Stat414; **FALL 2024**; Assessment Instructions

INSTRUCTIONS

I. You are expected to work on the assessments without the help of the discussion threads. You are expected to work on the assessments by yourself without any outside human or internet help. You will be signing an honor code to this effect.

II. Each assessment is based on two or three modules. Questions are based on the material supplied in the module folders. You may consult these materials while answering the questions.

III. You may have to use R (Rstudio plus any other package needed). R output along with the screenshot of your R session should be included in your submission.

IV. Once you open the submission link, you will have one hour to complete the assessment. Unlimited attempts are enabled. Please submit after completing each question. Your last attempt uploaded, containing ALL your completed work will be graded. (Only one upload per student to be graded by the grader)

V. You need to submit

- .Rmd file showing all your codes
- .pdf file showing the report generated by your .Rmd file
- .RData file which saves all commands and clicks of your session
- a video log of your entire session. Here's are three short videos showing exactly how to do this, along with your own face showing in the corner throughout the session. Click on the Zoom support link below if you are looking for more techincal details on how to do this.

  `https://youtu.be/Le216n4lNhw`

  `https://youtu.be/LfQYZ_3gCb4`

  `https://youtu.be/MLyW0JVlT98`

  Zoom Support for Local Recording

Stat414; **FALL 2024**; Assessment 01; 100 Points;

CHECKLIST

Item	Points	Included?
.Rmd	Your Code	
.pdf	Your output	
.RData	Your R session	
.mp4	Your Assessment session	

SCORE

Question	Points	check
Q1a	20	
Q1b	20	
Q1c	20	
Q2a	20	
Q2b	20	
Total	100	

1. Arsenic concentrations (ppb) collected quarterly at two groundwater monitoring wells. These data are stored in the data frame `EPA.92c.arsenic3.df`.

   (a) For each well, plot the observations by year. Do you see any major differences between years?

   (b) Compute summary statistics for each well (combine years).

   (c) Compare the observed distribution of arsenic at each well. Use whatever types of plots you wish. Does the compliance well appear to show any evidence of contamination? Why or why not?

2. Consider only the data from the Background wells for this part.

   (a) Does it appear that the background well data may be modeled as coming from a normal distribution? Support your conclusions with plots and tests of normality statistics.

   (b) Find the value of λ corresponding Box-Cox transformation of the data to normality. Plot a histogram of the transformed data and overlay the appropriate Normal curve to illustrate the quality of fit of this model. Repeat the above exercise with the original (non-transformed) data and illustrate how the transformation improves the normality fit.