3.4: Exercises

1. When copper metal and powdered sulfur are placed in a crucible and ignited, the product is a sulfide with an empirical formula of Cu_xS. The value of x is determined by weighing the Cu and the S before ignition and finding the mass of Cu_xS when the reaction is complete (any excess sulfur leaves as SO_2). The following table shows the Cu/S ratios from 62 such experiments (note that the values are organized from smallest-to-largest by rows). A copy of the data is available as a .csv file with data organized in a single column.

1.764	1.838	1.865	1.866	1.872	1.877
1.890	1.891	1.891	1.897	1.899	1.900
1.906	1.908	1.910	1.911	1.916	1.919
1.920	1.922	1.927	1.931	1.935	1.936
1.936	1.937	1.939	1.939	1.940	1.941
1.941	1.942	1.943	1.948	1.953	1.955
1.957	1.957	1.957	1.959	1.962	1.963
1.963	1.963	1.966	1.968	1.969	1.973
1.975	1.976	1.977	1.981	1.981	1.988
1.993	1.993	1.995	1.995	1.995	2.017
2.029	2.042				

(a) Construct a boxplot for this data and comment on your results.

(b) Construct a histogram and comment on your results.

2. Mizutani, Yabuki and Asai developed an electrochemical method for analyzing *l*-malate. As part of their study they analyzed a series of beverages using both their method and a standard spectrophotometric procedure based on a clinical kit purchased from Boerhinger Scientific. The following table summarizes their results. All values are in ppm.

Sample	Electrode	Spectrophotometric
Apple Juice 1	34.0	33.4
Apple Juice 2	22.6	28.4
Apple Juice 3	29.7	29.5
Apple Juice 4	24.9	24.8
Grape Juice 1	17.8	18.3
Grape Juice 2	14.8	15.4
Mixed Fruit Juice 1	8.6	8.5
Mixed Fruit Juice 2	31.4	31.9
White Wine 1	10.8	11.5
White Wine 2	17.3	17.6
White Wine 3	15.7	15.4
White Wine 4	18.4	18.3

Construct a scatterplot of this data, placing values for the electrochemical method on the *x*-axis and values for the spectrophotometric method on the *y*-axis. Use different symbols for the four types of beverages. The data in this problem are from Mizutani, F.; Yabuki, S.; Asai, M. *Anal. Chim. Acta* **1991**, *245*,145–150. A copy of the data is available as a .csv file.

3. Ten laboratories were asked to determine an analyte's concentration of in three standard test samples. Following are the results, in μg/ml.

Laboratory	Sample 1	Sample 2	Sample 3
1	22.6	13.6	16.0
2	23.0	14.2	15.9
3	21.5	13.9	16.9
4	21.9	13.9	16.9
5	21.3	13.5	16.7
6	22.1	13.5	17.4
7	23.1	13.5	17.5
8	21.7	13.5	16.8
9	22.2	12.9	17.2
10	21.7	13.8	16.7

(a) Construct a single plot that contains separate stripcharts for each of the three samples.

(b) Construct a single plot that contains separate boxplots for each of the three samples.

The data in this problem are adapted from Steiner, E. H. "Planning and Analysis of Results of Collaborative Tests," in *Statistical Manual of the Association of Official Analytical Chemists*, Association of Official Analytical Chemists: Washington, D. C., 1975. A copy of the data is available as a .csv file.

4. Real-time quantitative PCR is an analytical method for determining trace amounts of DNA. During the analysis, each cycle doubles the amount of DNA. A probe species that fluoresces in the presence of DNA is added to the reaction mixture and the increase in fluorescence is monitored during the cycling. The cycle threshold, C_t, is the cycle when the fluorescence exceeds a threshold value. The data in the following table shows C_t values for three samples using real-time quantitative PCR. Each sample was analyzed 18 times.

Sample X		Sample Y		Sample Z	
24.24	25.14	24.41	28.06	22.97	23.43
23.97	24.57	27.21	27.77	22.93	23.66
24.44	24.49	27.02	28.74	22.95	28.79
24.79	24.68	26.81	28.35	23.12	23.77
23.92	24.45	26.64	28.80	23.59	23.98
24.53	24.48	27.63	27.99	23.37	23.56
24.95	24.30	28.42	28.21	24.17	22.80
24.76	24.60	25.16	28.00	23.48	23.29
25.18	24.57	28.53	28.21	23.80	23.86

Use two or more methods to analyze this data visually and write a brief report on your conclusions. The data in this problem is from Burns, M. J.; Nixon, G. J.; Foy, C. A.; Harris, N. *BMC Biotechnol.* **2005**, *5:31* (open access publication). A copy of the data is available as a .csv file.

5. The file problem3_5.csv contains data for 1061 United States pennies organized into three columns: the year the penny was minted, the penny's mass (to four decimal places), and the location where the penny was minted (D = Denver and P = Philadelphia). Subset the data by year into three groups

- pennies minted before 1982
- pennies minted during 1982
- pennies minuted after 1982

Plot separate histograms for the masses of the pennies in each group and comment on your results. The data in this problem was collected by Jordan Katz at Denison University and is available at the Analytical Sciences Digital Library's Active Learning website.

6. Use the element data you created in Exercise 1.3.1 to create several visualizations of your choosing. At least one of your visualizations should be a scatterplot and one should be a boxplot.

7. Use the data set you created in Exercise 2.3.2 on the daily average NOX concentrations and daily average temperatures recorded at a roadside monitoring station located on Marlybone Road in Westminster. Use this data to prepare a scatterplot that shows the daily average NOX concentrations for January on the *y*-axis and the daily average temperature for January on the *x*-axis. Add to this plot, a second scatterplot that shows the daily average NOX concentrations for July on the *y*-axis and the daily average temperature for July on the *x*-axis. Comment on your results.

8. Use this link to access a case study on data analysis and complete the nine investigations included in Part II: Ways to Visualize Data.
