Stat13 Homework 1

(50 points, student scores will be converted to scores out of 100)
Suggested Solutions

<u>Problem 1</u>

• For each study, describe what treatment is being compared and what response is being measured to compare the treatments.

Study 1:	Treatment: hair color	(1 point)
	Response: pain tolerance level	(1 point)
Study 2:	Treatment: baking temperature	(1 point)
	Response: impact strength	(1 point)
Study 3:	Treatment: use of new diagnostic machine	(1 point)
	Response: service time	(1 point)
Study 4:	Treatment: gender	(1 point)
	Response: exam result	(1 point)
Study 5:	Treatment: different manufacturers	(1 point)
	Response: life span	(1 point)

• Which of the studies would be described as experiments and which would be described as observational studies?

Study 1:	Observational study	(1 point)
Study 2:	Experiment	(1 point)
Study 3:	Experiment	(1 point)
Study 4:	Observational study	(1 point)
Study 5:	Observational study	(1 point)

• For the studies that are observational, could an experiment have been carried out instead? If not, briefly explain why not.

Study 1: No. We have no control over which color we assign to each subject (assuming we are talking about natural hair color). (2 points, 1 for No and 1 for reason)

Study 4: No. We have no control over which gender we assign to each student. (2 points, 1 for No and 1 for reason)

Study 5: Yes. (1 point)

• For the studies that are experiments, briefly discuss what forms of blinding would be possible to be used.

Study 2: The person testing the impact strength of the particle boards in the laboratory does not know which temperature each board has been baked at. (1 point)

Study 3: possible answers (1 point)
a)The person analyzing the data is given a list of service times for each car, but does not know which list corresponds to each service method (i.e. using new diagnostic machine or not using new diagnostic machine). OR
b) The mechanic does not know whether they are using the new diagnostic machine or the old machine. The machines could have been disguised so that they both look the same.

• In which of the studies has blocking been used? Briefly describe what was blocked and why it was blocked.

Study 3: Blocking by mechanic has been used in this study. This was done because the difference in work experience between mechanics could also have an effect on the time for service. (2 points, 1 point for what was blocked and 1 point for why)

Study 4: Blocking by test subjects. This is because test subject may have an effect on the test grade difference, i.e. male and female students can perform differently at different subjects, e.g. male may be better at Math, whereas female may be better at linguistics. (2 points, 1 point for what was blocked and 1 point for why)

Study 5: Blocking by hard drive capacity has been used in this study. This was done because the retail chain wants to look at the life span of the computers for the 5 different manufacturers within each drive capacity. (2 points, 1 point for what was blocked and 1 point for why)

Problem 2
• Construct a stem-and-leaf plot of the data using an appropriate scale.(5 points, 2 for correct stems, 2 for correct leaves, 1 for key)

  Daily Number of Parking Tickets given out on UCLA campus over a period of 30 days
  3 | 244
  3 | 567788889
  4 | 01122234
  4 | 67799
  5 | 024
  5 | 8
  6 | key
  6 | 8   3|2 means 32
OR:
Daily Number of Parking Tickets given out on UCLA campus over a period of 30 days
 3 | 244567788889
  4 | 0112223467799
  5 | 0248   key
  6 | 8   3|2 means 32

• Calculate the five number summary for the data (min, lower quartile, median, upper quartile, and max).

Minimum: 32   (1 point)
Lower Quartile: 38   (1 point)
Median: 41.5   (1 point)
Upper Quartile: 47   (1 point)
Maximum: 68   (1 point)

• Draw a box plot for this set of data. Show your working. (8 points, 1 point for labeling axis correctly(either vertical or horizontal and the numbers could be labeled differently), 1 point for minimum, 1 point for Q1, 1 point for median, 1 point for Q3, 1 point for Q3+1.5(IQR) position, 1 point for the outlier, 1 point for drawing the plot correctly)

Parking Ticket Data at UCLA
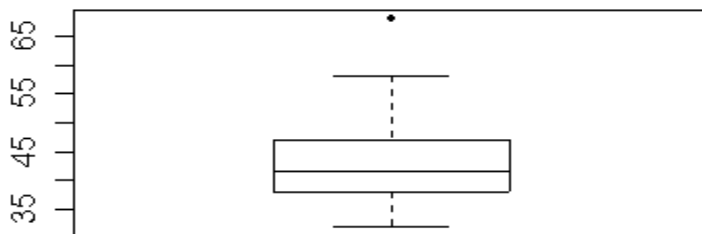
OR:

Parking Ticket Data at UCLA
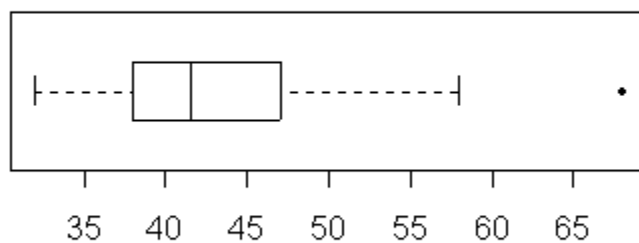
• Using your plots, in plain English, briefly comment on the data and the story the above summary/plots convey. (6 points, students may have slightly different answers)

From the analysis, it can be seen that the number of parking tickets given out at UCLA over a period of 30 days is about 42 per day, since the median is 41.5. The highest number of parking tickets given out on a single day during the period was 68, and the least number was 32. Most days had between 38 and 47 tickets given out.