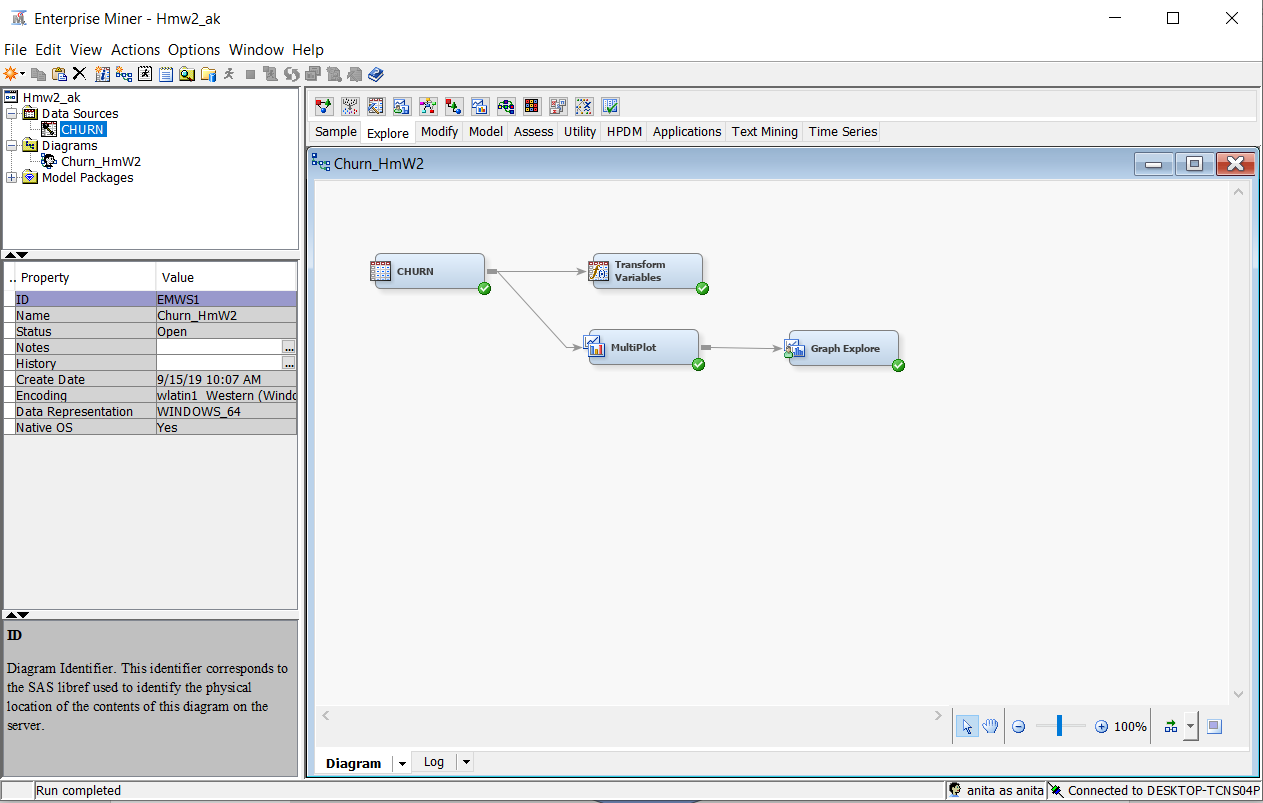
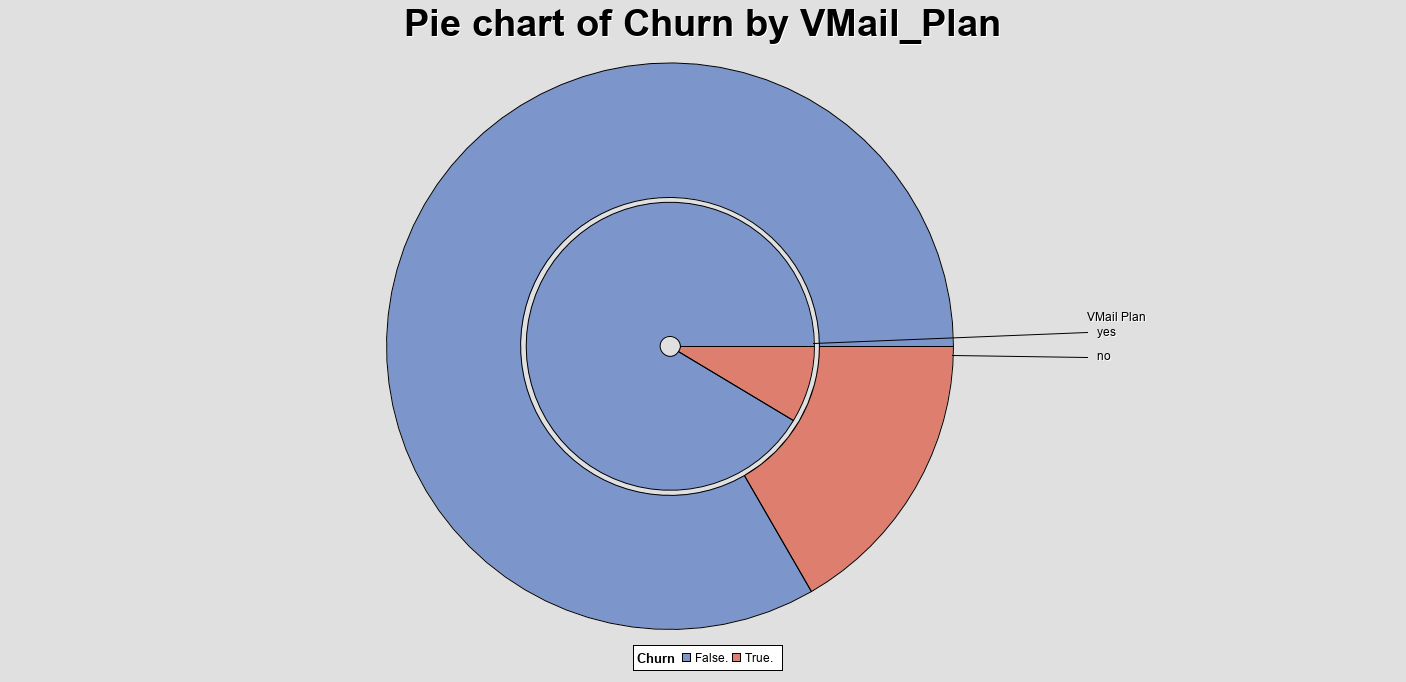
1. Develop appropriate EDA procedures, analyze and report the association between the target variable, *Churn*, and each of the following nominal/binary variables and interval variables:



* + ***Vmail\_plan:***



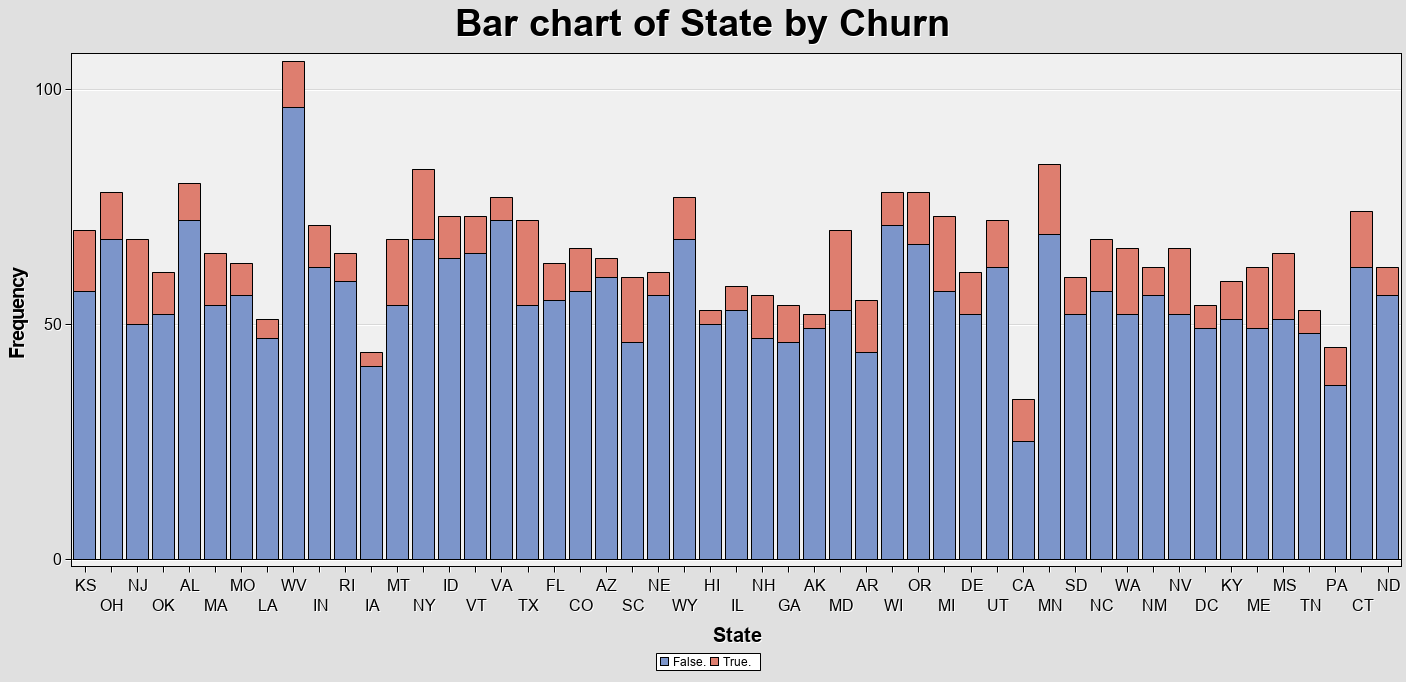
1. ***It appears that Vmail plan is associated with Churn. When the customer has Vmail plan, Churn is 8.6 % and when the customer doesn’t have Vmail plan, Churn is 16.7 %. So, we conclude that there is an association between voicemail plan and churn.***
2. ***Vmail\_plan has to be included as input in DM model, because it has a significant association on churn which is our target variable:***
   * ***Area code:***

******

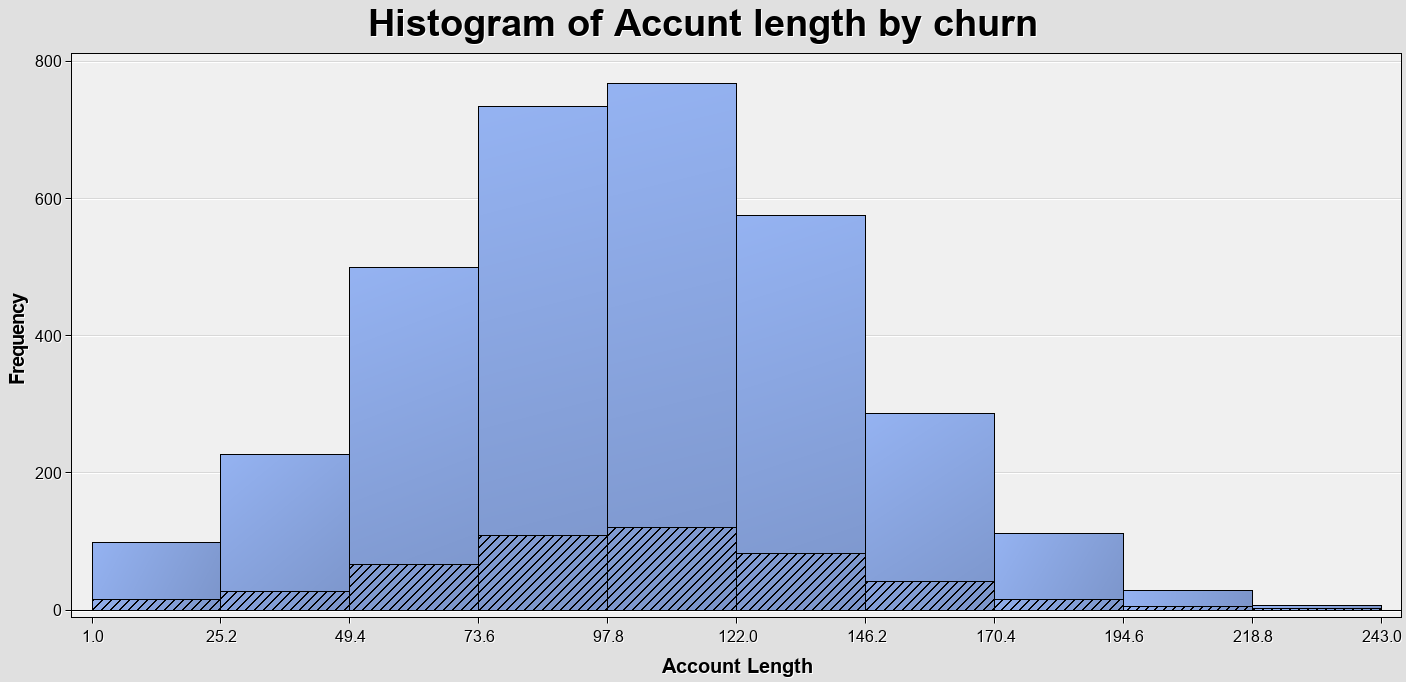
|  |  |
| --- | --- |
| ***Area code*** | ***Churn %*** |
| ***408*** | ***14.5*** |
| ***415*** | ***14.3*** |
| ***510*** | ***14.9*** |

***As we can see above there is just slightly difference between Churn percentage and different area codes so, we can conclude that there is no association between area code and churn.***

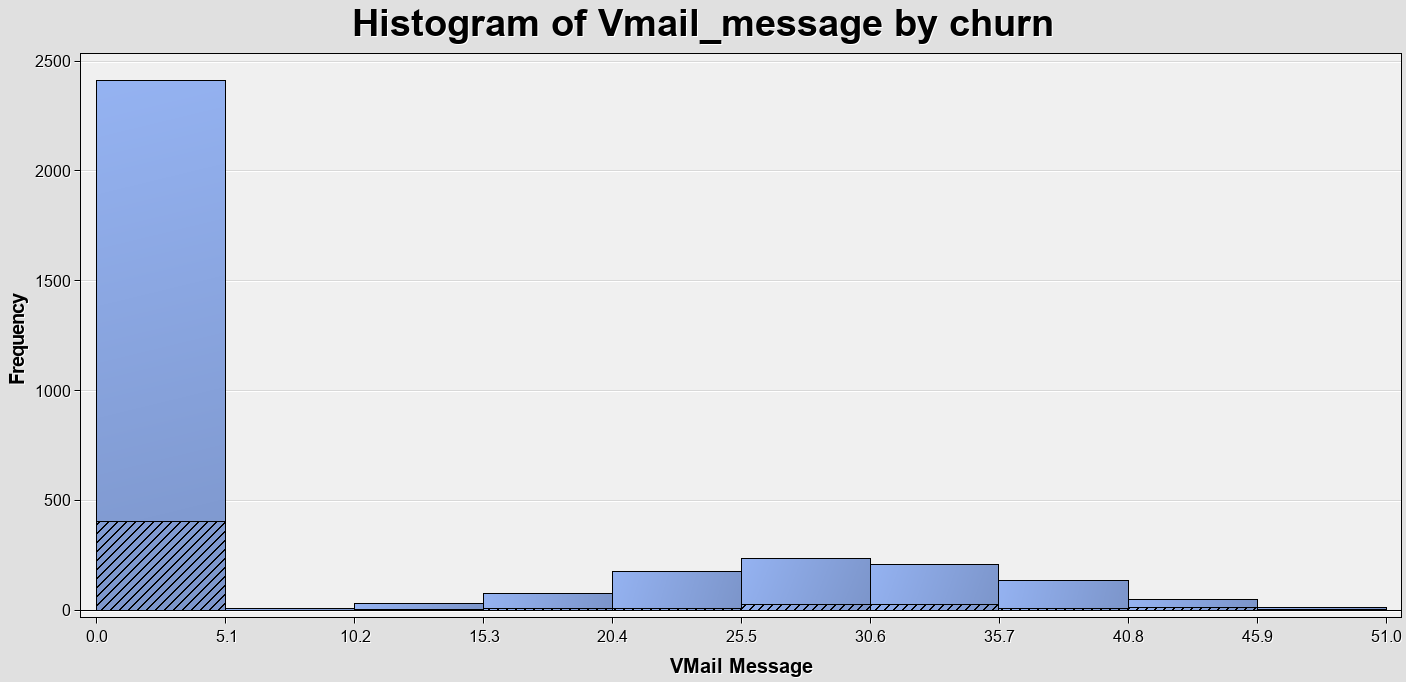
1. ***I’d rather not to include area code as an input in DM model not only because it’s not associated with Churn but also based on previous homework, we saw that only 3 area codes were used for 50 states which aren’t right. So I’ll reject area code variable***
   * ***State:***

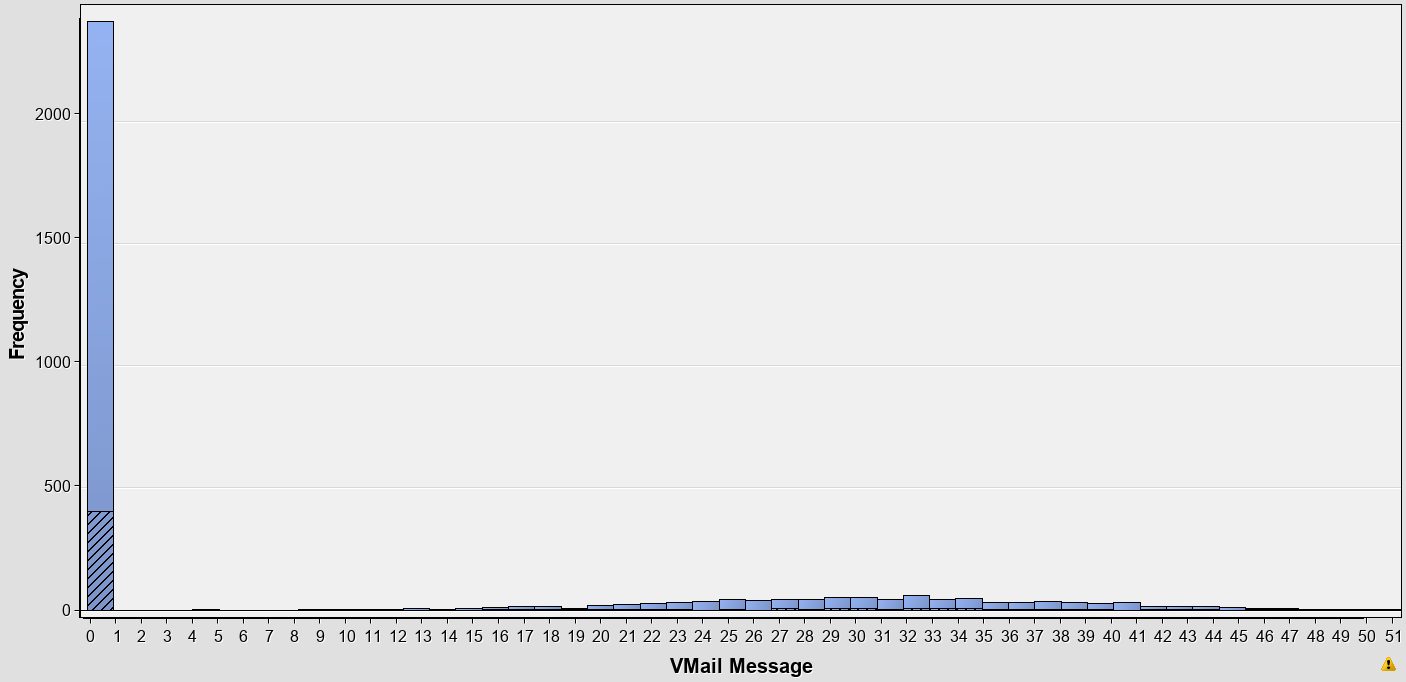
******

1. ***It’s drawn from the chart that in some states like Maryland (25%), customers are prone to churn more than other states such as Alaska (5.7 %).so the state variable appears to be associated with churn.***
2. ***I would include the state variable in the DM model, since it appears to have an association with churn rates.***
   * ***Account length:***

******

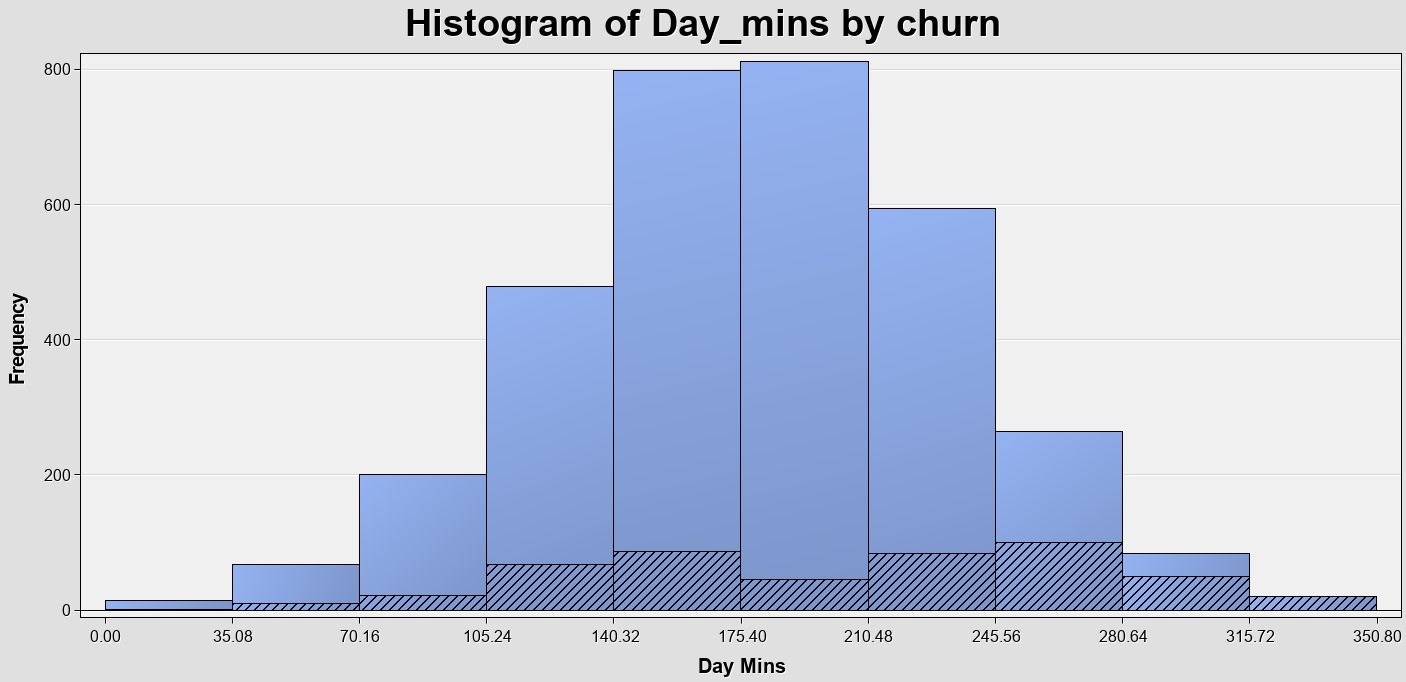
1. ***It seems there is no association between churn and account length. It might look that middle bars have more churn rates but It’s because the figure is not normalized. So, we can conclude that there is no association between churning and account length.***
2. ***I’d rather to include the account length in Dm model. It’s right that we can see any connection between churn and account length but account length might have some effects on other variables.so I’d rather to use it as an input.***
   * ***VMail Message:***

******

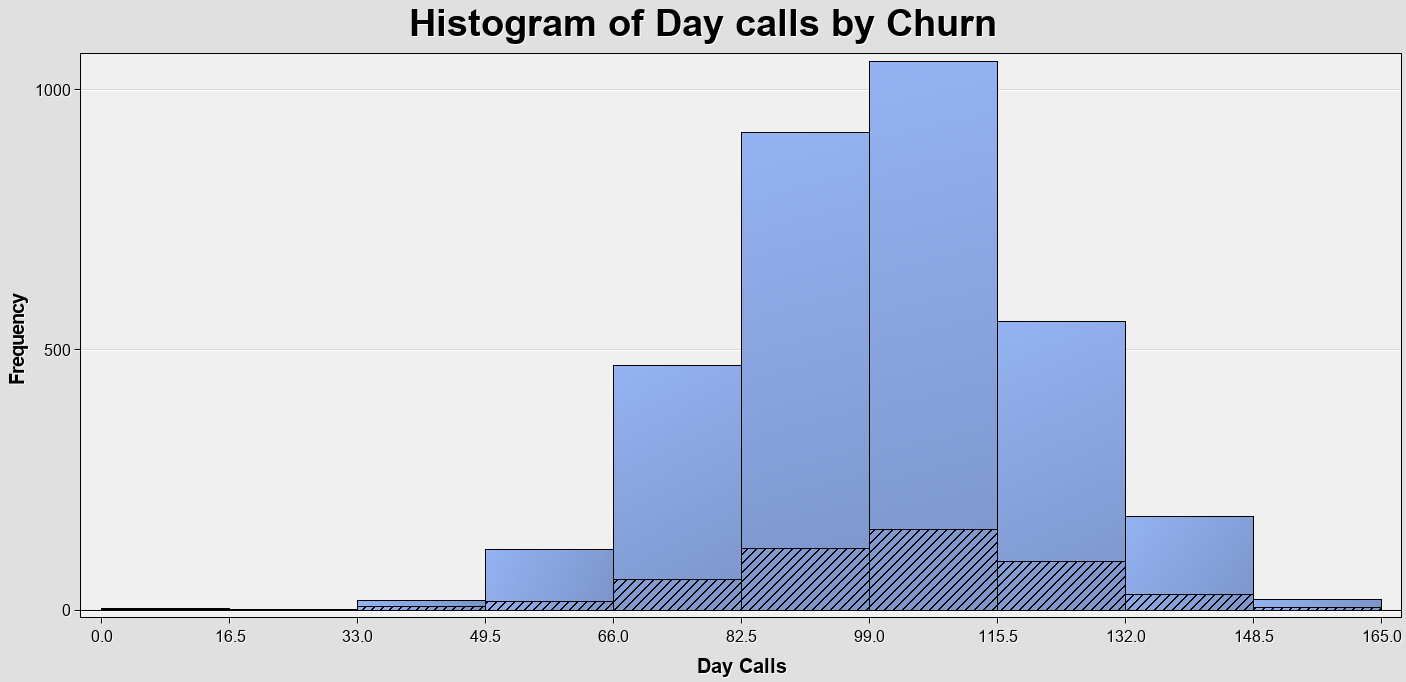
******

1. ***We can see that vast majority of customers don’t have any Vmail message and almost 20% of them have been churned So, there should be a significant association between Vmail message and churning.***

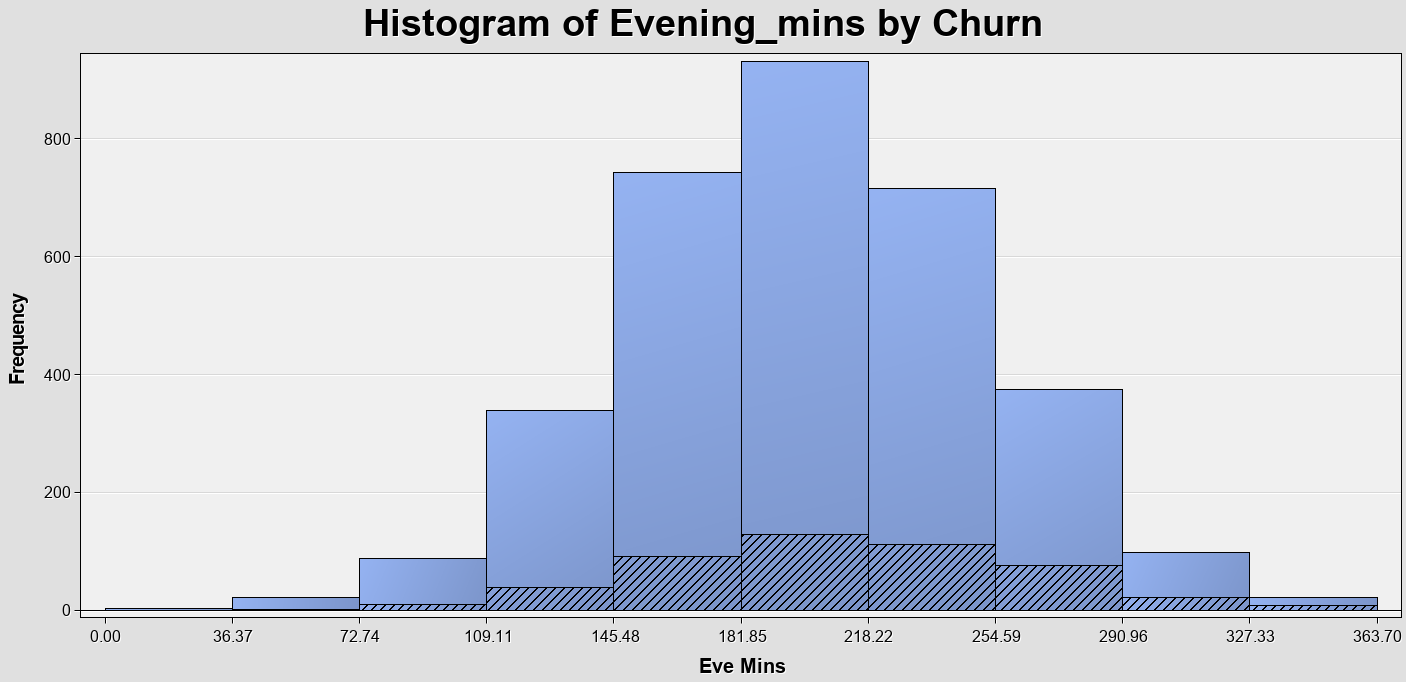
1. ***I would include the Vmail\_message variable in the DM model, since it appears to have an association with churn rates.***
   * ***Day Mins:***

******

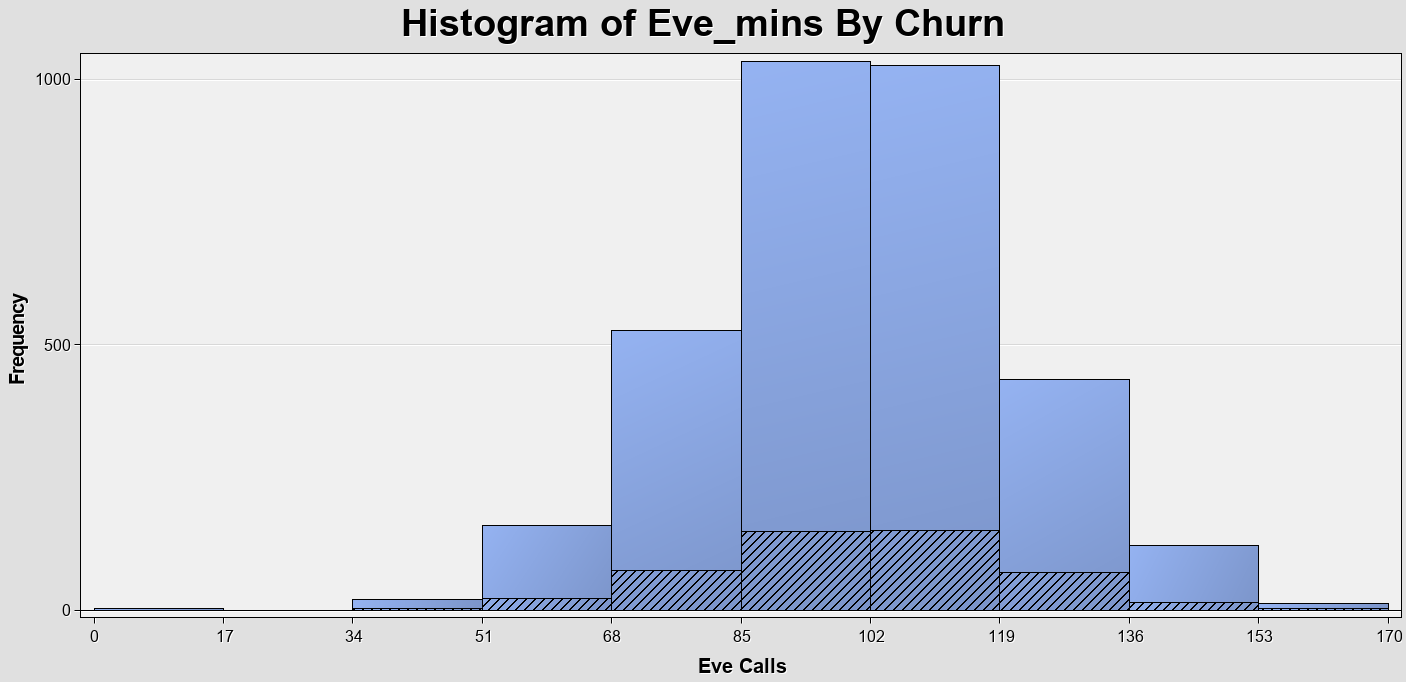
1. ***As the number of minutes of phone calls in the day increases, the percentage of churn increases too. And the customers with the highest number of minutes all seem to churn (100%). So, it must be an association between churn and Day\_mins.***
2. ***I would include the Day\_mins variable in the DM model, since it appears to have a significant association with churn rates***
   * ***Day Calls:***

******

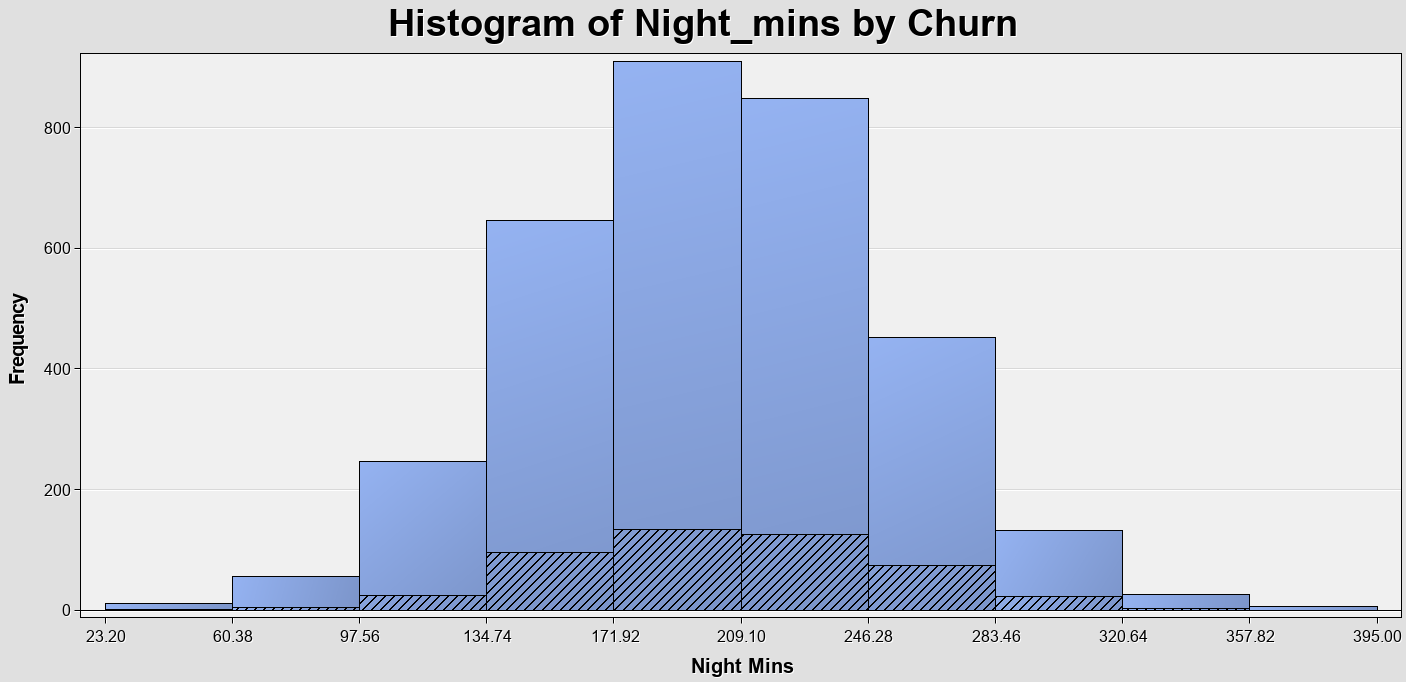
1. ***it seems that the churning rate is not changing by number of calls during the day. So, there is no association between our target value and day\_calls value.***
2. ***I’m still going to include this value as an input in the model while there is some relationships between day\_calls and other variables such as day\_mins.***
   * ***Eve Mins:***

******

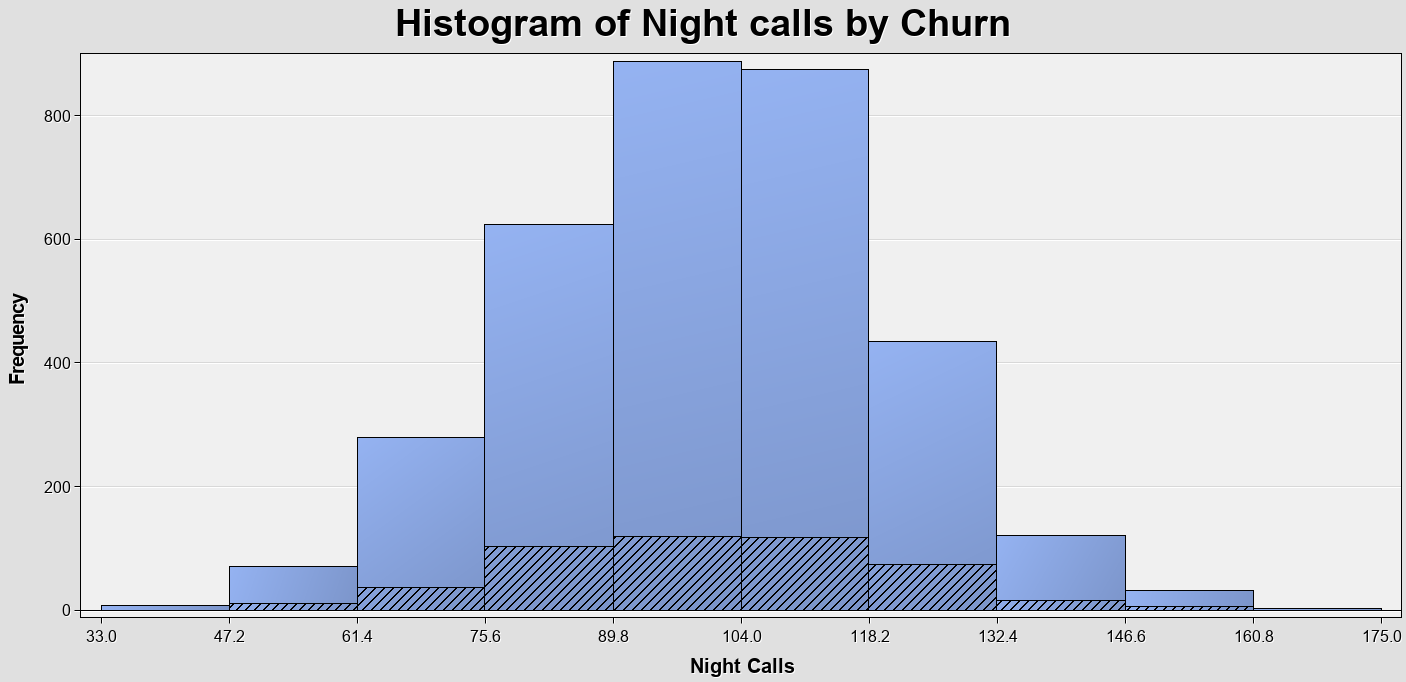
1. ***it appears to be an association between eve minutes and churn. It appears that customers with a high number of evening phone call minutes churn more than those with lower numbers of evening minutes.***
2. ***I would include the variable eve minutes in the model since there is an association between the two variables.***
   * ***Eve Calls:***

******

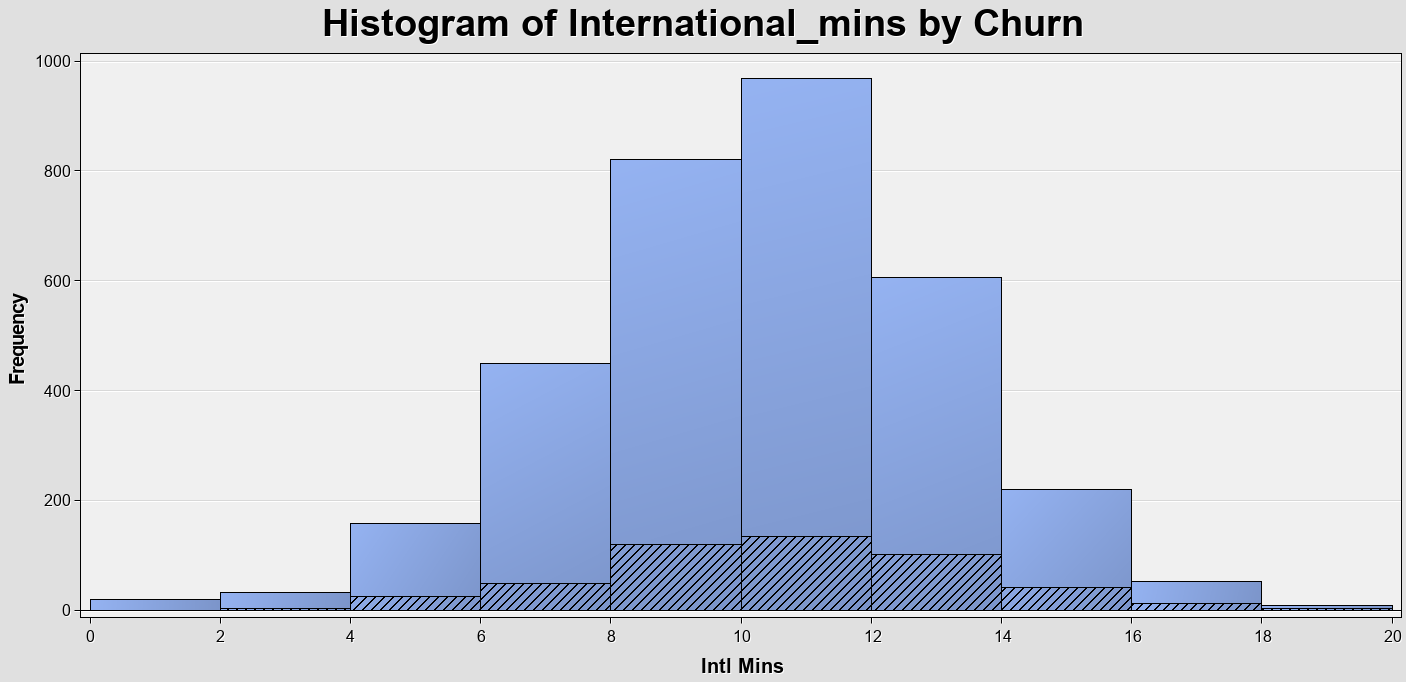
1. ***There is a slight difference in churn rate among different ranges .so it appears to be an association between Evening calls and our target value churn.***
2. ***I would include the variable as an input in the model since it has a relationship with the target variable.***
   * ***Night Mins:***

******

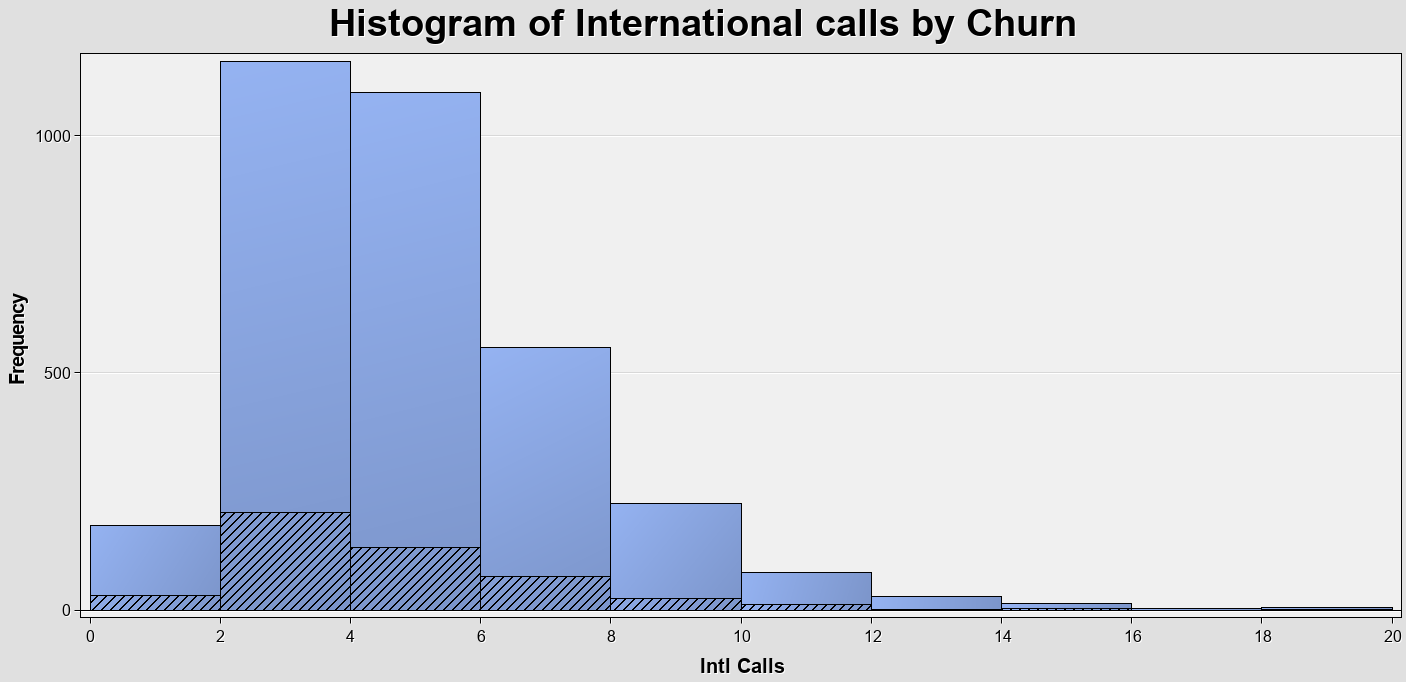
1. ***It doesn’t seem to be an association between the night minute’s variable and churn. churn rate appears to be almost constant while the number of minutes is changing at nights***
2. ***I would include night mins as an input variable in the model because the variable may have relationship with another variable that relates to the rate of churn***
   * ***Night Calls:***

******

1. ***It doesn’t seem to be an association between night calls and churn. The churn rate doesn’t seem to be changed much as the number of night calls increases.***
2. ***I would still rather to include night calls as an input variable because there may be some connections between night calls and other input variables that affect on Churn.***
   * ***Intl Mins:***

******

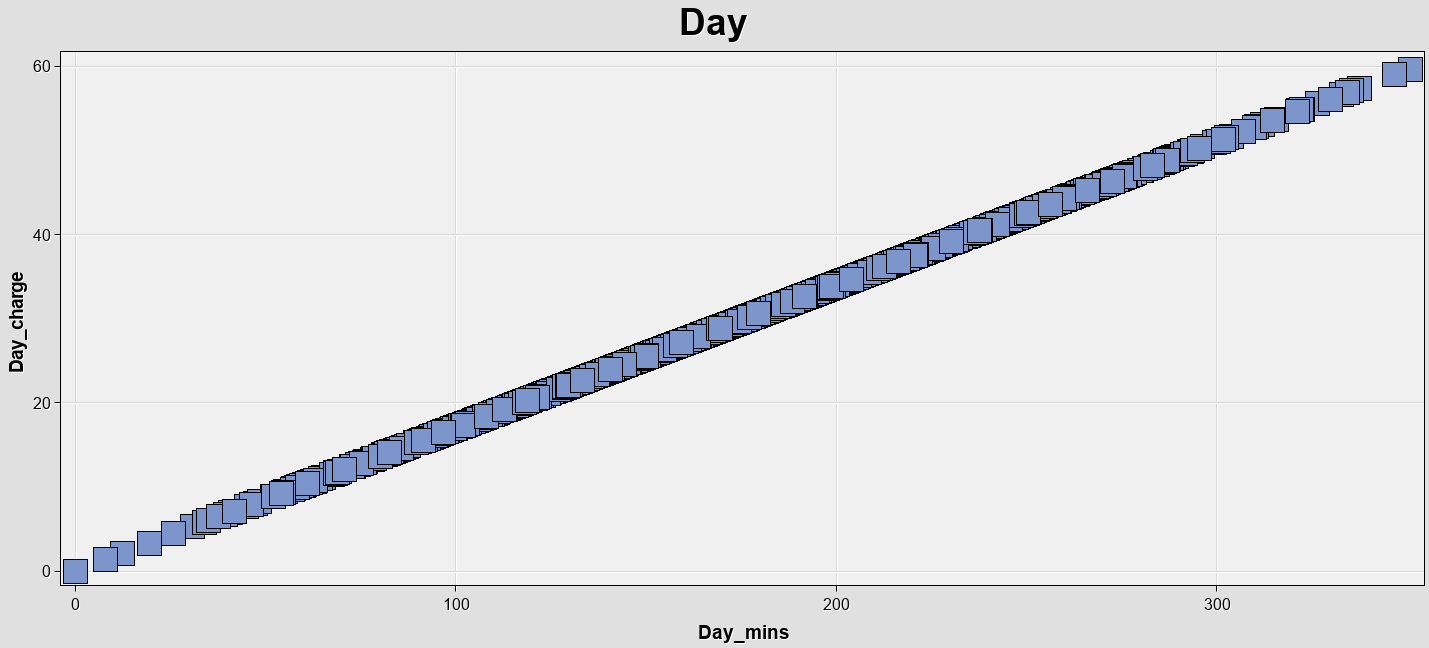
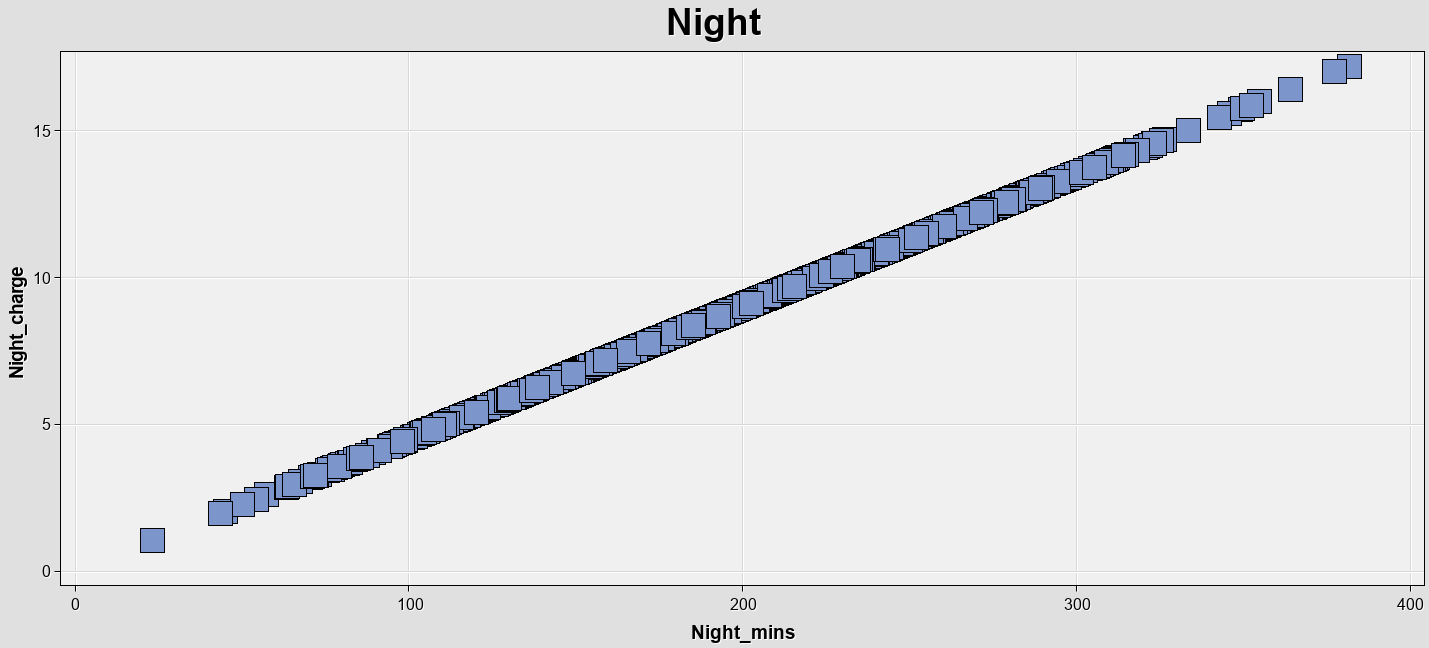
1. ***It doesn’t seem to be any relationship between the international minute’s variable and churn. churn rate appears to be almost constant in all ranges***
2. ***I would include international mins as an input variable in the model because the variable may have an association with another variable that relates to the rate of churn***
   * ***Intl Calls:***

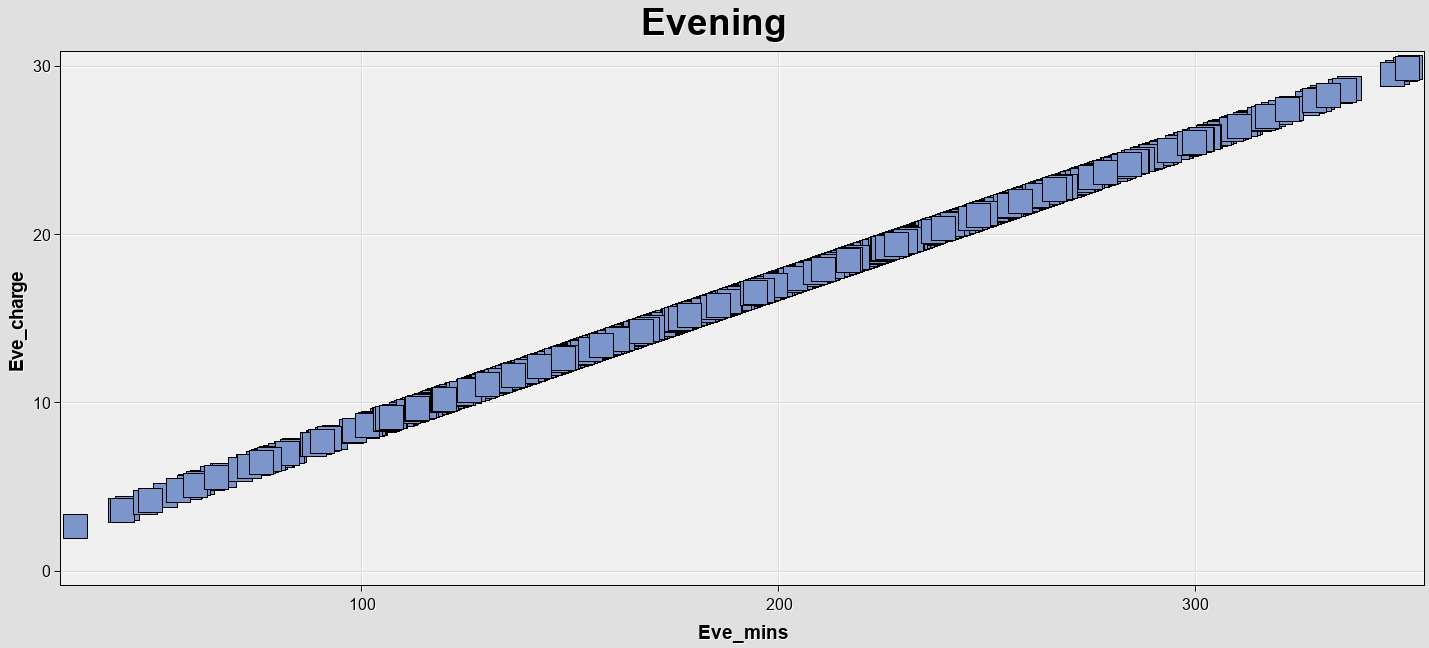
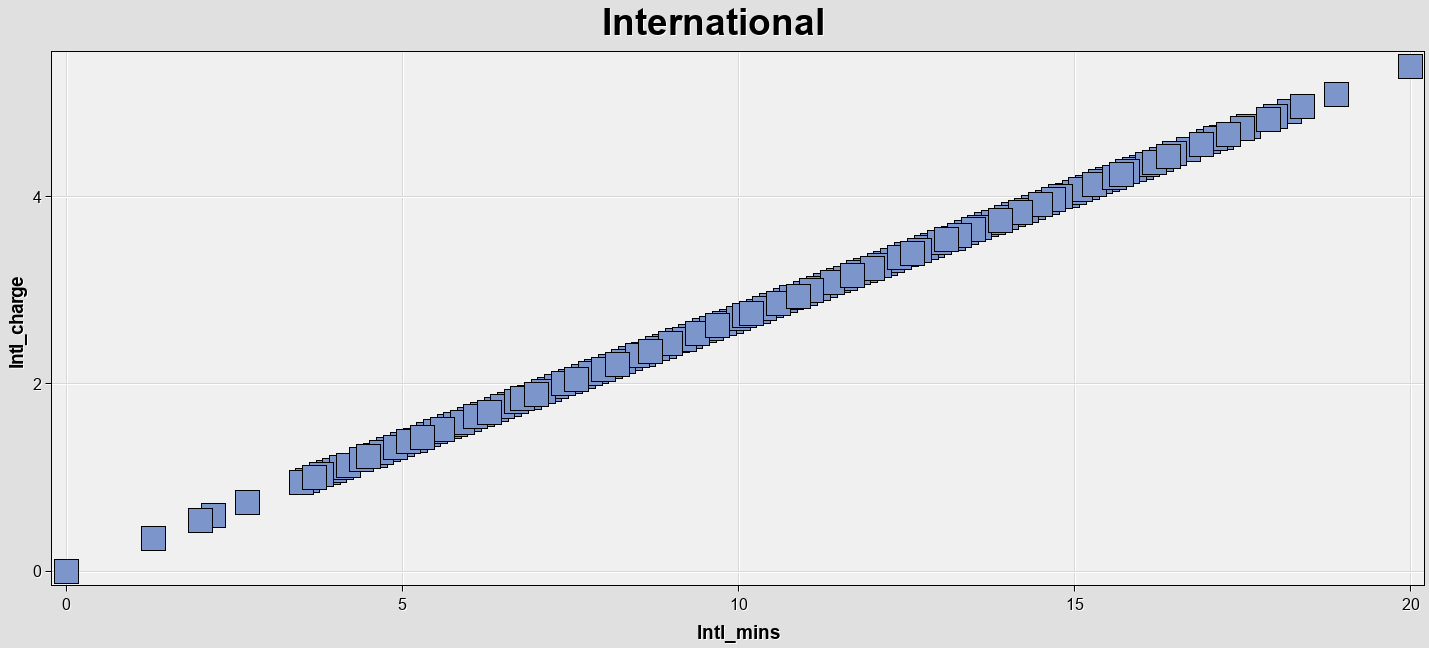
******

1. ***There appears to be an association between international calls and churn. Customers with very few international calls seem churn more than customers with more international calls.***
2. ***I would include the variable as an input in the model since it has a relationship with the target variable.***
3. Explain why you need not conduct the analyses above for the following variables:

*Day Charge, Eve Charge, Night Charge, Intl Charge*?

***Because all charge variables (Day, night, evening) are correlated with their minutes’ variables. As you can see in the scatter plots bellow, there is a linear relationship between minute and charge (charge is a function of minute). So we don’t need to do analysis on them***

*** ***

*** ***