Show **all** of your work on this assignment and answer each question fully in the given context. You will want to understand Exercise 1 from Section 2.1 before attempting the following questions. Your answers should be written in complete sentences. It is possible that a drawing or table may help make your thoughts more concrete or illustrate a concept that would be difficult to describe in words alone - if so I encourage you to use one.

Please staple your assignment!

- **Problem1: Chapter 2, Section 3, Exercise 1 (page 47)[5 pts]**
  Possible controlled variables are: operator, launch angle, launch force, paper clip size, paper manufacturer, plane constructor, distance measurer and wind.
  The response is Flight Distance.
  The experimental variables are Design, Paper Type, and Loading Condition.
  Concomitant variables might be wind speed and direction(if these cannot be controlled), ambient temperature, humidity and atmospheric pressure.

- **Problem2: Chapter 2, Section 3, Exercise 5 (page 47)[5 pts]**
  For the delta/ construction/ with clip condition(for example), flying the same plane twice would provide information about flight-to-flight variability for that particular plane. This would be useful if you are only interested in making conclusions about that particular plane. If you are interested in generalizing your conclusions to all delta design planes made with construction paper and loaded with a paper clip, then re-flying the same airplane does not provide much more information. But making and flying two planes for this condition would give you some idea of variability among different planes of this type, and would therefore validate any general conclusions made from the study. This argument would be true for all 8 conditions, and would also apply to comparisons made among the 8 conditions.

- **Problem3: Chapter 2, End of chapter exercise, Exercise 7 (page 65)[10 pts]**
  (a) Label the widgets $1, 2, \cdots, 354$. Select a random sample of size five either through software, random table or even on RANDOM.org. One possible sample is the widgets labeled 121, 91, 134, 313, and 249.

  (b) Label the 12 experimental runs as follows:

Labels	Level of A	Level of B
1, 2	1	1
3, 4	2	1
5, 6	1	2
7, 8	2	2
9, 10	1	3
11, 12	2	3

  Use the following coding for the test labels:

Table Number	Test Label
01-05	1
06-10	2
11-15	3
16-20	4
21-25	5
26-30	6
31-35	7
36-40	8
41-45	9
46-50	10
51-55	11
56-60	12

Now, moving through Table B-1 choosing two digits at a time, ignoring numbers between 61 and 00, and those corresponding t runs that have already been picked. Order runs in the same order that their corresponding 2 digit number are picked from the table. Using this method, an starting from where I left off in part (a), the order would be 3, 2, 11, 7, 6, 1, 4, 9, 12, 5, 8, 10.

Note: The whole process of taking the sample can also be done through software or RANDOM.org as was done in part (a).

- **Problem4: Chapter 2, End of chapter exercise, Exercise 11 (page 65)[5 pts each part]**
(a) Control the extraneous variable heat by using only one bar for the entire study. This will eliminate any heat-to-heat variability.
(b) Advantage: may reduce baseline variation (background noise) in the response, making it easier to see any difference between the 2 brands.
Disadvantage: One heat may not be representative of all such material that the drills would be used on. Controlling makes the expriment more artificial- It will be harder to generalize conclusions from this heat to others.

- **Problem5: Chapter 3, Section 2, Exercise 1 (page 92)[5 pts each part]**

1. (a)

i	$\frac{i-.5}{10}$	$Q_L(\frac{i-.5}{10})$	$Q_C(\frac{i-.5}{10})$
1	.05	.84	.46
2	.15	.86	.67
3	.25	.87	.69
4	.35	.88	.73
5	.45	.88	.77
6	.55	.91	.78
7	.65	.92	.79
8	.75	.93	.80
9	.85	.95	.85
10	.95	1.15	.89
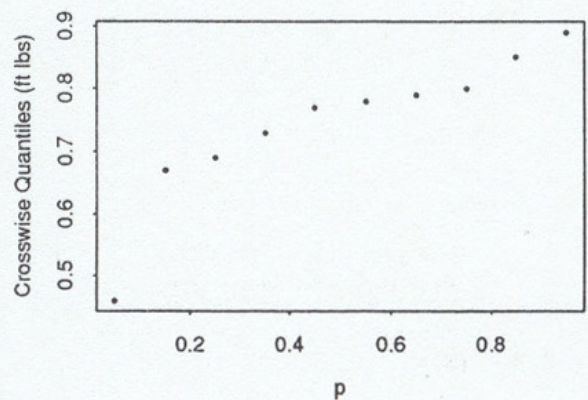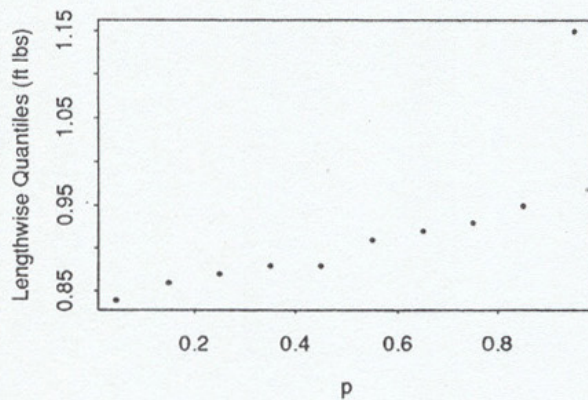
For the lengthwise sample:

$$Median = Q(.5) = \frac{.88 + .91}{2} = .895$$
$$1st\ Quartile = Q_1 = Q(.25) = .870$$
$$3rd\ Quartile = Q_3 = Q(.75) = .930$$
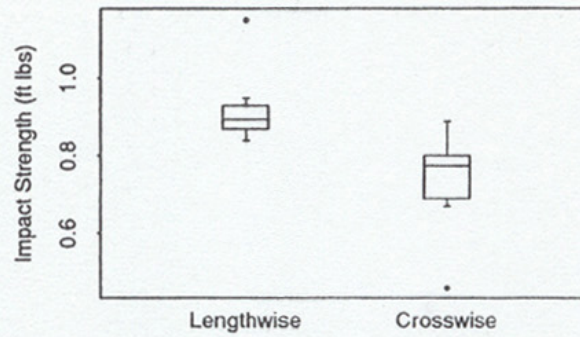$$Q(.37) = .880$$

For the crosswise sample:

$$Median = \frac{.77 + .78}{2} = .775$$
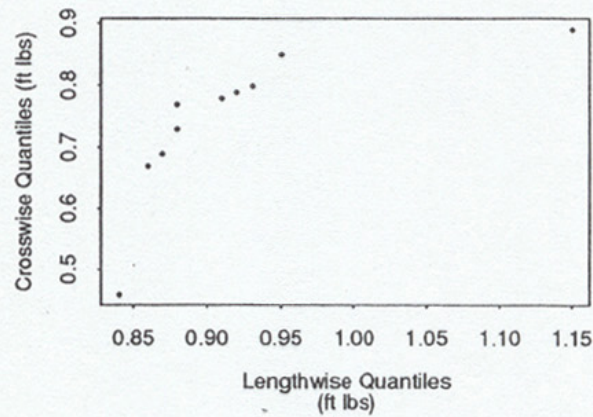$$Q_1 = .690$$
$$Q_3 = .800$$
$$Q(.37) = (.8)(.73) + (.2)(.77) = .738$$

(b)

On the whole, the impact strengths are larger and more consistent for lengthwise cuts. Each method produced an unusual impact strength value (outlier).

(c)

The non-linearity of the Q-Q plot indicates that the overall shapes of these two data sets are not the same. The lengthwise cuts had an unusually large data point ("long right tail"), whereas the crosswise cuts had an unusually small data point ("long left tail"). Without these two outliers, the data sets would have similar shapes, since the rest of the Q-Q plot is fairly linear.

- **Problem6: JMP Assignment. [5 pts each part]**

Computing is one of the most important parts of modern data analysis. A large part of data science simply wouldn't exist without the tools developed by scientists working at the intersections of computer science, mathematics, and statistics. Because of that, there will inevitably be parts of this course where a statistical computing tools are needed. SAS and R are the two main languages used by statisticians, though Python, Julia, F#, C++ and others are also common. SAS has a software called JMP ("Jump") that makes doing statistical analyses simpler - it is more powerful than Excel or your calculator but requires little in the sense of coding making the learning curve much lower. We will be using it this semester. There are labs in Snedecor Hall with the software pre-installed, but it is free for students and I encourage you to download a copy for yourself using the link below.

Download: https://www.stat.iastate.edu/statistical-software

In this problem, we will work through the tutorial found at http://web.utk.edu/~cwiek/201Tutorials/. Download and install `JMP` or find a computer with it already installed. Once you have done this, complete the following sections from the the tutorial linked above. For each part, print and include the table or graph produced. (Note: You can save the plots/tables and combine them in a single document).

NOTE: You should use the data set available on the tutorial to finish this exercise, but if you have any data sets of your own, I'd be more than happy to see you use JMP to make some summarizations on your data set.

1. Creating a JMP data table
2. Bar Chart
3. Pie Chart
4. Mosaic Plot
5. Histogram and Box Plot
6. Stem and Leaf Plot
7. Side-by-Side Box Plots
8. Calculating Summary Statistics of Quantitative Data

Total: 90 pts