

## LAB 1 ASSIGNMENT

### DISPLAYING AND DESCRIBING DISTRIBUTIONS

In this lab assignment, you will use Excel to display and describe observations on a single variable from several groups. In particular, you will use histograms and boxplots to display the data. Also, you will calculate summary statistics for the data like the mean, standard deviation, median, and interquartile range. Before you start working on the assignment questions, you should get familiar with the instructions provided in *Lab 1 Instructions*.

### Barbecue Grilling Time

A usual summer activity for several people, grilling food on a barbecue still needs to meet certain health standards while making cooking efficient, so the manufacturers of barbecues need to engineer appliances that function properly. Three brands were randomly selected by an independent research team: Greystone, Broil Queen, and BBBQ (the extra 'B' is for "biggest"). The research team then selected 17 pieces of steak for one barbecue of each brand, all the same thickness, but different masses due to their size, to measure the total cooking time for each piece. Each piece was cooked separately in the middle of the barbecue and each barbecue was turned off to cool down to room temperature before being turned on again for the next piece. Each piece was removed from the grill when it reached the desired internal temperature (145°F or 62.8°C), as measured with an instant-read thermometer. The total cooking time was recorded in minutes.

This dataset is available in the Data link located in the Lab 1 tab display in the Labs section on eClass. The data are not to be printed in your submission. Note that the data in the worksheet are provided in two alternative formats as one of those formats may be more suitable for some analyses. **Unless otherwise stated, numerical answers should be to three significant decimal places.**

The following is a description of the variables in the data file:

<u>Column</u>	<u>Name of Variable</u>	<u>Description of Variable</u>
1	Brand	BBQ brand (Greystone, Broil Queen, and BBBQ),
2	Time	total cooking time (in minutes),
3	Mass	mass (spaced evenly between 2.000 and 4.000 lbs.),
4	Bins	bin values to be used for histograms.

Use the data to answer the following questions:

1. Before carrying out statistical analysis for the data, examine the study design. Is this an observational study or an experiment? Are population inferences applicable? Causal inferences? Explain briefly.
2. First obtain the histograms of time for the three brands.
  - (a) Obtain a histogram of time for each of the three brand: Greystone, Broil Queen, and BBBQ. The format of each histogram should be the same as the format of the sample histogram in *Lab 1 Instructions* (title, names of axes, no gaps between bars). Histograms should initially be 17 cells high (each cell height = 13.20 pixels) and 7 cells wide (cell width = 8.11 pixels). You may reduce them (within reason) when pasting into your assignment. Use the same bins for all three histograms.
  - (b) Compare the shapes of the three histograms obtained in part (a). Are they unimodal or bimodal? Are they symmetric or skewed? If the latter, which way? Are there any outliers?
  - (c) Use the histograms to compare the centers (mode) and spreads (range) of the three distributions.
  - (d) Based on the histograms, what should be the relationship between the mean and the median for each distribution?

3. First obtain the boxplots of time for the three brands.
  - (a) Obtain side-by-side boxplots of time for the three brands. (See histogram dimensions in Question 2. Also, the minimum and maximum on the y-axis should be 9.4 and 11.5, respectively.) Paste the plot into your report.
  - (b) Compare the shapes of the three boxplots obtained in part (a). Are they symmetric or skewed? If the latter, which way? Are there any outliers? (Be careful if there are close/repeated values.)
  - (c) Use the boxplots to compare the centers (median) and spreads (IQR) of the three distributions.
  - (d) Are the conclusions for (b) and (c) consistent with the analysis from Question 2?
4. Now obtain some summary statistics for time for each of the three brands (Greystone, Broil Queen, and BBBQ) and for each of the 17 mass values. More precisely, do the following.
  - (a) Use the *Descriptive Statistics* tool (*Lab 1 Instructions*) to calculate the mean, standard deviation, mode(s), and range of time for each brand. Provide these values in a clear, organized table. Compare the means and standard deviations of the three distributions. Compare the modes and ranges of the three distributions. For the modes and ranges, are the conclusions consistent with the analysis from Question 2, part (c)?
  - (b) Use the *Insert Function* feature (*Quartile.Inc*) to compute the 5-number summary – minimum, first (lower) quartile, the second quartile (median), third (upper) quartile, and maximum – as well as the interquartile range of time for each brand. Provide these values in a clear, organized table. Does the 5-number summary for each brand show consistency with your conclusions about the shape of the corresponding distributions in Question 2 and 3? Explain briefly.
  - (c) Use the *Descriptive Statistics* tool to calculate the mean and standard deviation of time at each of the 17 mass values. Provide these values in a clear, organized table. How does the mean time change as mass increases? What are the highest and lowest values for mean time change? On average, by how much does time change as mass increases by 0.125 pounds? (Report the answer to five decimal places.)
5. What is the best brand? Answer the question (first in terms of variation and then by fastest time) by referring to the summary statistics and plots obtained in Questions 2-4. Does the trend identified in Question 4, part (c) impact the choice of best brand?

<b>LAB 1 ASSIGNMENT MARKING SCHEMA</b>
--

**Question 1 (6)**

Type of study: 2 points  
 Population inferences: 2 points  
 Causal inferences: 2 points

**Question 2 (28)**

- (a) Correctly formatted histogram: 4 points each (12 points total)
- (b) Analysis of the shape of each histogram: 3 points each (9 points total)
- (c) Comparing the centers and spreads: 2 points each (4 points total)
- (d) Relationship between the mean and the median: 1 point each (3 points total)

**Question 3 (16)**

- (a) Correctly formatted side-by-side boxplots of time versus brand: 4 points
- (b) Analysis of the shape of each boxplot: 2 points each (6 points total)
- (c) Comparing the centers and spreads: 2 points each (4 points total)
- (d) Consistency: 2 points

**Question 4 (51)**

- (a) Summaries (mean, std. dev., mode, range) for each of the three brands: 2 points each (6 total)  
Comparing the means and standard deviations: 2 points each (4 points total)  
Comparing the modes and ranges: 2 points each (4 points total)  
Consistency: 2 points each (4 points total)
- (b) Summaries (5-number summary and IQR) for each of the three brands: 3 points each (9 total)  
5-number summaries relating to shapes: 3 points
- (c) Table of means, standard deviations, and changes in mean time: 3 points per column (9 total)  
Effect of higher mass on mean time: 4 points  
Highest and lowest changes: 2 points each (4 points total)  
Average change in time as mass increases by 0.125: 4 points

**Question 5 (6)**

Best brand (variation, fastest): 2 points each (4 total)  
Impact of trend: 2 points

**TOTAL = 107**