**1. Go to blackboard and download the homework CSV file called:**

**SAS\_HMWK\_TEMP\_MISS\_DATASET.csv**

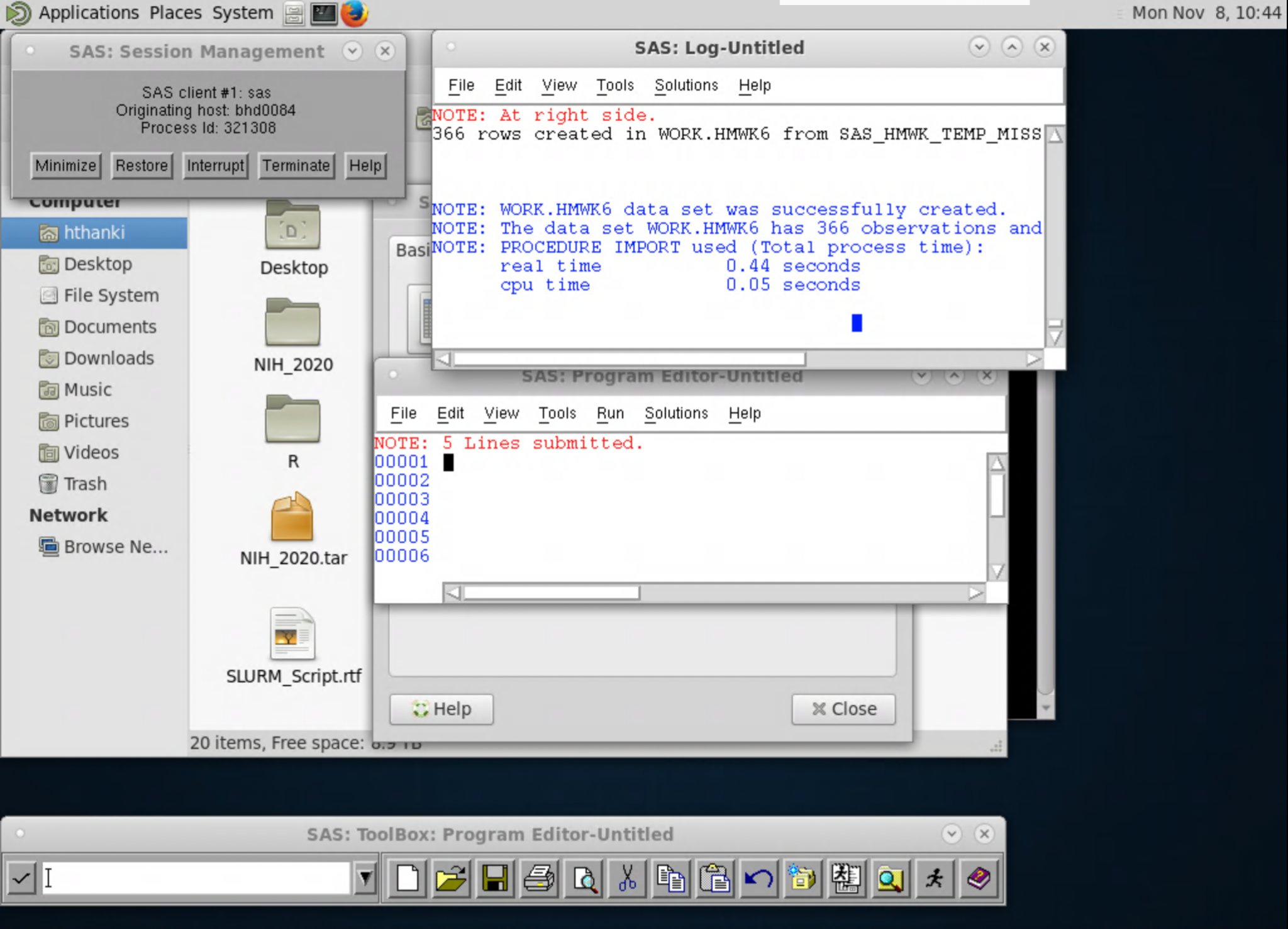
**Once the file is on your computer, transfer the file into your BlueHive environment. Then use the SAS import process presented during Lecture 15 and create a SAS dataset from the imported CSV file. Name the newly created SAS dataset as HMWK6. To receive full credit for this portion of the homework, show the SAS program code that was generated to import the CSV file and create the SAS dataset.**

proc import file=“\home\hthanki\SAS\_HMWK\_TEMP\_MISS\_DATASET.csv”

out= HMWK6

dbms=csv;

run;



**2. Perform a Proc Sort statement on the SAS dataset called HMWK6, such that a newly created SAS dataset called HMWK6SRT will be created. Note: you may want to look up the SAS documentation for a Proc Sort statement in order to see the syntax on how to create a sorted output dataset https://documentation.sas.com/doc/en/pgmsascdc/9.4\_3.5/proc/titlepage.htm.**

**Have the Proc Sort statement sort the dataset by location and then by month and then by date. Then print out the first 10 observations of the sorted dataset called HMWK6SRT. To receive full credit for this portion of the homework, show the SAS program code that was generated to perform the Proc Sort along with a screenshot of the SAS output showing the proc print statement for the first 10 observations of the newly sorted dataset.**

proc sort data= HMWK6;

by location;

out= HMWK6SRT

run;

proc sort data= HMWK6;

by month;

out= HMWK6SRT

run;

proc sort data= HMWK6;

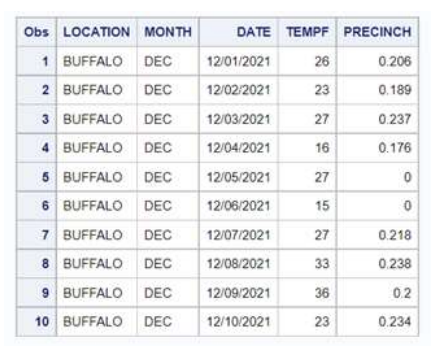
by date;

out= HMWK6SRT

run;

proc print data=HMWK6SRT.Class(obs=10);

run;



**3. Now take the sorted dataset called HMWK6SRT and perform a Proc Means statement in order to find the answers to the questions below. To receive full credit for this portion of the homework, show the SAS program code that was generated to perform the Proc Means along with a screenshot of the SAS output of the Proc Means that shows the answers to the questions below:**

**a) What is the average temperature of Buffalo in October?**

**b) What is the average rainfall of Rochester in December?**

**c) What is the average temperature of Syracuse in September?**

a) average temperature of Buffalo in October = 54.45

b) average rainfall of Rochester in December = 0.118

c) average temperature of Syracuse in September = 71.207

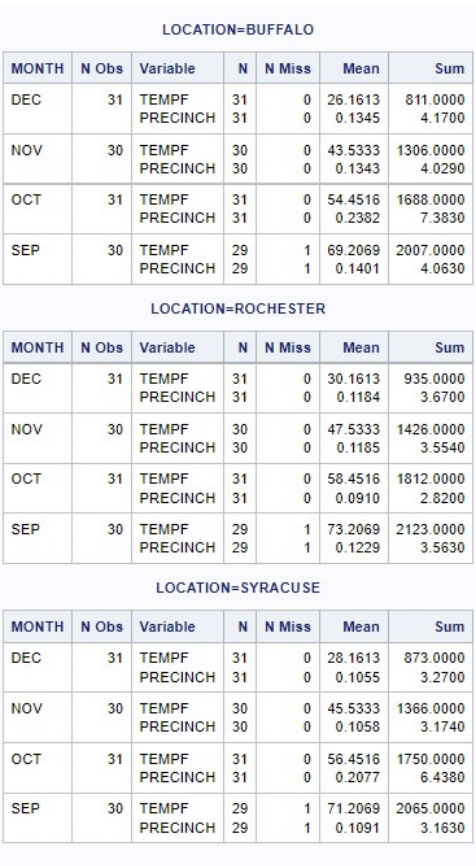
proc means data=WORK.HMWK6SRT n nmiss mean sum maxdec=4;

by location;

class month;

var tempf precinch;

run;



**4. Use the proc means command to show how many values are missing for the HMWK6SRT dataset. To receive full credit for this portion of the homework, show the SAS program code that was generated to perform the Proc Means along with a screenshot of the SAS output.**

1 TEMPF value (September ,Buffalo)

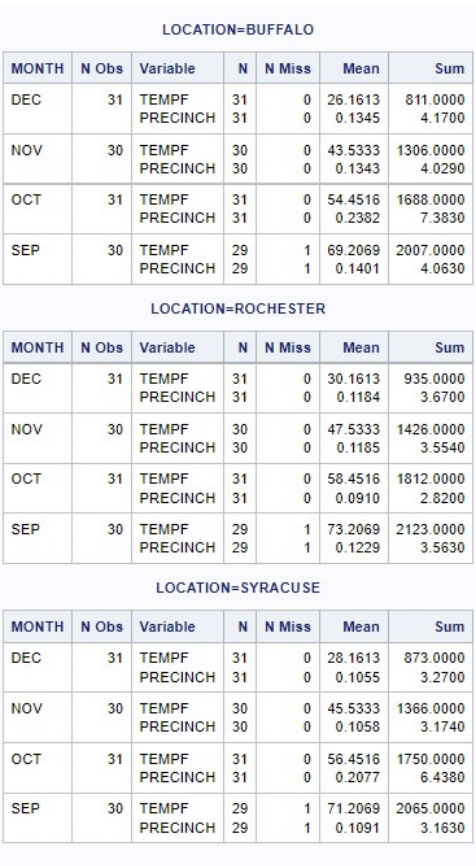
1 PRECINCH value (September, Buffalo)

1 TEMPF value (September , Rochester)

1 PRECINCH value ( September, Rochester)

1 TEMPF value (September ,Syracuse)

1 PRECINCH value (September, Rochester)

****

title "Average Temprature(tempf) and Average Rainfall(precinch) for each City, each Month";

proc means data=WORK.HMWK6SRT n nmiss mean sum maxdec=4;

by location;

class month;

var tempf precinch;

run;

The nmiss options gives us the number of missing values in a particular column.

**5. You will notice that the September 30th dates are missing values for the temperature and precipitation (in inches) for Buffalo, Rochester, and Syracuse because the monitoring equipment was broken that day. It turns out that the information was gathered at a different weather station and uploaded to the database the next day. Therefore, update your dataset such that the following values are entered in order to complete all of the values for the variables in the dataset:**

**LOCATION MONTH DATE TEMPF PRECINCH**

**Buffalo Sep 9/30/21 64 0.179**

**Rochester Sep 9/30/21 82 0.000**

**Syracuse Sep 9/30/21 80 0.000**

**Use these new values added to the original dataset to perform a new Proc Means statement to show the averages in questions a, b, and c below. Ensure you perform the Proc Means on the new dataset that has no missing values. In other words, to receive full credit for this portion of the homework, show the SAS program code that was generated to perform the Proc Means along with a screenshot of the SAS output of the Proc Means that shows the answers to the questions below:**

**a) What is the average temperature of Buffalo in September?**

**b) What is the average rainfall of Rochester in September?**

**c) What is the average temperature of Syracuse in September?**

a) average temperature of Buffalo in September = 69.033

b) average rainfall of Rochester in September = 0.119

c) average temperature of Syracuse in September = 71.50

data WORK.HMWK6;

set WORK.HMWK6;

if LOCATION = "BUFFALO" and missing(TEMPF) then TEMPF=64;

if LOCATION = "BUFFALO" and missing(PRECINCH) then PRECINCH=0.179;

if LOCATION = "ROCHESTER" and missing(TEMPF) then TEMPF=82;

if LOCATION = "ROCHESTER" and missing(PRECINCH) then PRECINCH=0.000;

if LOCATION = "SYRACUSE" and missing(TEMPF) then TEMPF=80;

if LOCATION = "SYRACUSE" and missing(PRECINCH) then PRECINCH=0.000;

run;

proc means data=WORK.HMWK6 n nmiss mean sum maxdec=4;

by location;

class month;

var tempf precinch;

run;

