**DAPT 621 Assignment 1: Old Faithful Geyser (parts 1 &2)**



This project will involve obtaining and analyzing data for the interval of time between eruptions of the Old Faithful Geyser in Yellowstone National Park. The first part will be done individually. It will verify your knowledge of prerequisite material and prepare for the second part that is to be done by each team. The team will respond to a scenario described on the second page.

Below is some information about the interval between eruptions that was found online from different sources on September 4, 2014.

“Its average interval between eruptions varies from 60 - 110 minutes. … It was named for its consistent performance by members of the Washburn Expedition in 1870. Although its average interval has lengthened through the years (due to earthquakes and vandalism), Old Faithful is still as spectacular and predictable as it was a century ago.” http://www.nps.gov/features/yell/tours/oldfaithful/oldfaith.htm

“Intervals between eruptions can range from 45 to 125 minutes, averaging 66.5 minutes in 1939, slowly increasing to an average of 90 minutes apart today. The time between eruptions has a bimodal distribution.” http://en.wikipedia.org/wiki/Old\_Faithful

**”There is no prediction for Old Faithful Geyser. Sometimes Rangers are unable to make predictions because of environmental factors. Geysers are also natural phenomena and change behavior. The next prediction will be made when it is possible to do so. We apologize for the inconvenience.”**

**Later in the day this site had a prediction time with a ±10 minutes.**

**http://www.nps.gov/features/yell/webcam/oldFaithfulStreaming.html**

Individually you are to find obtain appropriate data to perform the analyses indicated below (Your individual data can and probably will differ from the data the team chooses to use.):

1. Use your individual data to create a graphical display to show the distribution of the intervals between eruptions. For this graph place the interval between eruptions on the vertical axis and use the horizontal axis to indicate the order of observed eruptions.
2. Use your individual data to create a 95% confidence interval for the applicable interval between eruptions of the Old Faithful Geyser.
3. Use your data to test H0: µ = 90 minutes versus HA: µ ≠ 90 minutes that would test the claim from Wikipedia that the time between eruptions has an “average of 90 minutes apart today.”
4. Use the sample mean and sample standard deviation from your data and assume that the distribution of times between eruptions is normally distributed to find the probability of the interval being between 80 and 100 minutes.
5. Find the proportion of intervals between eruptions in your data that are between 80 and 100 minutes. Compare this proportion with the probability using the normal distribution and comment on any difference you observe between the two.

The individual assignment is to be submitted by **9 AM, Monday, September 15, 2014** in an Excel file with your data on the first tab and put each of the five parts above on a separate tab numbering them 1 through 5. Send an e-mail to **dapt.stat@gmail.com** attaching your Excel file that includes your last name in the file name. This individual work should provide preliminary information for your work on the team. All of the work submitted for parts 1 through 5 is to be your own work, understanding that you are obtaining data recorded by someone else. You may confer with others about how to do something but you are to do the work. You may have consultants about how to do the work, but **subcontracting the work is an Honor Code violation**.

**Scenario description and assignment for the team portion**

DAPT Tours runs bus tours that go through Yellowstone National Park. From June 1 through September they will operate a bus daily that will arrive at the Old Faithful Visitor Center at 5 PM and depart at 7 PM to take the group to their hotel for the night. At 11 AM the tour guide has to call to reserve a time for the group to eat dinner at the Old Faithful stop and schedule the dinner for either 5, 5:30, 6 or 6:30 PM. The room for the meal does not have a view of Old Faithful and it takes about 30 minutes for dinner leaving 90 minutes of free time for those on the tour. By going to <http://geysertimes.org/getGeyserInfo.php?geyserID=2> the tour guide is able to find the times of the most recent eruptions of Old Faithful. Your team is to serve a consulting role for DAPT and determine a procedure for using available information such as the most recent eruption times at geysertimes.org for selecting the best time for dinner from the four previously indicated times of 5, 5:30, 6 or 6:30 PM. Old Faithful may be reasonably regular with its eruptions but there is still a good deal of variation in its eruptions. You are to develop a way to calculate the probability for Old Faithful erupting during the 90 minutes of free time based on the dinner time you recommend choosing. DAPT Tours obviously wants those on the tour to have a good experience and seeing an Old Faithful eruption would be highly desirable.

Your team will use available information and make a recommendation to the DAPT CEO, who has no training or experience in data analysis or quantitative methods. On **Friday, October 3**, your team will make a verbal presentation of your proposal to the CEO (**5 minute maximum**). The purpose of this presentation is to provide a reasonable procedure for making the dinner decision each day along with convincing evidence to support your procedure. A portion of the support should be about estimating the probability of an Old Faithful eruption in the 90 minutes of free time if your recommended procedure is followed for selecting the dinner time. In addition to the presentation you will submit a written report to accompany the presentation. The first page will contain your group name and a one-page executive summary that the CEO will be able to read quickly and understand clearly. The following pages in this report will describe how you arrived at your recommended procedure for selecting the dinner time and the method for predicting that Old Faithful will erupt during 90 minutes of free time that would result by following your recommendation. A printed copy of the entire document will be submitted at the time of the presentation and the associated Word document will be submitted by your team on Blackboard.

On Friday, October 3, after the presentations for the DAPT CEO, each group will make a second more technically oriented presentation to the rest of the class to present the process steps and thinking your team used as you attacked this problem (**15 minute maximum**). The audience for this presentation is the other members of the class, including the instructor, who have an understanding of basic statistical analysis and thinking. This is clearly for a different audience from the DAPT CEO who has no training or experience in data analysis or quantitative methods.