Homework 1

Group 4

BUAN 6337 Predictive Analytics using SAS

Professor B.P.S. Murthi

Sept 4, 2020

Lin, Shih-Min

Devineni, Rajasekhar

Talreja, Vinita Govind

Zhou, Wei

Read the data

```
Data a1;
Infile 'D:\UTDallas\OneDrive - The University of Texas at
Dallas\0Coursework\0 BUAN 6337 - PREDICTIVE ANALYTICS USING
SAS\HW\car_insurance_19.csv' DLM=',' Missover firstobs=2;
Length Customer $ 12 State $ 12 Education $ 20 Employment_Status $ 20 Gender
$ 2 Policy_Type $ 15 Policy $ 15 Sales_Channel $ 12 Vehicle_Class $ 15;
Input Customer$ State$ Customer_Lifetime_Value Response$ Coverage$
Education$ Effective_To_Date$ Employment_Status$ Gender$ Income
Location_Code$ Marital_Status$ Monthly_Premium_Auto Months_Since_Last_Claim
Months_Since_Policy_Inception Number_of_Open_Complaints Number_of_Policies
Policy_Type$ Policy$ Renew_Offer_Type$ Sales_Channel$ Total_Claim_Amount
Vehicle_Class$ Vehicle_Size$ ;
Informat Effective_To_Date mmddyy10.;
Run;
```

9134 observations and 24 variable

1. What is the distribution of gender, vehicle size, and vehicle class?

```
Proc freq ;
table Gender Vehicle_Size Vehicle_Class;
Run;
Proc print;
Run;
```

The FREQ Procedure

Gender	Frequency	Percent	Cumulative Frequency	Cumulative Percent
F	4658	51.00	4658	51.00
M	4476	49.00	9134	100.00

Vehicle_ Size	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Large	946	10.36	946	10.36
Medsize	6424	70.33	7370	80.69
Small	1764	19.31	9134	100.00

The FREQ Procedure

Vehicle_Class	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Four-Door Car	4621	50.59	4621	50.59
Luxury Car	163	1.78	4784	52.38
Luxury SUV	184	2.01	4968	54.39
SUV	1796	19.66	6764	74.05
Sports Car	484	5.30	7248	79.35
Two-Door Car	1886	20.65	9134	100.00

- The dataset contains approximately equal number of female and male population.
- Medium sized cars are the most bought vehicles with over 70% share.
- Four door cars are popular among the vehicle class with 50% distribution.

2. What is the average customer lifetime value of each level of gender, vehicle size, and vehicle class?

```
Proc means;
var Customer_Lifetime_Value; class Gender;
Run;
Proc means;
var Customer_Lifetime_Value; class Vehicle_Size;
Run;
Proc means;
var Customer_Lifetime_Value; class Vehicle_Class;
Run;
```

The MEANS Procedure

Analysis Variable : Customer_Lifetime_Value

Gender	N Obs	N	Mean	Std Dev	Minimum	Maximum
F	4658	4658	8096.60	6956.06	1898.68	73225.96
M	4476	4476	7909.55	6780.74	1898.01	83325.38

The MEANS Procedure

Analysis Variable : Customer_Lifetime_Value

Vehicle_ Size	N Obs	N	Mean	Std Dev	Minimum	Maximum
Large	946	946	7545.00	6625.40	1940.98	60556.19
Medsize	6424	6424	8050.66	6833.10	1898.01	74228.52
Small	1764	1764	8085.10	7127.66	1898.68	83325.38

The MEANS Procedure

Analysis Variable : Customer_Lifetime_Value

Vehicle_Class	N Obs	N	Mean	Std Dev	Minimum	Maximum
Four-Door Car	4621	4621	6631.73	5164.94	1904.00	41787.90
Luxury Car	163	163	17053.35	12542.36	5886.22	83325.38
Luxury SUV	184	184	17123.00	12671.87	6383.61	73225.96
SUV	1796	1796	10443.51	7939.86	2864.82	58753.88
Sports Car	484	484	10750.99	8462.33	3074.11	67907.27
Two-Door Car	1886	1886	6671.03	5163.89	1898.01	38887.90

Classes	Level	Average Customer Lifetime Value
Gender	Female	8096.60
	Male	7909.55
Vehicle size	Large	7545.00
	Medsize	8050.66
	Small	8085.10
Vehicle class	Four-Door Car	6631.73
	Luxury Car	17053.35
	Luxury SUV	17123.00
	SUV	10443.51
	Sports Car	750.99
	Two-Door Car	6671.03

3. Do Large cars have a higher lifetime value than medsize cars. Do a ttest and report on your findings.

```
Proc ttest;
var Customer_Lifetime_Value;
class Vehicle_Size; where Vehicle_Size in ("Large", "Medsize");
Run;
```

The SAS System

The TTEST Procedure

Variable: Customer_Lifetime_Value

Vehicle_Size	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
Large		946	7545.0	6625.4	215.4	1941.0	60556.2
Medsize		6424	8050.7	6833.1	85.2540	1898.0	74228.5
Diff (1-2)	Pooled		-505.7	6806.8	237.0		
Diff (1-2)	Satterthwaite		-505.7		231.7		

Vehicle_Size	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
Large		7545.0	7122.3	7967.7	6625.4	6339.7	6938.2
Medsize		8050.7	7883.5	8217.8	6833.1	6717.0	6953.4
Diff (1-2)	Pooled	-505.7	-970.3	-40.9917	6806.8	6698.7	6918.5
Diff (1-2)	Satterthwaite	-505.7	-960.2	-51.1690			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	7368	-2.13	0.0329
Satterthwaite	Unequal	1259.7	-2.18	0.0292

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	6423	945	1.06	0.2183

H_0: Larger cars have lifetime value <u>less</u> than or equal to medium size cars.
H_1: Larger cars have lifetime value <u>greater</u> than medium size cars.

- Assumption: Null hypothesis is true.
- Since p > 0.05 in Equality of Variances, it means that the variances are not significantly different. Hence, the pooled method applies.
- p=0.0329 for pooled method. Since p < 0.05, we reject the H_0.
- Hence larger cars have a higher lifetime value than medium-sized cars.

4. Is there a significant difference between men and women in customer lifetime value?

```
Proc ttest;
var Customer_Lifetime_Value;
class Gender;
Run;
```

The SAS System

The TTEST Procedure

Variable: Customer_Lifetime_Value

Gender	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
F		4658	8096.6	6956.1	101.9	1898.7	73226.0
M		4476	7909.6	6780.7	101.4	1898.0	83325.4
Diff (1-2)	Pooled		187.1	6870.7	143.8		
Diff (1-2)	Satterthwaite		187.1		143.7		

Gender	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
F		8096.6	7896.8	8296.4	6956.1	6817.6	7100.3
M		7909.6	7710.9	8108.3	6780.7	6643.1	6924.2
Diff (1-2)	Pooled	187.1	-94.8477	468.9	6870.7	6772.5	6971.8
Diff (1-2)	Satterthwaite	187.1	-94.7043	468.8			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	9132	1.30	0.1934
Satterthwaite	Unequal	9130.1	1.30	0.1932

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	4657	4475	1.05	0.0847

H_0: No significant difference between men and women in customer lifetime value.
H_1: Has a significant difference between men and women in customer lifetime value.

- Equality of variance shows that p >0.05, which implies that we use pooled method.
- Pooled method shows a p-value > 0.05. We cannot reject the null hypothesis.
- There is no significant difference between men and women in customer lifetime value.

5. Use ANOVA to test whether there is difference in customer lifetime value across different sales channels. Which sales channel generates the highest lifetime value?

```
Proc ANOVA;
class Sales_Channel;
model Customer_Lifetime_Value=Sales_Channel;
Run;
```

The SAS System

The ANOVA Procedure

Class Level Information		
Class	Levels	Values
Sales_Channel	4	Agent Branch Call Cen Web

Number of Observations Read	9134
Number of Observations Used	9134

The ANOVA Procedure

Dependent Variable: Customer_Lifetime_Value

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	124717067.24	41572355.748	0.88	0.4503
Error	9130	431046001860	47212048.396		
Corrected Total	9133	431170718927			

R-Square	Coeff Var	Root MSE	Customer_Lifetime_Value Mean
0.000289	85.83577	6871.102	8004.940

Source	DF	Anova SS	Mean Square	F Value	Pr > F
Sales_Channel	3	124717067.2	41572355.7	0.88	0.4503

H_0: All the means of customer lifetime value of different sales channel are equal.

H_1: At least two means of customer lifetime value of different sales channel differ.

- Since $p > 0.05$, we do not reject the null hypothesis.
- Hence, we conclude that the means of customer lifetime value across different sales channel are equal. In other words, there isn't a specific sales channel that generates more lifetime value than the others.

6. What demographic factors (education, income, marital_status) affect customer lifetime value?

```
DATA A2;
SET a1;
if Marital_Status="Single" then status=0;
if Marital_Status="Married" then status=1;
if Marital_Status="Divorced" then status=2;
if Education="High School or Below" then Education_level=0;
if Education="College" then Education_level=1;
if Education="Bachelor" then Education_level=2;
if Education="Master" then Education_level=3;
if Education="Doctor" then Education_level=4;
run;
```

```
Proc REG data=a2;
model Customer_Lifetime_Value=Education_level Income status;
Run;
```

The SAS System

The REG Procedure
Model: MODEL1
Dependent Variable: Customer_Lifetime_Value

Number of Observations Read	9134
Number of Observations Used	9134

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	658807250	219602417	4.66	0.0030
Error	9130	4.305119E11	47153550		
Corrected Total	9133	4.311707E11			

Root MSE	6866.84425	R-Square	0.0015
Dependent Mean	8004.94047	Adj R-Sq	0.0012
Coeff Var	85.78258		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	7792.08974	158.18025	49.26	<.0001
Education_level	1	-142.54898	66.91589	-2.13	0.0332
Income	1	0.00462	0.00243	1.90	0.0575
status	1	252.75767	116.38893	2.17	0.0299

Considering the three demographic variables Income, Education_Level and Marital Status, we find that Education level and Marital Status affect Customer Lifetime Value in 95% of the cases, while Income variable affects Customer Lifetime Value in 93% of the cases.

7. Is there a relationship between renew_offer_type and response (use Chi-sq test)? Which offer type generates the highest response rate?

```
Proc freq ;
tables Renew_Offer_Type*Response/chisq;
Run;
```

The SAS System

The FREQ Procedure

Frequency Percent Row Pct Col Pct	Table of Renew_Offer_Type by Response		
		Response	
Renew_Offer_Type	No	Yes	Total
Offer1	3158 34.57 84.17 40.35	594 6.50 15.83 45.41	3752 41.08
Offer2	2242 24.55 76.62 28.65	684 7.49 23.38 52.29	2926 32.03
Offer3	1402 15.35 97.91 17.91	30 0.33 2.09 2.29	1432 15.68
Offer4	1024 11.21 100.00 13.08	0 0.00 0.00 0.00	1024 11.21
Total	7826 85.68	1308 14.32	9134 100.00

Statistics for Table of Renew_Offer_Type by Response

Statistic	DF	Value	Prob
Chi-Square	3	548.1645	<.0001
Likelihood Ratio Chi-Square	3	751.4675	<.0001
Mantel-Haenszel Chi-Square	1	242.3027	<.0001
Phi Coefficient		0.2450	
Contingency Coefficient		0.2379	
Cramer's V		0.2450	

Sample Size = 9134

- There is a relationship between renew offer type and response since the calculated chi-sq value (548.1645) is greater than the table chi-sq value at 95% confidence.
- Offer 2 generates the most "yes" responses compared to other offers with a value of 23.38%.

8. Do different renew_offer_types have different lifetime values? Which offer type is the best?

```
Proc ANOVA;
class Renew_Offer_Type;
model Customer_Lifetime_Value=Renew_Offer_Type;
means Renew_Offer_Type/TUKEY CLDIFF;
Run;
```

The SAS System

The ANOVA Procedure

Class Level Information		
Class	Levels	Values
Renew_Offer_Type	4	Offer1 Offer2 Offer3 Offer4

Number of Observations Read	9134
Number of Observations Used	9134

The ANOVA Procedure

Dependent Variable: Customer_Lifetime_Value

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	3629085924.8	1209695308.3	25.83	<.0001
Error	9130	427541633002	46828218.292		
Corrected Total	9133	431170718927			

R-Square	Coeff Var	Root MSE	Customer_Lifetime_Value Mean
0.008417	85.48614	6843.115	8004.940

Source	DF	Anova SS	Mean Square	F Value	Pr > F
Renew_Offer_Type	3	3629085925	1209695308	25.83	<.0001

The SAS System

The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for Customer_Lifetime_Value

Note: This test controls the Type I experimentwise error rate.

Alpha	0.05
Error Degrees of Freedom	9130
Error Mean Square	46828218
Critical Value of Studentized Range	3.63381

Comparisons significant at the 0.05 level are indicated by ***.

Renew_Offer_Type Comparison	Difference Between Means	Simultaneous 95% Confidence Limits		
Offer1 - Offer3	709.2	163.0	1255.4	***
Offer1 - Offer2	1310.3	876.7	1744.0	***
Offer1 - Offer4	1527.1	907.2	2147.1	***
Offer3 - Offer1	-709.2	-1255.4	-163.0	***
Offer3 - Offer2	601.1	34.1	1168.2	***
Offer3 - Offer4	817.9	98.3	1537.5	***
Offer2 - Offer1	-1310.3	-1744.0	-876.7	***
Offer2 - Offer3	-601.1	-1168.2	-34.1	***
Offer2 - Offer4	216.8	-421.6	855.2	
Offer4 - Offer1	-1527.1	-2147.1	-907.2	***
Offer4 - Offer3	-817.9	-1537.5	-98.3	***
Offer4 - Offer2	-216.8	-855.2	421.6	

H_0: All the means of lifetime value of different renew offer types are equal.

H_1: At least two means of lifetime value of different renew offer types are different.

- Since $p < 0.05$, we reject the null hypothesis. Hence, at least of two means of lifetime value of different renew offer types differ.
- Tukey HSD test is used to find the best among the offers. The difference between all possible combinations of different offers shows that Offer1 is the best.

9. Is the effectiveness of renew_offer_type different across different states with respect to lifetime value?

```
Proc ANOVA plots=none;
class Renew_Offer_Type State;
model Customer_Lifetime_Value=Renew_Offer_Type State;
means Renew_Offer_Type State/TUKEY CLDIFF;
Run;
```

The SAS System

The ANOVA Procedure

Class Level Information		
Class	Levels	Values
Renew_Offer_Type	4	Offer1 Offer2 Offer3 Offer4
State	5	Arizona California Nevada Oregon Washington

Number of Observations Read	9134
Number of Observations Used	9134

The ANOVA Procedure

Dependent Variable: Customer_Lifetime_Value

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	7	3680635641.4	525805091.63	11.22	<.0001
Error	9126	427490083286	46843094.815		
Corrected Total	9133	431170718927			

R-Square	Coeff Var	Root MSE	Customer_Lifetime_Value Mean
0.008536	85.49972	6844.202	8004.940

Source	DF	Anova SS	Mean Square	F Value	Pr > F
Renew_Offer_Type	3	3629085925	1209695308	25.82	<.0001
State	4	51549717	12887429	0.28	0.8942

Alpha	0.05
Error Degrees of Freedom	9126
Error Mean Square	46843095
Critical Value of Studentized Range	3.63381

Comparisons significant at the 0.05 level are indicated by ***.

Renew_Offer_Type Comparison	Difference Between Means	Simultaneous 95% Confidence Limits		
Offer1 - Offer3	709.2	162.9	1255.5	***
Offer1 - Offer2	1310.3	876.6	1744.1	***
Offer1 - Offer4	1527.1	907.1	2147.2	***
Offer3 - Offer1	-709.2	-1255.5	-162.9	***
Offer3 - Offer2	601.1	34.0	1168.3	***
Offer3 - Offer4	817.9	98.2	1537.7	***
Offer2 - Offer1	-1310.3	-1744.1	-876.6	***
Offer2 - Offer3	-601.1	-1168.3	-34.0	***
Offer2 - Offer4	216.8	-421.7	855.3	
Offer4 - Offer1	-1527.1	-2147.2	-907.1	***
Offer4 - Offer3	-817.9	-1537.7	-98.2	***
Offer4 - Offer2	-216.8	-855.3	421.7	

Alpha	0.05
Error Degrees of Freedom	9126
Error Mean Square	46843095
Critical Value of Studentized Range	3.85843

Comparisons significant at the 0.05 level are indicated by ***.

State Comparison	Difference Between Means	Simultaneous 95% Confidence Limits		
Oregon - Nevada	21.2	-706.4	748.8	
Oregon - Washington	56.4	-699.2	812.1	
Oregon - California	74.3	-420.5	569.0	
Oregon - Arizona	216.6	-365.5	798.6	
Nevada - Oregon	-21.2	-748.8	706.4	
Nevada - Washington	35.2	-877.1	947.5	
Nevada - California	53.1	-658.3	764.4	
Nevada - Arizona	195.4	-579.3	970.0	
Washington - Oregon	-56.4	-812.1	699.2	
Washington - Nevada	-35.2	-947.5	877.1	
Washington - California	17.8	-722.2	757.9	
Washington - Arizona	160.1	-640.9	961.2	
California - Oregon	-74.3	-569.0	420.5	
California - Nevada	-53.1	-764.4	658.3	
California - Washington	-17.8	-757.9	722.2	
California - Arizona	142.3	-419.3	703.9	
Arizona - Oregon	-216.6	-798.6	365.5	
Arizona - Nevada	-195.4	-970.0	579.3	
Arizona - Washington	-160.1	-961.2	640.9	
Arizona - California	-142.3	-703.9	419.3	

Higher the customer lifetime value, more effective is the renew offer. From the ANOVA test results, we conclude that the effectiveness of renew offer types is not significantly different across states.

10. What other interesting insights that are useful to the company in terms of action can be obtained from the data?
    a. Write any three (3) hypotheses. The hypotheses should be useful to the insurance firm.
    b. Do appropriate statistical tests or analysis.
    c. Report what you found in each case and also write how management can use this information to improve their operations.

    1) Hypotheses 1: Different sales channel have different lifetime values.

```
Proc ANOVA plots=none;
class Sales_Channel;
model Customer_Lifetime_Value=Sales_Channel;
means Sales_Channel/TUKEY CLDIFF;
Run;
```

The SAS System

The ANOVA Procedure

Class Level Information		
Class	Levels	Values
Sales_Channel	4	Agent Branch Call Cen Web

Number of Observations Read	9134
Number of Observations Used	9134

The ANOVA Procedure

Dependent Variable: Customer_Lifetime_Value

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	124717067.24	41572355.748	0.88	0.4503
Error	9130	431046001860	47212048.396		
Corrected Total	9133	431170718927			

R-Square	Coeff Var	Root MSE	Customer_Lifetime_Value Mean
0.000289	85.83577	6871.102	8004.940

Source	DF	Anova SS	Mean Square	F Value	Pr > F
Sales_Channel	3	124717067.2	41572355.7	0.88	0.4503

Alpha	0.05
Error Degrees of Freedom	9130
Error Mean Square	47212048
Critical Value of Studentized Range	3.63381

Comparisons significant at the 0.05 level are indicated by ***.

Sales_Channel Comparison	Difference Between Means	Simultaneous 95% Confidence Limits	
Branch - Call Cen	19.6	-526.3	565.6
Branch - Agent	162.0	-297.4	621.4
Branch - Web	339.9	-257.3	937.2
Call Cen - Branch	-19.6	-565.6	526.3
Call Cen - Agent	142.4	-373.6	658.4
Call Cen - Web	320.3	-321.5	962.1
Agent - Branch	-162.0	-621.4	297.4
Agent - Call Cen	-142.4	-658.4	373.6
Agent - Web	177.9	-392.1	747.9
Web - Branch	-339.9	-937.2	257.3
Web - Call Cen	-320.3	-962.1	321.5
Web - Agent	-177.9	-747.9	392.1

H_0: All the means of lifetime value of different sales channel are equal.
H1: At least two means of lifetime value of different sales channel are different.

- Since p > 0.05, we do not reject the null hypothesis.
- There are no significant differences between different sales channel. The insurance company can invest equally among different sales channel.

2) Hypotheses 2: Different auto policies have different lifetime values.

```
Proc ANOVA;
class Policy_Type;
model Customer_Lifetime_Value=Policy_Type;
means Policy_Type/TUKEY CLDIFF;
Run;
```

The SAS System

The ANOVA Procedure

Class Level Information		
Class	Levels	Values
Policy_Type	3	Corporat Personal Special

Number of Observations Read	9134
Number of Observations Used	9134

The ANOVA Procedure

Dependent Variable: Customer_Lifetime_Value

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	206127153.28	103063576.64	2.18	0.1127
Error	9131	430964591774	47197962.082		
Corrected Total	9133	431170718927			

R-Square	Coeff Var	Root MSE	Customer_Lifetime_Value Mean
0.000478	85.82297	6870.077	8004.940

Source	DF	Anova SS	Mean Square	F Value	Pr > F
Policy_Type	2	206127153.3	103063576.6	2.18	0.1127

The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for Customer_Lifetime_Value

Note: This test controls the Type I experimentwise error rate.

Alpha	0.05
Error Degrees of Freedom	9131
Error Mean Square	47197962
Critical Value of Studentized Range	3.31504

Comparisons significant at the 0.05 level are indicated by ***.

Policy_Type Comparison	Difference Between Means	Simultaneous 95% Confidence Limits	
Special - Personal	566.9	-284.2	1417.9
Special - Corporat	779.8	-124.5	1684.2
Personal - Special	-566.9	-1417.9	284.2
Personal - Corporat	213.0	-199.3	625.2
Corporat - Special	-779.8	-1684.2	124.5
Corporat - Personal	-213.0	-625.2	199.3

H_0: All the means of customer lifetime value of different auto policies are equal.

H1: At least two means of customer lifetime value of different auto policies are different.

- Since p > 0.05 we do not reject the null hypothesis.
- There are no significant differences between different auto policies. The insurance company can develop equally among different auto policies.

3) Hypotheses 3: Rural customers are less valuable than Urban customers

```
Proc ttest;
var Customer_Lifetime_Value;
class Location_Code; where Location_Code in ("Rural", "Urban");
Run;
```

The SAS System

The TTEST Procedure

Variable: Customer_Lifetime_Value

Location_Code	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
Rural		1773	7953.7	6595.5	156.6	2124.2	67907.3
Urban		1582	8064.1	6836.3	171.9	2242.1	51426.2
Diff (1-2)	Pooled		-110.4	6710.1	232.1		
Diff (1-2)	Satterthwaite		-110.4		232.5		

Location_Code	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
Rural		7953.7	7646.5	8260.9	6595.5	6385.3	6820.0
Urban		8064.1	7727.0	8401.3	6836.3	6606.1	7083.2
Diff (1-2)	Pooled	-110.4	-565.4	344.6	6710.1	6553.3	6874.7
Diff (1-2)	Satterthwaite	-110.4	-566.4	345.5			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	3353	-0.48	0.6342
Satterthwaite	Unequal	3279.4	-0.47	0.6349

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	1581	1772	1.07	0.1422

H_0: No significant difference between rural and urban customers in terms of customer lifetime value.

H_1: There is a significant difference between rural and urban customers in terms of customer lifetime value.

- Equality of variance shows that p >0.05 which means that it is a pooled variance.
- Pooled method shows a p-value > 0.05. We cannot reject the null hypothesis.
- We conclude that there is no significant difference in rural and urban customers in terms of customer lifetime value. Hence, the insurance company should focus on both rural and urban customers equally.