**Homework 4 Answer Sheet**

**STUDENT NAME**

## Section 2

**Question 2.1** In your own words, provide a description of what each of the following bash terminal commands

do (google is your friend):

• pwd

• ls

• cd

• touch

• rm

• cat

• cp

• mv

• mkdir

**Question 2.2** In your home directory on Clark, create a folder named hw4 and in it create a file name test.txt.

Screenshot the output of ls in this directory (Hint: windows users hold windows button + shift + s).

## Section 3

**Question 3.1** What are the dimensions of this dataset? What is represented on the x and y dimension?

(Hint: similar to hw3).

**Question 3.2** What is the beginning and end date of this time series?

**Question 3.3** What is the column range of the data? (Names begin with US. . . .) Use this information to

make a new data frame called rain.

**Question 3.4** How many missing data entries are there in the entirety of this dataset? (Hint: sum() and

is.na())

**Question 3.5** What percentage of the total data entries is NA?

## Section 4

**Question 4.1** How long did this take to take to run on Clark? (Do not try on your personal computer).

## Section 5

**Question 5.1** What is the output of getDoParWorkers()? What does this represent?

**Question 5.2** What is the elapsed time of this code? Is it faster than our earlier serial code? Why is this?

**Question 5.3** Run the same code with 4 cores, 6 cores, and 8 cores. Record and plot the output with cores

on the x axis and time on the y. Use any (R) plotting method you like.

**Question 5.4** Describe your graph.

## Section 6

**Question 6.1** Include a screenshot of the plot. Where would you begin your analysis?