

CSU101
Spring 2008
Lab Assignment 4

To complete this assignment you must submit an electronic copy to your lab instructor by the due date using Blackboard, and printed output handed-in in class on the first class after the due date. (Don't depend upon the printer in the lab at the last minute!)

Download the file **Lab4.xls**. Then rename the file as *yournameLab4.xls* and save a copy of it. The file contains records on a number of students in a class, with an ID, a height, and a weight, separated by sex. Our problem is to study the distribution of body mass index for men and for women in the population for which we have data, and to plot a histogram, a frequency polygon, and a cumulative frequency polygon, that will allow us to compare the two resulting distributions. A distribution tells you how many cases there are for various ranges of values of some characteristic.

Body Mass Index (BMI)

Body mass index, or BMI, is an index designed to allow physicians and researchers to study obesity. A person's BMI index is a function of his or her height and weight. The formula for calculation is:

$$\text{BMI} = (\text{weight} / \text{height}^2) * 703$$

where weight is expressed in pounds and height is in inches.

To calculate the risk associated with a BMI, researchers provide the following table:

Risk of Associated Disease According to BMI and Waist Size			
BMI		Waist less than or equal to 40 in. (men) or 35 in. (women)	Waist greater than 40 in. (men) or 35 in. (women)
18.5 or less	Underweight		
18.5 – 24.9	Normal		
25.0 – 29.9	Overweight	Increased	High
30.0 – 34.9	Obese	High	Very High
35.0 – 39.9	Obese	Very High	Very High
40 or greater	Extremely Obese	Extremely High	Extremely High

Notice the BMI ranges in this table are given to one decimal place. Since this table provides the only means of interpreting a BMI value, no greater precision is meaningful. Therefore, your spreadsheet should never use a BMI of greater precision. The table reflects current thinking about the significance of BMI as an index of health. You may find it interesting to calculate your own BMI and gauge the associated risks.

Our study proceeds in the following manner:

1. Define an expression to calculate each person's BMI. When you have the expression, copy it into E8:E107 for women, and K8:K98 for men. Round the values to only 1 decimal place. You have now calculated a BMI value for everyone in our population.
2. Assign names to ranges as follows:
 - a. E8:E107 FemaleBMI
 - b. K8:K98 MaleBMI
3. In D114:F122 you will find an array of cells in which you are asked to place functions that answer the following questions. Use the above range names in this process:
 - a. How many women and men, and how many total people in our study? This is often called the sample size.
 - b. What are the average and standard deviation of BMI for men, women, and combined in our population?
 - c. What are the minimum, maximum, and range of BMI for women, men and combined?
4. Since we want to find bin widths suitable for both men and women, we will calculate bin widths for the combined population only. Remember we are trying to divide the range for the combined population into 7 equal bins. The calculated bin width can be adjusted a bit to make it easier to think about. Your final value should go into F122.

Now construct a suitable bin table at B127:C133, where B127 contains the minimum BMI for our population and B133 contains one bin width down from the maximum in our population. Place suitable bin labels in C127:C133. For our purposes, suitable bin labels should be the mid-point of the corresponding interval, rounded to two decimal places, and should be calculated from its corresponding bin floor.

5. Assign the name BinTable to B127:C133, and use it in a VLOOKUP to assign bin values to each woman and man in our population. The women's values should appear in F8:F107, and the men's in L8:L98. Assign the range names FemaleBinVal and MaleBinVal, respectively, to these two ranges.
6. We will report our results in E127:J133. E127 should contain the number of cases in FemaleBinVal of bin values matching C127; E133 the same matching C133. F127 should contain the number of cases in MaleBinVal of values matching C127; likewise for F133 and C133. Remember to use appropriate functions to arrive at these numbers. When you have finished E127:F133, use these values to calculate the corresponding percentages. These should appear in G127:H133. Now calculate the cumulative percents for females and males. These go in I127:J133. Place suitable functions in E134:J134 to reflect the totals of their respective columns. This is a check to insure that values are not getting lost.

7. You are now ready to plot the histogram. Construct a chart, **a column chart to be placed in the spreadsheet**, using C127:C133 as your category labels, and using G127:G133 and H127:H133 as your two data series. Provide the title “Histogram of Female and Male BMI”, and label the data series “Women” and “Men”, respectively.
8. Now construct another chart of the same information, **a line chart with data points shown**. This time the chart title should be “Frequency Polygon of Female and Male BMI”. The x values should be the interval mid-points, i.e., the bin labels, and the y values should be the appropriate percentages. Remember, your chart should show two graphs.
9. Now construct another chart, a cumulative frequency polygon, which plots as its x values the interval endpoints, and as its y values the sum of the percentages up to and including the interval that x is in. Ranges have been laid out for this purpose in I127:J133. Use appropriate functions in those cells to calculate the cumulative percentages, so they can now be plotted. Your chart this time is **a line chart with data points shown**, and its title should be “Cumulative Frequency Polygon of Female and Male BMI”.
10. Finally, print your chart results as a four-page report, page 1 comprising A112:J134, page 2 your first chart, page 3 the second chart, and page 4 the third chart. Make sure your name and your instructor’s name appears in the footer of each page. The printed report can be handed in on the due date in class.