MA 315 – Lab – Analyzing Ozone Data

Find the EPA's air quality data site by googling: "epa air quality data"

Go to their "Download Data" page and click on "Download Daily Data"

For the year 2010, compare daily ozone levels in Salt Lake City, Utah with Portland, Oregon. For each location, choose data from the first "Monitor Site." Choose to "Exclude exceptional events data." The data we will analyze in this lab is the daily maximum 8-hour ozone concentration in parts per million (ppm). In your lab report, please round off all responses to three significant figures.

Part I – Ozone Pollution in Portland, Oregon, 2010.

1. Compute the mean, standard deviation, and sample size.
2. Compute a five number summary.
3. Create a histogram.
4. Create a normal probability quantile-quantile plot.
5. Could this data be normally distributed, or is it clearly not normally distributed? Explain and justify your conclusion.
6. Create a 95% confidence interval.
7. Write a sentence or two explaining the meaning of your confidence interval.

Part II – Ozone Pollution in Salt Lake City, Utah, 2010.

1. Compute the mean, standard deviation, and sample size.
2. Compute a five number summary.
3. Create a histogram.
4. Create a normal probability quantile-quantile plot.
5. Could this data be normally distributed, or is it clearly not normally distributed? Explain and justify your conclusion.
6. Create a 95% confidence interval.
7. Write a sentence or two explaining the meaning of your confidence interval.

Part III – Comparing Ozone Pollution, Portland vs S.L.C., 2010.

Assuming that alpha = 0.05, do a formal hypothesis test to determine which city had better ozone pollution levels in 2010.

8. Should we do a one sided or a two sided hypothesis test? Explain your thinking.
9. What is the null hypothesis?
10. What is our test statistic?
11. What are the critical t values?
12. What is the p value of our result?
13. Explain the meaning of this p value: What exactly is it the probability of?
14. Create a 95% confidence interval on the difference between the two cities.
15. Write a sentence or two explaining the meaning of your confidence interval.
16. Are the results statistically significant? Explain.
17. Do we reject the null hypothesis? Explain.
18. Draw a conclusion, explain what it means in this context, and explain how these three methods of analysis (test statistic, p value, and confidence interval) all lead you to the same conclusion.
19. The analysis we have performed in this lab assumes that the daily ozone levels in these cities give us a simple random sample. Discuss the validity of this assumption, and its impact on our conclusions.