Telecommunication Company's Customer Churn Analysis

1. Executive Summary

In the dynamic telecommunications industry, understanding customer behavior and managing churn rates are critical for sustainable growth. This report explores the association between monthly charges, payment methods, and churn status to develop targeted strategies for reducing churn rates and driving business growth.

2. Business objectives

The primary objective is to identify factors influencing churn in the telecommunications industry, aiming to develop targeted pricing and retention strategies for reducing churn rates, enhancing customer satisfaction, and driving sustainable growth.

3. Data Description

Data Definition	Data Source	Data Size	Rows/Columns
Telecom Customer Churn Data	Telco-Customer-Churn.xlsx	737 KB	(7043, 21)

The dataset used in this analysis comprises 21 columns and 7043 rows. We will focus on several key columns for our analysis, including:

- *Churn*: indicates whether a customer has churned, with two categories 'Yes' and 'No'.
- *MonthlyCharges*: records the customer's most recent monthly charges.
- *TotalCharges*: records the total charges incurred by the customer during their tenure.
- *Tenure*: indicates the customer's tenure in terms of the number of months.
- *PaymentMethod*: includes four categories and records the customer's payment method.

4. EDA – Exploratory Data Analysis

Based on our past experience, we have formulated several hypotheses and aim to verify their validity:

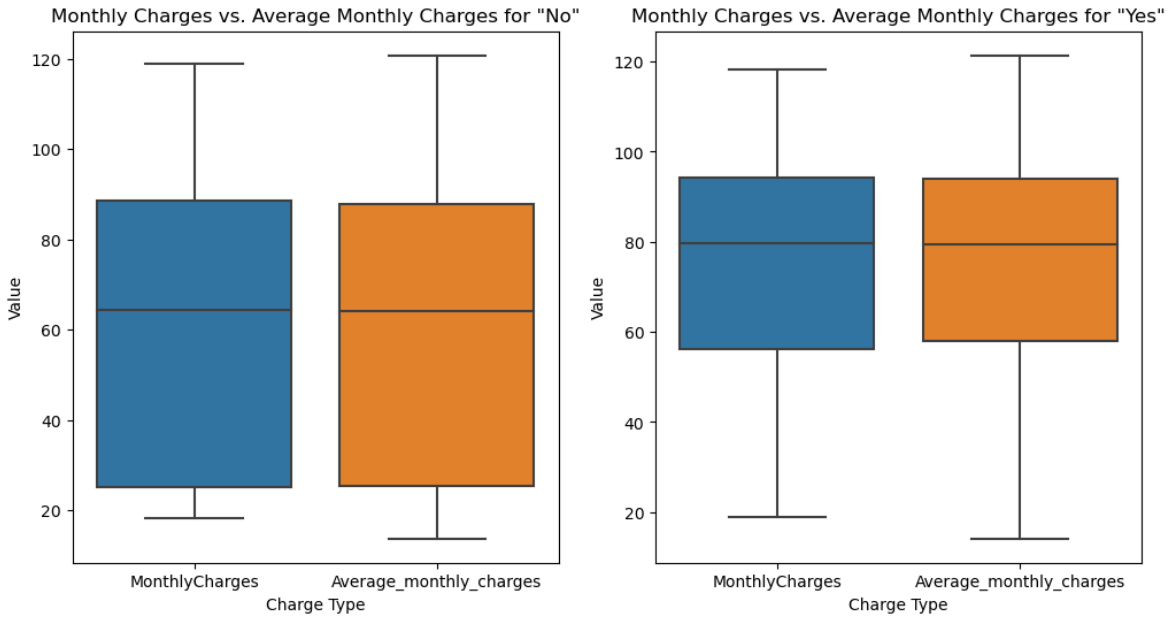
1) Customers may consider churning if their recent monthly charges are comparatively higher.
2) Customers who churned may prefer certain payment methods.
3) Customers' intent to churn may affect their monthly charges.
4) There may be variations in monthly charges between customers who churn and those who do not.

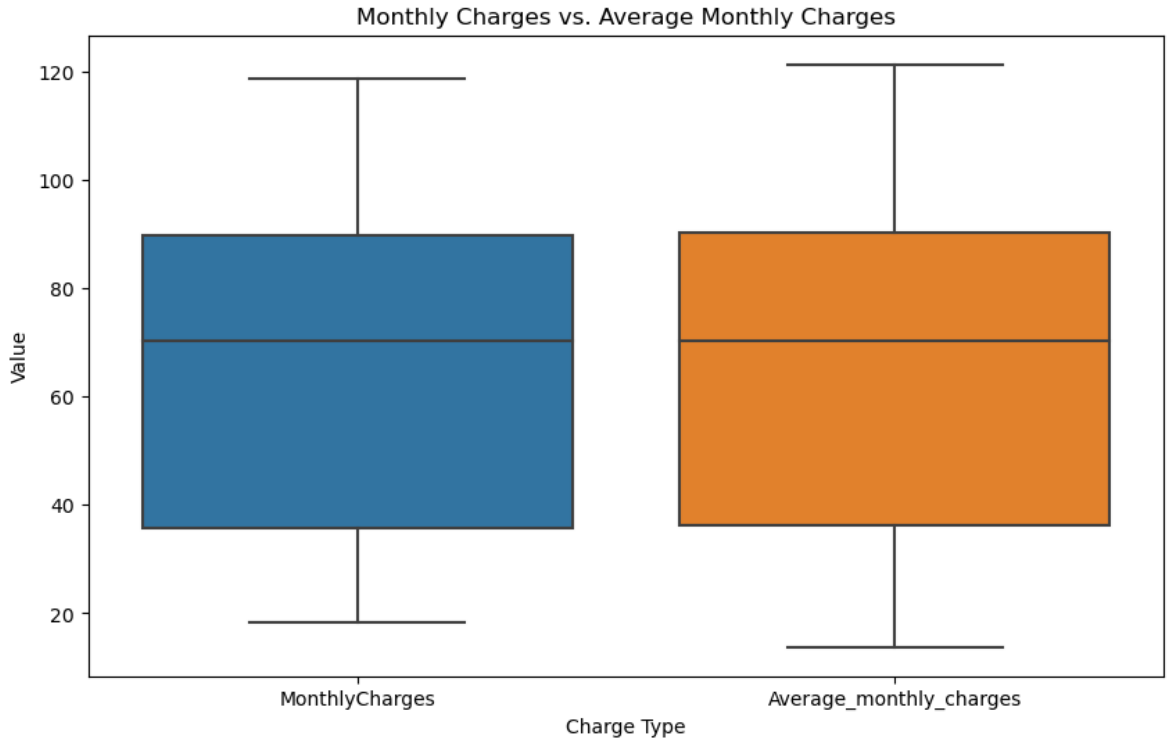
To investigate these assumptions, we will conduct exploratory data analysis (EDA) step by step.

Hypothesis 1: Monthly Charges and Churn

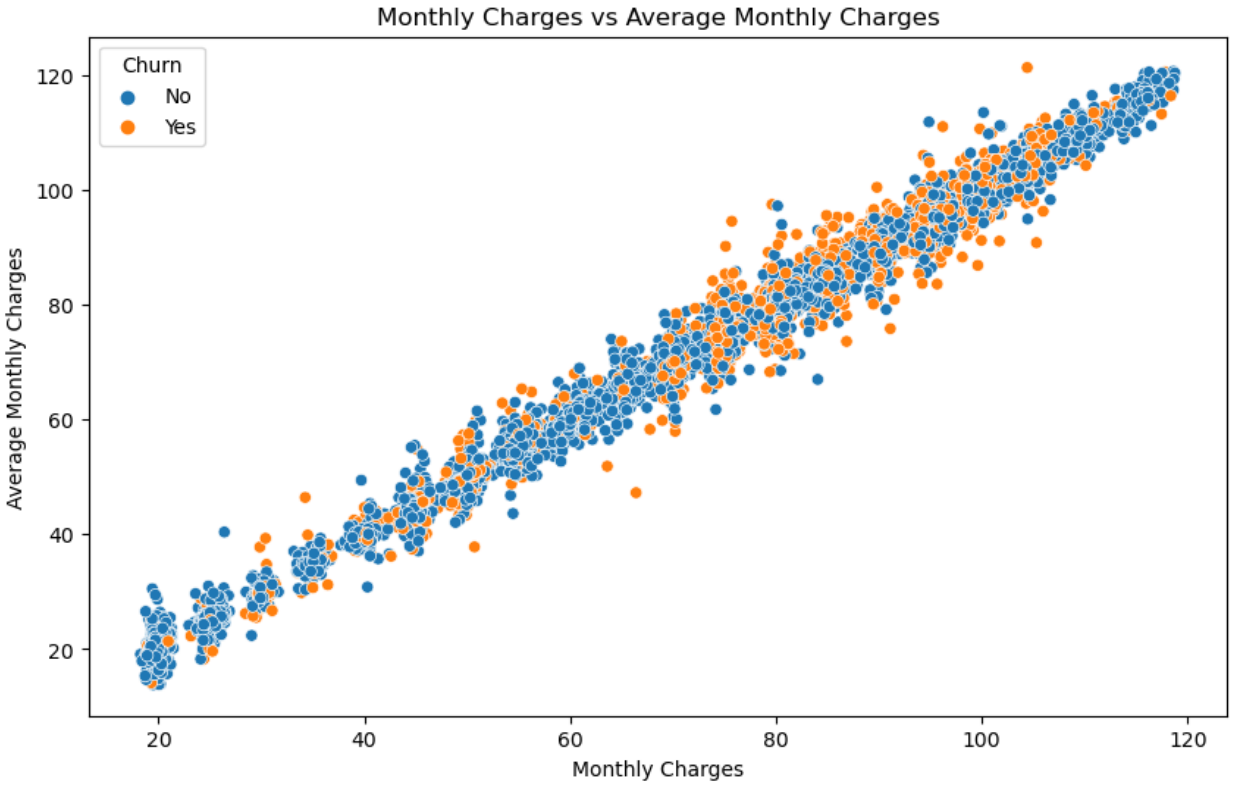
To verify this assumption, we compared the monthly charges with the average monthly charges, which are the total charges divided by the number of months of tenure.

First, we generated a chart that contains the vertical boxplots of both *MonthlyChages* and average monthly charges.





We also generated a scatterplot of *MonthlyCharges* versus average monthly charges.



To assess the distributions and relationship between *MonthlyCharges* and average monthly charges, we computed Pearson, Spearman's Rank-Order, Kendall's Tau-b, and Distance correlations between the two variables. The results are as follows:

- Pearson Correlation: 0.9962373123907768
- Spearman's Rank-Order Correlation: 0.9917562816538894
- Kendall's Tau-b Correlation: 0.9271331330292459
- Distance Correlation: 0.9954266667578495

Based on the above analysis, we concluded that high correlations, including Pearson, Spearman's rank-order, and Kendall's Tau-b, close to 1, suggest a robust linear and monotonic relationship between monthly charges and average monthly charges. Additionally, a distance correlation nearing 1 indicates a strong association, whether linear or nonlinear. The boxplots reveal a high similarity in the distribution of the two variables, as they significantly overlap. The scatterplot further supports a linear relationship, affirming the strong correlations observed, emphasizing both the linear and monotonic nature of the association between the variables.

Hypothesis 2: Churn and Payment Methods

To test the assumption that customers who churned may prefer certain payment methods, we establish the null hypothesis that there are no specific payment method preferences among churned customers. We set our significance level at 5% to determine the validity of this hypothesis.

We first cross-tabulate the payment method by the churn indicator. The cross-tabulation of Payment Method is as below:

PaymentMethod / Churn	No	Yes
Bank transfer (automatic)	1286	258
Credit card (automatic)	1290	232
Electronic check	1294	1071
Mailed check	1304	308

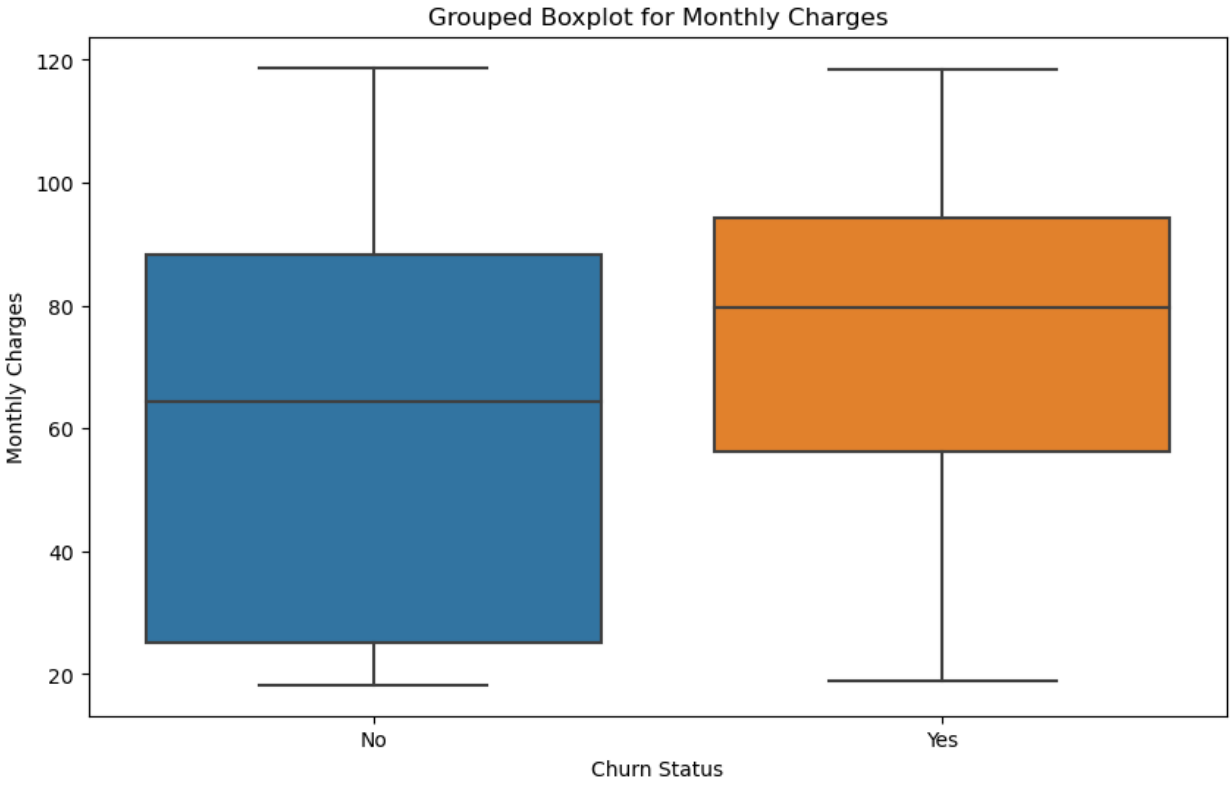
Afterwards, we perform the Pearson Chi-Square test to test if Churn is statistically independent of payment method. We calculated the test statistic value, the degree of freedom, the significance value, and the Cramer's V value.

- Chi-Square Value: 648.1423274814001
- Degrees of Freedom: 3
- P-value: 3.68e-140
- Cramer's V Value: 0.3033586255540706

The above analysis reveals a significant relationship between the *Churn* and *PaymentMethod*, as evidenced by the Chi-Square value of 648.14 with 3 degrees of freedom and an extremely low p-value of 3.68e-140. This indicates that the association is not likely due to chance. Furthermore, Cramer's V value of 0.30 suggests a moderate strength of association between the variables. In conclusion, the statistical findings strongly support the presence of a meaningful relationship between the two variables under consideration.

Hypothesis 3: Churn Status Impact on Monthly Charges

In order to find out if a customer's intent to churn will affect *MonthlyCharges*, we first generated a grouped boxplot for *MonthlyCharges* with *Churn* as the grouping variable.



We then calculated the Counts, the Means, and the Standard Deviations of *MonthlyCharges* within each *Churn* category. The results are as below:

Churn	**Count**	**Mean**	**Standard deviation**
No	5174	61.265124	31.092648
Yes	1869	74.441332	24.666053

The Eta-Squared statistic is 0.037386706020624216.

Based on the above analyses, we concluded that churning customers exhibit a higher average *MonthlyCharges* at $74.44, in contrast to non-churning customers with an average of $61.27. The standard deviation for churning customers is lower, indicating less spending variability. Although there is statistical evidence supporting a significant relationship between churn status and *MonthlyCharges*, the effect size is small, explaining only a 3.74% variation. While a statistically significant difference exists in monthly charges between customers planning to churn and those who don't, the impact of churn status on *MonthlyCharges* is relatively modest.

Hypothesis 4: Variations in Monthly Charges

Since we do not know if the variations of *MonthlyCharges* are the same between the two groups of customers, we would use the Welch's t-test for comparison. We set our significance level at 5% to determine the validity of this hypothesis.

The Welch's t statistic is 18.40752667641466, the Satterthwaite degree of freedom is 4135.795017764662, and the test significance value is 8.592449331549498e-73. Since the significance value is well below the 0.05 significance level, this leads to the rejection of the null hypothesis. Thus, there's strong evidence supporting a significant difference in average monthly charges between customers who churn and those who do not.

Moreover, we would also like to know if the two groups have the same population variances by performing the Snedecor-Cochran F test. The F Statistic is 1.5889714, degree of freedom is (5,173.0000000, 1,868.0000000), and the significance is 0.0000000. With a significance value below 0.05, the null hypothesis is rejected, indicating evidence that the two groups (churning and non-churning customers) exhibit different population variances in Monthly Charges.

5. Insights and Conclusions

Based on our analyses, we have identified four insights regarding associations among customers' churn status, monthly charges, and payment methods:

Hypothesis 1: Monthly Charges and Churn

- Robust linear and monotonic relationship is observed between monthly charges and average monthly charges, indicating a strong association between the variables.
- The strong correlation emphasizes the need for targeted pricing strategies.

Hypothesis 2: Churn and Payment Methods

- Significant relationship is found between churn status and preferred payment methods, supported by statistical analysis.
- The significant relationship highlights the importance of payment method preferences in customer retention.

Hypothesis 3: Churn Status Impact on Monthly Charges

- Churning customers exhibit higher average monthly charges, though the effect size is relatively small.
- The impact of churn status on monthly charges indicates potential opportunities for pricing adjustments.

Hypothesis 4: Variations in Monthly Charges

- Significant difference is observed in average monthly charges between churning and non-churning customers, supported by statistical tests.
- The difference in average monthly charges between churning and non-churning customers suggests the need for tailored retention strategies.

6. Recommendations

Based on our insights, we have provided several recommendations to reach our intended goal in reducing churn rates and driving business growth:

- Implement targeted pricing strategies based on customer preferences and behavior.
- Enhance retention efforts by offering personalized incentives and promotions.
- Continuously monitor churn rates and adjust strategies accordingly to maintain sustainable business growth.