Name Jerry

Student Number

STA 441s 2024 Quiz 7

1. (4 points) In your analysis of the Sales Data, you fit a model with *unequal* slopes. You tested whether the slope of the line for software package three (yes, 3) is significantly different from zero at the 0.05 level. Fill in the table below.

Must all be right

t, χ^2 or F Statistic (a number)	p-value (a number or range)	Reject H_0? (Yes or No)	Statistically Significant? (Yes or No)
$t = 5.54$	$p < 0.0001$	Yes	Yes

On your printout, circle the test statistic and write "Question 1" beside it.

2. (~~5 points~~ 8) In your analysis of the Diversity data, you arrived at a model with four explanatory variables, each of which was significantly related to job satisfaction when controlling for the other ones. This was based on the exploratory data.

1 pt (a) What were the four variables? Relations with colleagues at work, Relations with management, Fair opportunities for advancement, Visible minority status

1 pt (b) You then tested these four variables on the replication sample, with a Bonferroni correction. What p-value is necessary for statistical significance at the *joint* 0.05 level? The answer is a number between zero and one.

$0.05/4 = 0.0125$

2 pts (c) Please fill in the table below.

One mark off for each mistake up to a maximum of two.

Variable	t or F Statistic (a number)	Uncorrected p-value (a number or range)	Significant with Bonferroni Correction? (Yes or No)
relcol	$t = 0.66$	0.5123	No
relman	$t = 5.25$	$p < 0.0001$	Yes
fairad	$t = 12.23$	$p < 0.0001$	Yes
minority	$t = -2.77$	$p = 0.0059$	Yes

On your printout, circle the test statistics and p-values. Write "Question 2" beside them.

2 pts (d) In plain, non-statistical language, what do you conclude ~~from the regression analysis~~? Base your conclusions on what replicated, but do not mention the exploration-replication strategy.

about visible minority status and job satisfaction

Allowing for relations with colleagues, relations with management and perceived fair opportunities for advancement, visible minority employees had lower average job satisfaction.

No marks for conclusion without statistical results.

```
1          OPTIONS NONOTES NOSTIMER NOSOURCE NOSYNTAXCHECK;
72
73         /************************* 441s24HW7Sales.sas ******************/
74         title 'STA441S24 Homework 7 Check: Regression on the telephone sales data';
75
76         /* Read data directly from Excel spreadsheet */
77         proc import datafile="/home/u1407221/441s24/data/sales.data.xlsx"
78                     out=sales dbms=xlsx replace;
79                     getnames=yes;
80         /* Input data file is sales.data.xlsx
81            Ouput data set is called sales
82            dbms=xlsx      The input file is an Excel spreadsheet.
83                           Necessary to read an Excel spreadsheet directly under unix/linux
84                           Works in PC environment too except for Excel 4.0 spreadsheets
85                           If there are multiple sheets, use sheet="sheet1" or something.
86            replace      If the data set cars already exists, replace it.
87            getnames=yes   Use column names as variable names.                */
88
89         /* proc print; */
90
NOTE: The import data set has 36 observations and 4 variables.
NOTE: WORK.SALES data set was successfully created.
NOTE: PROCEDURE IMPORT used (Total process time):
      real time           0.01 seconds
      user cpu time       0.01 seconds
      system cpu time     0.00 seconds
      memory              3400.21k
      OS Memory           28156.00k
      Timestamp           02/26/2024 04:21:13 PM
      Step Count                        24  Switch Count  2
      Page Faults                       0
      Page Reclaims                     1026
      Page Swaps                        0
      Voluntary Context Switches        19
      Involuntary Context Switches      0
      Block Input Operations            0
      Block Output Operations           264


91         data phonesales;
92              set sales;
93              if software=1 then s1 = 1; else s1=0;
94              if software=2 then s2 = 1; else s2=0;
95              if software=3 then s3 = 1; else s3=0;
96              last1 = SalesLastQuarter*s1;
97              last2 = SalesLastQuarter*s2;
98              last3 = SalesLastQuarter*s3;
99
NOTE: There were 36 observations read from the data set WORK.SALES.
NOTE: The data set WORK.PHONESALES has 36 observations and 10 variables.
NOTE: DATA statement used (Total process time):
      real time           0.00 seconds
      user cpu time       0.00 seconds
      system cpu time     0.00 seconds
```

```
      memory                951.18k
      OS Memory             26284.00k
      Timestamp             02/26/2024 04:21:13 PM
      Step Count                        25  Switch Count  2
      Page Faults                       0
      Page Reclaims                     156
      Page Swaps                        0
      Voluntary Context Switches        12
      Involuntary Context Switches      0
      Block Input Operations            0
      Block Output Operations           264


100        proc means;
101             title2 'Q1a: Means';
102             class Software;
103             var SalesThisQuarter;
104

NOTE: There were 36 observations read from the data set WORK.PHONESALES.
NOTE: PROCEDURE MEANS used (Total process time):
      real time             0.04 seconds
      user cpu time         0.03 seconds
      system cpu time       0.01 seconds
      memory                11367.93k
      OS Memory             36796.00k
      Timestamp             02/26/2024 04:21:13 PM
      Step Count                        26  Switch Count  2
      Page Faults                       0
      Page Reclaims                     3389
      Page Swaps                        0
      Voluntary Context Switches        22
      Involuntary Context Switches      0
      Block Input Operations            0
      Block Output Operations           0


105        proc reg plots=none data = phonesales;
106             title2 'Q1b: Ignoring Sales Last Quarter';
107             model SalesThisQuarter = s1 s2 s3 /noint;
108             Overall: test s1=s2=s3;
109             Onevs2: test s1=s2;
110             Onevs3: test s1=s3;
111             Twovs3: test s2=s3;
112
113        /* Checking */

NOTE: PROCEDURE REG used (Total process time):
      real time             0.07 seconds
      user cpu time         0.08 seconds
      system cpu time       0.01 seconds
      memory                2952.96k
      OS Memory             33472.00k
      Timestamp             02/26/2024 04:21:13 PM
      Step Count                        27  Switch Count  2
      Page Faults                       0
      Page Reclaims                     916
      Page Swaps                        0
```

```
      Voluntary Context Switches        24
      Involuntary Context Switches      0
      Block Input Operations            0
      Block Output Operations           80


114        proc glm data=phonesales;
115             class software;
116             model SalesThisQuarter = Software;
117             lsmeans Software / pdiff tdiff; /* No correction */
118

NOTE: PROCEDURE GLM used (Total process time):
      real time           2.74 seconds
      user cpu time       0.24 seconds
      system cpu time     0.04 seconds
      memory              26559.37k
      OS Memory           54972.00k
      Timestamp           02/26/2024 04:21:15 PM
      Step Count                        28  Switch Count  2
      Page Faults                       0
      Page Reclaims                     8188
      Page Swaps                        0
      Voluntary Context Switches        1408
      Involuntary Context Switches      0
      Block Input Operations            0
      Block Output Operations           1720


119        proc reg plots=none data=phonesales;
120             title2 'Q1c: Controlling for Sales Last Quarter';
121             model SalesThisQuarter = s1 s2 s3 SalesLastQuarter / noint;
122             SoftwarePackage: test s1=s2=s3;
123             Onevs2: test s1=s2;
124             Onevs3: test s1=s3;
125             Twovs3: test s2=s3;
126

NOTE: PROCEDURE REG used (Total process time):
      real time           0.08 seconds
      user cpu time       0.08 seconds
      system cpu time     0.00 seconds
      memory              2449.46k
      OS Memory           55232.00k
      Timestamp           02/26/2024 04:21:16 PM
      Step Count                        29  Switch Count  2
      Page Faults                       0
      Page Reclaims                     301
      Page Swaps                        0
      Voluntary Context Switches        24
      Involuntary Context Switches      0
      Block Input Operations            0
      Block Output Operations           88


127        proc iml;
NOTE: IML Ready
128             a = 2*0.24/(32 + 2*0.24);
```

```
129             print "Proportion of remaining variation explained by software package:" a;
130
NOTE: Exiting IML.
NOTE: PROCEDURE IML used (Total process time):
      real time           0.01 seconds
      user cpu time       0.01 seconds
      system cpu time     0.00 seconds
      memory              594.65k
      OS Memory           53412.00k
      Timestamp           02/26/2024 04:21:16 PM
      Step Count                        30  Switch Count  0
      Page Faults                       0
      Page Reclaims                     169
      Page Swaps                        0
      Voluntary Context Switches        0
      Involuntary Context Switches      0
      Block Input Operations            0
      Block Output Operations           8


131         proc sgplot data=phonesales;
132              title2 'Q1d: Look at the regression lines';
133              reg x=SalesLastQuarter y=SalesThisQuarter / group = Software;
134

135         proc reg plots=none data = phonesales;
NOTE: PROCEDURE SGPLOT used (Total process time):
      real time           0.25 seconds
      user cpu time       0.07 seconds
      system cpu time     0.01 seconds
      memory              3469.62k
      OS Memory           55212.00k
      Timestamp           02/26/2024 04:21:16 PM
      Step Count                        31  Switch Count  2
      Page Faults                       0
      Page Reclaims                     980
      Page Swaps                        0
      Voluntary Context Switches        325
      Involuntary Context Switches      1
      Block Input Operations            0
      Block Output Operations           728

NOTE: There were 36 observations read from the data set WORK.PHONESALES.

136              title2 'Q1e: Interactions with intercept';
137              model SalesThisQuarter = s1 s2 SalesLastQuarter Last1 Last2;
138              Interaction: test last1=last2=0;
139              Slopes1vs2: test last1=last2;
140              Slope2: test SalesLastQuarter+last2=0;
141              EqualReg: test s1=s2=last1=last2=0;
142

NOTE: PROCEDURE REG used (Total process time):
      real time           0.07 seconds
      user cpu time       0.08 seconds
      system cpu time     0.00 seconds
      memory              2647.87k
      OS Memory           55744.00k
```

```
      Timestamp           02/26/2024 04:21:16 PM
      Step Count                        32  Switch Count  2
      Page Faults                       0
      Page Reclaims                     271
      Page Swaps                        0
      Voluntary Context Switches        20
      Involuntary Context Switches      0
      Block Input Operations            0
      Block Output Operations           88


143        proc iml;
NOTE: IML Ready
144             title2 'Q1e (ix): Remaining variation explained by unequal slopes';
145             a = 2*10.3/(30 + 2*10.3);
145      !                                print a;
146
NOTE: Exiting IML.
NOTE: PROCEDURE IML used (Total process time):
      real time           0.00 seconds
      user cpu time       0.01 seconds
      system cpu time     0.00 seconds
      memory              548.59k
      OS Memory           53924.00k
      Timestamp           02/26/2024 04:21:16 PM
      Step Count                        33  Switch Count  1
      Page Faults                       0
      Page Reclaims                     102
      Page Swaps                        0
      Voluntary Context Switches        10
      Involuntary Context Switches      0
      Block Input Operations            0
      Block Output Operations           0


147        proc iml;
NOTE: IML Ready
148             title2 'Q1e (x, xi and xii)';
149             slope1 = 1.58834+0.62387;
150             slope2 = 1.58834-1.23405;
151             slope3 = 1.58834;
152             print slope1 slope2 slope3;
153
NOTE: Exiting IML.
NOTE: PROCEDURE IML used (Total process time):
      real time           0.00 seconds
      user cpu time       0.01 seconds
      system cpu time     0.01 seconds
      memory              499.75k
      OS Memory           53668.00k
      Timestamp           02/26/2024 04:21:16 PM
      Step Count                        34  Switch Count  1
      Page Faults                       0
      Page Reclaims                     57
      Page Swaps                        0
      Voluntary Context Switches        10
      Involuntary Context Switches      0
      Block Input Operations            0
```

```
      Block Output Operations           0


154        proc reg plots=none data = phonesales;
155             title2 'Q1e: Interactions with no intercept: Cell means coding';
156             model SalesThisQuarter = s1 s2 s3 Last1 Last2 Last3 / noint;
157             Interaction: test last1=last2=last3;
158             Slopes1vs2: test last1=last2;
159             Slopes1vs3: test last1=last3;
160             Slopes2vs3: test last2=last3;
161             EqualReg: test s1=s2=s3, last1=last2=last3;
162
163        quit;

NOTE: PROCEDURE REG used (Total process time):
      real time           0.08 seconds
      user cpu time       0.09 seconds
      system cpu time     0.00 seconds
      memory              2393.56k
      OS Memory           55488.00k
      Timestamp           02/26/2024 04:21:16 PM
      Step Count                        35  Switch Count  2
      Page Faults                       0
      Page Reclaims                     261
      Page Swaps                        0
      Voluntary Context Switches        18
      Involuntary Context Switches      0
      Block Input Operations            0
      Block Output Operations           168


164
165
166
167
168
169        OPTIONS NONOTES NOSTIMER NOSOURCE NOSYNTAXCHECK;
181
```

STA441S24 Homework 7 Check: Regression on the telephone sales data
Q1a: Means

The MEANS Procedure

Analysis Variable : SalesThisQuarter SalesThisQuarter						
Software	N Obs	N	Mean	Std Dev	Minimum	Maximum
1	12	12	81.5833333	14.8596974	62.0000000	114.0000000
2	12	12	79.5833333	6.5983239	70.0000000	91.0000000
3	12	12	73.9166667	11.0983892	56.0000000	87.0000000

STA441S24 Homework 7 Check: Regression on the telephone sales data
Q1b: Ignoring Sales Last Quarter

The REG Procedure
Model: MODEL1
Dependent Variable: SalesThisQuarter SalesThisQuarter

Number of Observations Read	36
Number of Observations Used	36

Note: No intercept in model. R-Square is redefined.

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	221436	73812	571.41	<.0001
Error	33	4262.75000	129.17424		
Uncorrected Total	36	225699			

Root MSE	11.36548	R-Square	0.9811
Dependent Mean	78.36111	Adj R-Sq	0.9794
Coeff Var	14.50399		

Parameter Estimates						
Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
s1		1	81.58333	3.28093	24.87	<.0001
s2		1	79.58333	3.28093	24.26	<.0001
s3		1	73.91667	3.28093	22.53	<.0001

STA441S24 Homework 7 Check: Regression on the telephone sales data
Q1b: Ignoring Sales Last Quarter

The REG Procedure
Model: MODEL1

Test Overall Results for Dependent Variable SalesThisQuarter				
Source	DF	Mean Square	F Value	Pr > F
Numerator	2	189.77778	1.47	0.2448
Denominator	33	129.17424		

STA441S24 Homework 7 Check: Regression on the telephone sales data
Q1b: Ignoring Sales Last Quarter

The REG Procedure
Model: MODEL1

Test Onevs2 Results for Dependent Variable SalesThisQuarter				
Source	DF	Mean Square	F Value	Pr > F
Numerator	1	24.00000	0.19	0.6692
Denominator	33	129.17424		

STA441S24 Homework 7 Check: Regression on the telephone sales data
Q1b: Ignoring Sales Last Quarter

The REG Procedure
Model: MODEL1

Test Onevs3 Results for Dependent Variable SalesThisQuarter				
Source	DF	Mean Square	F Value	Pr > F
Numerator	1	352.66667	2.73	0.1080
Denominator	33	129.17424		

STA441S24 Homework 7 Check: Regression on the telephone sales data
Q1b: Ignoring Sales Last Quarter

The REG Procedure
Model: MODEL1

Test Twovs3 Results for Dependent Variable SalesThisQuarter				
Source	DF	Mean Square	F Value	Pr > F
Numerator	1	192.66667	1.49	0.2306
Denominator	33	129.17424		

STA441S24 Homework 7 Check: Regression on the telephone sales data
Q1b: Ignoring Sales Last Quarter

The GLM Procedure

Class Level Information		
Class	Levels	Values
Software	3	1 2 3

Number of Observations Read	36
Number of Observations Used	36

STA441S24 Homework 7 Check: Regression on the telephone sales data
Q1b: Ignoring Sales Last Quarter

The GLM Procedure

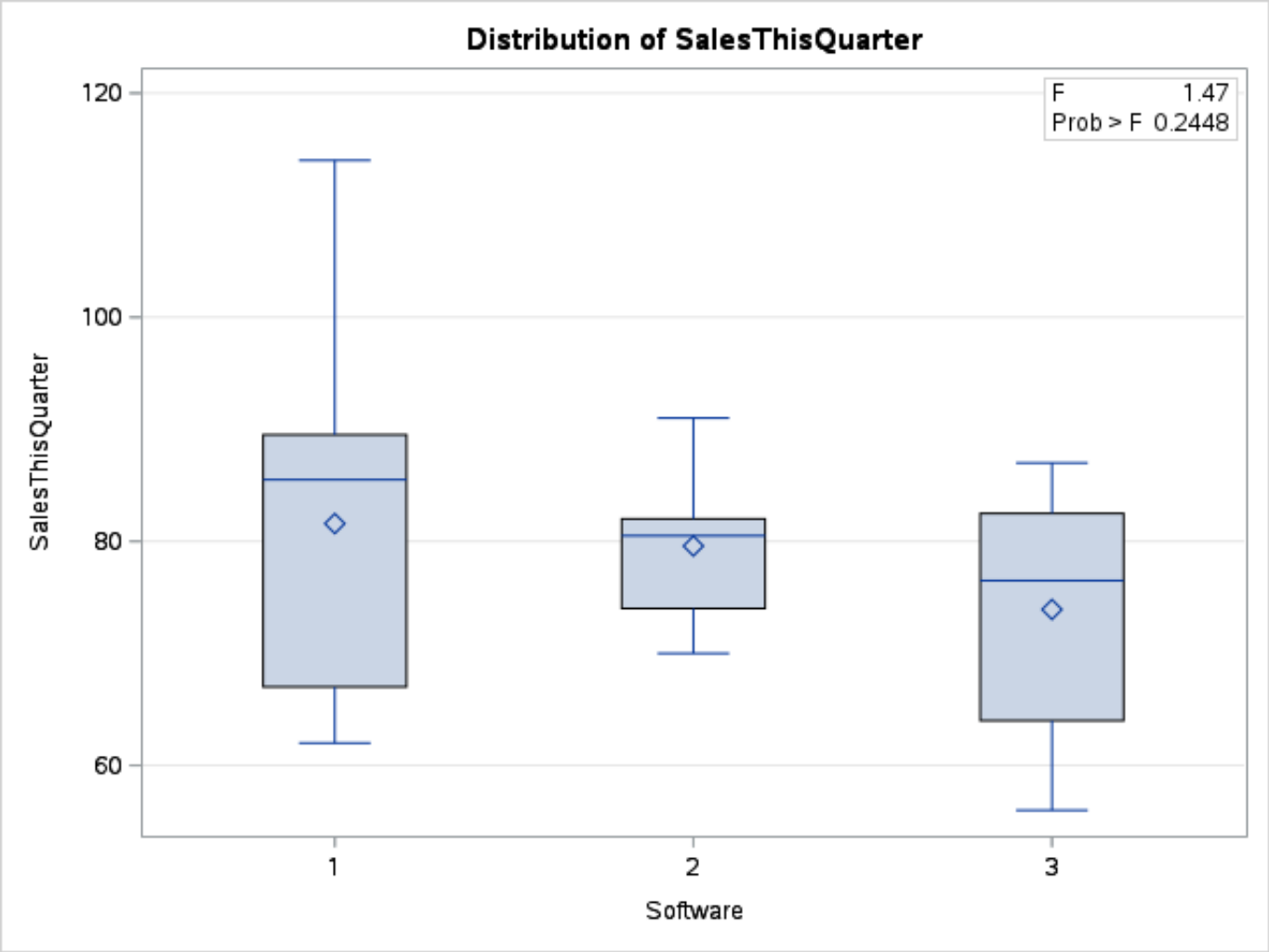
Dependent Variable: SalesThisQuarter SalesThisQuarter

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	379.555556	189.777778	1.47	0.2448
Error	33	4262.750000	129.174242		
Corrected Total	35	4642.305556			

R-Square	Coeff Var	Root MSE	SalesThisQuarter Mean
0.081760	14.50399	11.36548	78.36111

Source	DF	Type I SS	Mean Square	F Value	Pr > F
Software	2	379.5555556	189.7777778	1.47	0.2448

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Software	2	379.5555556	189.7777778	1.47	0.2448



STA441S24 Homework 7 Check: Regression on the telephone sales data
Q1b: Ignoring Sales Last Quarter

The GLM Procedure
Least Squares Means

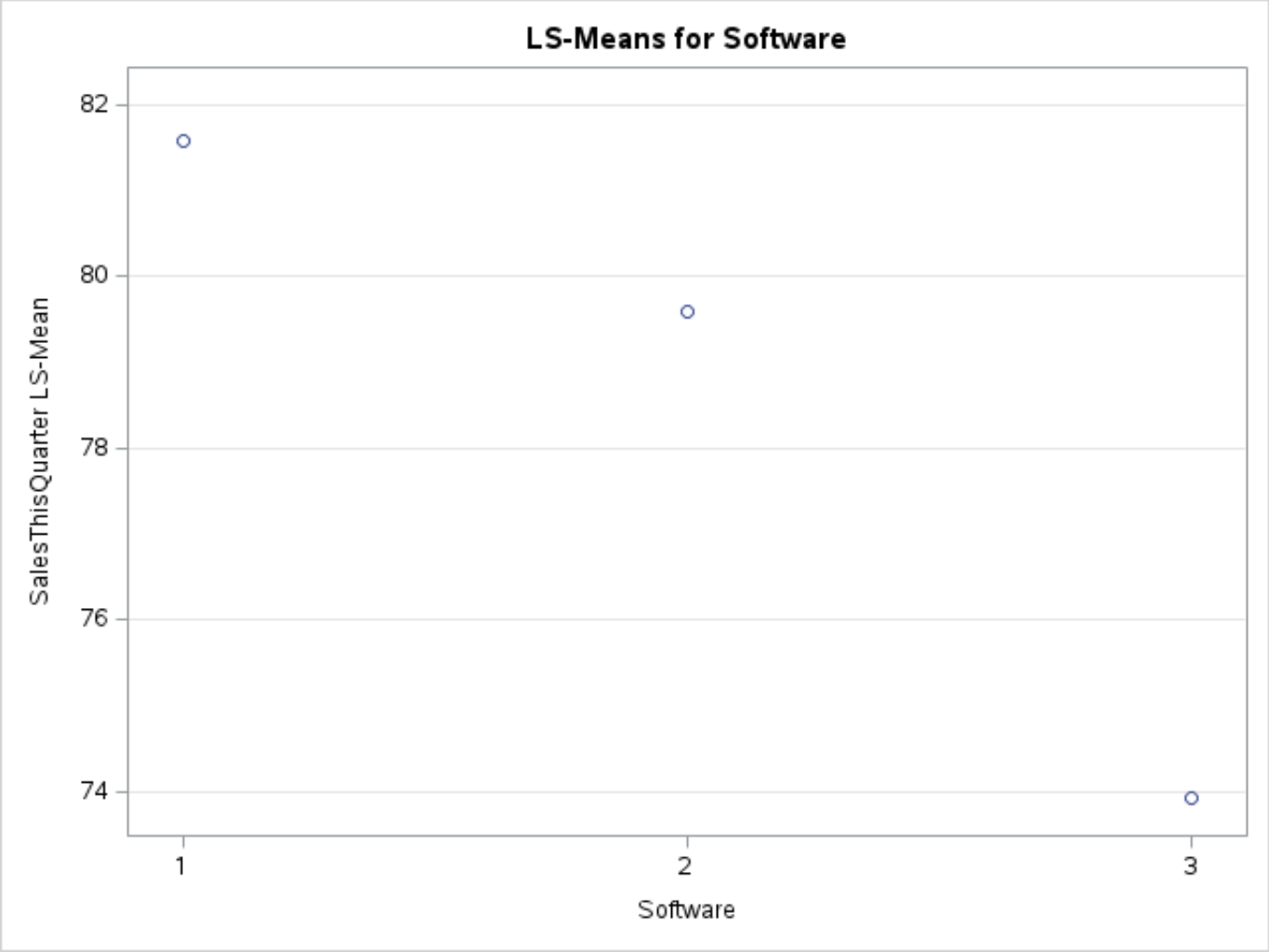
Software	SalesThisQuarter LSMEAN	LSMEAN Number
1	81.5833333	1
2	79.5833333	2
3	73.9166667	3

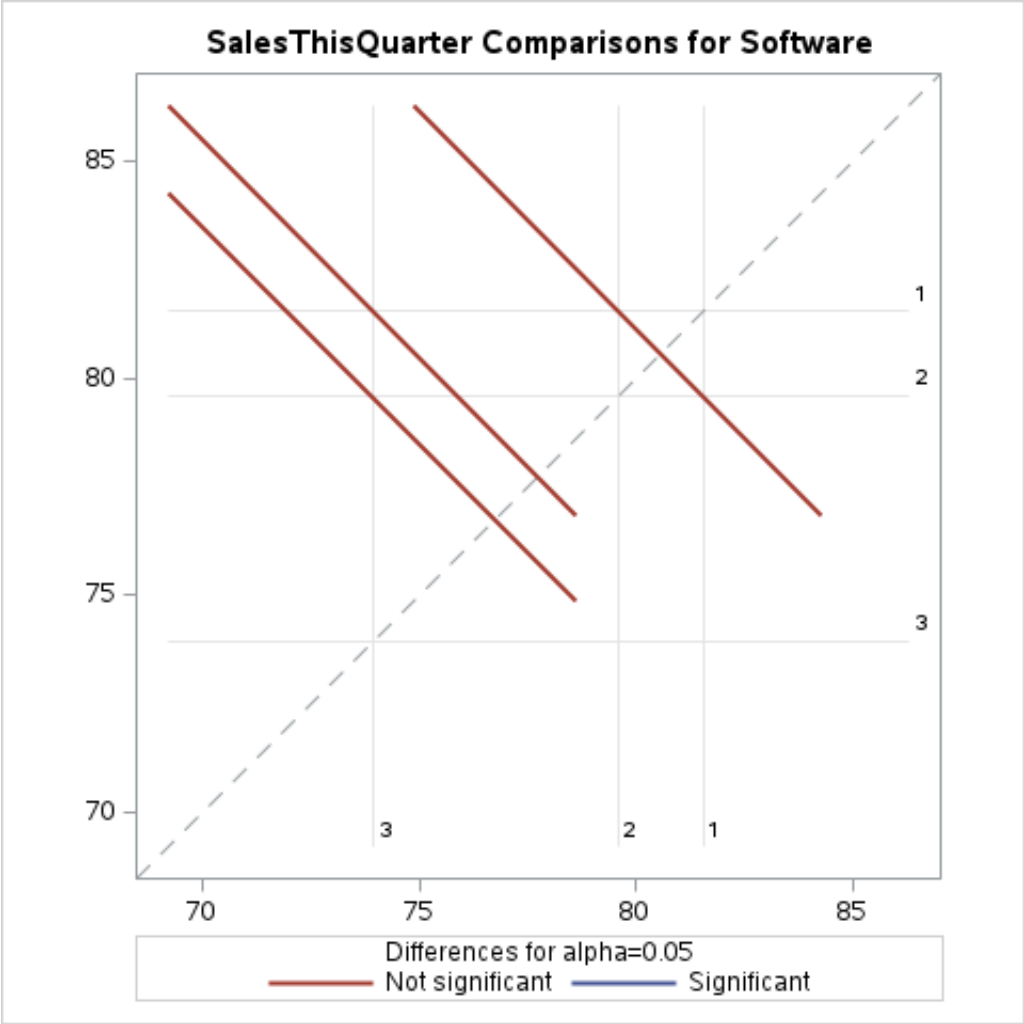
Least Squares Means for Effect Software
t for H0: LSMean(i)=LSMean(j) / Pr > |t|

Dependent Variable: SalesThisQuarter

i/j	1	2	3
1		0.43104 0.6692	1.65232 0.1080
2	-0.43104 0.6692		1.22128 0.2306
3	-1.65232	-1.22128	

	0.1080	0.2306	


LS-Means for Software
82
80
78
76
74
SalesThisQuarter LS-Mean
1
2
3
Software



Note: To ensure overall protection level, only probabilities associated with pre-planned comparisons should be used.

STA441S24 Homework 7 Check: Regression on the telephone sales data
Q1c: Controlling for Sales Last Quarter

The REG Procedure
Model: MODEL1
Dependent Variable: SalesThisQuarter SalesThisQuarter

Number of Observations Read	36
Number of Observations Used	36

Note: No intercept in model. R-Square is redefined.

Analysis of Variance					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	4	224084	56021	1110.10	<.0001
Error	32	1614.87069	50.46471		

Uncorrected Total	36	225699			

Root MSE	7.10385	**R-Square**	0.9928
Dependent Mean	78.36111	**Adj R-Sq**	0.9920
Coeff Var	9.06553		

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
s1		1	-36.44229	16.42229	-2.22	0.0337
s2		1	-35.68878	16.04521	-2.22	0.0333
s3		1	-37.72582	15.54837	-2.43	0.0211
SalesLastQuarter	SalesLastQuarter	1	1.50192	0.20734	7.24	<.0001

STA441S24 Homework 7 Check: Regression on the telephone sales data
Q1c: Controlling for Sales Last Quarter

The REG Procedure
Model: MODEL1

Test SoftwarePackage Results for Dependent Variable SalesThisQuarter

Source	DF	Mean Square	F Value	Pr > F
Numerator	2	12.22473	0.24	0.7863
Denominator	32	50.46471		

STA441S24 Homework 7 Check: Regression on the telephone sales data
Q1c: Controlling for Sales Last Quarter

The REG Procedure
Model: MODEL1

Test Onevs2 Results for Dependent Variable SalesThisQuarter

Source	DF	Mean Square	F Value	Pr > F
Numerator	1	3.34916	0.07	0.7984
Denominator	32	50.46471		

STA441S24 Homework 7 Check: Regression on the telephone sales data
Q1c: Controlling for Sales Last Quarter

The REG Procedure
Model: MODEL1

Test Onevs3 Results for Dependent Variable SalesThisQuarter

Source	DF	Mean Square	F Value	Pr > F

Numerator	1	9.04908	0.18	0.6748
Denominator	32	50.46471		

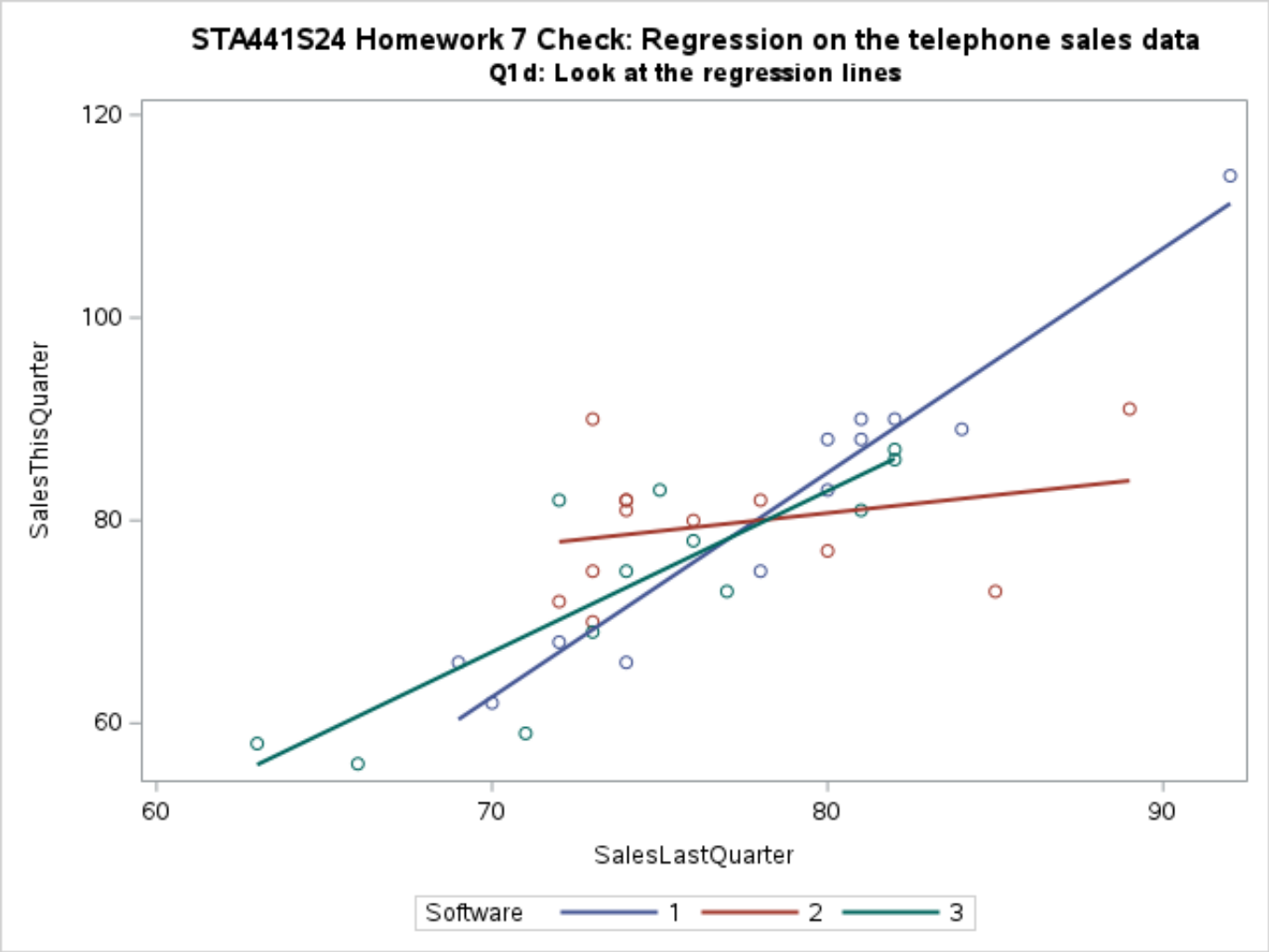
STA441S24 Homework 7 Check: Regression on the telephone sales data
Q1c: Controlling for Sales Last Quarter

The REG Procedure
Model: MODEL1

Test Twovs3 Results for Dependent Variable SalesThisQuarter				
Source	**DF**	**Mean Square**	**F Value**	**Pr > F**
Numerator	1	24.17536	0.48	0.4938
Denominator	32	50.46471		

STA441S24 Homework 7 Check: Regression on the telephone sales data
Q1c: Controlling for Sales Last Quarter

	a
Proportion of remaining variation explained by software package:	0.0147783



STA441S24 Homework 7 Check: Regression on the telephone sales data
Q1e: Interactions with intercept

The REG Procedure
Model: MODEL1
Dependent Variable: SalesThisQuarter SalesThisQuarter

Number of Observations Read	36
Number of Observations Used	36

Analysis of Variance					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	5	3685.05989	737.01198	23.10	<.0001
Error	30	957.24567	31.90819		
Corrected Total	35	4642.30556			

Root MSE	5.64873	**R-Square**	0.7938
Dependent Mean	78.36111	**Adj R-Sq**	0.7594
Coeff Var	7.20859		

Parameter Estimates						
Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	-44.14966	21.36069	-2.07	0.0475
s1		1	-48.10967	29.68078	-1.62	0.1155
s2		1	96.54191	32.39651	2.98	0.0057
SalesLastQuarter	SalesLastQuarter	1	1.58834	0.28652	5.54	<.0001
last1		1	0.62387	0.38786	1.61	0.1182
last2		1	-1.23405	0.42703	-2.89	0.0071

STA441S24 Homework 7 Check: Regression on the telephone sales data
Q1e: Interactions with intercept

The REG Procedure
Model: MODEL1

Test Interaction Results for Dependent Variable SalesThisQuarter				
Source	DF	Mean Square	F Value	Pr > F
Numerator	2	328.81251	10.30	0.0004
Denominator	30	31.90819		

STA441S24 Homework 7 Check: Regression on the telephone sales data
Q1e: Interactions with intercept

The REG Procedure
Model: MODEL1

Test Slopes1vs2 Results for Dependent Variable SalesThisQuarter				
Source	DF	Mean Square	F Value	Pr > F
Numerator	1	653.28548	20.47	<.0001
Denominator	30	31.90819		

STA441S24 Homework 7 Check: Regression on the telephone sales data
Q1e: Interactions with intercept

The REG Procedure
Model: MODEL1

Test Slope2 Results for Dependent Variable SalesThisQuarter				
Source	DF	Mean Square	F Value	Pr > F
Numerator	1	39.94521	1.25	0.2721
Denominator	30	31.90819		

STA441S24 Homework 7 Check: Regression on the telephone sales data
Q1e: Interactions with intercept

The REG Procedure
Model: MODEL1

Test EqualReg Results for Dependent Variable SalesThisQuarter				
Source	DF	Mean Square	F Value	Pr > F
Numerator	4	170.51862	5.34	0.0023
Denominator	30	31.90819		

STA441S24 Homework 7 Check: Regression on the telephone sales data
Q1e (ix): Remaining variation explained by unequal slopes

a
0.4071146

STA441S24 Homework 7 Check: Regression on the telephone sales data
Q1e (x, xi and xii)

slope1	slope2	slope3
2.21221	0.35429	1.58834

STA441S24 Homework 7 Check: Regression on the telephone sales data
Q1e: Interactions with no intercept: Cell means coding

The REG Procedure
Model: MODEL1
Dependent Variable: SalesThisQuarter SalesThisQuarter

Number of Observations Read	36
Number of Observations Used	36

Note: No intercept in model. R-Square is redefined.

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	224742	37457	1173.90	<.0001
Error	30	957.24567	31.90819		
Uncorrected Total	36	225699			

Root MSE	5.64873	R-Square	0.9958
Dependent Mean	78.36111	Adj R-Sq	0.9949

Coeff Var	7.20859		

Parameter Estimates						
Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
s1		1	-92.25933	20.60752	-4.48	0.0001
s2		1	52.39225	24.35683	2.15	0.0397
s3		1	-44.14966	21.36069	-2.07	0.0475
last1		1	2.21221	0.26142	8.46	<.0001
last2		1	0.35428	0.31664	1.12	0.2721
last3		1	1.58834	0.28652	5.54	<.0001

Question 1

STA441S24 Homework 7 Check: Regression on the telephone sales data
Q1e: Interactions with no intercept: Cell means coding

The REG Procedure
Model: MODEL1

Test Interaction Results for Dependent Variable SalesThisQuarter				
Source	DF	Mean Square	F Value	Pr > F
Numerator	2	328.81251	10.30	0.0004
Denominator	30	31.90819		

STA441S24 Homework 7 Check: Regression on the telephone sales data
Q1e: Interactions with no intercept: Cell means coding

The REG Procedure
Model: MODEL1

Test Slopes1vs2 Results for Dependent Variable SalesThisQuarter				
Source	DF	Mean Square	F Value	Pr > F
Numerator	1	653.28548	20.47	<.0001
Denominator	30	31.90819		

STA441S24 Homework 7 Check: Regression on the telephone sales data
Q1e: Interactions with no intercept: Cell means coding

The REG Procedure
Model: MODEL1

Test Slopes1vs3 Results for Dependent Variable SalesThisQuarter				
Source	DF	Mean Square	F Value	Pr > F
Numerator	1	82.55526	2.59	0.1182
Denominator	30	31.90819		

STA441S24 Homework 7 Check: Regression on the telephone sales data
Q1e: Interactions with no intercept: Cell means coding

The REG Procedure
Model: MODEL1

Test Slopes2vs3 Results for Dependent Variable SalesThisQuarter				
Source	DF	Mean Square	F Value	Pr > F
Numerator	1	266.46888	8.35	0.0071
Denominator	30	31.90819		

STA441S24 Homework 7 Check: Regression on the telephone sales data
Q1e: Interactions with no intercept: Cell means coding

The REG Procedure
Model: MODEL1

Test EqualReg Results for Dependent Variable SalesThisQuarter				
Source	DF	Mean Square	F Value	Pr > F
Numerator	4	170.51862	5.34	0.0023
Denominator	30	31.90819		

```
1          OPTIONS NONOTES NOSTIMER NOSOURCE NOSYNTAXCHECK;
72
73         /* 441s24HW7Diversity.sas */
74         title 'STA441 SAS Homework 7: Exploratory Diversity Data';
75
76         /* Read data from Excel spreadsheet */
77         proc import datafile="/home/u1407221/441s24/data/DiversityExplore.xlsx"
78                     out=explor dbms=xlsx replace;
79                     getnames=yes;
80
81         /* proc print; */
82         /* proc freq; tables _all_; */
83
84

NOTE: One or more variables were converted because the data type is not supported by the V9 engine. For more details, run with
      options MSGLEVEL=I.
NOTE: The import data set has 500 observations and 47 variables.
NOTE: WORK.EXPLOR data set was successfully created.
NOTE: PROCEDURE IMPORT used (Total process time):
      real time           0.21 seconds
      user cpu time       0.21 seconds
      system cpu time     0.01 seconds
      memory              3632.18k
      OS Memory           34812.00k
      Timestamp           02/26/2024 07:53:43 PM
      Step Count                        147  Switch Count  4
      Page Faults                       0
      Page Reclaims                     832
      Page Swaps                        0
      Voluntary Context Switches        30
      Involuntary Context Switches      0
      Block Input Operations            0
      Block Output Operations           520


85         proc format;
86              value ynfmt  0 = 'No'  1 = 'Yes';
NOTE: Format YNFMT is already on the library WORK.FORMATS.
NOTE: Format YNFMT has been output.
87              value sexfmt 0 = 'Male' 1 = 'Female';
NOTE: Format SEXFMT is already on the library WORK.FORMATS.
NOTE: Format SEXFMT has been output.
88              value marfmt 1 = 'Never married'
89                           2 = 'Married'
90                           3 = 'Divorced or separated'
91                           4 = 'Widowed';
NOTE: Format MARFMT is already on the library WORK.FORMATS.
NOTE: Format MARFMT has been output.
92

NOTE: PROCEDURE FORMAT used (Total process time):
      real time           0.00 seconds
      user cpu time       0.00 seconds
      system cpu time     0.00 seconds
      memory              259.06k
      OS Memory           32164.00k
      Timestamp           02/26/2024 07:53:43 PM
      Step Count                        148  Switch Count  0
      Page Faults                       0
      Page Reclaims                     15
      Page Swaps                        0
      Voluntary Context Switches        0
      Involuntary Context Switches      0
      Block Input Operations            0
      Block Output Operations           40


93         data diversity;
94              set explor;
95              /* Convert some character variables to numeric.
96                 Must have different names. First the dummy variables. */
97              sex = input(Gender, 8.); minority = input(VisMinority, 8.);
98              foreign = input(CAN_Foreign_Born, 8.);
99              education = input(EDUClevel, 8.); /* Numeric version of education level */
100             mstatus = input(MaritalStatus, 8.);
101             /* I really want age to have the same name. */
102             numericage = input(age, 8.);
```

```
103             drop age; rename numericage=age;
104             /* Indicator for Married */
105            if mstatus = . then married = .;
106                else if mstatus = 2 then married = 1;
107                else married = 0;
108             /* Compute scales */
109             comorg = sum(of com1-com10);
110             relcol = sum(of relc1-relc5);
111             relman = sum(of relm1-relm12);
112             fairad = sum(of fair1-fair6);
113             jobsat = sum(of sat1-sat4);
114             comdiv = sum(of sm1-sm3);
115             label
116                mstatus  = 'Marital Status'
117        comorg   = 'Commitment to the organization'
118             relcol   = 'Relations with colleagues at work'
119             relman   = 'Relations with magnagement'
120             fairad   = 'Fair opportunities for advancement'
121             jobsat   = 'Job satisfaction'
122             comdiv   = "Senior management's commitment to diversity"
123             minority = 'Visible minority'
124             foreign  = 'Born outside Canada'
125             educ     = 'Numeric EDUClevel';
126
127             format sex sexfmt.;
128             format minority foreign married ynfmt.;
129             format mstatus marfmt.;
130
131         /* Checking mstatus, married
132        proc freq; tables MaritalStatus married;
133        proc freq;
134             tables MaritalStatus * mstatus / norow nocol nopercent missing;
135             tables married * mstatus / norow nocol nopercent missing;
136        */
137
```

```
NOTE: Variable educ is uninitialized.
NOTE: There were 500 observations read from the data set WORK.EXPLOR.
NOTE: The data set WORK.DIVERSITY has 500 observations and 59 variables.
NOTE: DATA statement used (Total process time):
      real time           0.00 seconds
      user cpu time       0.01 seconds
      system cpu time     0.00 seconds
      memory              1390.15k
      OS Memory           32684.00k
      Timestamp           02/26/2024 07:53:43 PM
      Step Count                        149  Switch Count  2
      Page Faults                       0
      Page Reclaims                     128
      Page Swaps                        0
      Voluntary Context Switches        10
      Involuntary Context Switches      0
      Block Input Operations            0
      Block Output Operations           520
```

```
138        proc freq;
139             title2 'Q2e: Frequency distributions of categorial variables';
140             tables sex minority foreign EDUCLevel mstatus married;
141
```

```
NOTE: There were 500 observations read from the data set WORK.DIVERSITY.
NOTE: PROCEDURE FREQ used (Total process time):
      real time           0.07 seconds
      user cpu time       0.07 seconds
      system cpu time     0.00 seconds
      memory              3627.00k
      OS Memory           33196.00k
      Timestamp           02/26/2024 07:53:43 PM
      Step Count                        150  Switch Count  3
      Page Faults                       0
      Page Reclaims                     264
      Page Swaps                        0
      Voluntary Context Switches        18
      Involuntary Context Switches      0
      Block Input Operations            0
      Block Output Operations           296
```

```
142        proc means;
143             title2 'Q2e: Basic descriptive statistics on quantitative variables';
144             var comorg -- comdiv age education;
145
```

```
NOTE: There were 500 observations read from the data set WORK.DIVERSITY.
NOTE: PROCEDURE MEANS used (Total process time):
      real time           0.03 seconds
      user cpu time       0.03 seconds
      system cpu time     0.01 seconds
      memory              6759.53k
      OS Memory           38588.00k
      Timestamp           02/26/2024 07:53:43 PM
      Step Count                        151  Switch Count  2
      Page Faults                       0
      Page Reclaims                     1471
      Page Swaps                        0
      Voluntary Context Switches        27
      Involuntary Context Switches      0
      Block Input Operations            0
      Block Output Operations           8
```

```
146        proc freq;
147             title2 'Q2f: Sex and Married: Yes or No';
148             tables sex * married / nocol nopercent chisq;
149
```

```
NOTE: There were 500 observations read from the data set WORK.DIVERSITY.
NOTE: PROCEDURE FREQ used (Total process time):
      real time           0.03 seconds
      user cpu time       0.04 seconds
      system cpu time     0.00 seconds
      memory              1622.71k
      OS Memory           33968.00k
      Timestamp           02/26/2024 07:53:43 PM
      Step Count                        152  Switch Count  5
      Page Faults                       0
      Page Reclaims                     200
      Page Swaps                        0
      Voluntary Context Switches        29
      Involuntary Context Switches      0
      Block Input Operations            0
      Block Output Operations           544
```

```
150        proc corr;
151             title2 'Q2g: Correlation matrix of quantitative variables';
152             var comorg -- comdiv age education;
153
```

```
NOTE: PROCEDURE CORR used (Total process time):
      real time           0.09 seconds
      user cpu time       0.09 seconds
      system cpu time     0.00 seconds
      memory              1389.18k
      OS Memory           33448.00k
      Timestamp           02/26/2024 07:53:43 PM
      Step Count                        153  Switch Count  1
      Page Faults                       0
      Page Reclaims                     57
      Page Swaps                        0
      Voluntary Context Switches        8
      Involuntary Context Switches      0
      Block Input Operations            0
      Block Output Operations           48
```

```
154        proc means;
155             title2 'Q2h: Education broken down by minority status';
156             class minority; var education;
157
```

```
NOTE: There were 500 observations read from the data set WORK.DIVERSITY.
NOTE: PROCEDURE MEANS used (Total process time):
      real time           0.02 seconds
      user cpu time       0.01 seconds
      system cpu time     0.01 seconds
      memory              9282.71k
```

```
      OS Memory           41916.00k
      Timestamp           02/26/2024 07:53:43 PM
      Step Count                        154  Switch Count  2
      Page Faults                       0
      Page Reclaims                     2003
      Page Swaps                        0
      Voluntary Context Switches        19
      Involuntary Context Switches      0
      Block Input Operations            0
      Block Output Operations           8


158        proc glm plots=none;
159             title2 'Q2h: Minority status and education with proc glm';
160             class minority;
161             model education = minority;
162
NOTE: PROCEDURE GLM used (Total process time):
      real time           0.04 seconds
      user cpu time       0.05 seconds
      system cpu time     0.00 seconds
      memory              1943.78k
      OS Memory           35768.00k
      Timestamp           02/26/2024 07:53:43 PM
      Step Count                        155  Switch Count  3
      Page Faults                       0
      Page Reclaims                     222
      Page Swaps                        0
      Voluntary Context Switches        22
      Involuntary Context Switches      0
      Block Input Operations            0
      Block Output Operations           312


163        proc reg plots=none;
164             title2 'Minority status and education with proc reg';
165             model education = minority;
166
NOTE: PROCEDURE REG used (Total process time):
      real time           0.03 seconds
      user cpu time       0.03 seconds
      system cpu time     0.00 seconds
      memory              2553.81k
      OS Memory           36032.00k
      Timestamp           02/26/2024 07:53:43 PM
      Step Count                        156  Switch Count  2
      Page Faults                       0
      Page Reclaims                     261
      Page Swaps                        0
      Voluntary Context Switches        21
      Involuntary Context Switches      0
      Block Input Operations            0
      Block Output Operations           56


167        proc reg plots=none;
168             title2 'Q2i: Big regression on Job Satisfaction';
169             model jobsat = relcol relman fairad comdiv sex minority education
170                            married age foreign;
171             nonsig: test comdiv=sex=education=married=age=foreign = 0;
172
NOTE: PROCEDURE REG used (Total process time):
      real time           0.05 seconds
      user cpu time       0.05 seconds
      system cpu time     0.00 seconds
      memory              2639.25k
      OS Memory           35776.00k
      Timestamp           02/26/2024 07:53:43 PM
      Step Count                        157  Switch Count  2
      Page Faults                       0
      Page Reclaims                     259
      Page Swaps                        0
      Voluntary Context Switches        18
      Involuntary Context Switches      0
      Block Input Operations            0
      Block Output Operations           80
```

```
173        proc iml;
NOTE: IML Ready
174             title2 'Q2j: Proportion of remaining variation with proc iml';
175             print "Proportion explained by non-significant vars";
176             s = 6;
176      !              F = 1.12;
176      !                      NminusP = 427;
177             a = s*F/(NminusP + s*F);
178             print a;
179
NOTE: Exiting IML.
NOTE: PROCEDURE IML used (Total process time):
      real time           0.01 seconds
      user cpu time       0.01 seconds
      system cpu time     0.00 seconds
      memory              519.40k
      OS Memory           33956.00k
      Timestamp           02/26/2024 07:53:43 PM
      Step Count                        158  Switch Count  1
      Page Faults                       0
      Page Reclaims                     56
      Page Swaps                        0
      Voluntary Context Switches        8
      Involuntary Context Switches      0
      Block Input Operations            0
      Block Output Operations           0


180        proc reg plots=none data=diversity;
181             title2 'Q2k: Targeted Regression on Job Satisfaction';
182             model jobsat = relcol relman fairad minority;
183
184        /* I tried lots of other analyses. See 441s18Assignment3Work.txt
185           I also tried diagnostics. The question and SAS code are in a
186           comment statement at the end of the assignment.
187        */
188
189
190        /*********** Bonferroni-corrected cross-validation ***********/
191
192        /* Read data from Excel spreadsheet */

NOTE: PROCEDURE REG used (Total process time):
      real time           0.03 seconds
      user cpu time       0.04 seconds
      system cpu time     0.00 seconds
      memory              2557.12k
      OS Memory           36032.00k
      Timestamp           02/26/2024 07:53:43 PM
      Step Count                        159  Switch Count  2
      Page Faults                       0
      Page Reclaims                     295
      Page Swaps                        0
      Voluntary Context Switches        27
      Involuntary Context Switches      0
      Block Input Operations            0
      Block Output Operations           64


193        proc import datafile="/home/u1407221/441s24/data/DiversityReplic.xlsx"
194                    out=replic dbms=xlsx replace;
195                    getnames=yes;
196

NOTE: One or more variables were converted because the data type is not supported by the V9 engine. For more details, run with
      options MSGLEVEL=I.
NOTE: The import data set has 500 observations and 47 variables.
NOTE: WORK.REPLIC data set was successfully created.
NOTE: PROCEDURE IMPORT used (Total process time):
      real time           0.22 seconds
      user cpu time       0.22 seconds
      system cpu time     0.00 seconds
      memory              3669.18k
      OS Memory           36860.00k
      Timestamp           02/26/2024 07:53:44 PM
      Step Count                        160  Switch Count  4
      Page Faults                       0
```

```
      Page Reclaims                     671
      Page Swaps                        0
      Voluntary Context Switches        28
      Involuntary Context Switches      1
      Block Input Operations            0
      Block Output Operations           520


197         data diversity2; /* Same code as creation of data set diversity */
198              set replic;
199              /* Convert some character variables to numeric.
200                 Must have different names. First the dummy variables. */
201              sex = input(Gender, 8.); minority = input(VisMinority, 8.);
202              foreign = input(CAN_Foreign_Born, 8.);
203              education = input(EDUClevel, 8.); /* Numeric version of education level */
204              mstatus = input(MaritalStatus, 8.);
205              /* I really want age to have the same name. */
206              numericage = input(age, 8.);
207              drop age; rename numericage=age;
208              /* Indicator for Married */
209             if mstatus = . then married = .;
210                 else if mstatus = 2 then married = 1;
211                 else married = 0;
212              /* Compute scales */
213              comorg = sum(of com1-com10);
214              relcol = sum(of relc1-relc5);
215              relman = sum(of relm1-relm12);
216              fairad = sum(of fair1-fair6);
217              jobsat = sum(of sat1-sat4);
218              comdiv = sum(of sm1-sm3);
219              label
220                 mstatus  = 'Marital Status'
221         comorg   = 'Commitment to the organization'
222              relcol   = 'Relations with colleagues at work'
223              relman   = 'Relations with magnagement'
224              fairad   = 'Fair opportunities for advancement'
225              jobsat   = 'Job satisfaction'
226              comdiv   = "Senior management's commitment to diversity"
227              minority = 'Visible minority'
228              foreign  = 'Born outside Canada'
229              educ     = 'Numeric EDUClevel';
230
231              format sex sexfmt.;
232              format minority foreign married ynfmt.;
233              format mstatus marfmt.;
234

NOTE: Variable educ is uninitialized.
NOTE: There were 500 observations read from the data set WORK.REPLIC.
NOTE: The data set WORK.DIVERSITY2 has 500 observations and 59 variables.
NOTE: DATA statement used (Total process time):
      real time           0.00 seconds
      user cpu time       0.00 seconds
      system cpu time     0.00 seconds
      memory              1345.62k
      OS Memory           34732.00k
      Timestamp           02/26/2024 07:53:44 PM
      Step Count                        161  Switch Count  2
      Page Faults                       0
      Page Reclaims                     98
      Page Swaps                        0
      Voluntary Context Switches        14
      Involuntary Context Switches      0
      Block Input Operations            0
      Block Output Operations           528


235         proc freq data=diversity2;
236              title2 'Replication data: Frequency distributions of categorial variables';
237              tables sex minority foreign EDUCLevel mstatus married;
238

NOTE: There were 500 observations read from the data set WORK.DIVERSITY2.
NOTE: PROCEDURE FREQ used (Total process time):
      real time           0.05 seconds
      user cpu time       0.06 seconds
      system cpu time     0.00 seconds
      memory              1166.18k
      OS Memory           34732.00k
```

```
      Timestamp           02/26/2024 07:53:44 PM
      Step Count                        162  Switch Count  3
      Page Faults                       0
      Page Reclaims                     130
      Page Swaps                        0
      Voluntary Context Switches        20
      Involuntary Context Switches      0
      Block Input Operations            0
      Block Output Operations           296


239         proc means data=diversity2;
240              title2 'Replication data: Basic descriptive statistics on quantitative variables';
241              var comorg -- comdiv age education;
242
NOTE: There were 500 observations read from the data set WORK.DIVERSITY2.
NOTE: PROCEDURE MEANS used (Total process time):
      real time           0.03 seconds
      user cpu time       0.03 seconds
      system cpu time     0.01 seconds
      memory              6652.15k
      OS Memory           39612.00k
      Timestamp           02/26/2024 07:53:44 PM
      Step Count                        163  Switch Count  2
      Page Faults                       0
      Page Reclaims                     1364
      Page Swaps                        0
      Voluntary Context Switches        25
      Involuntary Context Switches      0
      Block Input Operations            0
      Block Output Operations           16


243         proc reg plots=none data=diversity2;
244              title2 'Q2n: Replicate Regression on Job Satisfaction';
245              title3 'Bonferroni Correction: 0.05/4 = 0.0125';
246              model jobsat = relcol relman fairad minority;
247
248         quit;
NOTE: PROCEDURE REG used (Total process time):
      real time           0.03 seconds
      user cpu time       0.03 seconds
      system cpu time     0.00 seconds
      memory              2558.71k
      OS Memory           36032.00k
      Timestamp           02/26/2024 07:53:44 PM
      Step Count                        164  Switch Count  3
      Page Faults                       0
      Page Reclaims                     262
      Page Swaps                        0
      Voluntary Context Switches        27
      Involuntary Context Switches      0
      Block Input Operations            0
      Block Output Operations           120


249
250
251         OPTIONS NONOTES NOSTIMER NOSOURCE NOSYNTAXCHECK;
263
```

STA441 SAS Homework 7: Exploratory Diversity Data
Q2e: Frequency distributions of categorial variables

The FREQ Procedure

sex	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Male	288	58.66	288	58.66
Female	203	41.34	491	100.00
Frequency Missing = 9				

Visible minority				
minority	Frequency	Percent	Cumulative Frequency	Cumulative Percent
No	320	67.94	320	67.94
Yes	151	32.06	471	100.00
Frequency Missing = 29				

Born outside Canada				
foreign	Frequency	Percent	Cumulative Frequency	Cumulative Percent
No	357	72.41	357	72.41
Yes	136	27.59	493	100.00
Frequency Missing = 7				

EDUCLevel				
EDUCLevel	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	1	0.20	1	0.20
2	49	9.86	50	10.06
3	114	22.94	164	33.00
4	201	40.44	365	73.44
5	68	13.68	433	87.12
6	2	0.40	435	87.53
7	62	12.47	497	100.00
Frequency Missing = 3				

Marital Status				
mstatus	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Never married	76	15.29	76	15.29
Married	383	77.06	459	92.35
Divorced or separated	34	6.84	493	99.20
Widowed	4	0.80	497	100.00
Frequency Missing = 3				

married	Frequency	Percent	Cumulative Frequency	Cumulative Percent
No	114	22.94	114	22.94
Yes	383	77.06	497	100.00
Frequency Missing = 3				

STA441 SAS Homework 7: Exploratory Diversity Data
Q2e: Basic descriptive statistics on quantitative variables

The MEANS Procedure

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
comorg	Commitment to the organization	500	42.2100000	6.7860150	10.0000000	50.0000000
relcol	Relations with colleagues at work	500	21.1180000	3.4163458	5.0000000	25.0000000
relman	Relations with magnagement	500	45.9440000	10.7110371	12.0000000	60.0000000
fairad	Fair opportunities for advancement	500	20.3020000	5.0272600	7.0000000	30.0000000
jobsat	Job satisfaction	500	14.7880000	3.7076339	4.0000000	20.0000000
comdiv	Senior management's commitment to diversity	500	12.1240000	3.7593736	3.0000000	18.0000000
age		477	42.0607966	9.0782357	20.0000000	64.0000000
education		497	4.0865191	1.3834270	1.0000000	7.0000000

STA441 SAS Homework 7: Exploratory Diversity Data
Q2f: Sex and Married: Yes or No

The FREQ Procedure

Frequency Row Pct	Table of sex by married		
	married		
sex	No	Yes	Total
Male	52 18.18	234 81.82	286
Female	61 30.20	141 69.80	202
Total	113	375	488
Frequency Missing = 12			

Statistics for Table of sex by married

Statistic	DF	Value	Prob
Chi-Square	1	9.6064	0.0019
Likelihood Ratio Chi-Square	1	9.4927	0.0021
Continuity Adj. Chi-Square	1	8.9430	0.0028
Mantel-Haenszel Chi-Square	1	9.5867	0.0020
Phi Coefficient		-0.1403	
Contingency Coefficient		0.1389	
Cramer's V		-0.1403	

Fisher's Exact Test	
Cell (1,1) Frequency (F)	52
Left-sided Pr <= F	0.0015
Right-sided Pr >= F	0.9993
Table Probability (P)	0.0008
Two-sided Pr <= P	0.0023

Sample Size = 488
Frequency Missing = 12

STA441 SAS Homework 7: Exploratory Diversity Data
Q2g: Correlation matrix of quantitative variables

The CORR Procedure

8 Variables:	comorg relcol relman fairad jobsat comdiv age education

Simple Statistics							
Variable	N	Mean	Std Dev	Sum	Minimum	Maximum	Label
comorg	500	42.21000	6.78602	21105	10.00000	50.00000	Commitment to the organization
relcol	500	21.11800	3.41635	10559	5.00000	25.00000	Relations with colleagues at work
relman	500	45.94400	10.71104	22972	12.00000	60.00000	Relations with magnagement
fairad	500	20.30200	5.02726	10151	7.00000	30.00000	Fair opportunities for advancement
jobsat	500	14.78800	3.70763	7394	4.00000	20.00000	Job satisfaction
comdiv	500	12.12400	3.75937	6062	3.00000	18.00000	Senior management's commitment to diversity
age	477	42.06080	9.07824	20063	20.00000	64.00000	
education	497	4.08652	1.38343	2031	1.00000	7.00000	

Pearson Correlation Coefficients Prob > |r| under H0: Rho=0 Number of Observations	comorg	relcol	relman	fairad	jobsat	comdiv	age	education
comorg Commitment to the organization	1.00000 500	0.42422 <.0001 500	0.45668 <.0001 500	0.54485 <.0001 500	0.44917 <.0001 500	0.30542 <.0001 500	0.07948 0.0829 477	0.00031 0.9945 497
relcol Relations with colleagues at work	0.42422 <.0001 500	1.00000 500	0.53683 <.0001 500	0.52031 <.0001 500	0.44149 <.0001 500	0.24976 <.0001 500	0.06598 0.1502 477	-0.01058 0.8141 497
relman Relations with magnagement	0.45668 <.0001 500	0.53683 <.0001 500	1.00000 500	0.65108 <.0001 500	0.53481 <.0001 500	0.29619 <.0001 500	-0.03239 0.4803 477	-0.02447 0.5863 497
fairad Fair opportunities for advancement	0.54485 <.0001 500	0.52031 <.0001 500	0.65108 <.0001 500	1.00000 500	0.63800 <.0001 500	0.31199 <.0001 500	0.03228 0.4819 477	-0.03062 0.4958 497
jobsat Job satisfaction	0.44917 <.0001 500	0.44149 <.0001 500	0.53481 <.0001 500	0.63800 <.0001 500	1.00000 500	0.27995 <.0001 500	0.04522 0.3244 477	0.02424 0.5897 497
comdiv Senior management's commitment to diversity	0.30542 <.0001 500	0.24976 <.0001 500	0.29619 <.0001 500	0.31199 <.0001 500	0.27995 <.0001 500	1.00000 500	0.01638 0.7213 477	-0.02708 0.5470 497

age	0.07948 0.0829 477	0.06598 0.1502 477	-0.03239 0.4803 477	0.03228 0.4819 477	0.04522 0.3244 477	0.01638 0.7213 477	1.00000 477	-0.15320 0.0008 475
education	0.00031 0.9945 497	-0.01058 0.8141 497	-0.02447 0.5863 497	-0.03062 0.4958 497	0.02424 0.5897 497	-0.02708 0.5470 497	-0.15320 0.0008 475	1.00000 497

STA441 SAS Homework 7: Exploratory Diversity Data
Q2h: Education broken down by minority status

The MEANS Procedure

Analysis Variable : education						
Visible minority	**N Obs**	**N**	**Mean**	**Std Dev**	**Minimum**	**Maximum**
No	320	319	3.9435737	1.3929114	1.0000000	7.0000000
Yes	151	149	4.4026846	1.2835733	2.0000000	7.0000000

STA441 SAS Homework 7: Exploratory Diversity Data
Q2h: Minority status and education with proc glm

The GLM Procedure

Class Level Information		
Class	**Levels**	**Values**
minority	2	No Yes

Number of Observations Read	500
Number of Observations Used	468

STA441 SAS Homework 7: Exploratory Diversity Data
Q2h: Minority status and education with proc glm

The GLM Procedure

Dependent Variable: education

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	21.4075170	21.4075170	11.59	0.0007
Error	466	860.8232522	1.8472602		
Corrected Total	467	882.2307692			

R-Square	Coeff Var	Root MSE	education Mean
0.024265	33.23288	1.359140	4.089744

Source	DF	Type I SS	Mean Square	F Value	Pr > F
minority	1	21.40751704	21.40751704	11.59	0.0007

Source	DF	Type III SS	Mean Square	F Value	Pr > F
minority	1	21.40751704	21.40751704	11.59	0.0007

STA441 SAS Homework 7: Exploratory Diversity Data
Minority status and education with proc reg

The REG Procedure
Model: MODEL1
Dependent Variable: education

Number of Observations Read	500
Number of Observations Used	468
Number of Observations with Missing Values	32

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	21.40752	21.40752	11.59	0.0007
Error	466	860.82325	1.84726		
Corrected Total	467	882.23077			

Root MSE	1.35914	R-Square	0.0243
Dependent Mean	4.08974	Adj R-Sq	0.0222
Coeff Var	33.23288		

Parameter Estimates						
Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	3.94357	0.07610	51.82	<.0001
minority	Visible minority	1	0.45911	0.13486	3.40	0.0007

STA441 SAS Homework 7: Exploratory Diversity Data
Q2i: Big regression on Job Satisfaction

The REG Procedure
Model: MODEL1
Dependent Variable: jobsat Job satisfaction

Number of Observations Read	500
Number of Observations Used	438
Number of Observations with Missing Values	62

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F

Model	10	2620.26055	262.02606	33.88	<.0001
Error	427	3302.25314	7.73361		
Corrected Total	437	5922.51370			

Root MSE	2.78094	**R-Square**	0.4424
Dependent Mean	14.92466	**Adj R-Sq**	0.4294
Coeff Var	18.63318		

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	2.14976	1.20091	1.79	0.0741
relcol	Relations with colleagues at work	1	0.11125	0.04780	2.33	0.0204
relman	Relations with magnagement	1	0.05362	0.01818	2.95	0.0034
fairad	Fair opportunities for advancement	1	0.32499	0.03767	8.63	<.0001
comdiv	Senior management's commitment to diversity	1	0.02410	0.03826	0.63	0.5290
sex		1	0.23226	0.27566	0.84	0.4000
minority	Visible minority	1	-0.75657	0.36642	-2.06	0.0395
education		1	0.14196	0.09904	1.43	0.1525
married		1	0.53030	0.33461	1.58	0.1137
age		1	0.00601	0.01531	0.39	0.6948
foreign	Born outside Canada	1	-0.37406	0.36972	-1.01	0.3122

STA441 SAS Homework 7: Exploratory Diversity Data
Q2i: Big regression on Job Satisfaction

The REG Procedure
Model: MODEL1

Test nonsig Results for Dependent Variable jobsat

Source	DF	Mean Square	F Value	Pr > F
Numerator	6	8.62399	1.12	0.3525
Denominator	427	7.73361		

STA441 SAS Homework 7: Exploratory Diversity Data
Q2j: Proportion of remaining variation with proc iml

Proportion explained by non-significant vars

a
0.0154939

STA441 SAS Homework 7: Exploratory Diversity Data
Q2k: Targeted Regression on Job Satisfaction

The REG Procedure
Model: MODEL1
Dependent Variable: jobsat Job satisfaction

Number of Observations Read	500
Number of Observations Used	471
Number of Observations with Missing Values	29

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	2818.95457	704.73864	90.58	<.0001
Error	466	3625.81401	7.78072		
Corrected Total	470	6444.76858			

Root MSE	2.78939	**R-Square**	0.4374
Dependent Mean	14.78769	**Adj R-Sq**	0.4326
Coeff Var	18.86295		

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	3.14157	0.83473	3.76	0.0002
relcol	Relations with colleagues at work	1	0.11140	0.04600	2.42	0.0158
relman	Relations with magnagement	1	0.06051	0.01685	3.59	0.0004
fairad	Fair opportunities for advancement	1	0.33468	0.03597	9.30	<.0001
minority	Visible minority	1	-0.94482	0.28033	-3.37	0.0008

STA441 SAS Homework 7: Exploratory Diversity Data
Replication data: Frequency distributions of categorial variables

The FREQ Procedure

sex	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Male	272	55.51	272	55.51
Female	218	44.49	490	100.00
Frequency Missing = 10				

Visible minority

minority	Frequency	Percent	Cumulative Frequency	Cumulative Percent
No	343	72.06	343	72.06

Yes	133	27.94	476	100.00
Frequency Missing = 24				

Born outside Canada				
foreign	Frequency	Percent	Cumulative Frequency	Cumulative Percent
No	372	75.61	372	75.61
Yes	120	24.39	492	100.00
Frequency Missing = 8				

EDUCLevel				
EDUCLevel	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	4	0.80	4	0.80
2	76	15.29	80	16.10
3	107	21.53	187	37.63
4	177	35.61	364	73.24
5	58	11.67	422	84.91
6	4	0.80	426	85.71
7	71	14.29	497	100.00
Frequency Missing = 3				

Marital Status				
mstatus	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Never married	82	16.50	82	16.50
Married	377	75.86	459	92.35
Divorced or separated	35	7.04	494	99.40
Widowed	3	0.60	497	100.00
Frequency Missing = 3				

married	Frequency	Percent	Cumulative Frequency	Cumulative Percent
No	120	24.14	120	24.14
Yes	377	75.86	497	100.00
Frequency Missing = 3				

STA441 SAS Homework 7: Exploratory Diversity Data
Replication data: Basic descriptive statistics on quantitative variables

The MEANS Procedure

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
comorg	Commitment to the organization	500	42.3940000	6.5480163	20.0000000	50.0000000

relcol	Relations with colleagues at work	500	21.2120000	3.2870451	6.0000000	25.0000000
relman	Relations with magnagement	500	46.6160000	10.9326845	12.0000000	60.0000000
fairad	Fair opportunities for advancement	500	20.3340000	5.1481502	6.0000000	30.0000000
jobsat	Job satisfaction	500	14.6680000	3.8568805	4.0000000	20.0000000
comdiv	Senior management's commitment to diversity	500	12.0540000	3.9321760	3.0000000	18.0000000
age		475	41.8863158	9.2158666	22.0000000	64.0000000
education		497	4.0160966	1.5291992	1.0000000	7.0000000

STA441 SAS Homework 7: Exploratory Diversity Data
Q2n: Replicate Regression on Job Satisfaction
Bonferroni Correction: 0.05/4 = 0.0125

The REG Procedure
Model: MODEL1
Dependent Variable: jobsat Job satisfaction

Number of Observations Read	500
Number of Observations Used	476
Number of Observations with Missing Values	24

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	3391.58068	847.89517	108.92	<.0001
Error	471	3666.63781	7.78479		
Corrected Total	475	7058.21849			

Root MSE	2.79012	**R-Square**	0.4805
Dependent Mean	14.66387	**Adj R-Sq**	0.4761
Coeff Var	19.02721		

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	2.85342	0.85490	3.34	0.0009
relcol	Relations with colleagues at work	1	0.03110	0.04743	0.66	0.5123
relman	Relations with magnagement	1	0.07900	0.01506	5.25	<.0001
fairad	Fair opportunities for advancement	1	0.37973	0.03104	12.23	<.0001
minority	Visible minority	1	-0.79359	0.28663	-2.77	0.0059

Question 2