Homework for Lesson 1

In this assignment we will generate some simple statistics in an exploratory data analysis (EDA) by hand and with SAS, and produce a means (or interval) plot using Minitab.

In the hypothetical greenhouse example there were six pots that were assigned (at random) to receive a fertilizer treatment. Each of these pots represents an *experimental unit*, which by definition is that which receives a treatment. The response variable is plant height (cm). The data is presented below in ‘unstacked’ format, a very common way that we encounter experimental data.

|  |  |  |  |
| --- | --- | --- | --- |
| F1 | F2 | F3 | Control |
| 32 | 22.5 | 28 | 21 |
| 30.5 | 26 | 27.5 | 19.5 |
| 25 | 28 | 31 | 22.5 |
| 27.5 | 27 | 29.5 | 21.5 |
| 28 | 26.5 | 30 | 20.5 |
| 28.6 | 25.2 | 29.2 | 21 |
|  |  |  |  |

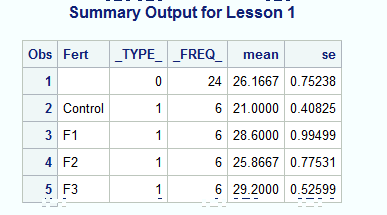
\*\*\* **Note that we are not asking for an ANOVA in this assignment**. \*\*\*

1) Compute the overall, or grand mean, of all observations (N=24), disregarding treatment levels. This can be done by hand, using a calculator, or using Excel.

Grand mean: 26.16666667

2) Compute the mean for each treatment level.

We get the SAS output for this data



So from the output we can say:

Treatment means are:

Control 21.0000

F1 28.6000

F2 25.8667

F3 29.2000

3) Compute the variance, standard deviation, and sample standard error for each treatment level.

Using output from excel and minitab:

Excel:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | F1 | F2 | F3 | Control |  |  |  |
|  | 32 | 22.5 | 28 | 21 |  |  |  |
|  | 30.5 | 26 | 27.5 | 19.5 |  |  |  |
|  | 25 | 28 | 31 | 22.5 |  |  |  |
|  | 27.5 | 27 | 29.5 | 21.5 |  |  |  |
|  | 28 | 26.5 | 30 | 20.5 |  |  |  |
|  | 28.6 | 25.2 | 29.2 | 21 |  |  |  |
| Mean | 28.6 | 25.86667 | 29.2 | 21 |  | Grand Mean | 26.16667 |
| Variance | 5.94 | 3.606667 | 1.66 | 1 |  |  |  |
| SD | 2.437212 | 1.899123 | 1.28841 | 1 |  |  |  |
| SE | 0.994987 | 0.775314 | 0.525991 | 0.408248 |  |  |  |

Minitab:

**Descriptive Statistics: F1, F2, F3, Control**

Variable N N\* Mean SE Mean StDev Minimum Q1 Median Q3 Maximum

F1 6 0 28.600 0.995 2.437 25.000 26.875 28.300 30.875 32.000

F2 6 0 25.867 0.775 1.899 22.500 24.525 26.250 27.250 28.000

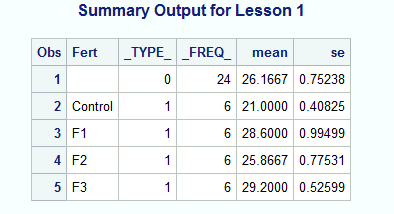
F3 6 0 29.200 0.526 1.288 27.500 27.875 29.350 30.250 31.000

Control 6 0 21.000 0.408 1.000 19.500 20.250 21.000 21.750 22.500

The highlighted area above has the answer i.e:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | F1 | F2 | F3 | Control |
| Variance | 5.94 | 3.606667 | 1.66 | 1 |
| SD | 2.437212 | 1.899123 | 1.28841 | 1 |
| SE | 0.994987 | 0.775314 | 0.525991 | 0.408248 |

4) Open SAS and use the code provided in this lesson (see SAS code for Lesson 1.docx) to run the Summary procedure to produce the means and standard errors for the treatment levels. The SAS code can be cut and pasted from the Word document into the Program Editor Window in SAS and simply run without modification. Note that the data is re-formatted in the SAS code – this is the ‘stacked’ format. You can extract the output entitled Summary Output for Lesson 1 from the Output Window, and then paste it into your HW1 document.



5) Make an interval plot using Minitab or a Bar chart using Excel. The bar chart or interval plot will be similar to the ones shown in the Lesson 1 Notes, except that a) the plot shown in the Lesson 1 notes was produced by SAS following the ANOVA, and b) you won’t be labelling the means with lettering (a, ab, b, c) from the mean comparison procedure. You should include this graph in your submission.



What to turn in for the Homework Assignment Drop Box: Please use the Drop Box provided in this Lesson to submit ONE document (preferably a Word document, but a pdf file or scan will also work).