Customer Value Analysis and Customer Retention Strategy

GROUP 8

SREELEKHA PULIPATI
HARSHADEEP SHINDE
DIVYANK SINGH
SOUNDHARYAA TRICHY VENKATESAN
PRAYASH

Keeping an existing customer is far cheaper than going out and finding a new one.

Customers make their decision to buy a product or service based on their perception of value. This is the underlying assumption behind the measurement of Customer Value Added (CVA). Studies have indicated that the value of a company's product or service, as perceived by its customers, is the single best predictor of future changes in that company's market share. Customer satisfaction with a product or service results from providing superior value - the right combination of quality and price.

Insurance companies make their money by selling insurance policies and collecting premiums from consumers, and they pay out policy claims when necessary. Retaining customers not only makes the insurance company profitable, it helps offset new customer recruitment advertising and marketing costs.

The Customer Value Analysis Dashboard focuses on various aspects of insurance coverage for varied employment status and helps to analyze the value of the future cash flows attributed to the customer during his/her entire relationship with the company.

We are using dataset of car insurance company which has 9134 observation and 24 variables. We will be doing Customer Value Analysis and finding out the factors though which we can retain old customers as well as improve the business though data visualization.

The main KPIs we are looking here are geographical region, gender, marital status, vehicle class/type, claims history, coverage, response, employment status, education, monthly premium, policy type and income.

Data Acquisition:

We are using customer dataset of car insurance company for data visualization.
Original Dataset: Original Dataset CVA
Cleaned Dataset: Cleaned Dataset CVA
Source: https://www.kaggle.com/pankajjsh06/ibm-watson-marketing-customer-value-data

Data Manipulation:

We have analyzed our dataset through R Studio and done cleaning through excel.

1)Summary of the Dataset:

Our Dataset contains 24 variables 9134 records

RStudio: Notebook Output

```
'data.frame':   9134 obs. of  24 variables:
 $ Customer                     : Factor w/ 9134 levels "AA10041","AA11235",..: 601 5947 97 8017 2489 4948 8434 756 1352 548 ...
 $ State                        : Factor w/ 5 levels "Arizona","California",..: 5 1 3 2 5 4 4 1 4 4 ...
 $ Customer.Lifetime.Value      : num  2764 6980 12887 7646 2814 ...
 $ Response                     : Factor w/ 2 levels "No","Yes": 1 1 1 1 1 2 2 1 2 1 ...
 $ Coverage                     : Factor w/ 3 levels "Basic","Extended",..: 1 2 3 1 1 1 1 3 1 2 ...
 $ Education                    : Factor w/ 5 levels "Bachelor","College",..: 1 1 1 1 1 1 2 5 1 2 ...
 $ Effective.To.Date            : Factor w/ 59 levels "1/1/11","1/10/11",..: 48 25 42 13 53 18 48 10 19 40 ...
 $ EmploymentStatus             : Factor w/ 5 levels "Disabled","Employed",..: 2 5 2 5 2 2 2 5 3 2 ...
 $ Gender                       : Factor w/ 2 levels "F","M": 1 1 1 2 2 1 1 2 2 1 ...
 $ Income                       : int  56274 0 48767 0 43836 62902 55350 0 14072 28812 ...
 $ Location.Code                : Factor w/ 3 levels "Rural","Suburban",..: 2 2 2 2 1 1 2 3 2 3 ...
 $ Marital.Status               : Factor w/ 3 levels "Divorced","Married",..: 2 3 2 2 3 2 2 3 1 2 ...
 $ Monthly.Premium.Auto         : int  69 94 108 106 73 69 67 101 71 93 ...
 $ Months.Since.Last.Claim      : int  32 13 18 18 12 14 0 0 13 17 ...
 $ Months.Since.Policy.Inception: int  5 42 38 65 44 94 13 68 3 7 ...
 $ Number.of.Open.Complaints    : int  0 0 0 0 0 0 0 0 0 0 ...
 $ Number.of.Policies           : int  1 8 2 7 1 2 9 4 2 8 ...
 $ Policy.Type                  : Factor w/ 3 levels "Corporate Auto",..: 1 2 2 1 2 2 1 1 1 3 ...
 $ Policy                       : Factor w/ 9 levels "Corporate L1",..: 3 6 6 2 4 6 3 3 3 8 ...
 $ Renew.Offer.Type             : Factor w/ 4 levels "Offer1","Offer2",..: 1 3 1 1 1 2 1 1 1 2 ...
 $ Sales.Channel                : Factor w/ 4 levels "Agent","Branch",..: 1 1 1 3 1 4 1 1 1 2 ...
 $ Total.Claim.Amount           : num  385 1131 566 530 138 ...
 $ Vehicle.Class                : Factor w/ 6 levels "Four-Door Car",..: 6 1 6 5 1 6 1 1 1 1 ...
 $ Vehicle.Size                 : Factor w/ 3 levels "Large","Medsize",..: 2 2 2 2 2 2 2 2 2 2 ...
```

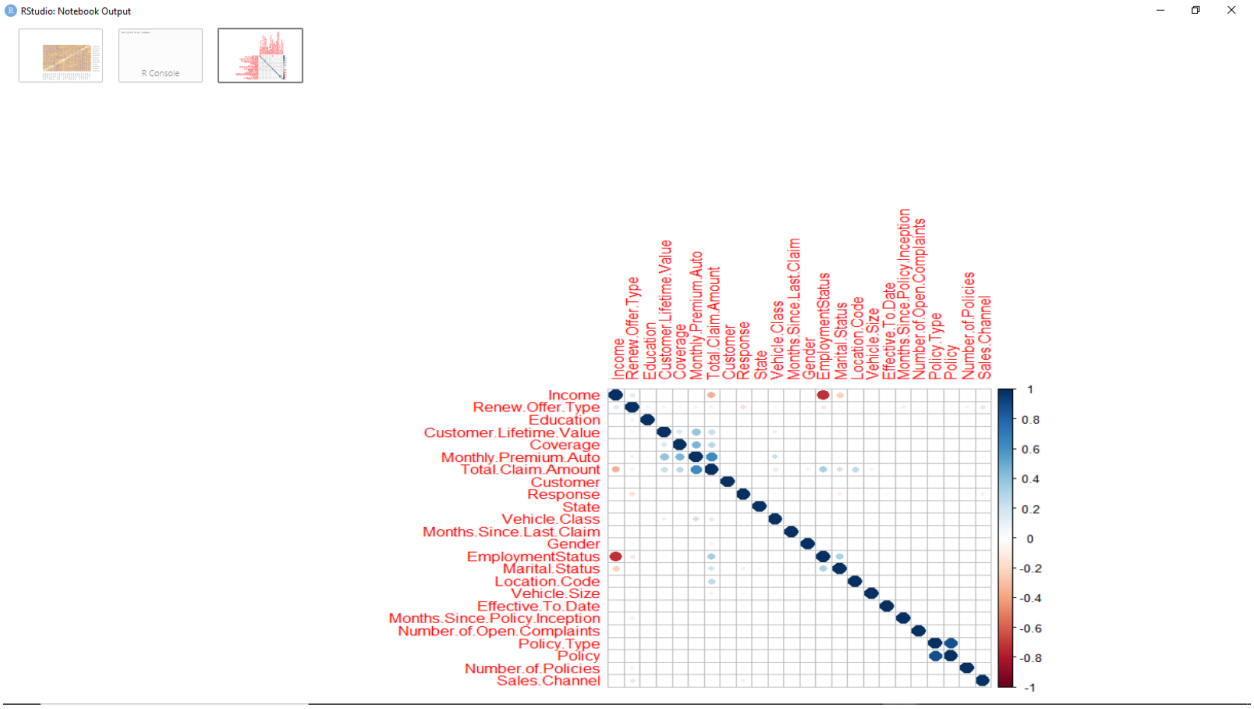
2) Cleaning

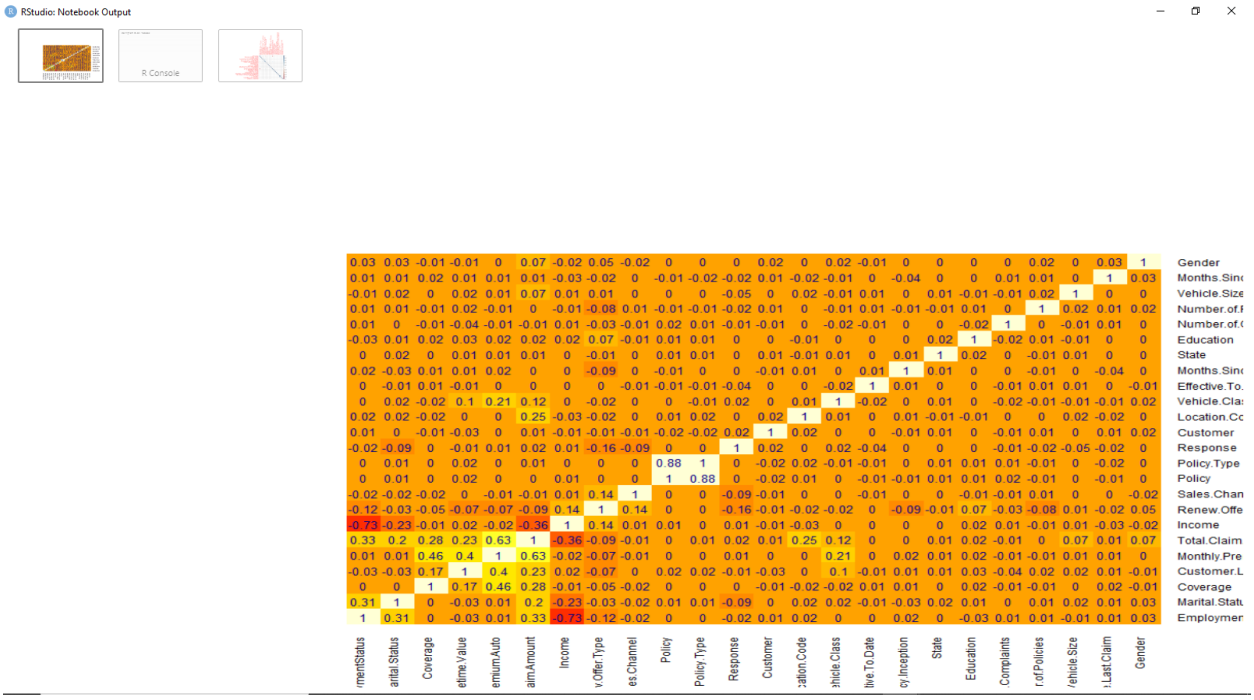
There are over-all 4 null values in the dataset. The record number 9083,9120, 9124 and 9129 doesn't have value in Gender column. Since null values constitute 0.044% of the total record which is less than 2% and thus, we can delete those records.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Custom	State	Customer Lifetime Valu	Respon	Covera	Educati	Effective To Da	EmploymentSta	Gender	Income	Locatio	Marital	Monthly Premium Aut
9083	RX14132	California	7228.005386	No	Basic	High Scho	1/6/2011	Employed		25926	Suburban	Married	61
9120	OT56964	California	9424.256842	No	Basic	Master	2/10/2011	Employed		46897	Urban	Married	118
9124	FH43628	California	25464.82059	Yes	Extended	College	2/1/2011	Retired		13663	Suburban	Single	97
9129	TF56202	California	5032.165498	No	Basic	College	2/12/2011	Employed		66367	Suburban	Divorced	64

3) Correlation between variables:

We have used 2 packages to display heatmap to find the correlation between variables.





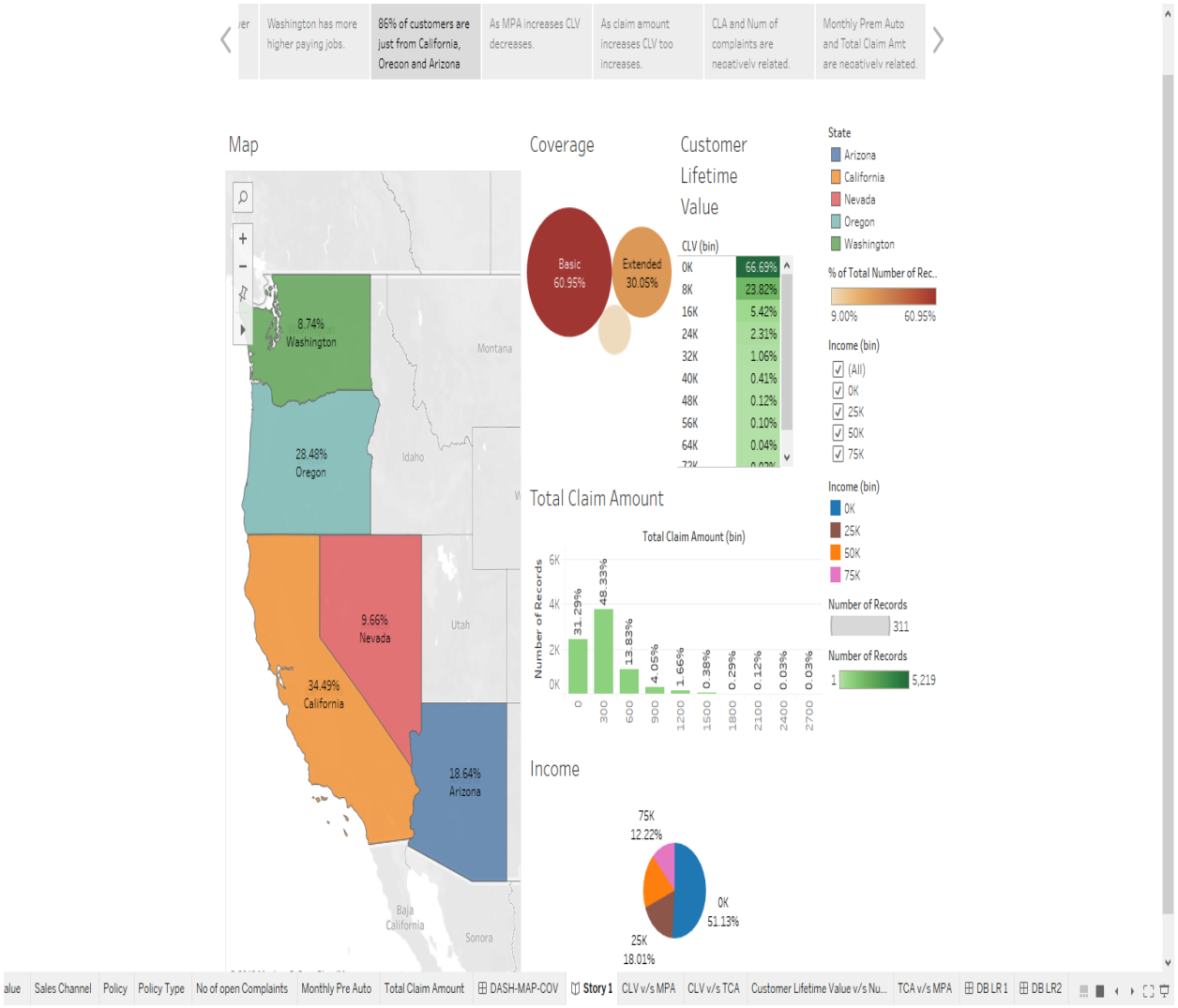
4)Here we can say that our primary variables:

Customer Lifetime Variable is dependent on Total Claim Amount, Monthly Premium Amount and Coverage.

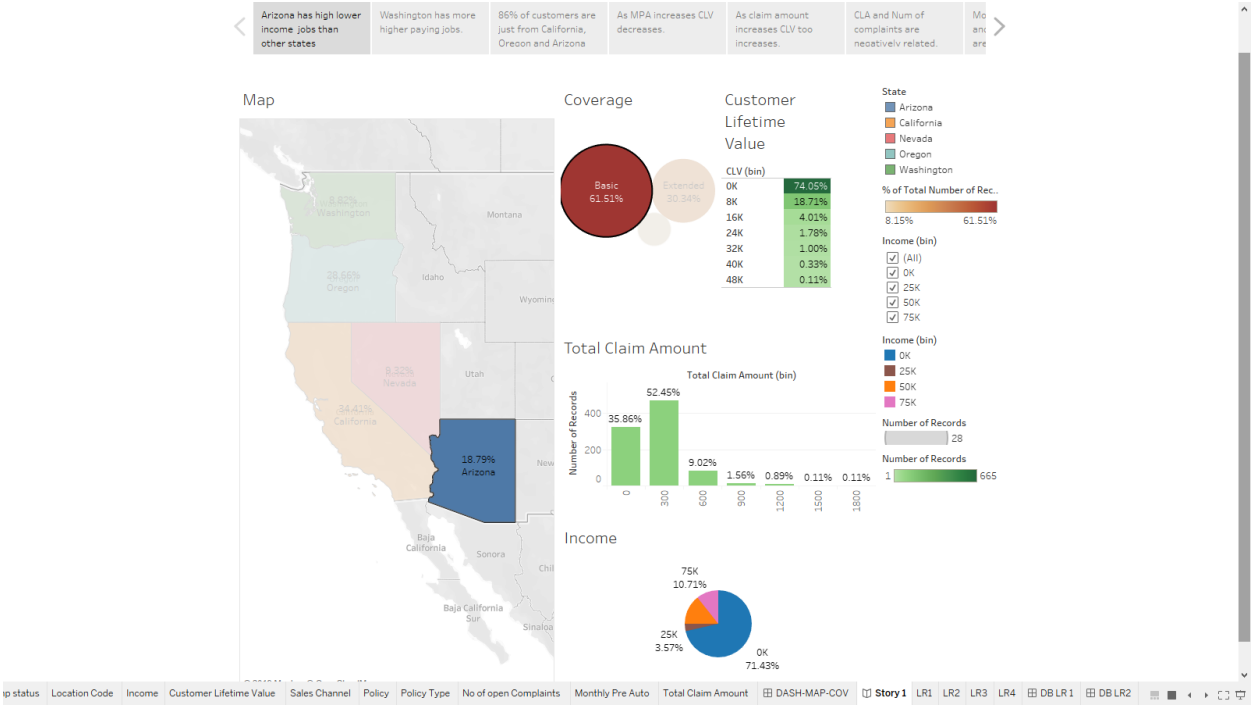
Response is dependent on Renew Offer Type, Marital Status and Sales Channel.

Data Visualization: *Insights*

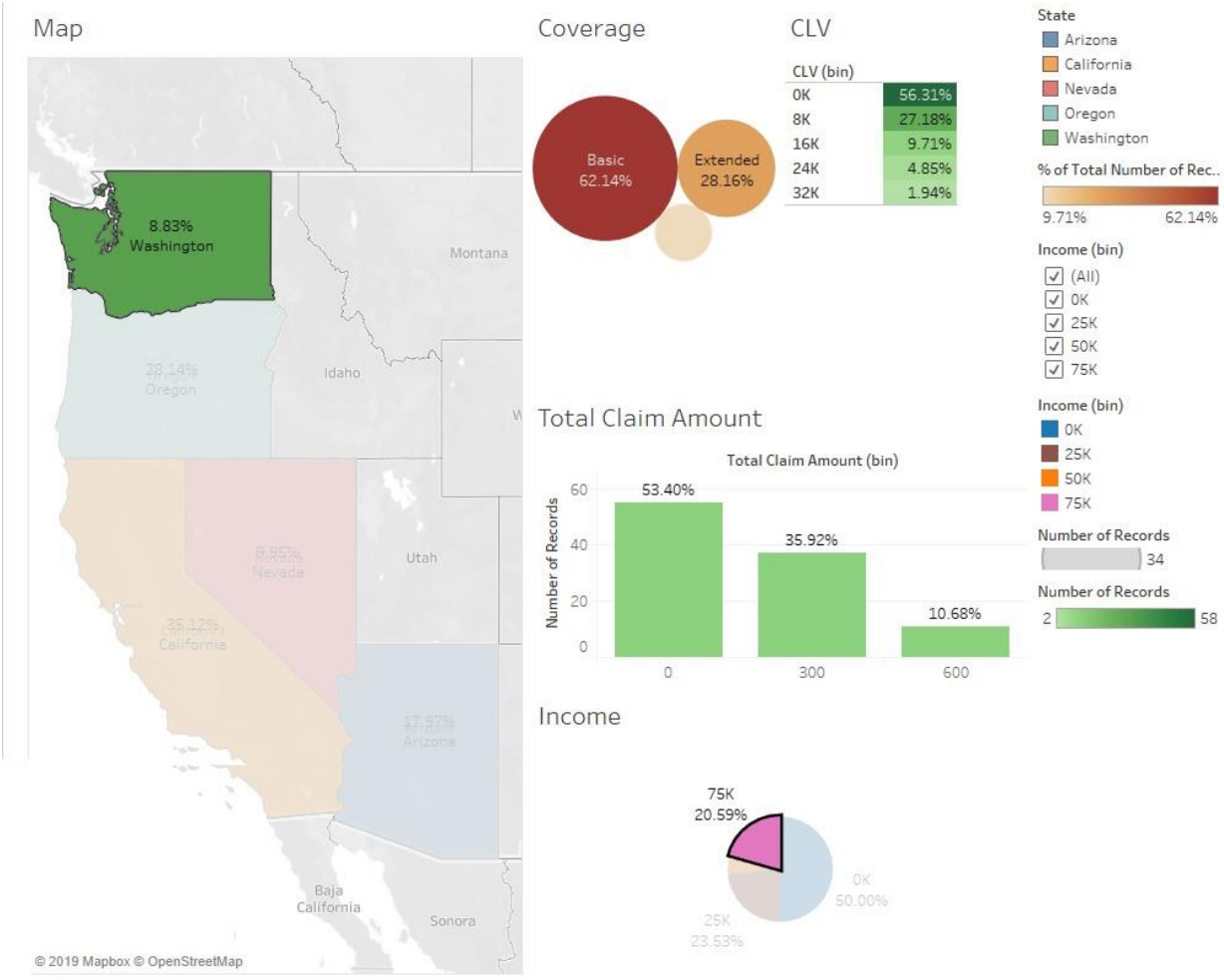
1. 86 % of our customers are from Arizona, Washington and Oregon. So, we must focus on retaining the customers from these states. We must focus on the other regions too, to maximize the business.



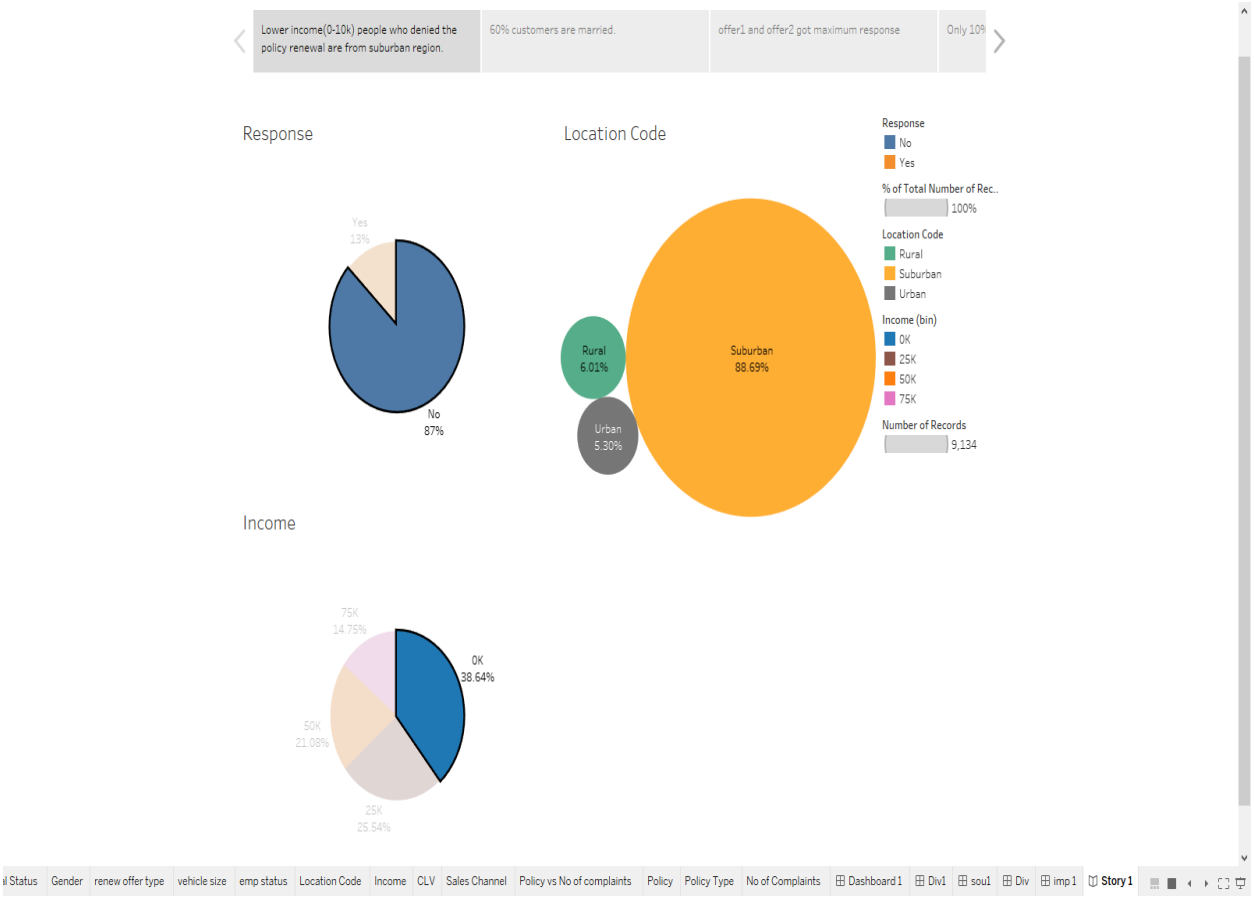
2. Arizona has high percentage of lower income people than compared to other states, its more that 71%. Hence, company should promote basic coverage policies in this state in order to retain maximum of the customers.



3. Washington has more high paying jobs, which indicates more potential customers for premium insurance policy. But we can see out of these higher income customers 60% are using basic coverage policy whereas only 10% are using premium policy. Hence, insurance company should focus on upgrading this class to premium policies.

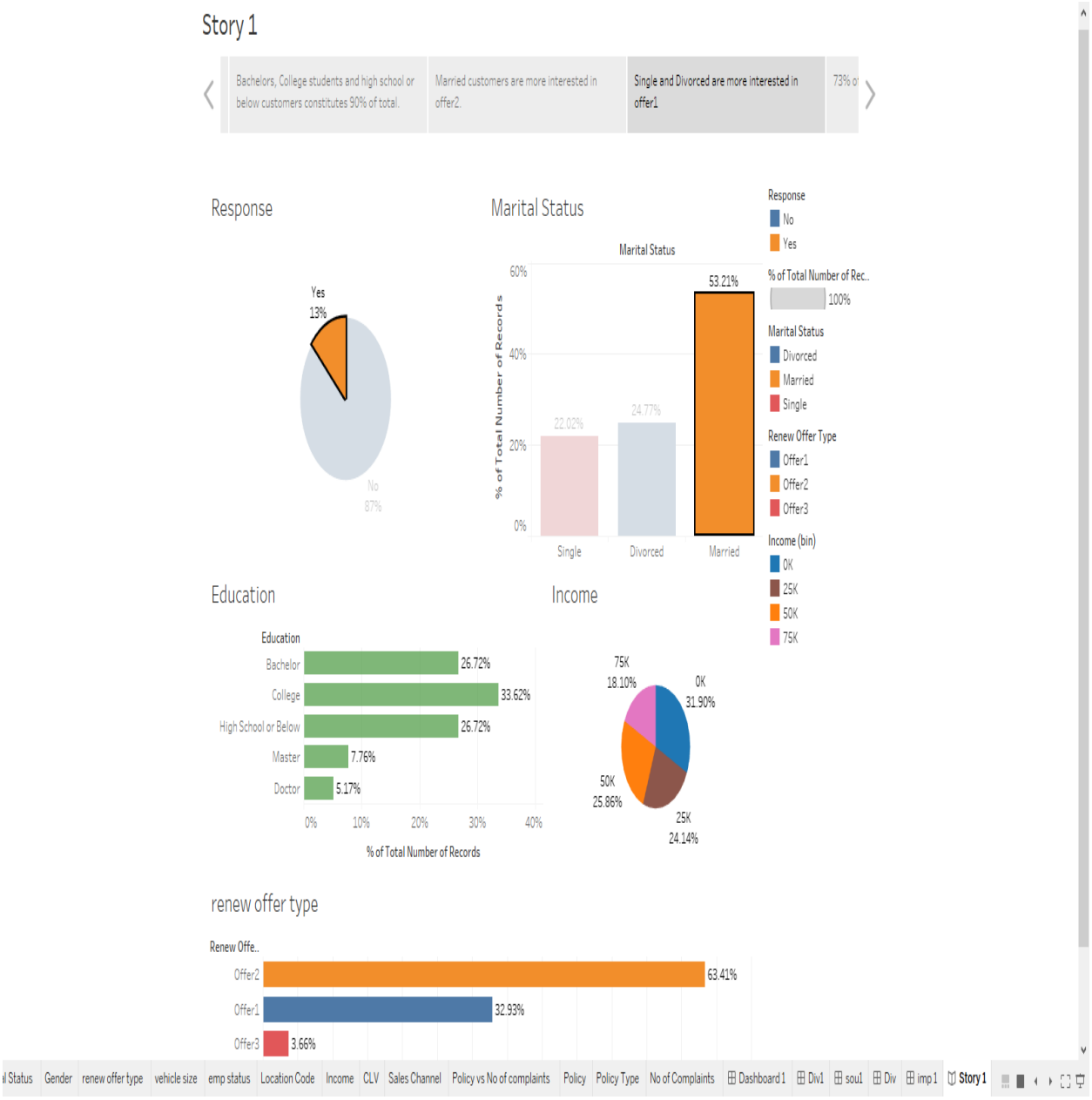


4. 89% customers are from suburbs who denied the offer are in lower income class. So, company needs to modify the basic coverage policy in-order to retain their maximum customers.



5. Offer 1 and 2 have received the maximum response from customers. The insurance company should focus on marketing these offers only to their customers. Majority of the company's customers are married. And so, the policies should attract more families in-order to retain the maximum customers.

6. Offer 2 received maximum attraction from married customers and offer 1 received maximum attraction from single and divorced people. In order to maximize the profits and bring in more customers, the company should promote offer 2 to married customers and offer 1 to single customers. Instead of confusing the customers with all the offers.



Story 1

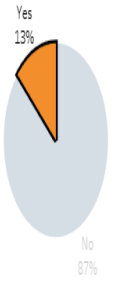
Bachelors, College students and high school or below customers constitutes 90% of total.

Married customers are more interested in offer2.

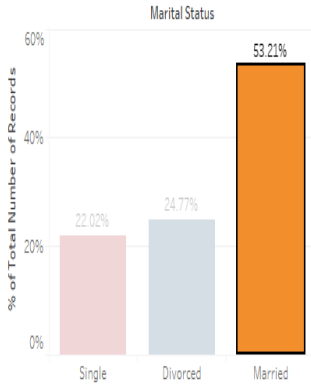
Single and Divorced are more interested in offer1

73% o

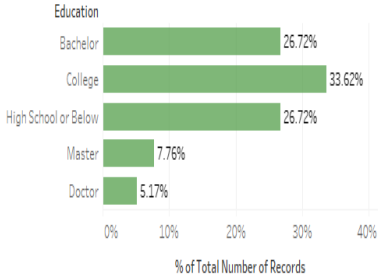
Response



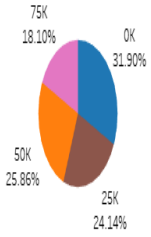
Marital Status



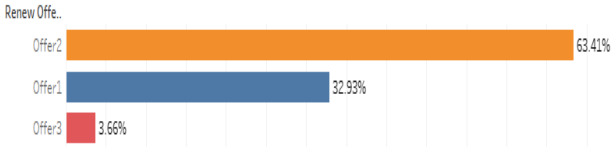
Education



Income



renew offer type

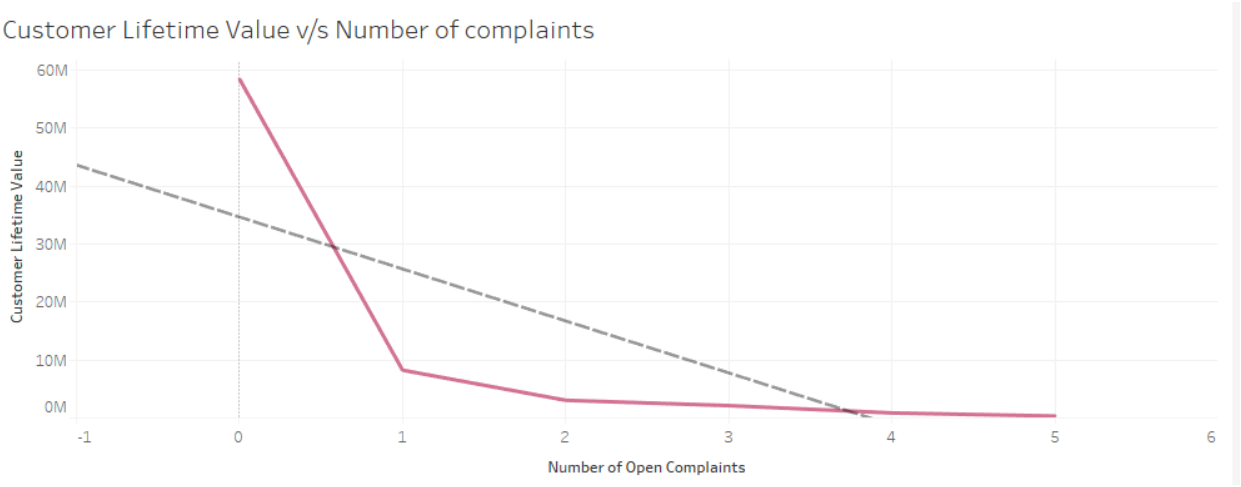


7. When we correlate between the number of open complaints and customer lifetime value (CLV), we can see that as the number of open complaints increases, the customer lifetime value decreases. For this to show, we have drawn the visualization below and plotted the linear regression line for it.

   This is equation for it

   Customer Lifetime value =

   (-8.9575 * 10^6)* (Number of Open complaints) + 3.4581 * 10^7

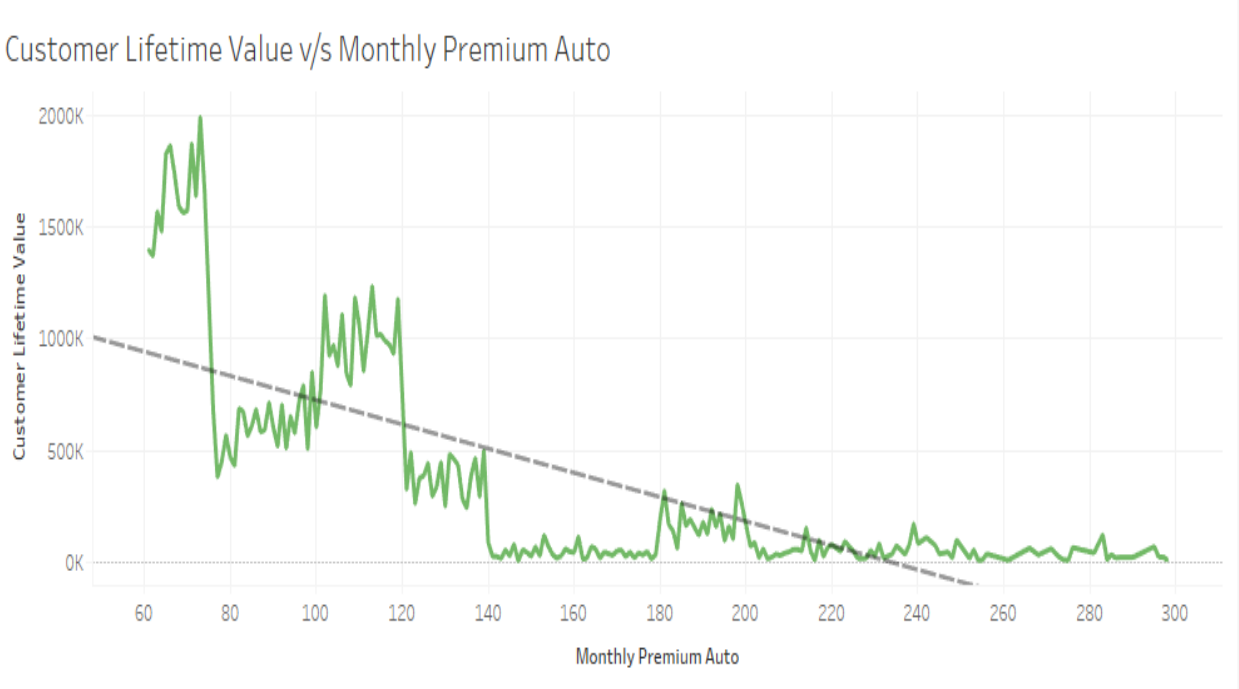


The above graph and linear regression equation provides us sufficient evidence to say that, customer lifetime value negatively relates with number of open complaints. Which also seems reasonable, customers won't stay if the number of open complaints are more. So, company should focus on solving the complaints for the customer to stay and act on such complaints as soon as possible.

8. In this visualization monthly premium auto is negatively related with customer lifetime value. This means that increase in monthly premium auto will decrease the customer lifetime value. Here we have used regression in tableau. The visualization implies that as the premium amount decreases there are higher chances that customers will leave the policy and thus their Lifetime Value decreases and will cause loss to the company. So, we need to keep the monthly premium lower with exciting offers in order to make customers continue availing company's insurance policies.

The equation provided by linear regression is mentioned here:

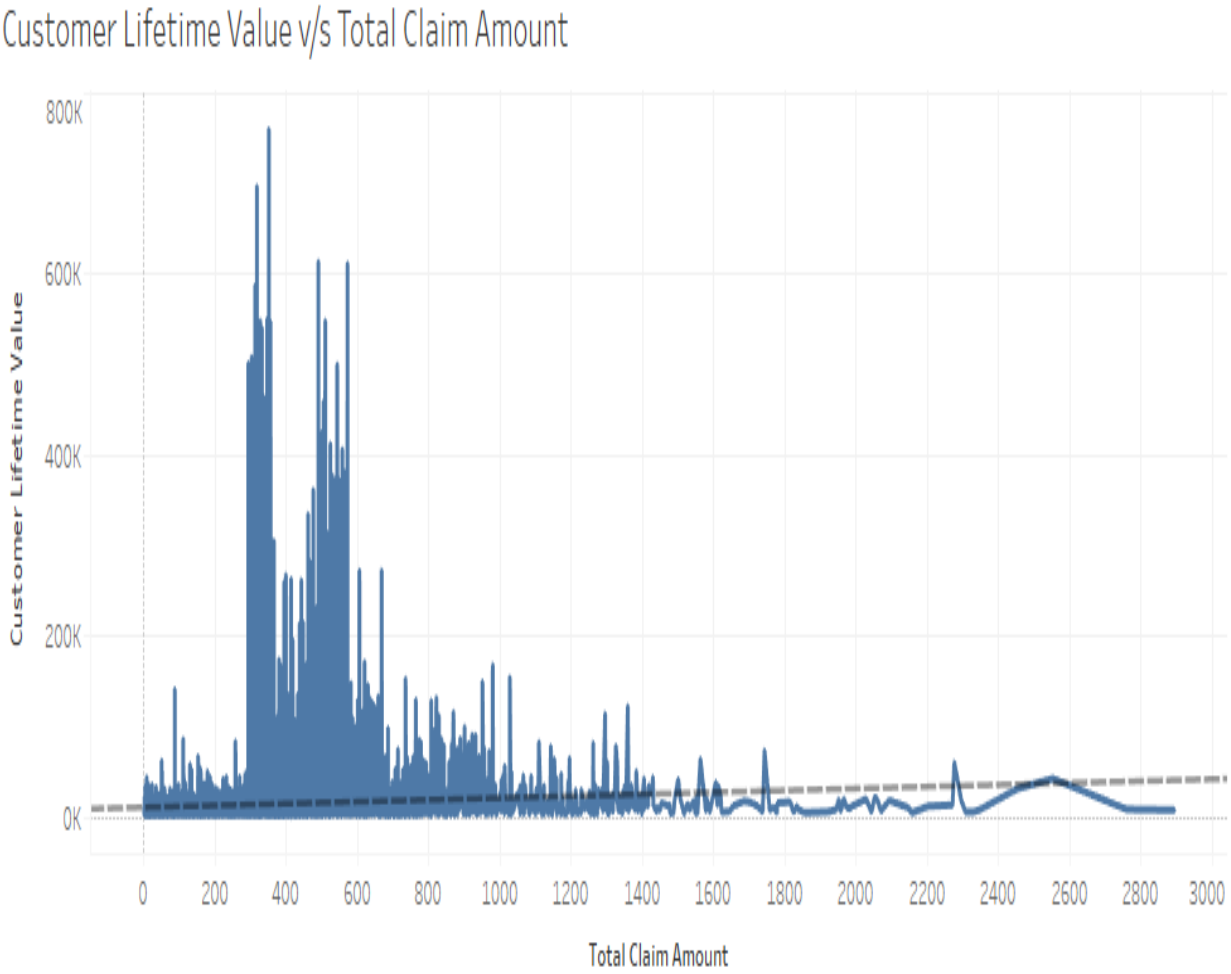
Customer Lifetime value = -5428.53 * (Monthly Premium Auto) + 1.266 * 10^6



9. In this graph, we can see the relation of total claim amount with customer lifetime value. Here we have also applied the linear regression for best fit of the values and the equation obtained is:

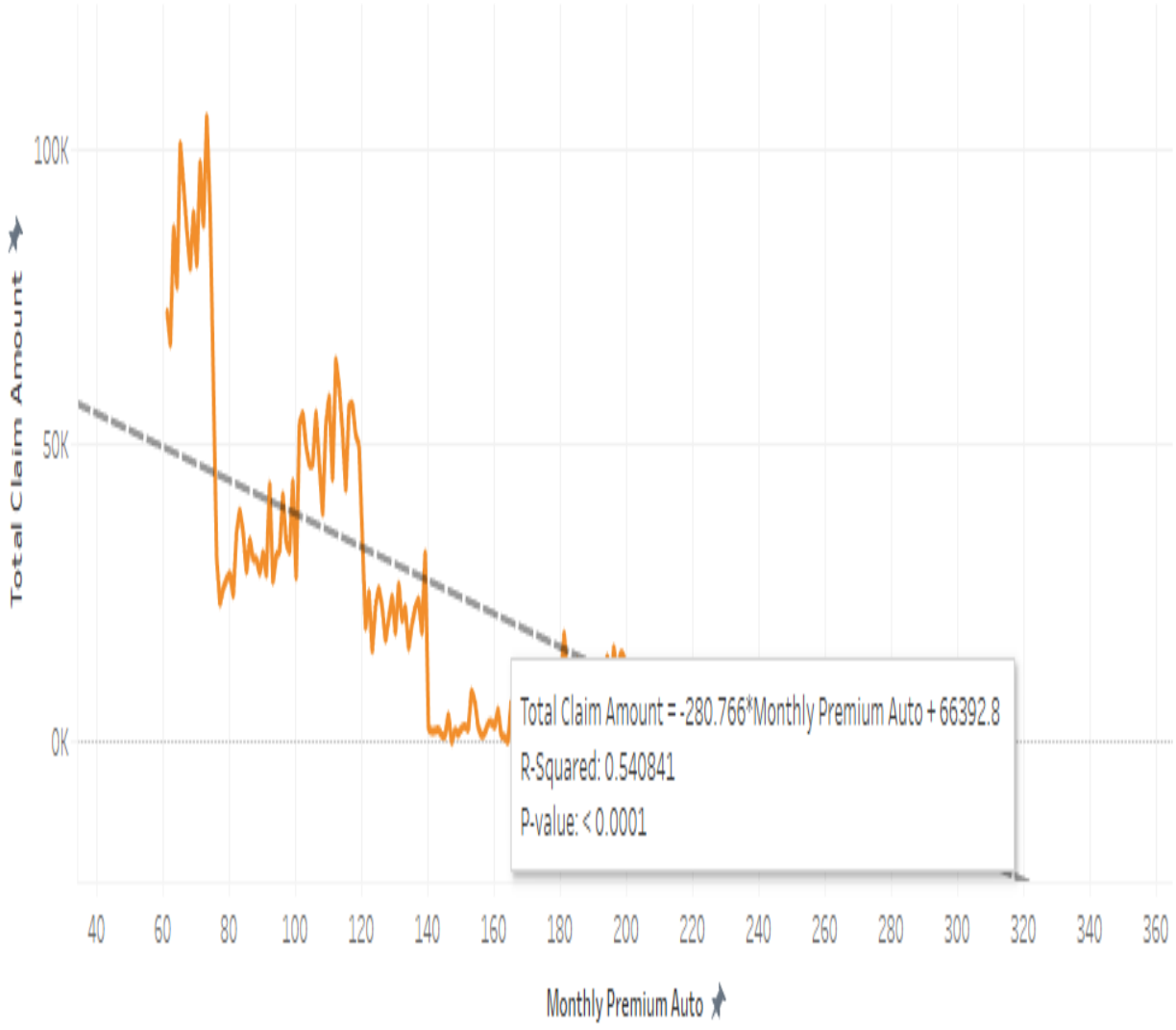
   Customer Lifetime Value = 10.3591 * (Total Claim Amount) + 10121.3

   With this graph and linear regression equation we can conclude that, one unit increase in total claim amount will increase the customer lifetime value by almost 10 units. In general, this is a good factor as people will a stay if they will get claim after any accident or damage. The customers are able to claim a higher amount in this insurance company and so the company can retain the customers for a longer time.



10. In this visualization we can see that Monthly Premium Amount and Total claim amount are negatively correlated. This shows that higher is the monthly premium, lower are the chances of customer claiming an amount. But company has more than 60% basic coverage customers. So, the company should focus on converting basic coverage customers into premium to get maximum profit.

    Equation: Total Claim Amount= -280.766*Monthly Premium Amount + 66392.8



Conclusion:

Thus, in order to retain the customers and maximize the profit of the insurance company, we need to offer relevant policy coverages to customers based on their income levels, suggest appropriate offers to customers based on marital status and the company should focus on improving the customer support department in order to resolve the customer complaints as soon as possible.