LAST NAME (Please Print): ________________________________

FIRST NAME (Please Print): ________________________________

HONOR PLEDGE (Please Sign): ________________________________

Statistics 101

Homework 6

You are allowed to discuss problems with other students, but the final answers must be your own work.

For all problems that require calculation, YOU MUST ATTACH SEPARATE PAGES, NEATLY WRITTEN, THAT SHOW YOUR WORK.

Please mark your answer in the space provided. As a general rule, each blank counts for one point. If necessary work is not shown, or if that work is substantially wrong, then you will not get credit even if the answer is correct. (The obvious purpose of this seemingly draconian policy is to prevent people from mindlessly copying each other’s answers.)

Report all numerical answers to at least two correct decimal places.

Upload to Gradescope by 3:30 p.m. on Tuesday, March 1.
1. Use the data that are available at:

http://pages.stat.wisc.edu/ acovington/303/4/beef.txt

Read the description. If last name begins with a letter in A-F, do a linear regression that predicts the price of beef from the consumption of beef, the price of pork, the retail food price index adjusted by the CPI, and the disposable income per capita index. If your last name begins with a letter in G-S, do a linear regression that predicts the price of pork from the price of beef, the consumption of pork, the retail food price index adjusted by the CPI, and the disposable income per capita index. If your last name begins with a letter in T-Z, do a linear regression that predicts the retail food price index adjusted by the CPI from the consumption of beef, the consumption of pork, the price of pork, and the disposable income per capita index.

What is your prediction for a year in which the price of beef is 60 cents/lb, the price of pork is 50 cents/lb, the consumption of beef per capita is 55 pounds, the consumption of pork per capita is 58 pounds, the RFP index adjusted by the CPI is 850, and the disposable income per capita index is 35?

What is the significance probability for deciding whether disposable income per capita index is a useful variable in your regression?

What distribution is used for determining the significance probability above. (Include the df if appropriate.)

What is the strongest correlation among two of the explanatory variables in your model? Which variables are these?

Plot the residuals for your fitted model against the disposable income per capita index (1 point). Comment upon any issues that you see in it below.
Interpret the sign of your most significant coefficient. Identify the variable, indicate the significance probability associated with it, and try to explain whether it makes sense (3 points, 1 for each part).

2. Use the data that are available on our FAQ site in the file named PopularKids.csv. If your last name begins with A-G, determine whether gender is related to the child’s goal. If it begins with H-S, determine whether the school is related to goal. If it begins with T-Z, determine whether race is related to goal. Note that the file uses “looks” and “popularity” as synonyms.

In words, what is your null hypothesis?

In words, what is your alternative hypothesis?

What is the value of your test statistic?

What distribution does this follow? (Include df if appropriate.)

What is your significance probability (or P-value)?
In words, what is your conclusion? Assume $\alpha = 0.05$.

--------------------------------------------------

--------------------------------------------------

3. Google “Brink Finkelstein Levin parking meter theft” to find the website which has the url that begins “www3.nd.edu/ busiforc/handouts” and read the story. Assume the city workers are always honest.

Use regression to estimate the amount of money that should have been collected in time period 20 or 21 or 22, depending on whether your last name begins with A-G, H-S, or T-Z, respectively.

Set a 95% prediction interval on the amount that should have been collected in that time period.

$L=\underline{\text{\hspace{2cm}}} \quad U=\underline{\text{\hspace{2cm}}}$

Suppose that the city workers are not completely honest. What effect would this have on your analysis?

--------------------------------------------------

--------------------------------------------------

One of the values in the data looks like an odd. Can you explain it?

--------------------------------------------------

--------------------------------------------------