STAT 305 D Exam 1
Show all your work.

1. (20 points) Caustic stress corrosion cracking of iron and steel has been studied because of failures around rivets in steel boilers and failures of steam rotors. A new teflon coating may reduce the corrosion behind this cracking. 10 steel bars were taken from manufacturer A, which uses the teflon coating, and 10 were taken from manufacturer B, which does not. Constant load stress corrosion tests (with constant load and constant stress) were applied for the same length of time to each bar. The length of the longest crack in μm was measured for each.

    a. (3 points) Identify the sample (or samples).

    b. (3 points) Identify the population (or populations).

    c. (3 points) Identify and classify all the variables.

    d. (3 points) Is this study an experimental study or an observational study?

    e. (4 points) Suppose the teflon-coated bars corrode and crack less for both steel and iron bars. Can we say that the teflon *causes* this reduction in corrosion and cracking? Why or why not?

    f. (4 points) Suppose the teflon-coated bars corroded and cracked less than the ones that did not receive the teflon. Can we say that the teflon prevents corrosion and cracking? Why or why not?

2. (20 points)

    a. (10 points) Using the table of random digits below, select a simple random sample of 10 steel bars from a shipment of 100 from manufacturer A (see question 1). Also, select a simple random sample of 10 steel bars from a shipment of 100 from manufacturer B, continuing in the table of random digits from where you left off from the steel bars. Carefully describe how you did this.

    | 27252 | 37875 | 53679 | 01889 | 35714 | 63534 | 63791 | 76342 | 47717 | 73684 |
    | 93259 | 74585 | 11863 | 78985 | 03881 | 46567 | 93696 | 93521 | 54970 | 37601 |
    | 84068 | 43759 | 75814 | 32261 | 12728 | 09636 | 22336 | 75629 | 01017 | 45503 |
    | 68582 | 97054 | 28251 | 63787 | 57285 | 18854 | 35006 | 16343 | 51867 | 67979 |
    | 60646 | 11298 | 19680 | 10087 | 66391 | 70853 | 24423 | 73007 | 74958 | 29020 |

    b. (10 points) Suppose you want to vary the temperature (high or low) and see how the bars respond to the stress tests. Using a different table of random digits (below), randomize the 10 bars from manufacturer A to the two levels of temperature (5 high, 5 low).

Then, do the same with the 10 bars from manufacturer B, continuing in the table of random digits from where you left off from the steel bars. Carefully describe how you did this.
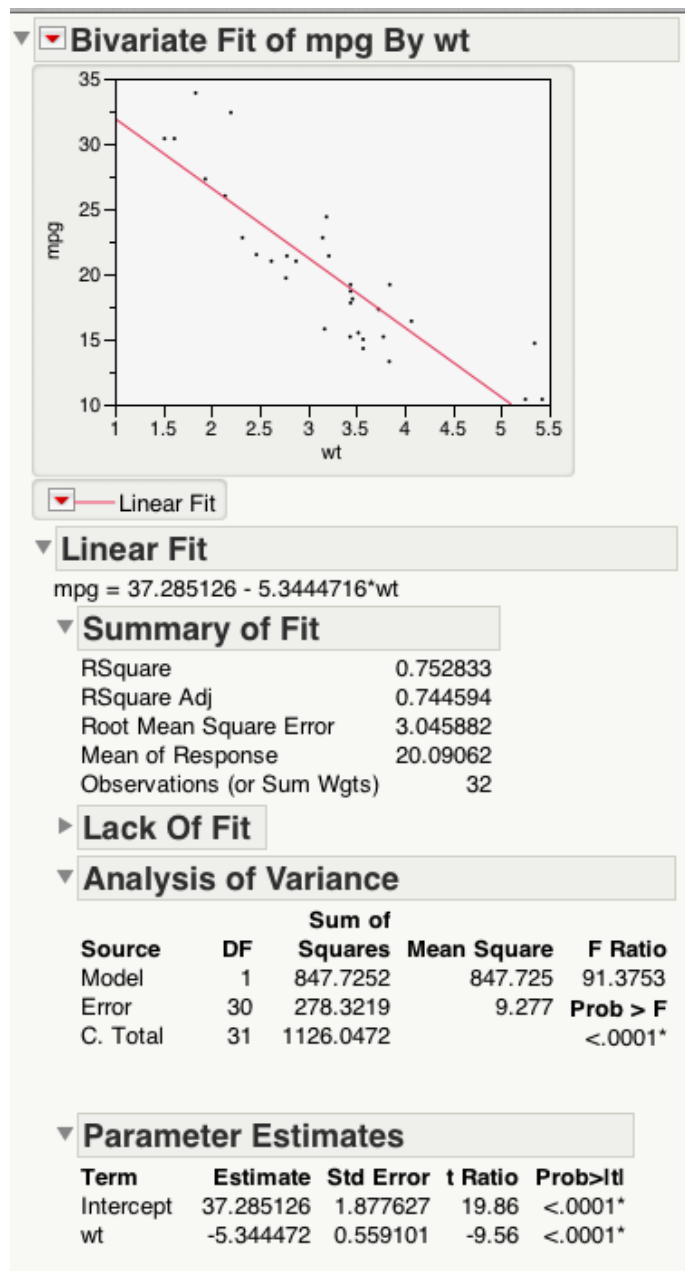
Random Digits

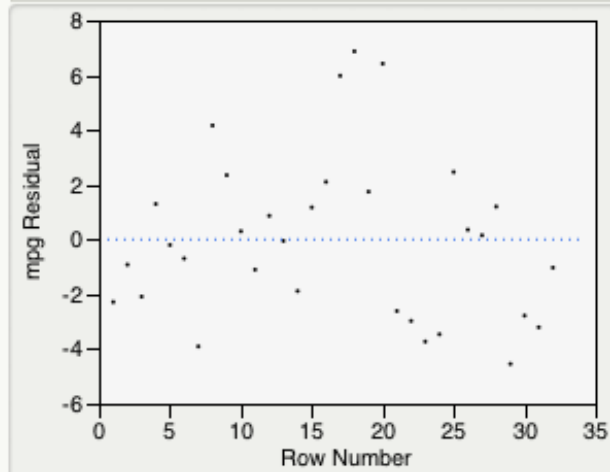|									
12159	66144	05091	13446	45653	13684	66024	91410	51351	22772
30156	90519	95785	47544	66735	35754	11088	67310	19720	08379
59069	01722	53338	41942	65118	71236	01932	70343	25812	62275
54107	58081	82470	59407	13475	95872	16268	78436	39251	64247
99681	81295	06315	28212	45029	57701	96327	85436	33614	29070

3. (20 points) Revisit the cars data from class. Recall:

- mpg is the fuel economy of the cars in miles per gallon.
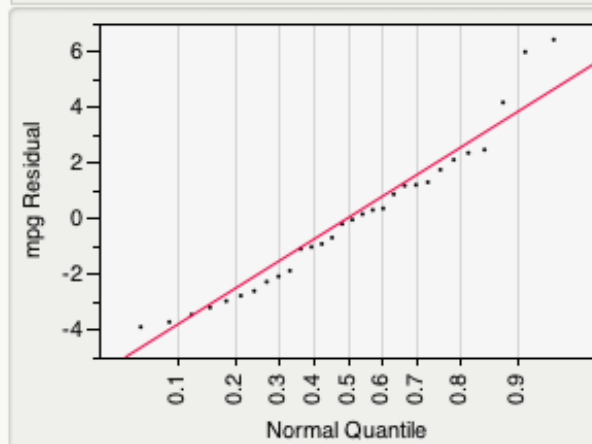- wt is the weight of the cars in tons.

Below, I fit a regression line of mpg on wt.

▼ ▼ Bivariate Fit of mpg By wt

▼ ── Linear Fit

▼ Linear Fit

mpg = 37.285126 - 5.3444716*wt

▼ Summary of Fit

RSquare	0.752833
RSquare Adj	0.744594
Root Mean Square Error	3.045882
Mean of Response	20.09062
Observations (or Sum Wgts)	32

▶ Lack Of Fit

▼ Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	847.7252	847.725	91.3753
Error	30	278.3219	9.277	Prob > F
C. Total	31	1126.0472		<.0001*

▼ Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	37.285126	1.877627	19.86	<.0001*
wt	-5.344472	0.559101	-9.56	<.0001*

## Residual by Row Plot

## Residual Normal Quantile Plot

a. (4 points) Identify and interpret the slope.

b. (4 points) Identify and interpret the intercept.

c. (4 points) What is problematic about your above interpretation of the intercept in practice?

d. (4 points) Based on the residual plot, comment on the validity of the model.

e. (4 points) Based on the normal quantile (normal QQ) plot, do the residuals look bell-shaped (normally-distributed)?
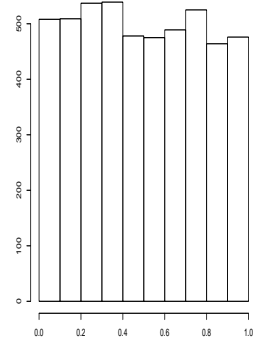
4. (20 points)

a. (10 points) What is the difference between a histogram and a bar plot?

b. (10 points) Identify the following distributional shapes.
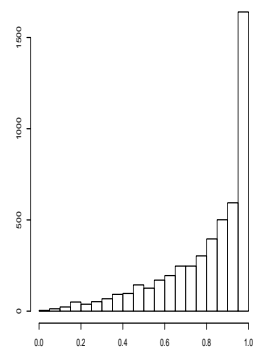
  i. (2.5 points)
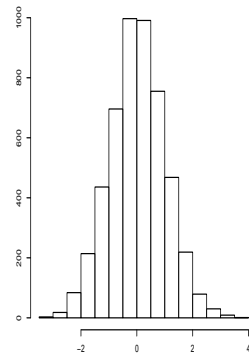
  ii. (2.5 points)
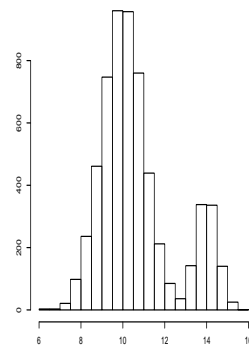
  iii. (2.5 points)

    iv. (2.5 points)

5. (20 points) One study reported on a study of strength properties of high-performance concrete obtained by using superplasticizers and certain binders. The compressive strength of such concrete had previously been investigated, but not much was known about flexural strength (a measure of ability to resist failure in bending). Below is part of the flexural strength data in MegaPascals (MPa):

   7.0   7.4   7.7   7.9   8   8.1   8.7   9.0   9.7   11.3   11.8   14

   a. (10 points) Find $Q(0.25)$ and $Q(0.75)$ of the data.
   b. (10 points) Make a boxplot of the data. Is the distribution symmetric? Are there any outliers?