

Instructions

- Complete solutions will include good explanation of the process used to find the solution, and complete sentences for any discussion of the answer.
- For questions involving Excel files, email me the Excel file that shows your solution. Clearly label with appropriate headings the cells in the Excel file so I can understand your solution!

The Excel file you submit should be your own original creation.

Remember—when you are using Excel, use new sheets where appropriate (it is usually a good idea to leave original data unaltered in the first sheet, and make modifications on other sheets). Focus on making your Excel file easy for a reader to understand, labeling sheets in a meaningful manner, deleting unnecessary sheets or data, and use AutoShapes for text you wish to include. Also, provide headings or labels to cells that you add, and use color to improve readability.

Questions

1. (20 marks) There have been a series of four robberies of jewel stores recently in a police precinct. The police suspect the robberies may be the actions of two or more separate sets of thieves acting independently.

Mud was found at the scene of each robbery, believed to come from the shoes of the thieves.

In an effort to determine if two or more separate groups are responsible for the robberies, different samples of mud from each robbery is collected, and analyzed for chemical content. The results of this analysis are provided in the following table:

	% Aluminum Oxide	% Iron Oxide	% Magnesium Oxide
Robbery 1	14.2, 13.1, 14.5, 14.1, 13.9, 13.7, 14.3, 14.0, 13.7, 13.3	5.5, 2.3, 2.6, 0.9, 4.8, 6.4, 2.1, 2.8, 1.9, 1.8	3.5, 2.9, 2.6, 1.9, 3.8, 3.4, 2.7, 2.8, 2.3, 1.8
Robbery 2	14.0, 13.4, 14.2, 14.0, 13.6, 13.2, 14.1, 14.5, 14.3, 13.5	2.5, 2.5, 2.9, 1.9, 2.8, 6.3, 2.1, 2.8, 1.9, 1.8	2.1, 2.1, 1.9, 2.2, 2.0, 1.3, 1.5, 1.3, 1.6, 1.4
Robbery 3	13.8, 13.2, 14.2, 14.6, 14.8, 14.1, 14.5, 14.0, 13.9, 13.2	5.2, 4.3, 4.6, 4.9, 4.7, 5.9, 4.5, 3.9, 4.9, 3.8	3.2, 3.3, 2.6, 1.9, 2.7, 3.9, 2.5, 2.9, 2.9, 1.8
Robbery 4	13.8, 14.4, 14.3, 14.6, 13.8, 13.3, 14.2, 14.0, 13.6, 13.9	2.2, 2.4, 2.8, 2.0, 2.7, 0.3, 2.5, 2.3, 1.6, 1.4	2.2, 2.2, 2.1, 2.0, 1.7, 1.3, 1.5, 1.3, 1.6, 1.4

By constructing appropriate box plots from this data, determine if separate groups of thieves are responsible for the heists. Because you will be comparing box plots, it is important that they are sketched accurately on a well laid out scale. If there are more than one group, which groups did which heists?

This is a small data set, so you may be able to figure this out from staring at the data for awhile. Finding the solution this way would be more difficult for larger data sets. Regardless, the box plots will provide a good way to present your conclusions in a convincing way to others!

You can either create the box plots by hand, or use Excel to answer this question. If you do solve the problem by hand, make sure your plots are to scale, and drawn neatly.

2. (40 marks) (**Excel Based**) Salary data can be obtained from a variety of sources on the internet. One source is <http://www.deed.state.mn.us/lmi/tools/oes/default.aspx> where Minnesota salaries can be examined.

Let's say we want to learn more about a particular class of occupations, and what types of salaries these jobs earn in Minnesota compared to the USA as a whole.

Do **EITHER** Management **OR** Education, Training, and Library, **NOT BOTH**.

- **Management Occupations**

Go to the Department of Employment and Economic Development and download, using the OES Data Tool, the salary data for all the subcategories of SOC code 11 (Management Occupations) in Minnesota. You are looking for 31 job categories, including Chief Executives, General and Operations Managers, Legislators, etc.

Once you have the data, you can use Excel to compare the distribution of the earnings of the various occupations. The two distributions we will choose to look at are the

- median hourly income for Minnesota
- median hourly income for USA

of the various occupations under SOC Code 11, Management Occupations.

Note that the number of elements in these two distributions is different.

- **Education, Training, and Library Occupations**

Go to the Department of Employment and Economic Development and download, using the OES Data Tool, the salary data for all the subcategories of SOC code 25 (Education, Training, and Library Occupations) in Minnesota. You are looking for about 60 job categories, including Business Teachers Postsecondary, Computer Science Teachers Postsecondary, etc.

Once you have the data, you can use Excel to compare the distribution of the earnings of the various occupations. The two distributions we will choose to look at are the

- mean annual income for Minnesota
- mean annual income for USA

of the various occupations under SOC Code 25, Education, Training, and Library Occupations.

Note that the number of elements in these two distributions is different.

- a) Using Excel, find the five number summary of the the two distributions.
- b) Construct (by hand, or Excel if you are adventurous) the boxplots for the two distributions on the same plot.
- c) Using Excel, find the mean and standard deviation of the distributions.
- d) Using Excel, construct histograms of the distributions.
- e) Describe each of the distributions—are there outliers, are they skewed to the left or right, are they symmetric, etc. Compare the distribution for Minnesota to the distribution for the USA. Can these distributions be described by the mean and standard deviation, or would the five-number summary be better? Explain you answer.