Everyone,

The first intellectual festival (AKA first exam) will cover all of the material in the first five chapters of the book and will have questions from the following list:

1) 18 multiple choice questions. These tend to be questions related to specific facts (e.g., definitions & symbols) or very short calculations.

2) an IVPPSS.

3) calculate (by hand, showing your work) the standard deviation for a short (~6) list of numbers.

4) calculate (by hand, showing your work) the median/IQR for a short (~20) list of numbers.

5) A univariate EDA for categorical or quantitative data. This will require you to extract information from R output that I will provide -- including a histogram and summary statistics for the quantitative case and a frequency and percentage table for the categorical case. You will NOT need to use R commands to produce histograms, summary statistics, boxplots, frequency tables, percentage tables, or bar charts.

6) Four normal distribution questions (one background, four questions related to that background). These questions will require you to be able to use `distrib()` in RStudio to produce output to answer the questions. I will provide a reminder of the arguments to `distrib()` that will look like that shown further below.

7) A bivariate EDA for quantitative data. This will require you to extract information from R output that I will provide -- including a scatterplot and correlation coefficient. You will NOT need to use R commands to produce scatterplots or correlation coefficients.

8) Two (out of four) short answer (paragraph-length) questions. These are generally related to major concepts, several of which I have rather pointedly hinted at during class.

```
library(NCStats)
distrib(x,mean=##,sd=##,lower.tail=XXXXX,type="X")
```

where **x** is replaced with the value of the quantitative variable or the area

mean=## has ## replaced by the value of the mean

sd=## has ## replaced by the value of the standard deviation

lower.tail=XXXXX has XXXXX replaced with TRUE (the default) for a "left-of" and FALSE for a "right-of" calculation

type="X" has X replaced with p (the default) for a forward and q for a reverse question

---

The exam is closed book, closed notes, etc. You should bring a calculator and you MUST use a pencil. Exams written in red or purple ink will not be accepted. In full disclosure -- there will be multiple versions of the exam so please do not embarrass yourself, and earn an "F" for the class, by cheating from your neighbor. Your RStudio must open with either no scripts or a blank script – i.e., the upper left pane must not have any previously entered R code in it. Opening any other software or previous script will result in an automatic "F" for the class. In addition, I may monitor computer usage and you will only be allowed to open RStudio.

The exam will start promptly at the beginning of class (on the hour) and will end promptly 1 hour and 50 minutes later. Please let me know ASAP if you have any conflict with staying 20 minutes past our usual class ending time.

To give you a feel for the exam, the directions for several major sections of the exam are below.

Multiple Choice [XX pts] -- choose the ONE BEST answer for each question by writing the corresponding letter in the blank to the left of the question

Answer the following question in the space provided. Please be as specific as possible.

Answer the following two questions in the space provided with your final answer clearly identified (e.g., circled). You must show all of your work to receive full credit (i.e., just providing the final answer will not receive full credit).

Use `distrib()` in RStudio to produce the result(s) needed to answer the next question. Write your answers with complete sentences with the code used to produce the result below your sentence.

Complete a thorough EDA that is appropriate to the type and number of variables in each of the following three questions. Your answer should be written with complete sentences.

Short (Paragraph) Answers -- Answer <u>TWO</u> of the following questions with complete sentences. Make sure to clearly label which questions you chose to answer. Each question is worth 3 points.