Question 1-

Code-

**PROC** **IMPORT** DATAFILE = "\\Client\C$\Users\Cmirwin\Desktop\SMU Data Sets\case0101.csv"

OUT=ARTDATA

DBMS=csv

REPLACE;

GETNAMES=YES;

**RUN**;

**PROC** **MEANS** DATA = ARTDATA;

CLASS TREATMENT;

VAR SCORE;

**RUN**;

**proc** **surveyselect** data=ARTDATA samprate=**0.50** out=Sample outall method=srs noprint;

**run**;

**PROC** **MEANS** DATA = SAMPLE;

CLASS SELECTED;

VAR SCORE;

**RUN**;

**proc** **surveyselect** data=ARTDATA samprate=**0.50** out=Sample2 outall method=srs noprint;

**run**;

**PROC** **MEANS** DATA = SAMPLE2;

CLASS SELECTED;

VAR SCORE;

**RUN**;

**proc** **surveyselect** data=ARTDATA samprate=**0.50** out=Sample3 outall method=srs noprint;

**run**;

**PROC** **MEANS** DATA = SAMPLE3;

CLASS SELECTED;

VAR SCORE;

**RUN**;

**proc** **surveyselect** data=ARTDATA samprate=**0.50** out=Sample4 outall method=srs noprint;

**run**;

**PROC** **MEANS** DATA = SAMPLE4;

CLASS SELECTED;

VAR SCORE;

**RUN**;

**proc** **surveyselect** data=ARTDATA samprate=**0.50** out=Sample5 outall method=srs noprint;

**run**;

**PROC** **MEANS** DATA = SAMPLE5;

CLASS SELECTED;

VAR SCORE;

**RUN**;

Results – After running the SurveySelect Proc 5 additional times I found that all the remaining averages had differences that were less than 1.

Question 2

Code –

**DATA** BLOODCONCENTRATION;

INPUT GROUPNAME $ MEASUREMENT;

DATALINES;

A 1.31

A 1.45

A 1.12

A 1.16

A 1.30

A 1.50

A 1.20

A 1.22

A 1.42

A 1.14

A 1.23

A 1.59

A 1.11

A 1.10

A 1.53

A 1.52

A 1.17

A 1.49

A 1.62

A 1.29

B 1.13

B 1.71

B 1.39

B 1.15

B 1.33

B 1.00

B 1.03

B 1.68

B 1.76

B 1.55

B 1.34

B 1.47

B 1.74

B 1.74

B 1.19

B 1.15

B 1.20

B 1.59

B 1.47

;

**PROC** **SORT** DATA = BLOODCONCENTRATION OUT = BLOODCONCENTRATIONSORT;

BY GROUPNAME MEASUREMENT;

**RUN**;

**PROC** **BOXPLOT** DATA=BLOODCONCENTRATIONSORT;

PLOT MEASUREMENT \* GROUPNAME;

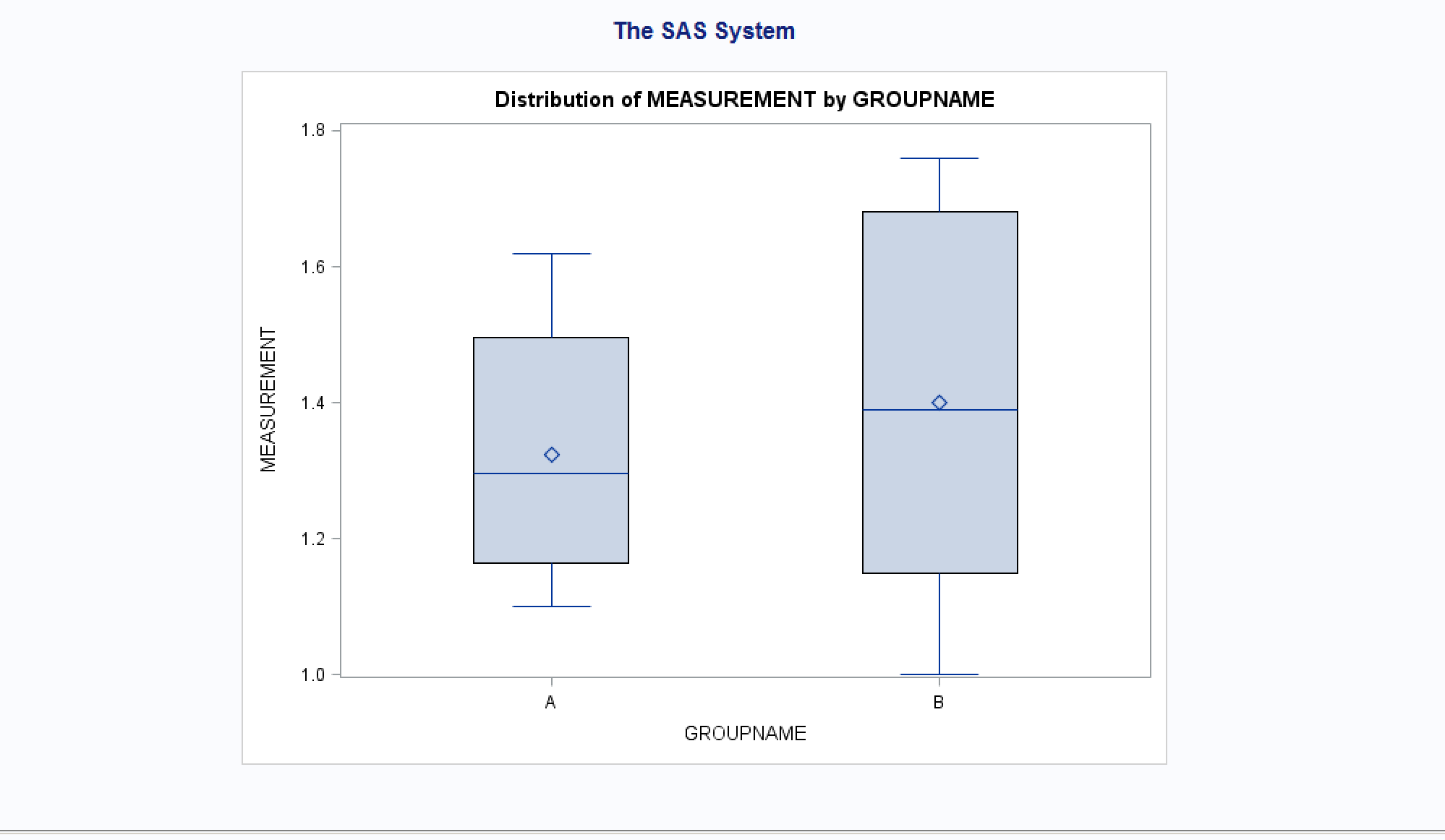
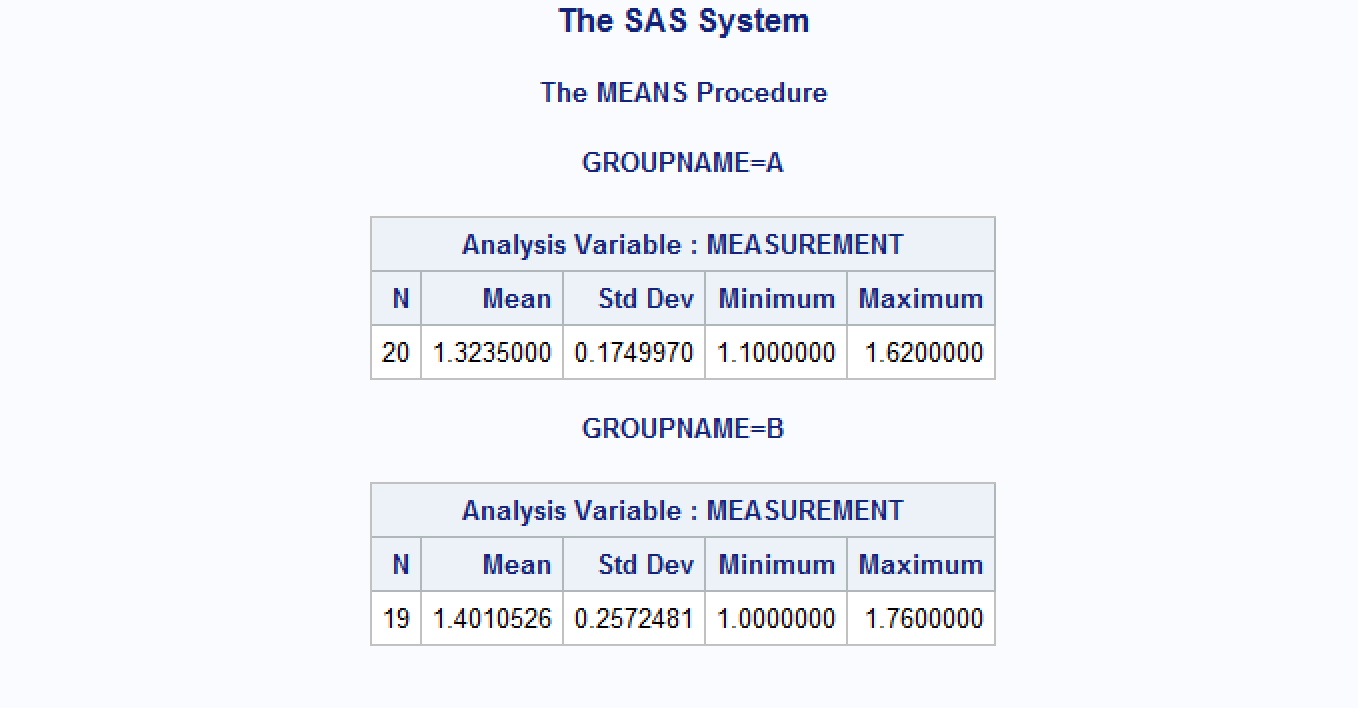
**RUN**;

**PROC** **MEANS** DATA=BLOODCONCENTRATIONSORT;

BY GROUPNAME;

**RUN**;

Results -



Conclusion – In conclusion Group A has a tighter overall distribution as shown by the smaller standard deviation when compared to Group B. Also Group B has a higher average measurement of 1.40 compared to Group A of 1.32.