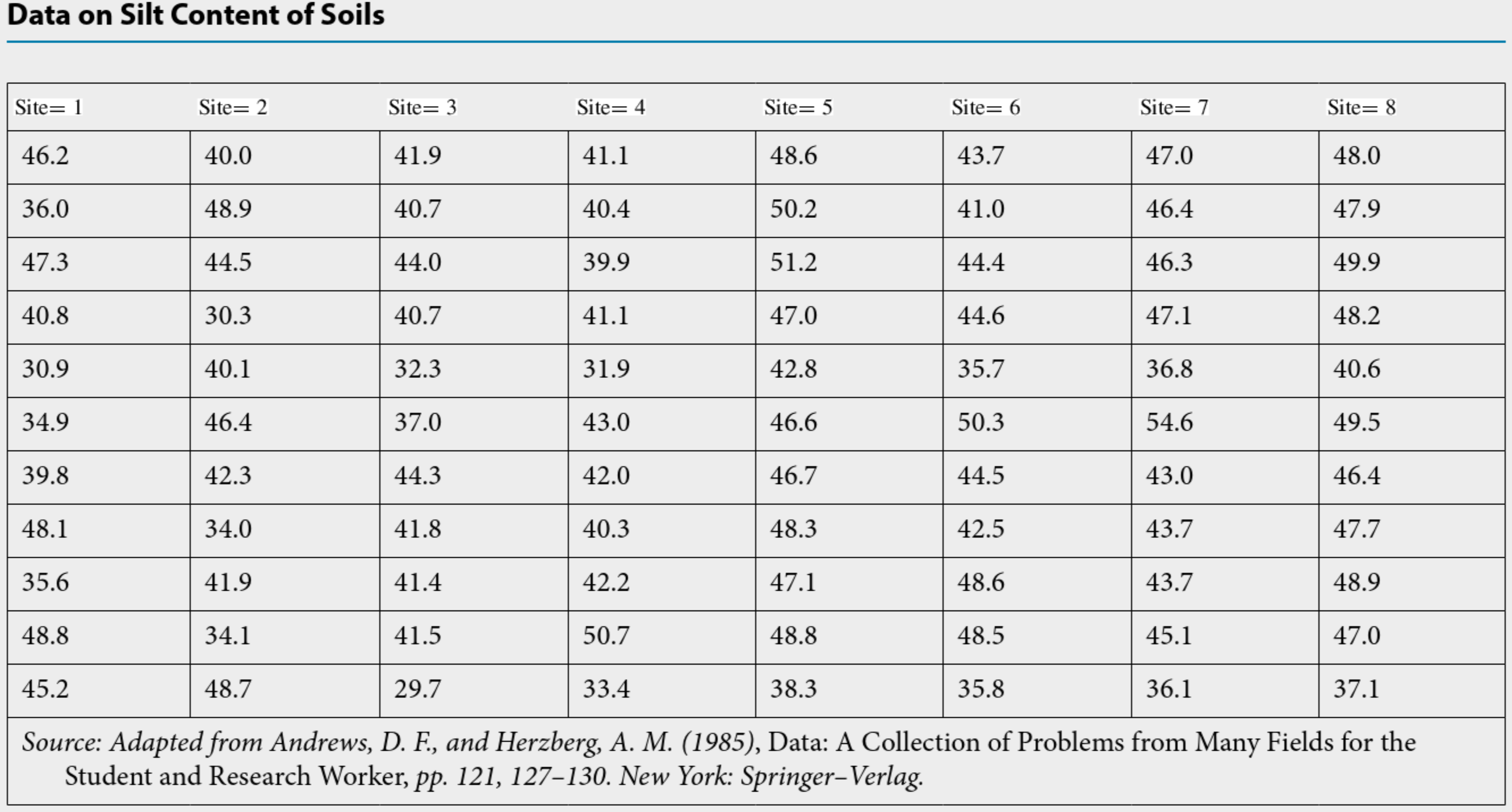
**STAT 441/541 Statistical Methods II**

**Silt Content of Soils**

**Scenario**

A study was done to compare soil mapping units based on their lateral variabilities for a single property, silt content. The study area consisted of a sequence of eight contiguous sites extending over the crest and flank of a low rise in a valley plain underlain by marl near Albudeite in the province of Murcia, Spain. The geomorphological sites were the primary mapping units adopted and were small areas of ground surface of uniform shape. Following the delimitation of the sites, soil samples were obtained in each site at 11 random points within a 10m x 10m area centered on the midpoint of the site. All samples were taken from the same depth. The soil property considered was the silt content, expressed as percentages of the total silt, clay, and sand content. One of the objectives is to determine if there is a difference in silt content among the soils from different sites.



**The R file is: Silt Content of Soils R Code.R**

**The dataset is Excel file: Silt Content of Soils.xlsx**

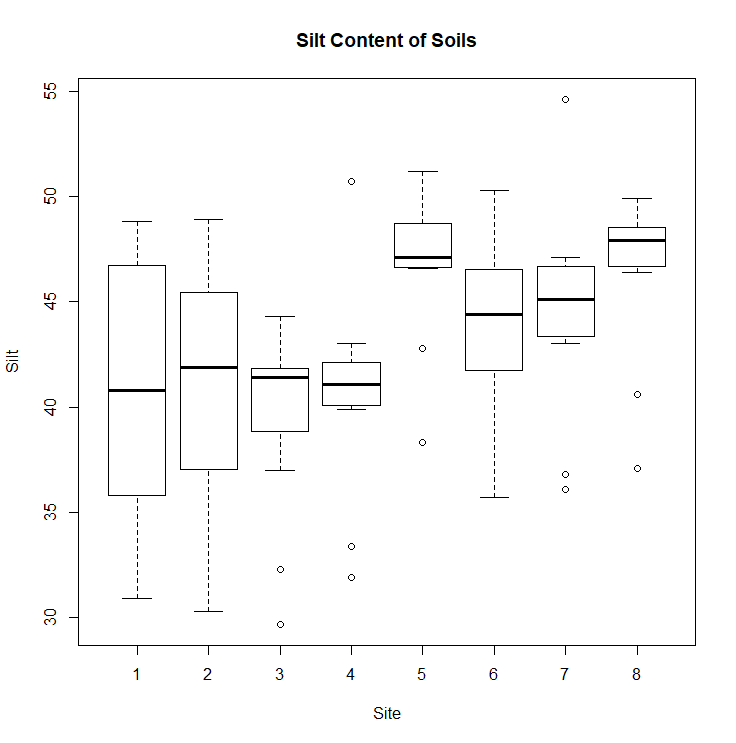
For this scenario, state the treatment effects model, describe all terms in the model, and give values for all subscripts:

Construct a table that gives the sample mean, variance, and number of observations for each site.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Site (Population)** | | | | | | | |
|  | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** |
| **Sample Mean** |  |  |  |  |  |  |  |  |
| **Sample Variance** |  |  |  |  |  |  |  |  |
| **Number of Observations** |  |  |  |  |  |  |  |  |

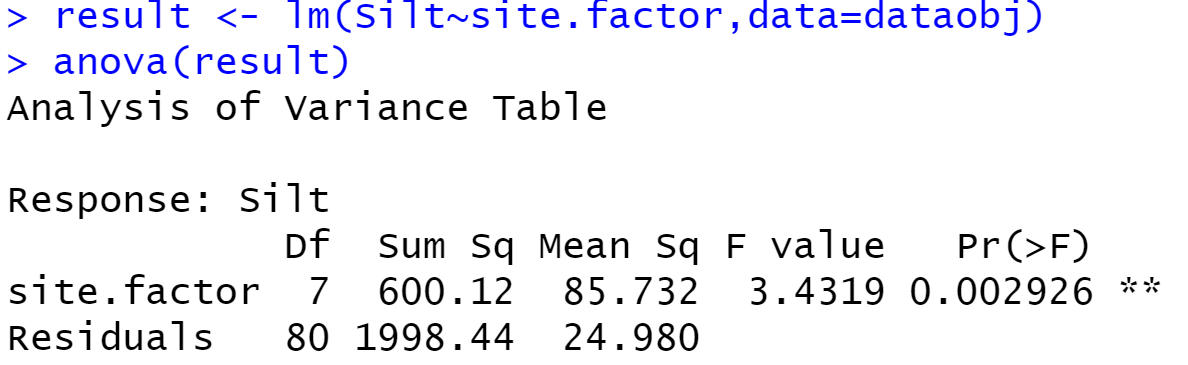
Why is it useful to examine summary statistics?

Next, let’s examine a boxplot of the data:



Why is it useful to examine a boxplot of the data?

Paste the Analysis of Variance table from R output and complete the following table:



Before completing the table, what is:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Source** | **df=degrees of freedom** | **Sum of Squares** | **Mean Square** | **F test** | **p-value** |
| Between Sites  (site.factor in R output) |  |  |  |  |  |
| Within Sites  (Residuals in R output) |  |  |  |  |  |
| Total (Corrected) |  |  |  |  |  |

Is there a difference in average silt content for the sites? Justify your answer using and use one-way ANOVA to test treatment effects. Perform the hypothesis test for the treatment effects model:

Hypotheses:

Test Statistic:

P-value:

Decision about the null hypothesis:

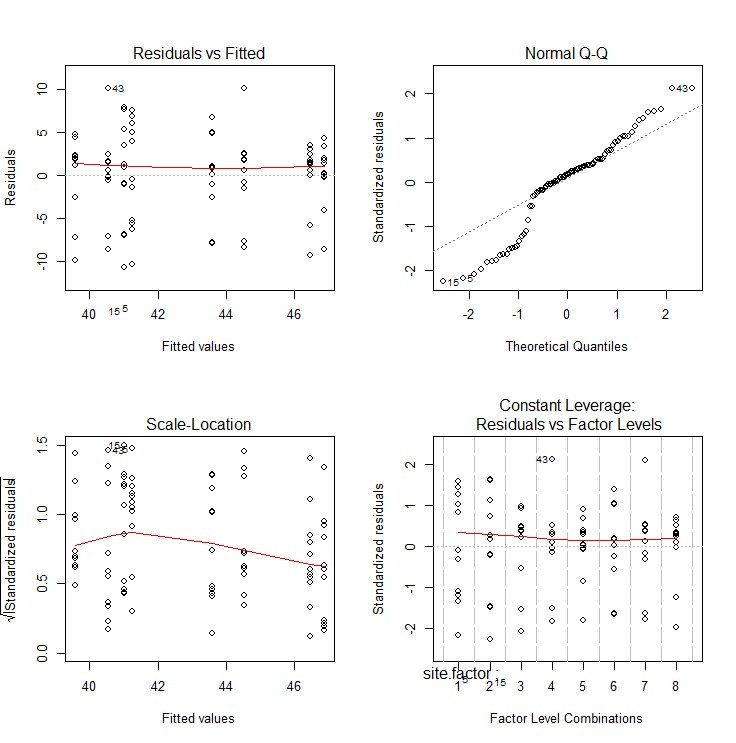
Conclusion:

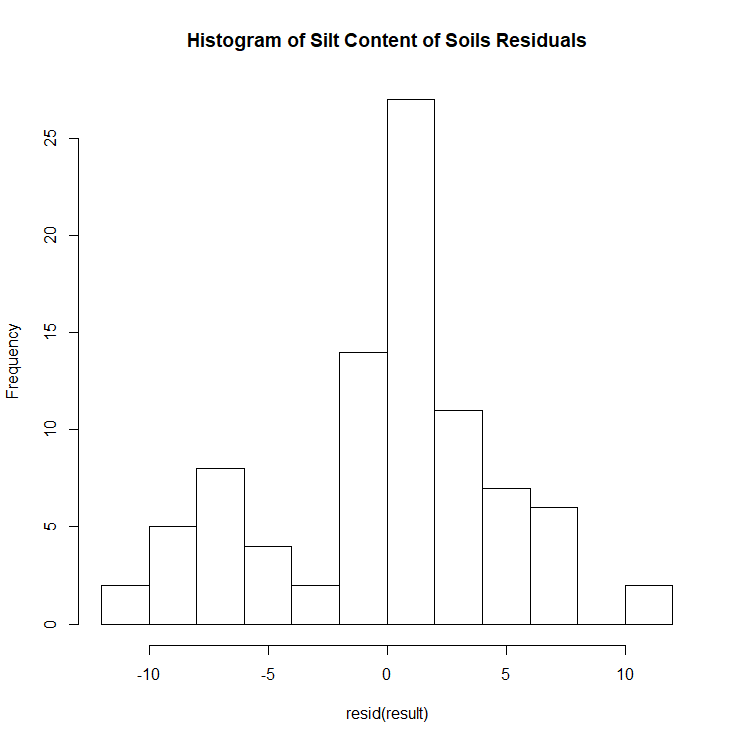
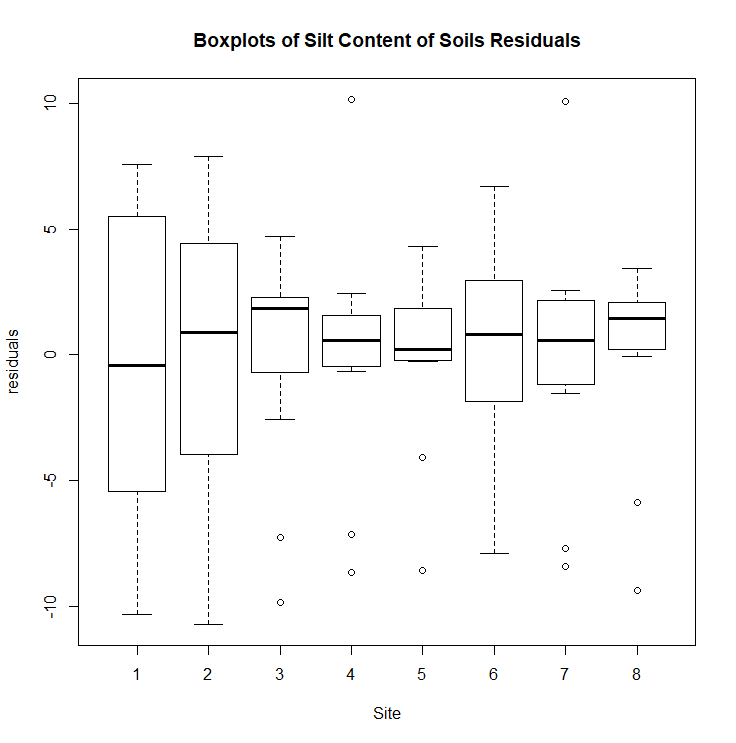
Check conditions for ANOVA based on the following assumptions:

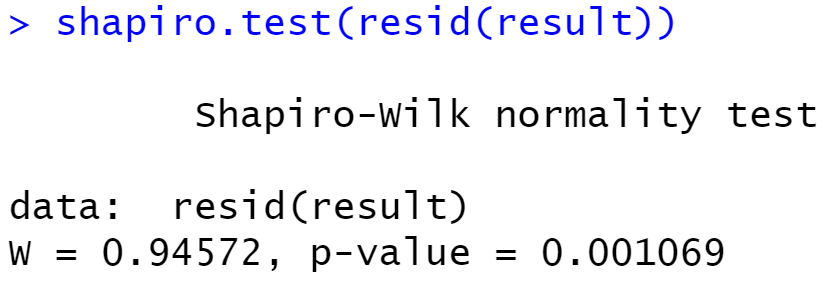
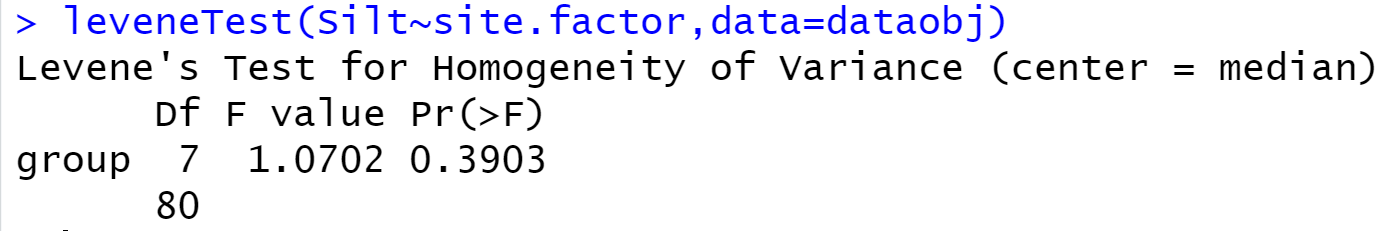
Errors are normally distributed

There are no outliers

The variance of the errors, , is the same for all treatments (common variance)



Have the assumptions been met? What is our overall conclusion?