FRQ Practice: This is a 40 point test grade. Show your work and work with each other. NAME _____

1. An important method for controlling the spread of the H6N2 influenza (bird flu) virus in chickens is having a procedure to determine whether chickens are infected with the virus. It is common to apply a procedure, called an ELISA test, to measure the concentration of anti–bird flu antibodies in a blood sample taken from a chicken. If the ELISA test reveals a high-enough concentration of antibodies, the chicken is said to test positive, and it is classified as infected with the virus. Otherwise, the chicken is said to test negative, and it is classified as not infected. However, the ELISA test is a complex procedure that is not always accurate. One type of mistake, a false positive result, occurs when the ELISA test gives a positive result for a chicken that is not infected with the virus. A second type of mistake, a false negative result, occurs when the ELISA test gives a negative result for an infected chicken.

Considering the possibility of false positives and false negatives for tests on individual chickens, veterinarians have developed the following procedure for determining if the H6N2 virus is present in a large flock of chickens.

- Randomly select 10 chickens from the flock.
- Perform the ELISA test on a blood sample from each of the 10 chickens.
- Conclude that the H6N2 virus is present in the flock if at least 3 out of the 10 chickens have positive ELISA test results.

Suppose a veterinarian applies the procedure to a flock of 100,000 chickens at a commercial egg production farm. The ELISA test is known to have probability 0.05 of producing a false positive result and probability 0.10 of producing a false negative result for a single chicken.

a. If no chicken in the flock is infected with the H6N2 virus, what is the probability that the veterinarian will conclude that the H6N2 virus is not present in the flock? Show how you found your answer. (Need 2 or less false positives)

$10C0(.05)^0(.95)^{10} + 10C1(.05)^1(.95)^9 + 10C2(.05)^2(.95)^8 = \boxed{.9885}$

b. If no chicken in the flock is infected with the H6N2 virus, what is the probability that the veterinarian will conclude that the H6N2 virus is present in the flock? Show how you found your answer.

$1 - .9885 = \boxed{.0115}$ because mutually exclusive.

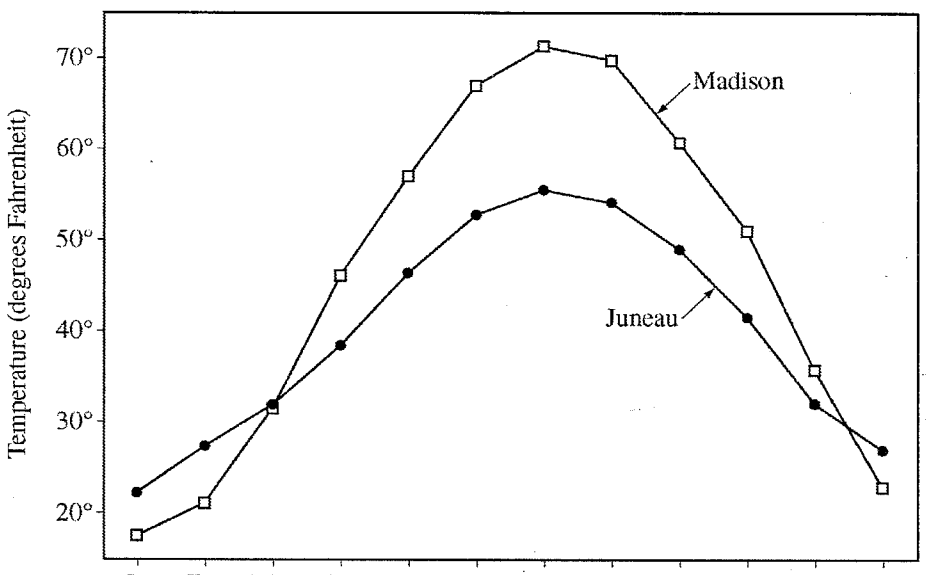
c. If every chicken in the flock is infected with the H6N2 virus, what is the probability that the veterinarian will conclude that the H6N2 virus is present in the flock? Show how you found your answer. Need atleast 3 positive so remain from 100%

$1 - [10C0(.9)^0(.1)^{10} + 10C1(.9)^1(.1)^9 + 10C2.9^2.1^8] =$

d. If 20 percent of the chickens in the flock are infected with the H6N2 virus and the other 80 percent are <u>not</u> infected, what is the probability that the veterinarian will conclude that the H6N2 virus is present in the flock? Show how you found your answer. $P(+)\,P(\text{inf chik}) + P(F+)\,P(\text{NI}c) = .22$

$1 - [10C0(.22)^0(.78)^{10}\ \ 10C1(.22)^1(.78)^9\ \ 10C2(.22)^2(.78)^8] = \boxed{.383}$

2. The graph below shows the monthly average temperatures, in degrees Fahrenheit, for two cities–Madison,Wisconsin, and Juneau, Alaska–in the United States.



a. Based on the graph, compare the two cities with respect to the monthly average temperatures over the year. Address both similarities and differences in the overall patterns.
Similarities: Colder in winter warmer in Summer
Differences: Madison has higher highs & lower lows.

b. For which of the two cities is the standard deviation of the 12 monthly average temperatures greater? Justify your answer without performing any calculations.
Madison has a greater σ because they have higher highs & lower lows. In other words they have greater variance from mean.

3. At a certain university, students who live in the dormitories eat at a common dining hall. Recently, some students have been complaining about the quality of the food served there. The dining hall manager decided to do a survey to estimate the proportion of students living in the dormitories who think that the quality of the food should be improved. One evening, the manager asked the first 100 students entering the dining hall to answer the following question.

> Many students believe that the food served in the dining hall needs improvement. Do you think that the quality of food served here needs improvement, even though that would increase the cost of the meal plan?
>
> ____ Yes ____ No ____ No opinion

a. In this setting, explain how bias may have been introduced based on the way this convenience sample was selected **and** suggest how the sample could have been selected differently to avoid that bias. *Convenience sample may add bias because first 100 students may not accurately represent all dorm residents
*Randomly select dorm residents

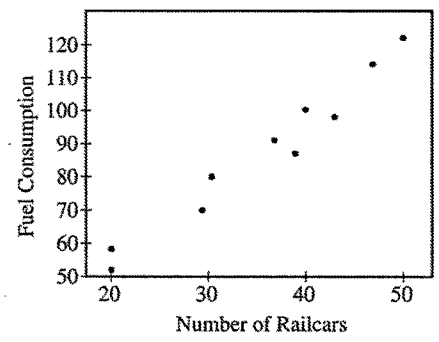
b. In this setting, explain how bias may have been introduced based on the way the question was worded **and** suggest how it could have been worded differently to avoid that bias. - The 1st part of question is leading because it says "many students" don't like food
- The last part talks about ↑ prices. This could lead people
"Do you think the food in the Caf. needs improvement"

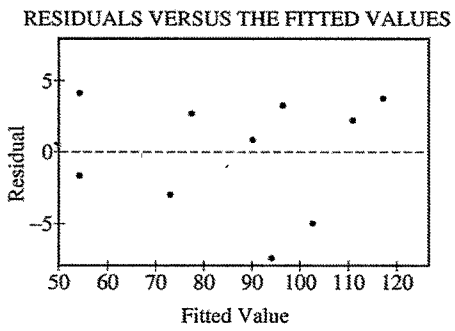
4. The Great Plains Railroad is interested in studying how fuel consumption is related to the number of railcars for its trains on a certain route between Oklahoma City and Omaha.

A random sample of 10 trains on this route has yielded the data in the table below.

Number of Railcars	Fuel Consumption (units/mile)
20	58
20	52
37	91
31	80
47	114
43	98
39	87
50	122
40	100
29	70

A scatterplot, a residual plot, and the output from the regression analysis for these data are shown below.





The regression equation is
Fuel Consumption = 10.7 + 2.15 Railcars

Predictor	Coef	StDev	T	P
Constant	10.677	5.157	2.07	0.072
Railcar	2.1495	0.1396	15.40	0.000

S = 4.361 R-Sq = 96.7% R-Sq(adj) = 96.3%

a. Is a linear model appropriate for modeling these data? Clearly explain your reasoning.

A linear model is appropriate.
It shows a strong positive correlation
The residual shows a random scatter of points

b. Suppose the fuel consumption cost is $25 per unit. Give a point estimate (single value) for the change in the average cost of fuel per mile for each additional railcar attached to a train. Show your work.

2.15 additional units per car
× 25
53.75 for each rail car

c. Interpret the value of r^2 in the context of this problem.

96.7% of fuel consumption is explained by the # of railcars

d. Would it be reasonable to use the fitted regression equation to predict the fuel consumption for a train on this route if the train had 65 railcars? Explain.

Anything past 50 cars would be extrapolation so 65 railcars is too high.

5. When a tractor pulls a plow through an agricultural field, the energy needed to pull that plow is called the draft. The draft is affected by environmental conditions such as soil type, terrain, and moisture.

A study was conducted to determine whether a newly developed hitch would be able to reduce draft compared to the standard hitch. (A hitch is used to connect the plow to the tractor.) Two large plots of land were used in this study. It was randomly determined which plot was to be plowed using the standard hitch. As the tractor plowed that plot, a measurement device on the tractor automatically recorded the draft at 25 randomly selected points in the plot.

After the plot was plowed, the hitch was changed from the standard one to the new one, a process that takes a substantial amount of time. Then the second plot was plowed using the new hitch. Twenty-five measurements of draft were also recorded at randomly selected points in this plot.

a. What was the response variable in this study? draft
Identify the treatments. Standard hitch & new hitch
What were the experimental units? 2 plots of land

b. Given that the goal of the study is to determine whether a newly developed hitch reduces draft compared to the standard hitch, was randomization used properly in this study? Justify your answer. The 2 hitches were randomly assigned to the 2 plots of land so YES

c. Given that the goal of the study is to determine whether a newly developed hitch reduces draft compared to the standard hitch, was replication used properly in this study? Justify your answer. There was NOT a good use of replication because each hitch was only used on one plot. For replication to be present we need to use each hitch on different plots.

d. Plot of land is a confounding variable in this experiment. Explain why. Each hitch is only used on one plot we do not know if difference is due to hitch variance or plot variance
The treatments are confounded with the plots.