Project 3 ASU AI Course: Graphing Page

Summary stats: dataframe info

Summary stats: shape and value_counts

The Request: Create the python code to create an interactive scatter chart with the linear regression line overlay in streamlit. There should be two selectionbox's where the x and y columns from df_initial can be selected. The chart would then update to show a scatter plot and regression line for the selected x and y columns. use sklearn for the regression.. Add the regression equation and a computed R squared value as chart annotations. Use R2 as the R squared label.

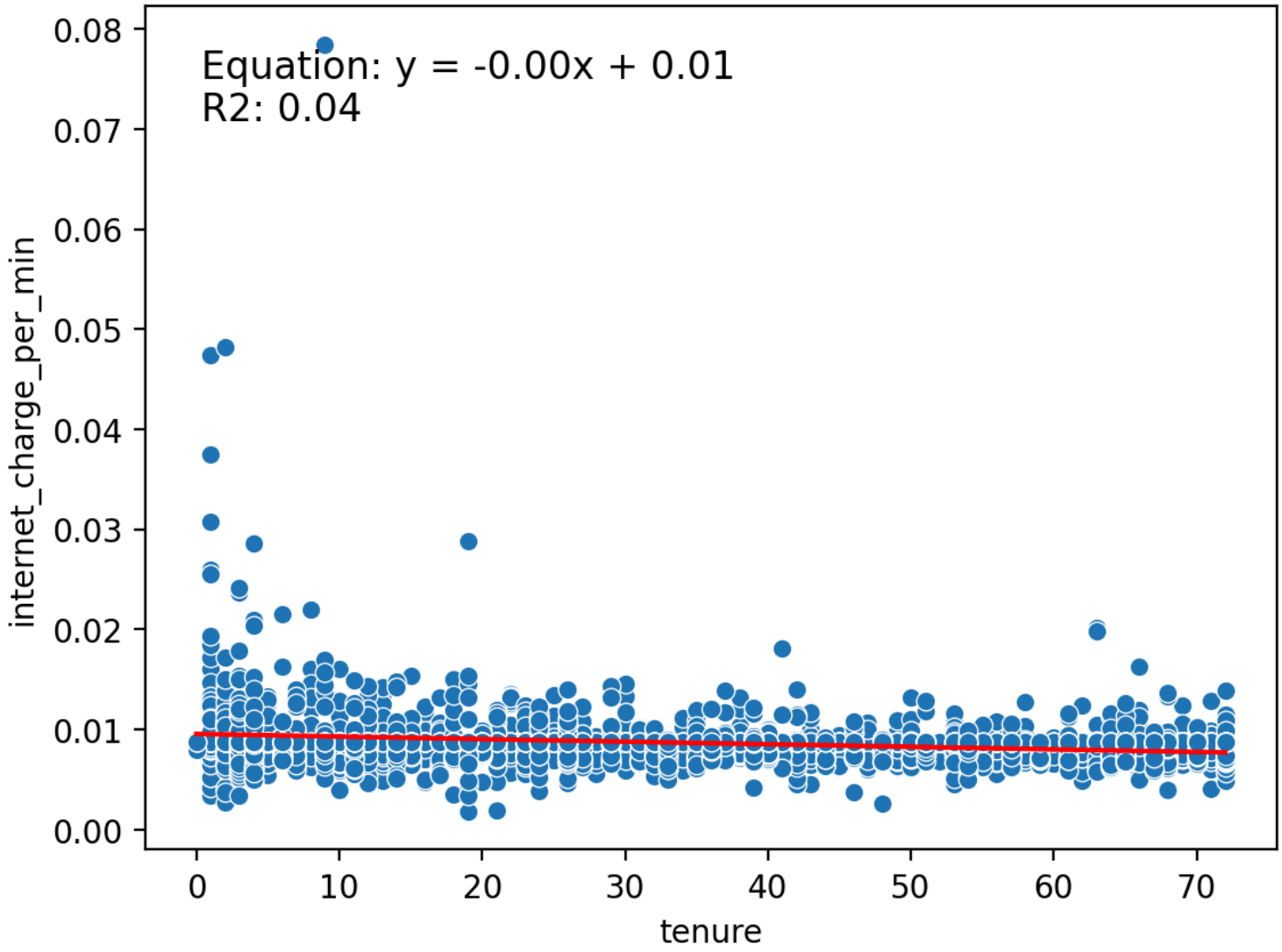
Interactive Scatter Plot with Regression Line

Select the x-axis column

tenure

Select the y-axis column

internet_charge_per_min



The Request: Create the python code to create an interactive scatter chart with the linear regression line overlay in streamlit. There should be two selectionbox's where the x and y columns from df_initial can be selected. The chart would then update to show a scatter plot and regression line for the selected x and y columns. use sklearn for the regression.. Add the regression equation and a computed R squared value as chart annotations. Use R2 as the R squared label.

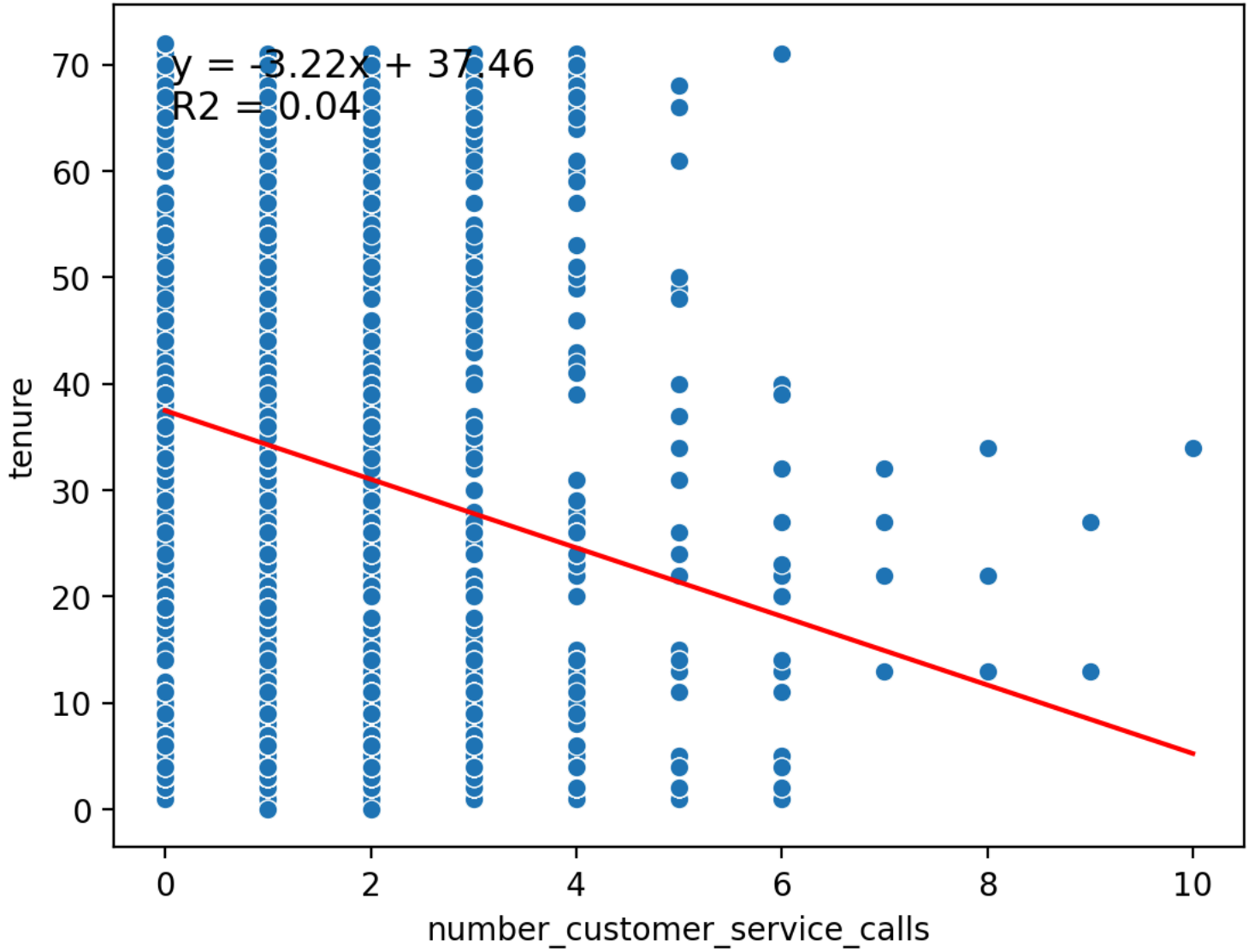
Interactive Scatter Plot with Regression Line

Select X-axis column

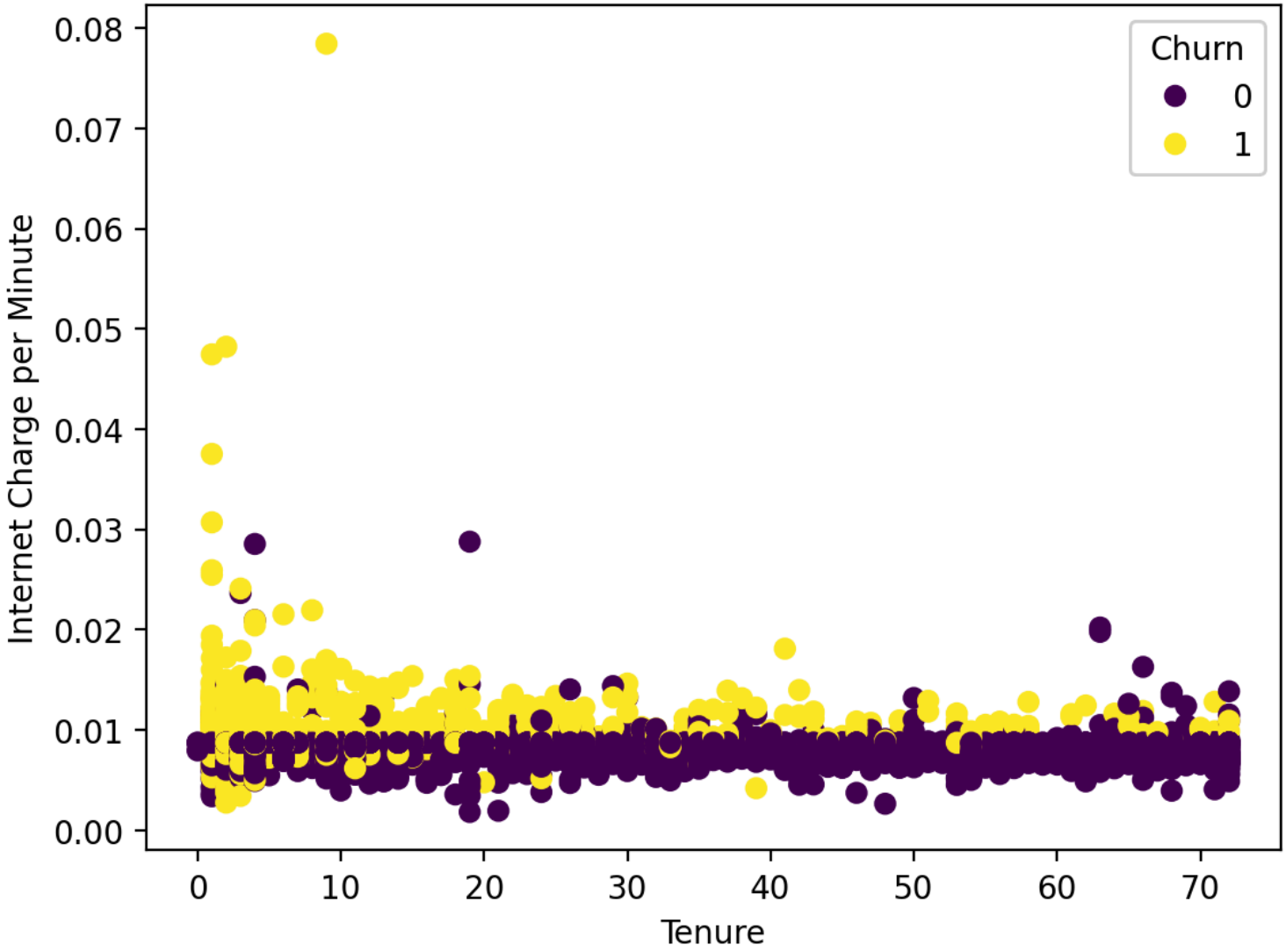
number_customer_service_calls

Select Y-axis column

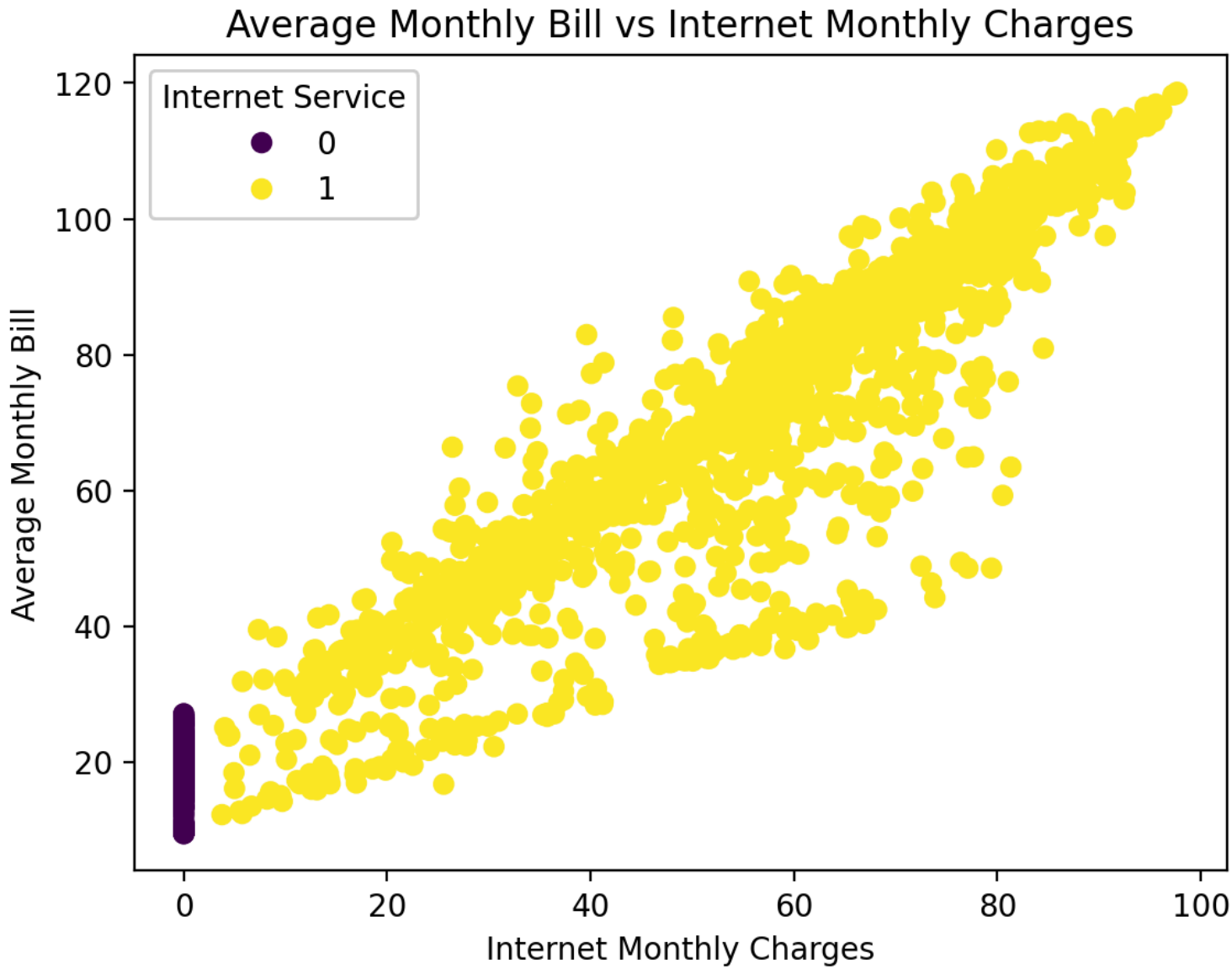
tenure



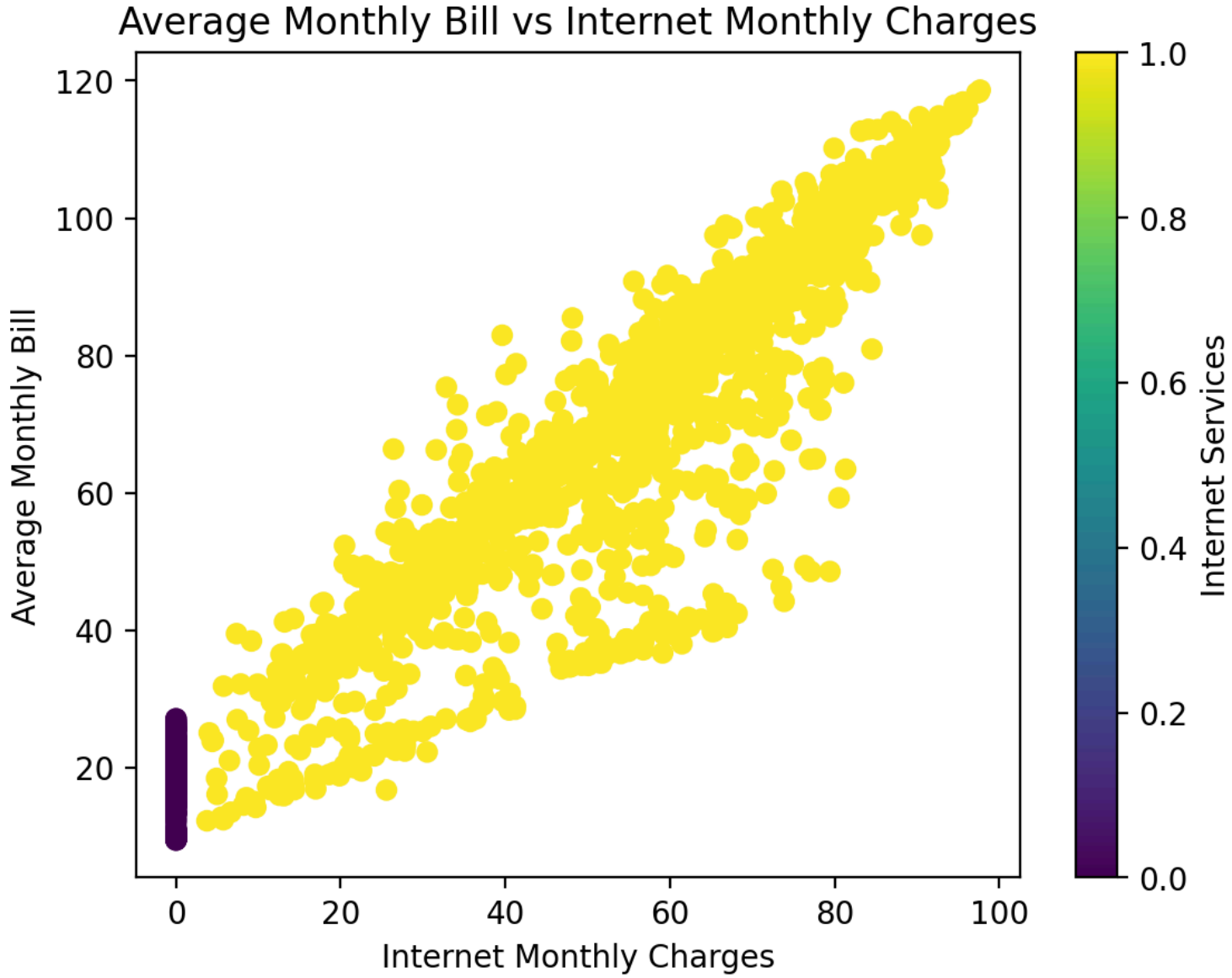
The Request: Create a plot of x=tenure, y=internet_charge_per_min and color it by churn. Create x and y labels. Add a legend for churn color.



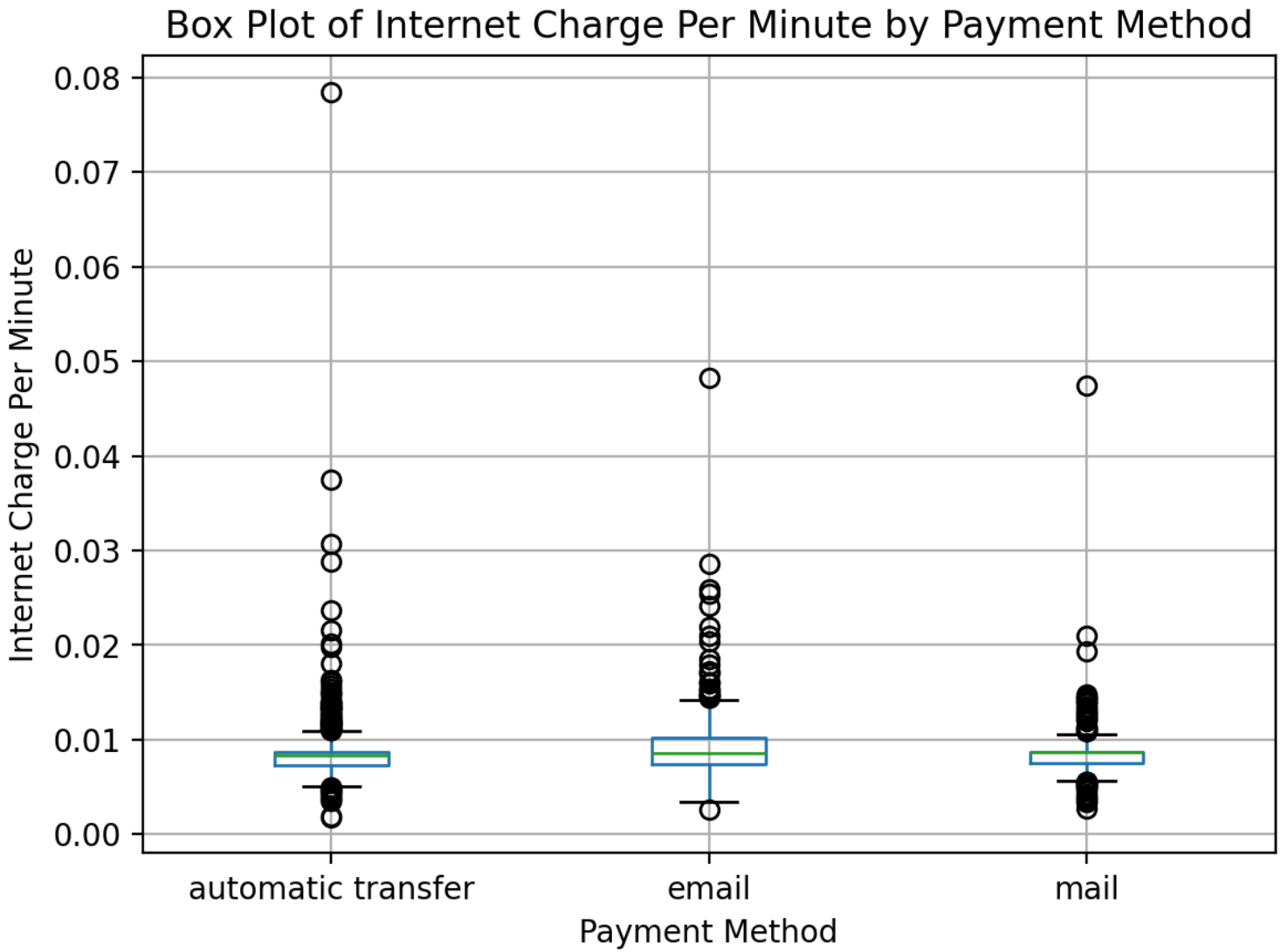
The Request: Plot avg_monthly_bill versus internet_monthly_charges. Color the points by internet_service. Add a legend for the color for internet_service.



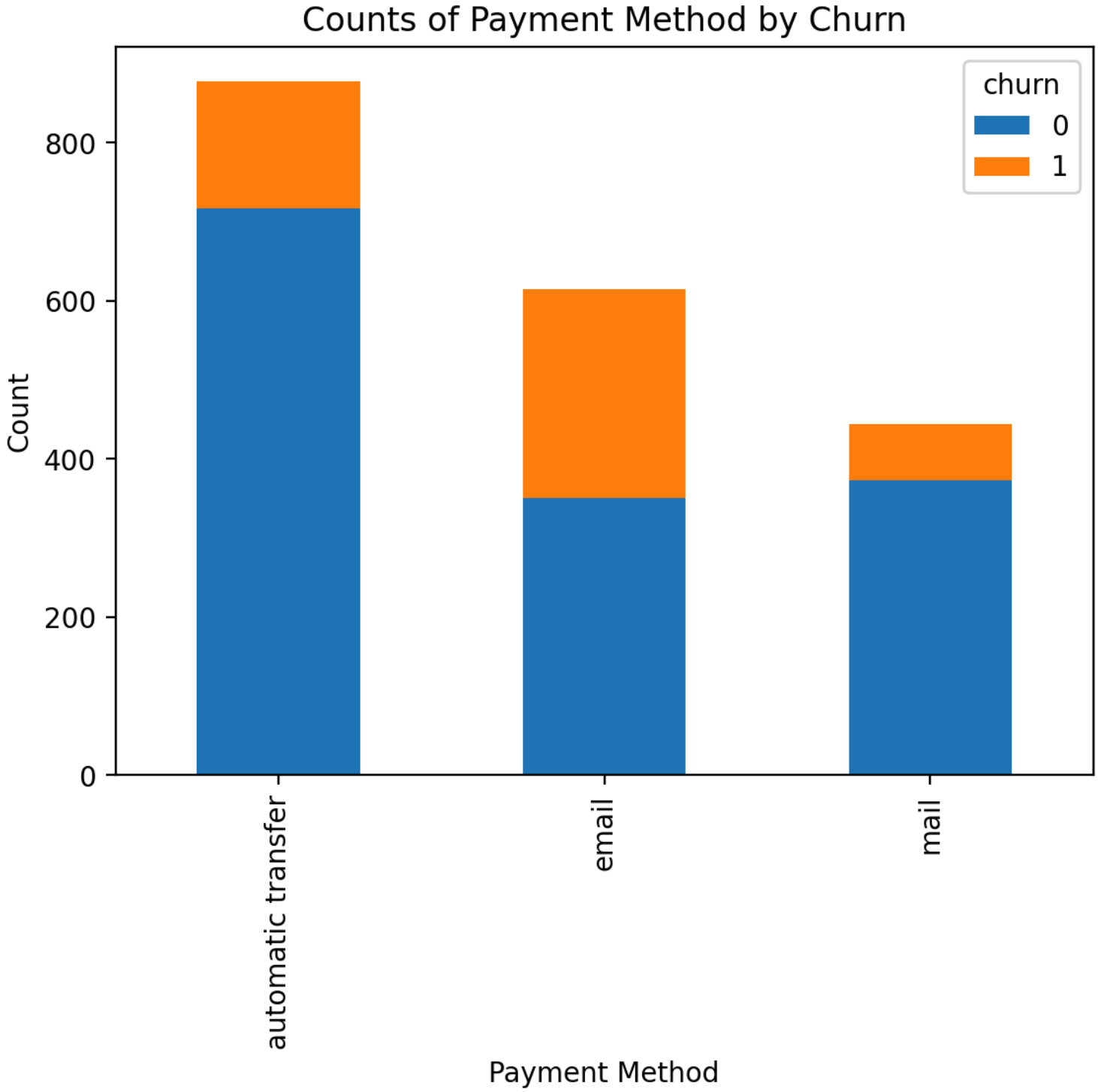
The Request: Plot avg_monthly_bill versus internet_monthly_charges. Color the points by internet_service. Add a legend for the color for internet_service.



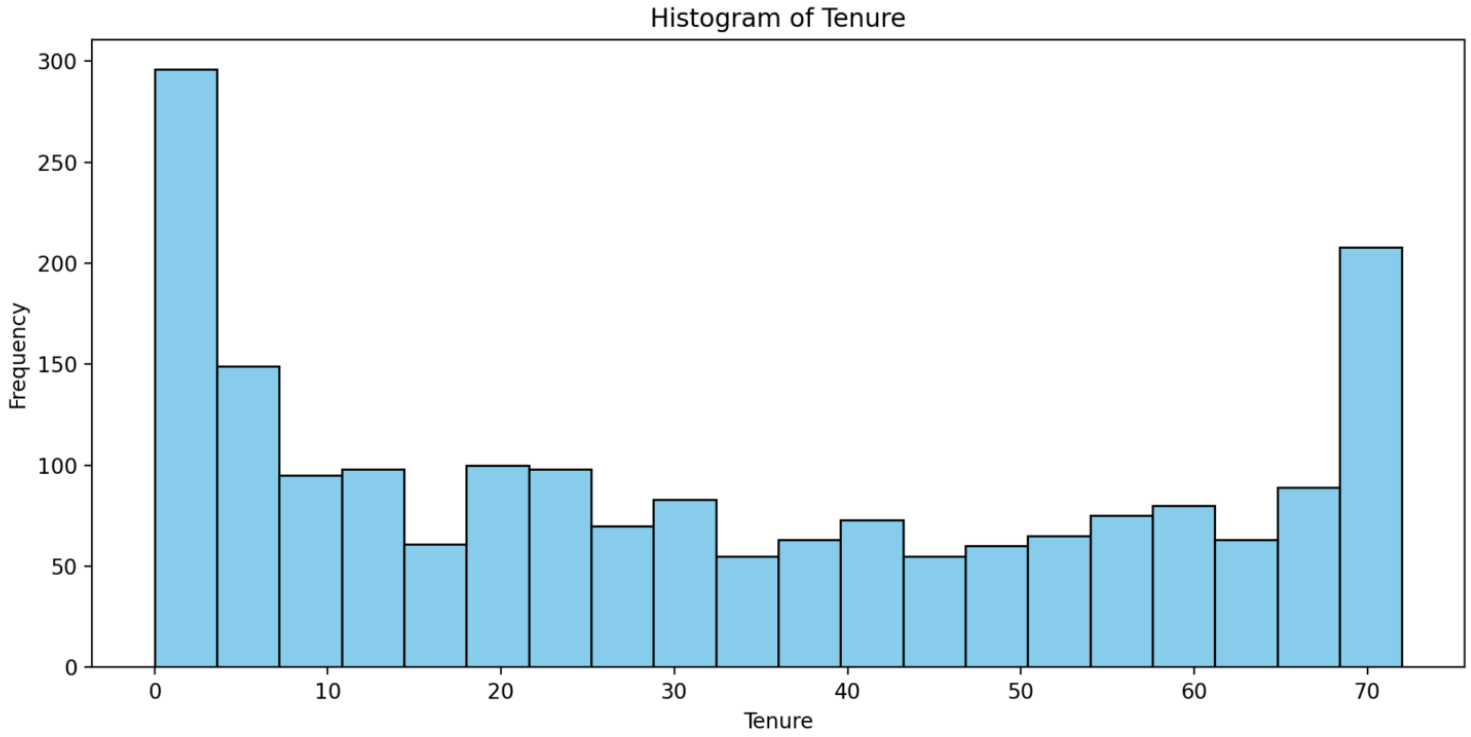
The Request: Create a box plot of internet_charge_per_min with payment_method on the x-axis

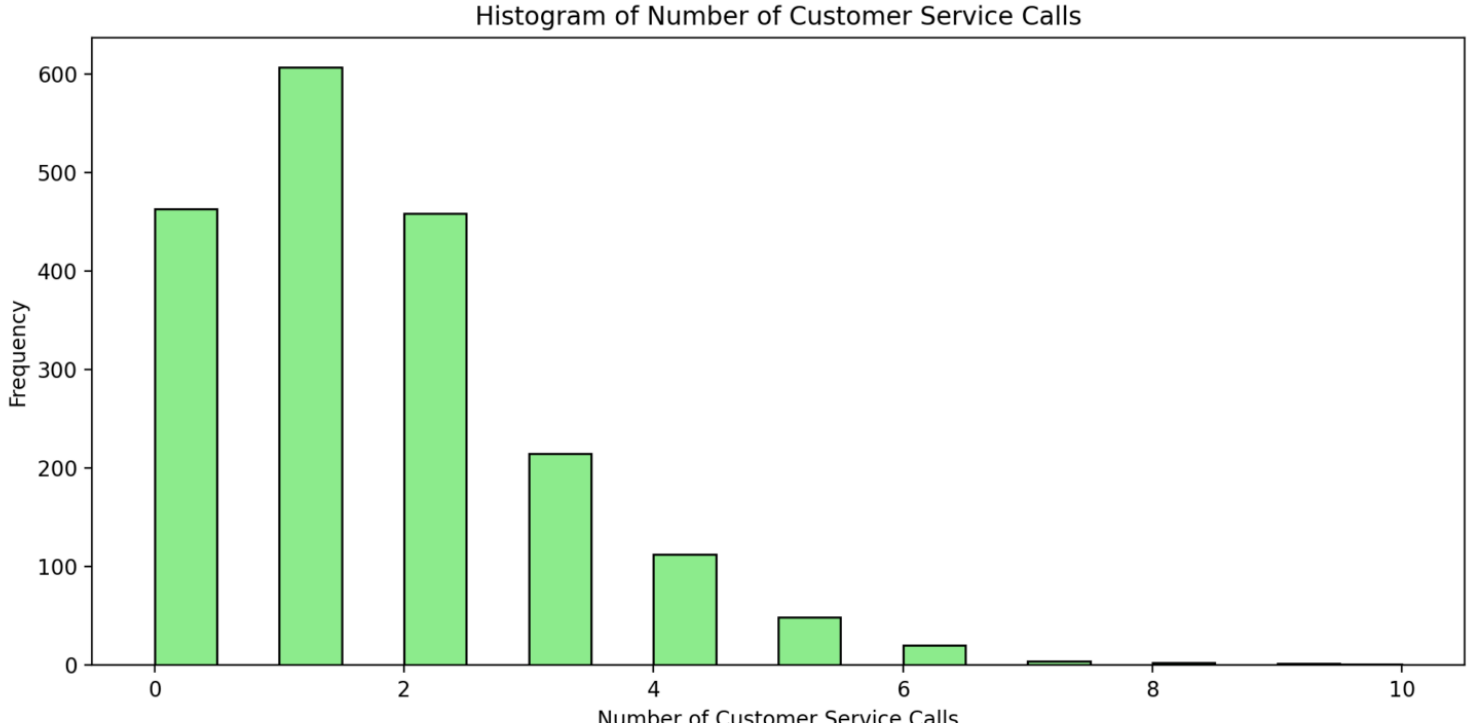


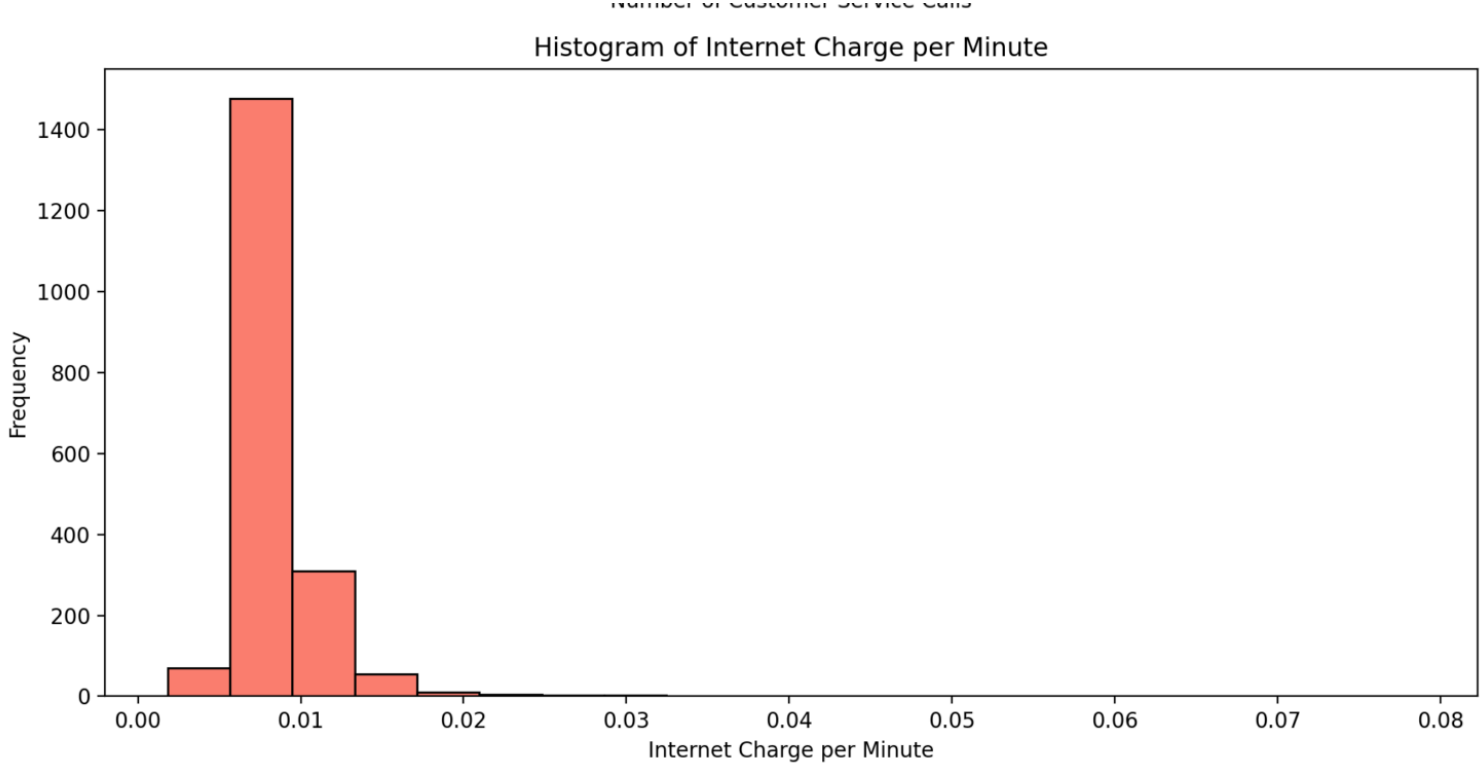
The Request: Generate a stacked bar chart of counts of payment_method. The bottom part of the stacked bar should be for churn=0 versus the top of the stacked bar should be for churn=1.


Counts of Payment Method by Churn
churn
0
1
Count
0
200
400
600
800
automatic transfer
email
mail
Payment Method

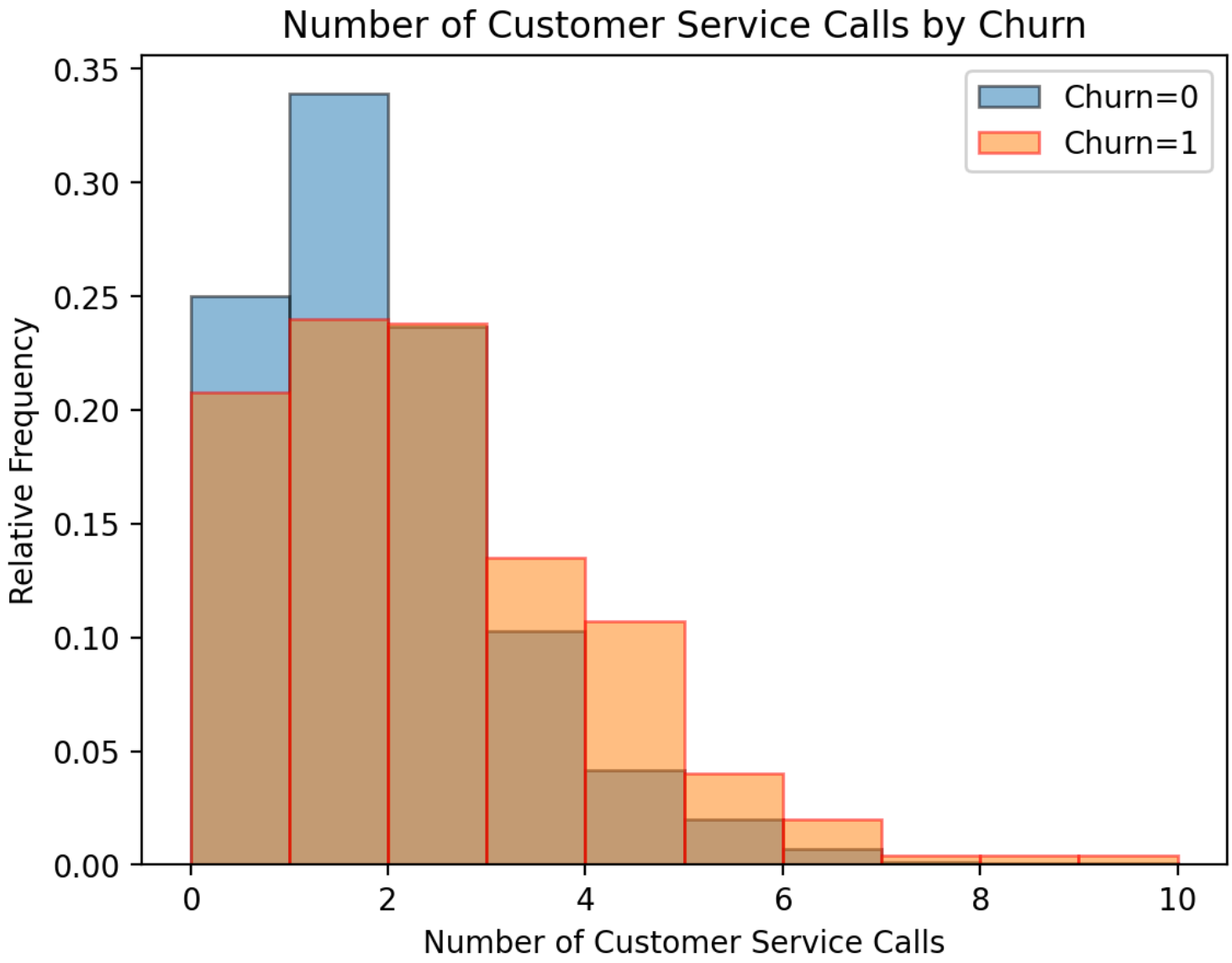
The Request: Create histograms for tenure, number_customer_service_calls and internet_charge_per_min


Histogram of Tenure
Frequency
Tenure
0
50
100
150
200
250
300
0
10
20
30
40
50
60
70

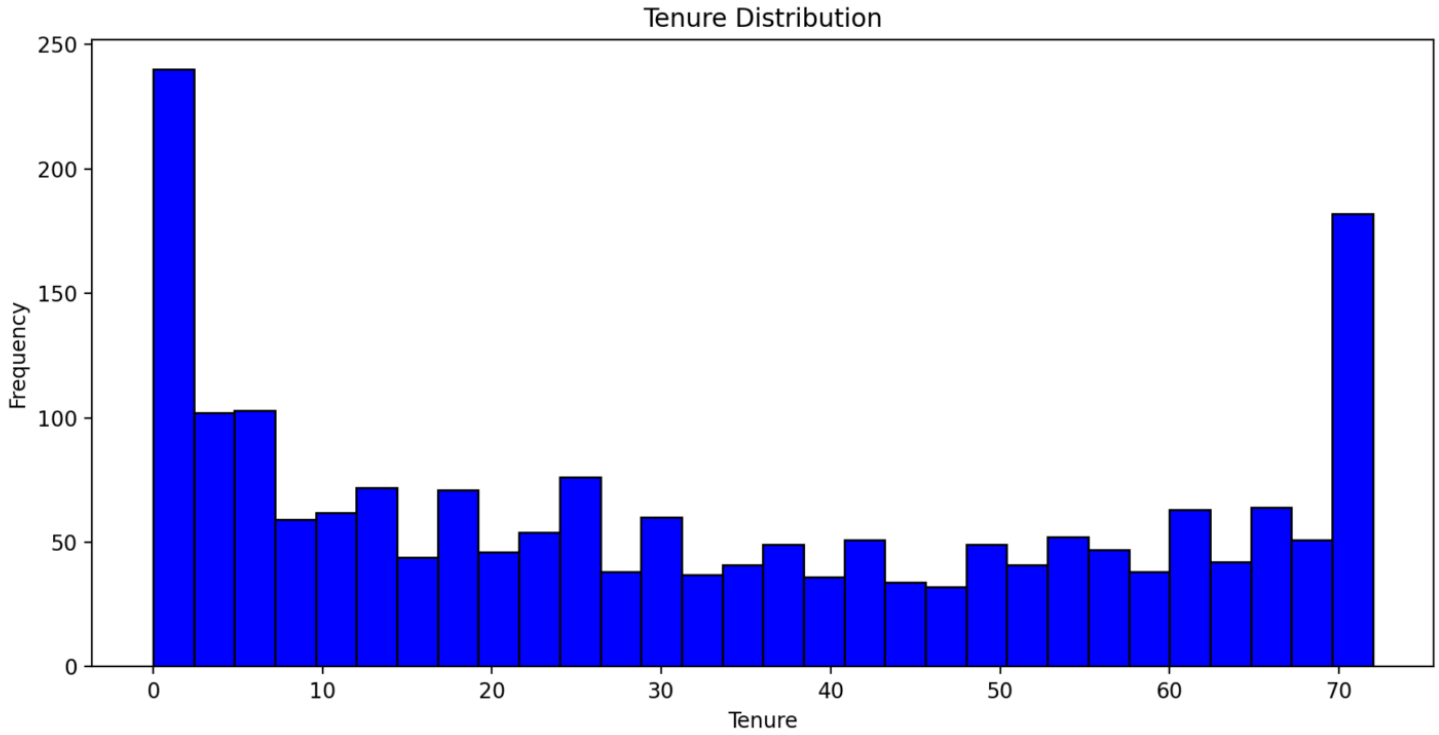

Histogram of Number of Customer Service Calls
Frequency
Number of Customer Service Calls
0
100
200
300
400
500
600
0
2
4
6
8
10

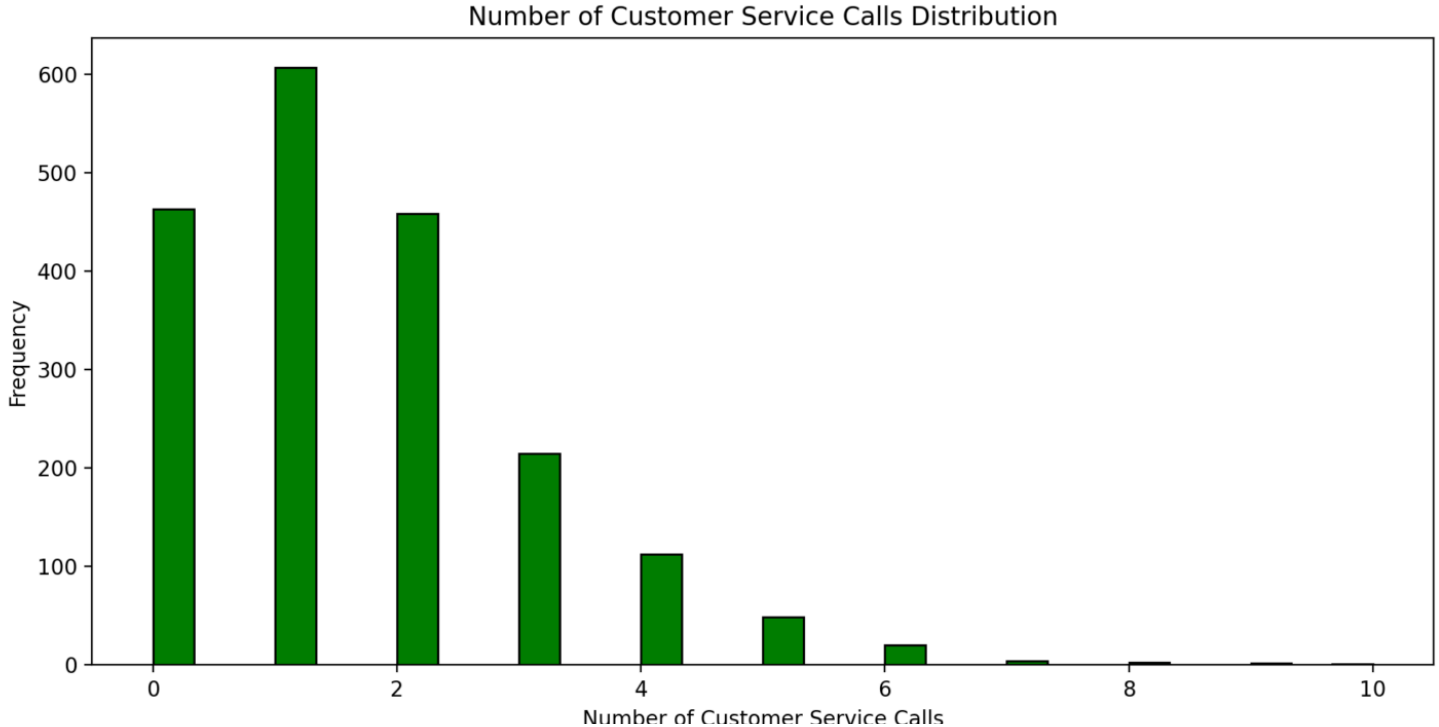


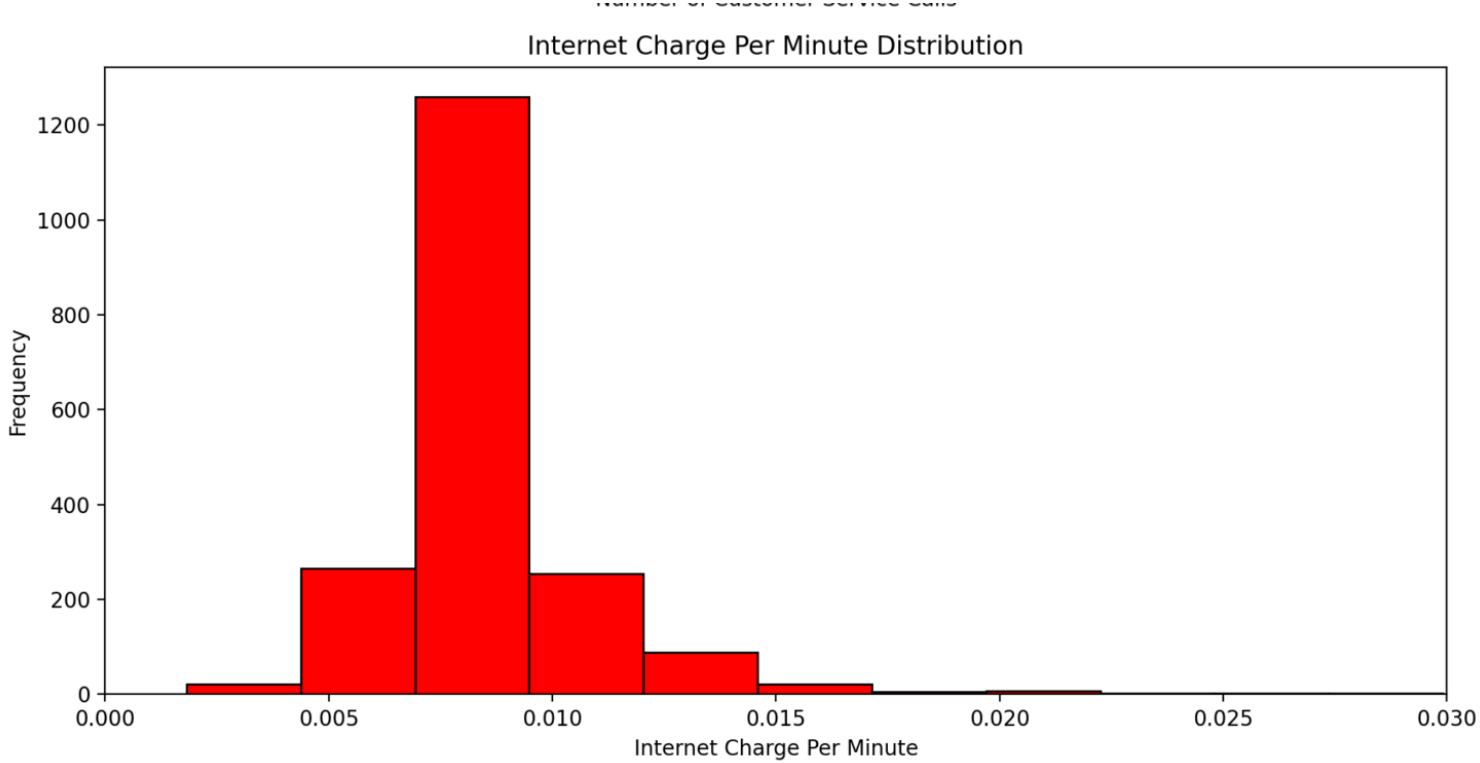
The Request: Create overlapping histograms for number_customer_service_calls for churn=0 versus churn=1. Since the are many more data rows where churn=0, convert the data to relative frequency.



The Request: Create histograms for tenure, number_customer_service_calls and internet_charge_per_min. For the internet_charge_per_min, limit the x axis to 0.03.


Tenure Distribution
Frequency
Tenure
0
50
100
150
200
250
0
10
20
30
40
50
60
70


Number of Customer Service Calls Distribution
Frequency
Number of Customer Service Calls
0
100
200
300
400
500
600
0
2
4
6
8
10


Internet Charge Per Minute Distribution
Frequency
0
200
400
600
800
1000
1200
0.000
0.005
0.010
0.015
0.020
0.025
0.030
Internet Charge Per Minute