Name: _____

Math 355 - Spring 2004 Assignment #9 - Part 1

Due: Friday, April 23, 2004

4 points

Soon you will receive a data file from me. This will also be available in the Math 355 subdirectory of my directory at the Download Course Data pick from the Network Menu OR at our web site `www.uwsp.edu/math/wetzel/Math355/assignme.htm`. The name of the file is FRIES.MTW. In this file, each row corresponds to a small and large fry purchased at the same time. There is one extra small fry. All of the mass measurements were supposed to be in GRAMS.

- `sm-tmass` is the total mass of the small fries and the bags.
- `sm-bmass` is the mass of just the small bags.
- `sm-num` is the number of French fries in the small order.
- `lg-tmass` is the total mass of the large fries and the bags.
- `lg-bmass` is the mass of just the large bags.
- `lg-num` is the number of French fries in the large order.
- `day` is the date in April of the purchase.
- `sm-time` is the time of the purchase of the small fry. `a` is the afternoon and `e` is the evening.
- `lg-time` is the time of the purchase of the large fry. `a` is the afternoon and `e` is the evening.

The main goal of assignment 9 is to completely analyze the data and to write a brief report to summarize our analysis.

For Part 1, we need to identify the outliers.

1. Look at a histogram for each column. Use these to identify possible outliers. You do not need to print out the histograms. Which rows contain outliers? Which column(s) in that row contain the outlier(s)? Why do you think is that value an outlier? In your opinion should we delete that value, keep it or change it? Why and/or how? I have made one decision already as an example.

row	column(s) in that row	Why is this an outlier	delete, keep or change, why / how?
19	day = 14	the rest are all 15th or 16th	keep it, because it really was collected on the 14th.

2. Do a boxplot for each column. Why are most of the outliers not detected by the boxplots?

Turn in this sheet with your answers.