# Case Study 1 – Data Visualization

## Problem Statement:

Consider yourself to be Sam who is a data scientist. He has been approached by a telecom company to build some aesthetic graphs to make better sense of the customer data.

## Tasks To Be Performed:

1. Sam has to build a bar-plot for the 'Contract' column:

- a. Set the x-axis label to be 'Contract Type of customer'
- b. Set the y-axis label to be 'Count'
- c. Set the title of the plot to be 'Distribution of Contract'
- d. Assign 'orange' color to all the bars

2. Sam has to build a histogram for the 'MonthlyCharges' column:

- a. Set the x-axis label to be 'Monthly Charges Incurred'
- b. Set the y-axis label to be 'Count'
- c. Set the title of the plot to be 'Distribution of Monthly Charges'
- d. Assign 'forestgreen' color to the bins

3. Sam has to build a scatter-plot between 'TotalCharges' and 'tenure'. 'TotalCharges' should be on the y-axis and 'tenure' should be on the x-axis.

- a. Set the x-axis label to be 'Tenure of the customer'
- b. Set the y-axis label to be 'Total chargesIncurred'
- c. Set the title of the plot to be 'Total Charges vs Tenure'
- d. Assign 'indigo' color to the points

4. Sam has to build a box-plot between 'MonthlyCharges' and 'PaymentMethod'. 'MonthlyCharges' should be on the y-axis and 'PaymentMethod' should be on the x-axis.

- a. Set the x-axis label to be 'Payment Method of customer'
- b. Set the y-axis label to be 'Monthly ChargesIncurred'
- c. Set the title of plot to be 'Monthly Charges vs. Payment Method'
- d. Assign 'olive' color to the box-plots

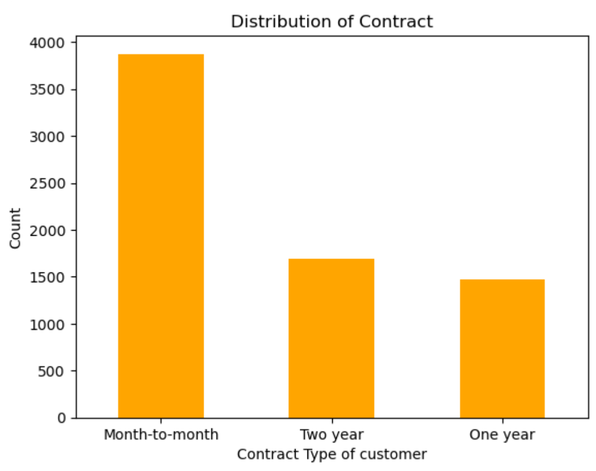
In [1]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
%matplotlib inline
import seaborn as sns

df = pd.read_csv(r"csv files\customer_churn.csv")
df.head()
```

Out[1]:

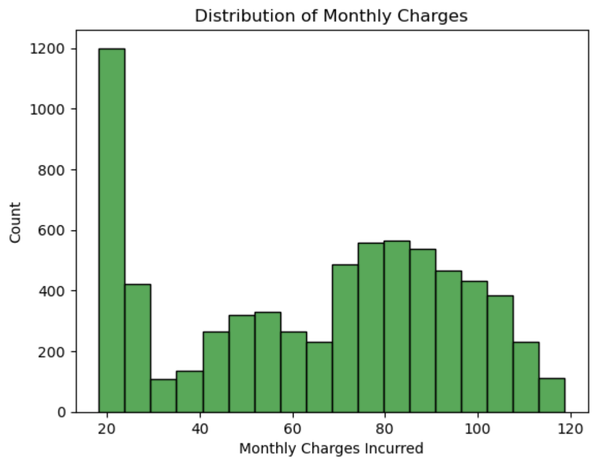
| | customerID | gender | SeniorCitizen | Partner | Dependents | tenure | PhoneService | MultipleLines | InternetService | OnlineSecurity | ... | DeviceProtection | TechSupport | StreamingTV | StreamingMovies | Contract | PaperlessBilling | PaymentMethod | MonthlyCharges | TotalCharges | Churn |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 7590-VHVEG | Female | 0 | Yes | No | 1 | No | No phone service | DSL | No | ... | No | No | No | No | Month-to-month | Yes | Electronic check | 29.85 | 29.85 | No |
| 1 | 5575-GNVDE | Male | 0 | No | No | 34 | Yes | No | DSL | Yes | ... | Yes | No | No | No | One year | No | Mailed check | 56.95 | 1889.5 | No |
| 2 | 3668-QPYBK | Male | 0 | No | No | 2 | Yes | No | DSL | Yes | ... | No | No | No | No | Month-to-month | Yes | Mailed check | 53.85 | 108.15 | Yes |
| 3 | 7795-CFOCW | Male | 0 | No | No | 45 | No | No phone service | DSL | Yes | ... | Yes | Yes | No | No | One year | No | Bank transfer (automatic) | 42.30 | 1840.75 | No |
| 4 | 9237-HQITU | Female | 0 | No | No | 2 | Yes | No | Fiber optic | No | ... | No | No | No | No | Month-to-month | Yes | Electronic check | 70.70 | 151.65 | Yes |

5 rows × 21 columns

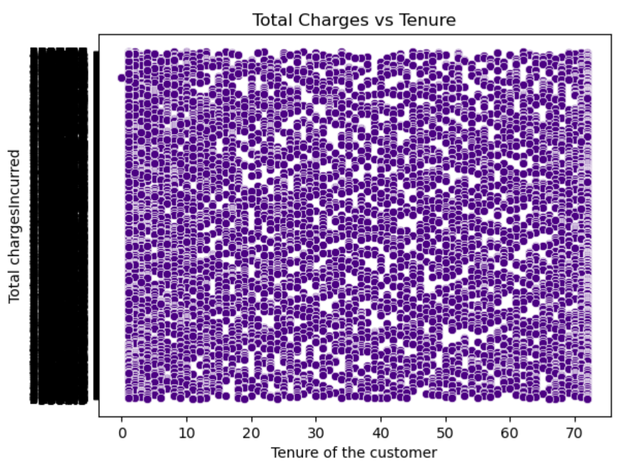
In [31]:
```
df['Contract'].value_counts().plot.bar(color='orange')
plt.xticks(rotation=0)
plt.xlabel('Contract Type of customer')
plt.ylabel('Count')
plt.title('Distribution of Contract')
plt.show()
```



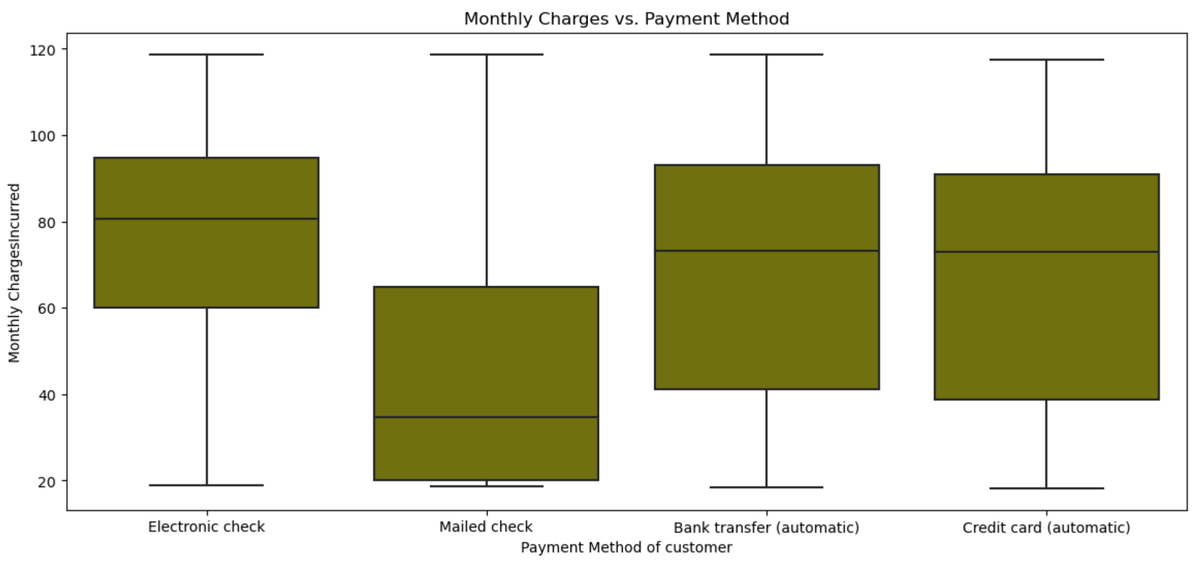

In [32]:
```
sns.histplot(df['MonthlyCharges'], color='forestgreen')
plt.xlabel('Monthly Charges Incurred')
plt.ylabel('Count')
plt.title('Distribution of Monthly Charges')
plt.show()
```




In [35]:
```
sns.scatterplot(y=df['TotalCharges'], x=df['tenure'], color='indigo')
plt.xlabel('Tenure of the customer')
plt.ylabel('Total chargesIncurred')
plt.title('Total Charges vs Tenure')
plt.show()
```




In [38]:
```
plt.figure(figsize=(14,6))
sns.boxplot(y=df['MonthlyCharges'], x=df['PaymentMethod'], color='olive')
plt.xlabel('Payment Method of customer')
plt.ylabel('Monthly ChargesIncurred')
plt.title('Monthly Charges vs. Payment Method')
plt.show()
```




In [ ]:
```
```