Reproducible Comprehensive Report Toy Example

PI: Dr. Jane Doe
Analyst: Lauren Gunn-Sandell
Date of report: 20DEC2023
Updates:

1. Added in the regression interpretation.
2. Determined t test was no longer applicable to main analysis.

Note this (comprehensive) report will be very similar to the collaborator report, differing only by inclusion of code syntax at the end.

We are interested in comparing weight of vehicles with their gas mileage. We hypothesize that vehicles that weigh more than 3,500 lbs will have worse gas mileage than the vehicles that do not.

Analysis plan initially included a t test as seen below.

The TTEST Procedure

Variable: MPG_City (MPG (City))

wt_cat	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
3500lbs or lighter		226	22.9115	5.5394	0.3685	12.0000	60.0000
Heavier than 3500lbs		202	16.8713	2.1402	0.1506	10.0000	23.0000
Diff (1-2)	Pooled		6.0402	4.2858	0.4150		
Diff (1-2)	Satterthwaite		6.0402		0.3981		

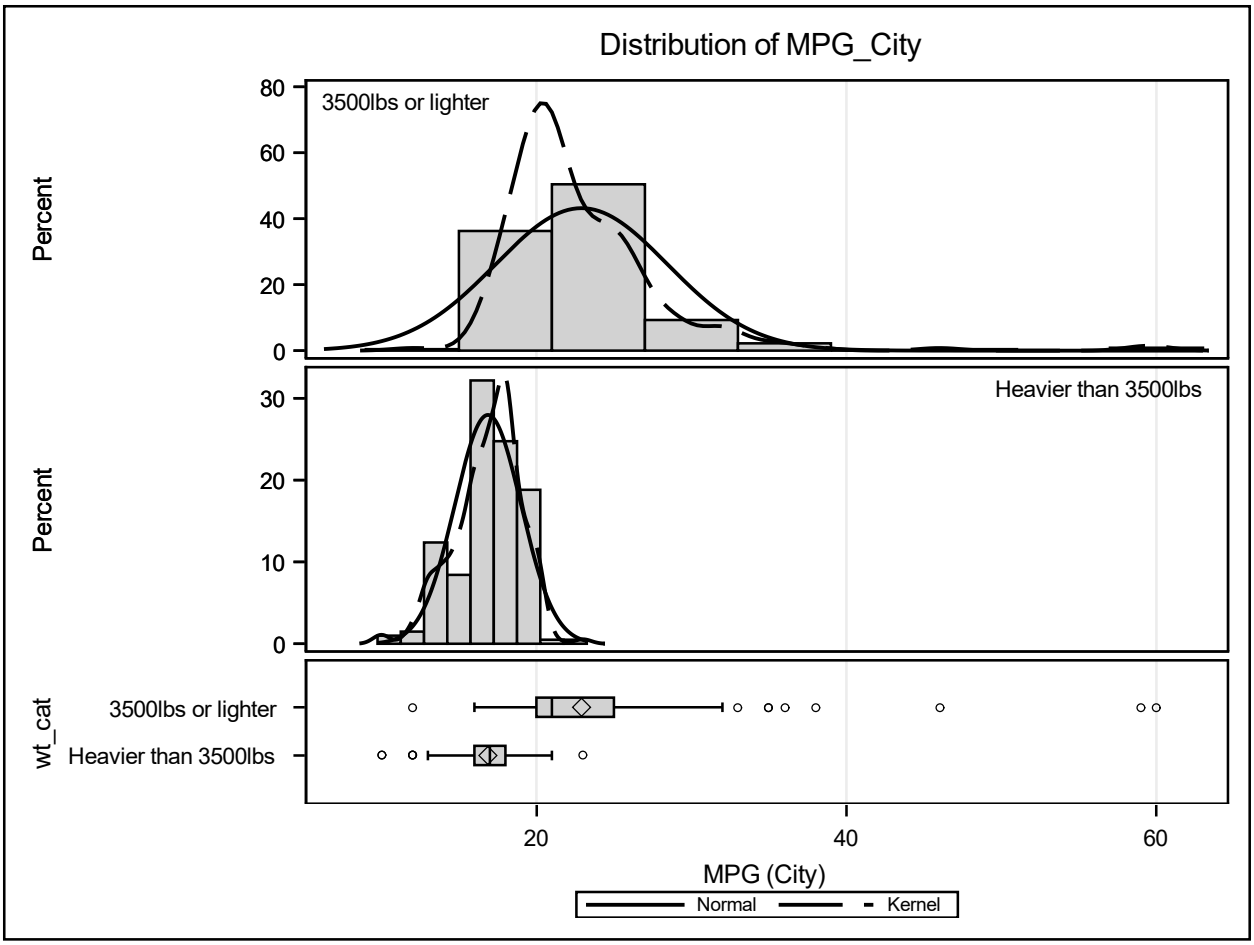
wt_cat	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
3500lbs or lighter		22.9115	22.1854	23.6376	5.5394	5.0715	6.1033
Heavier than 3500lbs		16.8713	16.5744	17.1682	2.1402	1.9498	2.3720
Diff (1-2)	Pooled	6.0402	5.2246	6.8559	4.2858	4.0164	4.5943
Diff (1-2)	Satterthwaite	6.0402	5.2568	6.8236			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	426	14.56	<.0001
Satterthwaite	Unequal	297.15	15.17	<.0001

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	225	201	6.70	<.0001

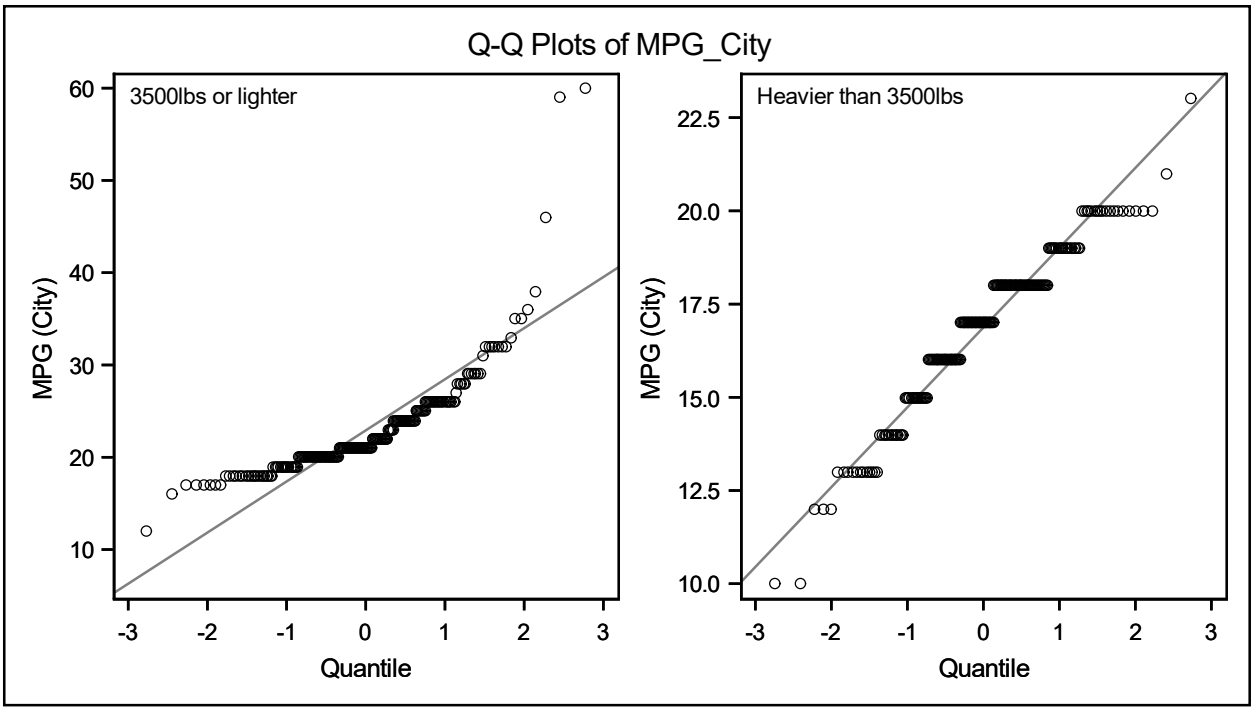
The TTEST Procedure

Variable: MPG_City (MPG (City))



The TTEST Procedure

Variable: MPG_City (MPG (City))



The t test comparing mean city gas mileage between cars that weigh over 3500lbs and those that are equal to and less than 3500 have a significant difference (6.70, p<.001).

Variable	*Class*	*N*	*Mean*	*LowerCLMean*	*UpperCLMean*
MPG_City	3500lbs or lighter	226	22.91	22.19	23.64
MPG_City	Heavier than 3500lbs	202	16.87	16.57	17.17

The REG Procedure
Model: MODEL1
Dependent Variable: MPG_City MPG (City)

Number of Observations Read	428
Number of Observations Used	428

Analysis of Variance

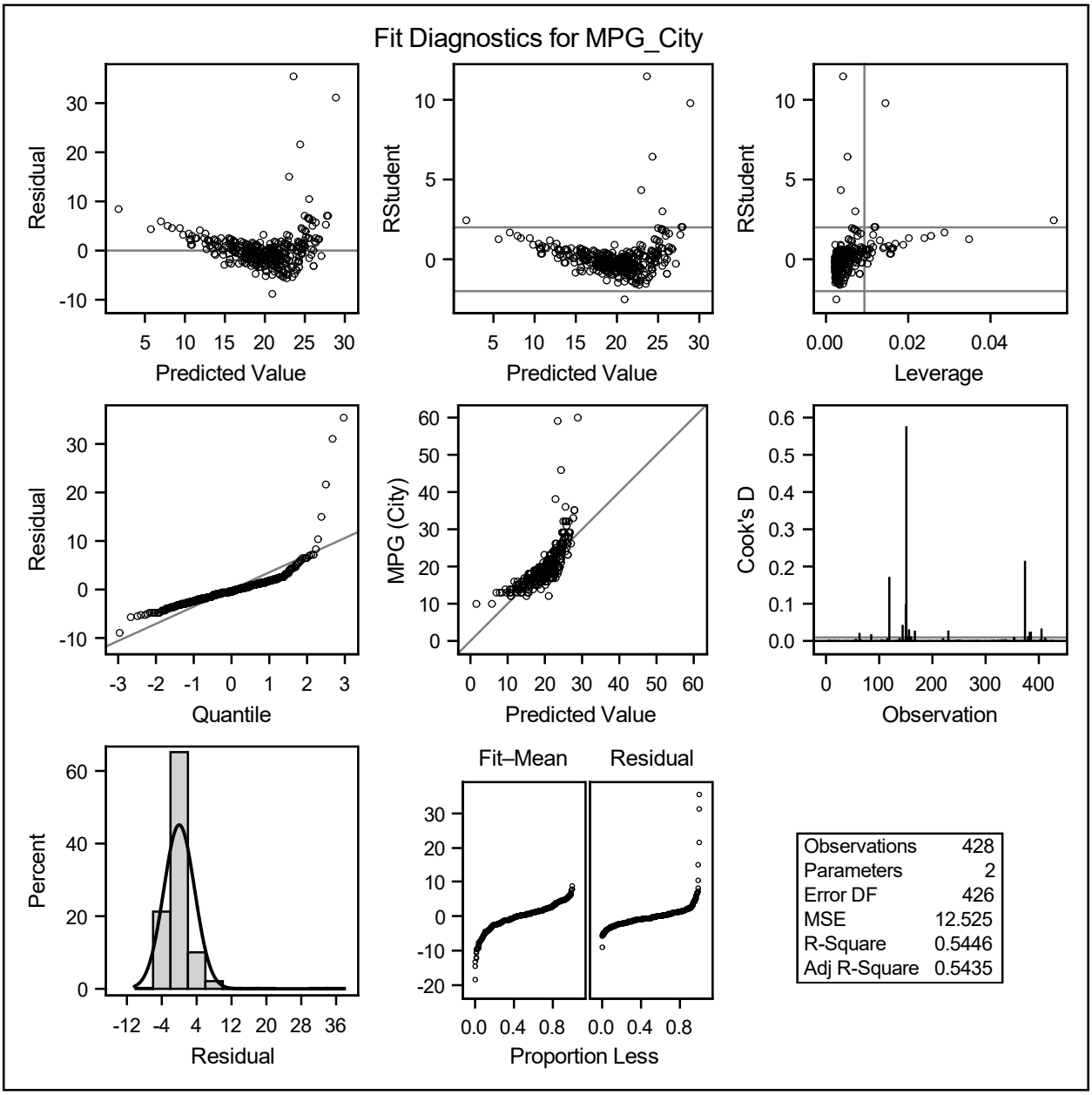
Source	*DF*	*Sum of Squares*	*Mean Square*	*F Value*	*Pr > F*
Model	1	6380.69016	6380.69016	509.43	<.0001
Error	426	5335.73040	12.52519		
Corrected Total	427	11716			

Root MSE	3.53909	*R-Square*	0.5446
Dependent Mean	20.06075	*Adj R-Sq*	0.5435
Coeff Var	17.64189		

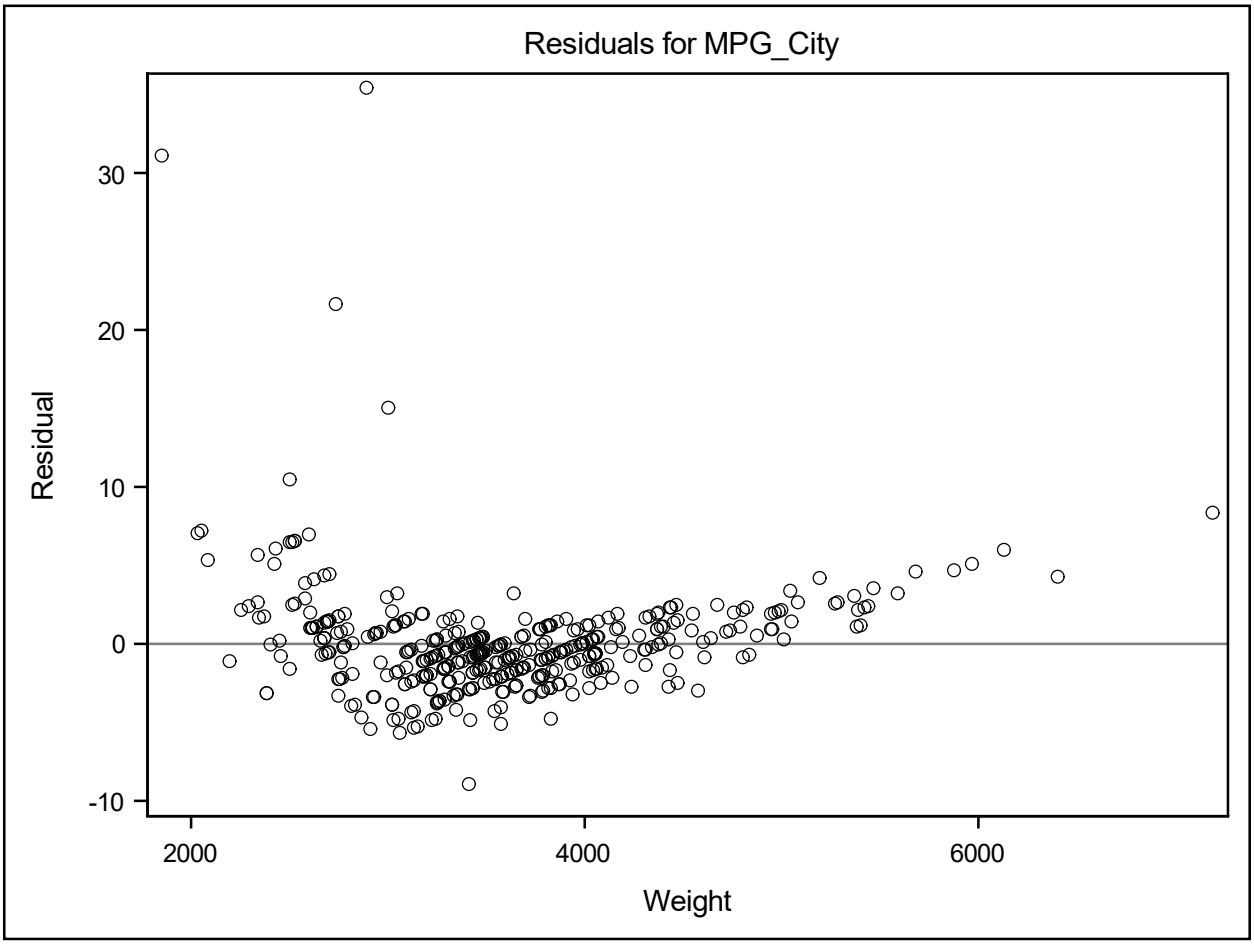
Parameter Estimates

Variable	*Label*	*DF*	*Parameter Estimate*	*Standard Error*	*t Value*	*Pr > |t|*
Intercept	Intercept	1	38.28385	0.82531	46.39	<.0001
Weight	Weight	1	-0.00509	0.00022566	-22.57	<.0001

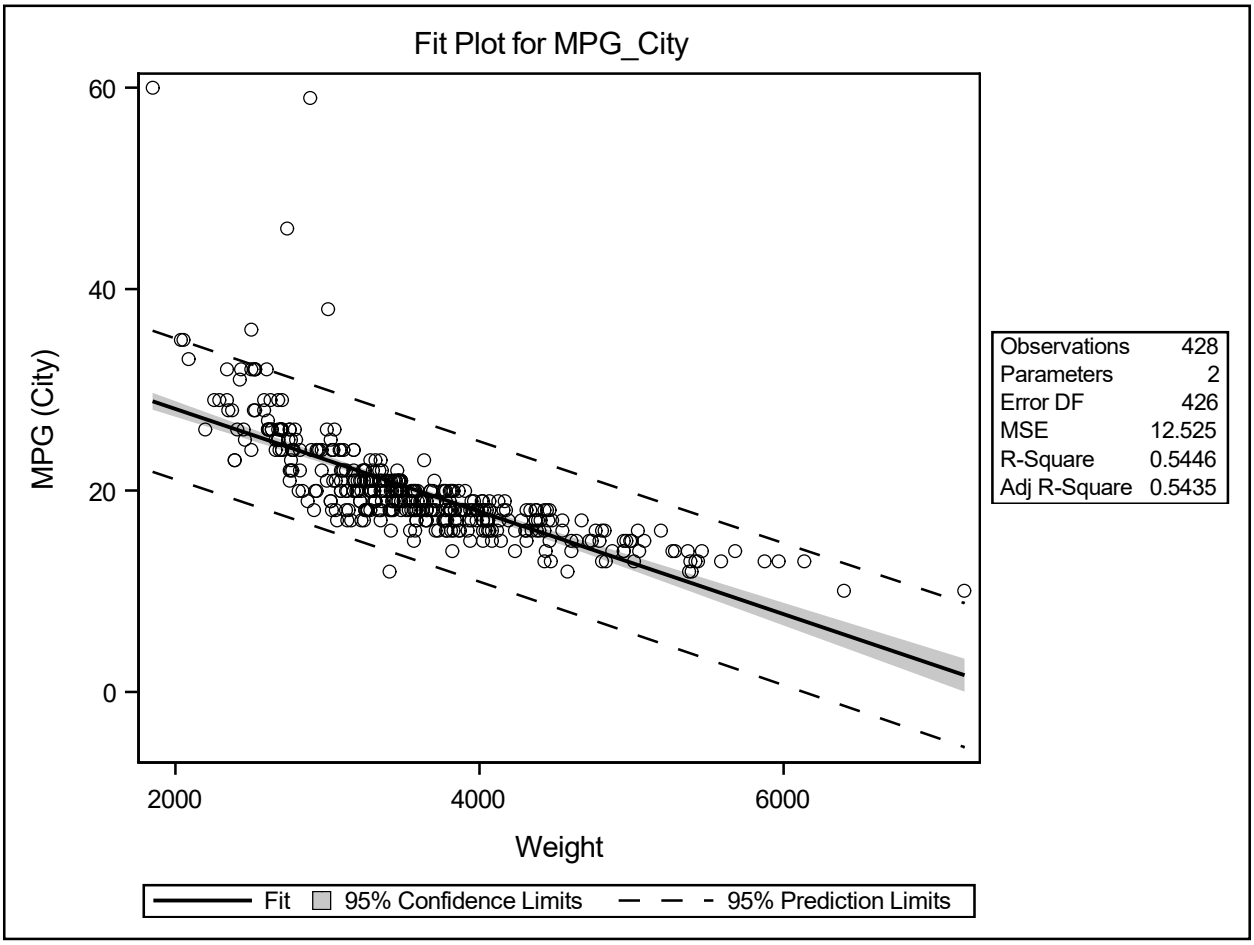
The REG Procedure
Model: MODEL1
Dependent Variable: MPG_City MPG (City)



The REG Procedure
Model: MODEL1
Dependent Variable: MPG_City MPG (City)



The REG Procedure
Model: MODEL1
Dependent Variable: MPG_City MPG (City)



Interpretation:

The linear regression assessing mean city gas mileage shows a significant relationship with car weight. For every one-thousand pound increase in weight, the expected MPG (city) changes by -5.09 ($p<.001$).

Session Information
Date: 20DEC2023
Operating System: WIN, X64_10PRO
SAS Version: 9.4

```
************************************************************************ ;
* SET WORKING DIRECTORY -- REQUIRES USER INPUT!!! *;
************************************************************************ ;

** THESE NEED TO BE MANUALLY CHANGED;
%let root   =  C:\Users\hochheic\OneDrive - The University of Colorado
Denver\Reproducibility_Leadership;
* Define the SAS program to be printing the syntax for (i.e., this SAS
program for the comprehensive report);
%let sasprgm = C:\Users\hochheic\OneDrive - The University of Colorado
Denver\Reproducibility_Leadership\Code\SAS_Version\Comprehensive.sas;

*-----------------------------------------------------------------------*;
* PROJECT TITLE:  Reproducibility and Leadership in
       Statistics and Research
* AUTHOR/ANALYST:  Lauren Gunn-Sandell
* INVESTIGATORS:  ...
* DATE CREATED:   09DEC2023
* DATE LAST REVISED:  15DEC2023
* VERSION DETAILS: SAS v 9.4 on Windows 64
* MODIFICATIONS:  10Dec2023 - Added regression analysis
*-----------------------------------------------------------------------*;
run;

* REPORT TYPE:
/*  A. Comprehensive report - Keeping code/syntax, otherwise the same as
collaborator report. */


*  Project Summary:
*  We are interested in comparing weight of vehicles with their gas
mileage.
  We hypothesize that vehicles that weigh more than 3,500 lbs will have
worse
  gas mileage than the vehicles that do not. ;


************************************************************************ ;
* START REPORT *;
************************************************************************ ;

* Create SAS libraries;
libname dataraw   "&root/DataRaw";
libname code  "&root/Code";
libname reports  "&root/Reports";

options topmargin=1in bottommargin=1in
       leftmargin=1in rightmargin=1in;
options nodate;
```

```
ods escapechar=' ^';

ods pdf file="&root/Reports/SAS_Version/Comprehensive_SAS.pdf"
  nogtitle nogfootnote
  startpage = no
  style = journal;

title1 font='Times'  bold  height=14pt justify=L "Reproducible
Comprehensive Report Toy Example";
title2 font='Times'    height=12pt justify=L "PI: Dr. Jane Doe";
title3 font='Times'    height=12pt justify=L "Analyst: Lauren
Gunn-Sandell";
title4 font='Times'    height=12pt justify=L "Date of report: &sysdate9";
title5 font='Times'    height=12pt justify=L "Updates: ";

ods text = " 1. Added in the regression interpretation.";
ods text = " 2. Determined t test was no longer applicable to main analysis.";
ods text = " ";
ods text = " Note this (comprehensive) report will be very similar to the collaborator report,
differing only by inclusion of code syntax at the end.";

ods text = " ";
ods text = " We are interested in comparing weight of vehicles with their gas mileage. We
hypothesize that vehicles that weigh more than 3,500 lbs will have worse gas mileage than the
vehicles that do not.  ";
ods text = " ";


************************************************************************ ;
* READ IN DATA *;
************************************************************************ ;
title "Data Management";
* we are using sashelp.cars dataset;
data work.cars;
 set sashelp.cars;
run;
title;


************************************************************************ ;
* DATA MANAGEMENT *;
************************************************************************ ;

* Create categorical weight variable:
  * name variables appropriately and consistently such as
 using similar style (snake_case, camelCase or PascalCase);
  * this program will use snake_case;

title "Create a new binary variable for vehicle weight.";
```

```
data work.cars2;
 set work.cars;

 if weight > 3500 then wt_cat = "Heavier than 3500lbs";
  else if weight <= 3500 then wt_cat = "3500lbs or lighter";

 label wt_cat = "Weight Categories";
run;
title;


************************************************************************ ;
* ANALYSIS *;
************************************************************************ ;
ods text = " ";
ods text = " Analysis plan initially included a t test as seen below.";
ods text = " ";


* Perform t-test;
ods proclabel = 'T-Test Analysis';
title "T-test Analysis";
ods output  TTests = ttest_variance
   statistics = ttest_results
   Equality = test_results2;
proc ttest data = cars2;
 class wt_cat;
 var MPG_City;
run;

* Assign macro values for t-test statistic and p-value for report;
data _null_;
 set test_results2;
 call symputx ("test_stat", put(Fvalue, 5.2));
 call symputx ("pval", put(ProbF, pvalue6.3));
run;




* T-test Results Interpretation;
ods text = " ";
ods text = " The t test comparing mean city gas mileage between cars that weigh over
3500lbs and those that are equal to and less than 3500 have a significant difference (&test_stat,
p&pval).";
ods text = " ";

data mean_diff;
```

```
 set ttest_results(where = (method = " "));
 keep variable class n mean LowerCLMean UpperCLMean;
run;

ods proclabel = "Mean MPG by Weight Class";
title  "Mean MPG by Weight Class";
proc print data = mean_diff noobs;
 format mean LowerCLMean UpperCLMean 5.2;
run;
title;


ods pdf startpage=now; *insert page break for the new regression output;


ods proclabel = 'Regression Analysis';
title "Regression Analysis";
ods output  ParameterEstimates =  ParameterEstimates;
proc reg data = cars2; label;
 model MPG_City = weight;
 label  weight = "Weight"
   MPG_City = "MPG (City)";
quit;

* Convert estimate to reflect a 1000 pound increase rather than 1 pound;
* Assign macro values for parameter estimate and p-value for reporting;
data _null_;
 set ParameterEstimates;
 where variable = "Weight";
 estimate_1000 = estimate*1000;
 call symputx ("estimate", put(estimate_1000, 5.2));
 call symputx ("pvalreg", put(ProbT, pvalue6.3));
run;

* Regression Results Interpretation;
* Regression Results Interpretation;
ods text = " ";
ods text = " Interpretation:";
ods text = "^{style [font_face=Times fontsize=12pt]The linear regression
assessing mean city gas mileage shows a significant relationship with car
weight. For every one-thousand pound increase in weight, the expected MPG
(city) changes by &estimate (
ods text = " ";

* Session Info;
ods text = " Session Information";
ods text = " Date: &SYSDATE9";
ods text = " Operating System: &SYSSCP, &SYSSCPL";
ods text = " SAS Version: &SYSVER";
```

```
ods text = " ";



********************************************************************** ;
* CODE TO INCLUDE SYNTAX OF THIS PROGRAM AT END OF REPORT   *;
********************************************************************** ;

data thepgm;
 length linenum 8 pgmline $250;
 infile "&sasprgm." length=linelen lrecl=256 missover;
 input @;
 linenum = _n_;
 if linelen = . then do;
 ** have a blank line and want to keep it in the output;
  linelen=0;
  pgmline = ' ';
 end;
 if linelen gt 0 then do;
 ** have a program line and want to read it;
 input @1 pgmline $varying. linelen;
 end;
 keep linenum pgmline linelen;
run;

title j=c "Program: &sasprgm";
ods proclabel = "Code used to generate results.";
proc report data=thepgm nowd noheader missing
 style(report) ={font_face='Courier New' font_size=10pt rules=none
frame=void
 cellspacing=0 cellpadding=2 asis=on outputwidth=100%}
 style(column) ={font_face='Courier New' font_size=10pt asis=on just=l} ;
 column linenum pgmline;
 define linenum / order noprint;
 define pgmline / display;
run;


ods text = " End of report.";

title;
ods pdf close;
```

```
* End of program;
quit;
; *'; *"; */;
```

End of report.