STAT 448 HW #

Firstname Lastname

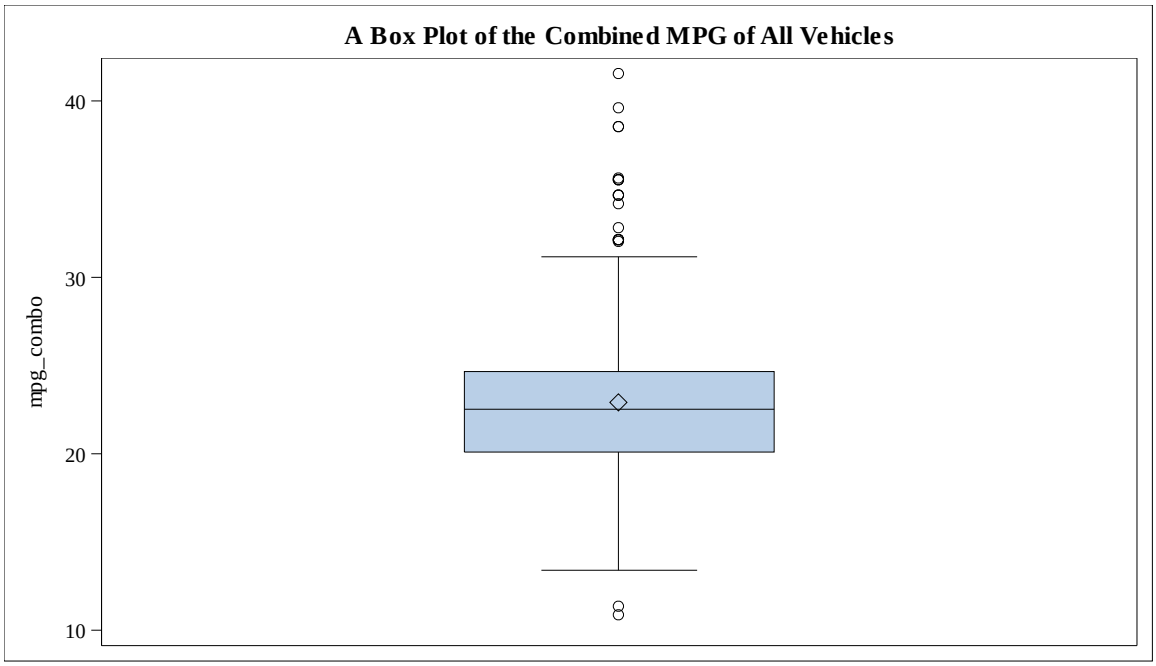
Current Date

Problem 1a

Use complete sentences for your answers to the problems. Use proper English grammar and writing rules. Times New Roman or the Roman (serif) are good fonts to use for the report in general. The font size should be 12 point, and the default document spacing should be double spaced. Suppose the problem 1a asks you to comment on what a box plot tells us about fuel efficiencies. Begin your response with words not code. Think of this as a report. Next, you should insert the SAS code that produces the box plot (excluding the DATA step), formatted by indenting and using single spacing and a monospaced Courier font such as Courier New.

```
proc sgplot data=cars;
 title "A Box Plot of the Combined MPG of All Vehicles";
 vbox mpg_combo;
run;
```

Your description of the box plot would appear before you insert the box plot. The description might mention that the box plot of the combined MPG has several outliers Ignoring those outliers, the mean and median are fairly close together (roughly 22 combined miles per gallon) and the spread of the distribution is not wide, rather low variability. Most vehicles get between 20 and 25 miles per gallon (combined). Now insert and center the box plot below keeping the image no wider than 6 inches (when possible).



If you still have more to say that is relevant to the question and answers it, you may do so here below the image.

Problem 1b

We would treat a table or series of tables the same way we treat images. Suppose problem 1b asked us to show more than one table. The code that you need to insert will contain the code used to produce the tables (see problem 1b), unless that code is calling on two separate procedures that produce two separate tables (see problem 1c).

```
proc univariate data=cars normaltest;
 var mpg_combo ;
 id model;
 ods exclude TestsforLocation Quantiles FitQuantiles
GoodnessofFit ParameterEstimates;
run;
```

Now you can describe the tables, then insert them below. Keep in mind that you may have to extend the sizes of the tables to take advantage of spacing and keeping information neat.

Moments			
N	425	**Sum Weights**	425
Mean	22.8847059	**Sum Observations**	9726
Std Deviation	4.6358734	**Variance**	21.4913221
Skewness	0.71081	**Kurtosis**	1.25934384
Uncorrected SS	231688.97	**Corrected SS**	9112.32059
Coeff Variation	20.2575179	**Std Error Mean**	0.22487289

Basic Statistical Measures			
Location		**Variability**	
Mean	22.88471	**Std Deviation**	4.63587
Median	22.50000	**Variance**	21.49132
Mode	21.15000	**Range**	30.70000
		Interquartile Range	4.55000

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.967372	**Pr < W**	<0.0001
Kolmogorov-Smirnov	**D**	0.103925	**Pr > D**	<0.0100
Cramer-von Mises	**W-Sq**	0.696027	**Pr > W-Sq**	<0.0050
Anderson-Darling	**A-Sq**	3.776247	**Pr > A-Sq**	<0.0050

Extreme Observations					
Lowest			Highest		
Value	Model	Obs	Value	Model	Obs
10.90	H2	165	35.7	Echo 2dr auto	381
11.35	Excursion 6.8 XLT	119	38.6	Echo 2dr manual	380
13.45	G500	250	38.6	Echo 4dr	382
13.80	Discovery SE	215	39.6	Civic HX 2dr	154
13.80	Range Rover HSE	214	41.6	Jetta GLS TDI 4dr	402

Problem 1c

Now suppose you have tables needed for one single problem but they come from two different procedures. Then, you should discuss them separately using one block of code and table inserted at a time.

```
proc univariate data=cars normaltest;
 var mpg_combo invoice;
 ods exclude TestsforLocation;
run;
```

Then description of the first table using this PROC. Then insert the table.

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.959591	Pr < W	0.0002
Kolmogorov-Smirnov	D	0.105193	Pr > D	<0.0100
Cramer-von Mises	W-Sq	0.315702	Pr > W-Sq	<0.0050
Anderson-Darling	A-Sq	1.837931	Pr > A-Sq	<0.0050

Then insert the different PROC.

```
proc freq data=cars order=data;
  table mpg*income/chisq expected nocol nopercent ;
  weight count;
  ods select ChiSq;
run;
```

Then describe and insert the corresponding table from that PROC. Also, keep in mind that additional spacing may be needed to keep a table together on one page instead of splitting a single table across pages.

Statistic	DF	Value	Prob
Chi-Square	1	719.3494	<.0001
Likelihood Ratio Chi-Square	1	789.6714	<.0001
Continuity Adj. Chi-Square	1	716.7068	<.0001
Mantel-Haenszel Chi-Square	1	718.9147	<.0001
Phi Coefficient		0.6593	
Contingency Coefficient		0.5504	
Cramer's V		0.6593	

Repeat this report styling for additional homework problems. This formatting makes grading easier and ensures neatness in your writing. Companies will probably not want your reports to include code, rather descriptions of your process and the analysis you performed, which may include images and tables. It would be wise to start on the homework early so that you don't get bogged down with formatting minutes before the deadline. This document was made using Microsoft Word and saved as a pdf.