**Homework 3 Answer Sheet**

**STUDENT NAME**

## Section 1

**Question 1.1** What variables are on the x and y axis of this dataset?

**Question 1.2** How many missing values are there for each station?

**Question 1.3** If we started our analysis on day 2, only including days 2-5, how many missing values are there for each station?

**Question 1.4** Day 3?

**Question 1.5** What will addNA(test,1) return? What does this mean?

**Question 1.6** What will addNA(test,4) return? What does this mean?

**Question 1.7** What are the two types of loops in R? What should we use for this purpose?

**Question 1.8** What is the output of 1:dim(test)[1]? What does it mean?

**Question 1.9** What will the dimensions of the output be?

**INCLUDE YOUR PLOT HERE!**

**Question 1.10** What do the different plots represent?

**Question 1.11** What variables are on the x and y axis?

## Section 2

**Question 2.1** What are the dimensions of this data? *Hint: dim().*

**Question 2.2** What are the column names of this data? *Hint: colnames().*

**Question 2.3** What is the column range of the data? (Names begin with US. . . .) Use this information to make a new data frame called rain.

**INCLUDE YOUR PLOT HERE!**

**Question 2.4** Describe the missing data trends in a few of the stations.

**Question 2.5** Describe how this algorithm works.

## Section 3

**Question 3.1** What is the elapsed time of this code block?

**Question 3.2** What is the output of getDoParWorkers()? This is most likely the number of processors your computer has.

**Question 3.3** What is the elapsed time of this code? Is it faster than our earlier serial code? Why is this?