In [1]:
```
!pip install pandas openpyxl
```

```
Requirement already satisfied: pandas in c:\users\donbo\anaconda3\lib\site-packages
(2.0.3)
Requirement already satisfied: openpyxl in c:\users\donbo\anaconda3\lib\site-packages
(3.0.10)
Requirement already satisfied: python-dateutil>=2.8.2 in c:\users\donbo\anaconda3\lib
\site-packages (from pandas) (2.8.2)
Requirement already satisfied: pytz>=2020.1 in c:\users\donbo\anaconda3\lib\site-pack
ages (from pandas) (2023.3.post1)
Requirement already satisfied: tzdata>=2022.1 in c:\users\donbo\anaconda3\lib\site-pa
ckages (from pandas) (2023.3)
Requirement already satisfied: numpy>=1.21.0 in c:\users\donbo\anaconda3\lib\site-pac
kages (from pandas) (1.24.3)
Requirement already satisfied: et_xmlfile in c:\users\donbo\anaconda3\lib\site-pack
es (from openpyxl) (1.1.0)
Requirement already satisfied: six>=1.5 in c:\users\donbo\anaconda3\lib\site-packages
(from python-dateutil>=2.8.2->pandas) (1.16.0)
```

In [2]:
```
import pandas as pd
```

In [5]:
```
# Load the Excel file
data = pd.read_excel("Research.xlsx")

# Display the first few rows of the data
print(data.head(10))
```

```
      Travel Frequency Preferred Airline Overall Satisfaction Punctuality  \
0     2-5 times a year           Vistara            Satisfied        Good
1     2-5 times a year            IndiGo            Satisfied     Average
2  Once a year or less         Air India            Satisfied        Good
3  Once a year or less           Vistara            Satisfied        Good
4     2-5 times a year            IndiGo            Satisfied   Excellent
5     2-5 times a year           Vistara       Very Satisfied        Good
6     2-5 times a year            IndiGo              Neutral        Good
7     2-5 times a year         Air India            Satisfied        Good
8     2-5 times a year            IndiGo              Neutral     Average
9     2-5 times a year            IndiGo              Neutral     Average

  Seating Comfort In-flight Service Booking Satisfaction Value for Money  \
0       Satisfied              Good            Satisfied         Average
1    Dissatisfied           Average            Satisfied            Good
2       Satisfied              Good            Satisfied            Good
3       Satisfied              Good            Satisfied         Average
4       Satisfied           Average            Satisfied            Good
5       Satisfied              Good            Satisfied            Good
6    Dissatisfied           Average              Neutral         Average
7       Satisfied              Good            Satisfied            Good
8         Neutral           Average              Neutral         Average
9       Satisfied           Average            Satisfied         Average

  Baggage Issues                       Key Factor Recommendation
0            Yes  Flight schedule and punctuality    Very Likely
1             No              Price of the ticket         Likely
2             No              Price of the ticket         Likely
3             No                 Customer service         Likely
4            Yes                    Safety record         Likely
5             No  Flight schedule and punctuality    Very Likely
6            Yes              Price of the ticket        Neutral
7             No              Price of the ticket         Likely
8            Yes                 Customer service        Neutral
9             No  Flight schedule and punctuality        Neutral
```

In [6]:

```
# Get an overview of the data
print(data.info())
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 66 entries, 0 to 65
Data columns (total 11 columns):
 #   Column                Non-Null Count  Dtype 
---  ------                --------------  ----- 
 0   Travel Frequency      66 non-null     object
 1   Preferred Airline     66 non-null     object
 2   Overall Satisfaction  66 non-null     object
 3   Punctuality           66 non-null     object
 4   Seating Comfort       66 non-null     object
 5   In-flight Service     66 non-null     object
 6   Booking Satisfaction  66 non-null     object
 7   Value for Money       66 non-null     object
 8   Baggage Issues        66 non-null     object
 9   Key Factor            66 non-null     object
 10  Recommendation        66 non-null     object
dtypes: object(11)
memory usage: 5.8+ KB
None
```

In [7]:

```python
# Check for missing values
print(data.isnull().sum())
```

```
Travel Frequency        0
Preferred Airline       0
Overall Satisfaction    0
Punctuality             0
Seating Comfort         0
In-flight Service       0
Booking Satisfaction    0
Value for Money         0
Baggage Issues          0
Key Factor              0
Recommendation          0
dtype: int64
```

In [9]:

```python
# Check for duplicate rows
duplicate_rows = data[data.duplicated()]

# Display duplicate rows
print(duplicate_rows)
```

```
Empty DataFrame
Columns: [Travel Frequency, Preferred Airline, Overall Satisfaction, Punctuality, Sea
ting Comfort, In-flight Service, Booking Satisfaction, Value for Money, Baggage Issue
s, Key Factor, Recommendation]
Index: []
```

In [10]:

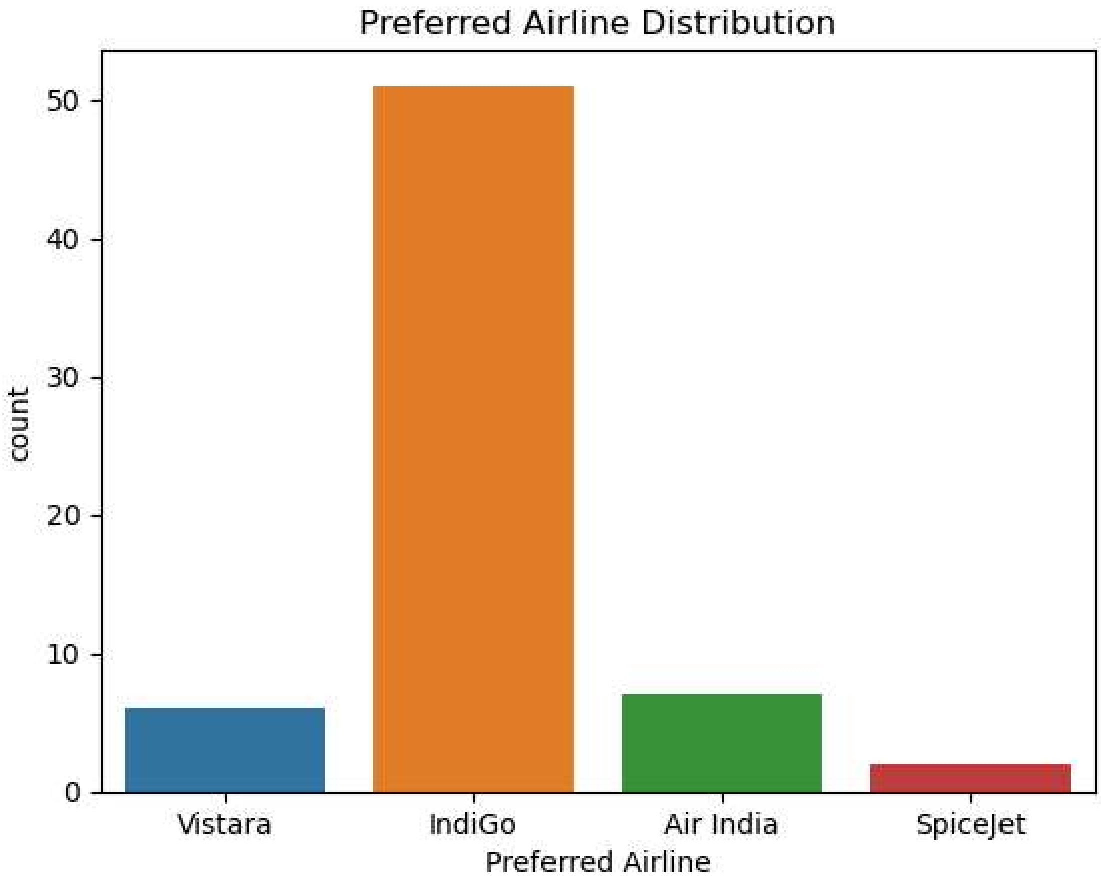
```python
# Count total number of duplicate rows
num_duplicates = data.duplicated().sum()
print(f"Number of duplicate rows: {num_duplicates}")
```

```
Number of duplicate rows: 0
```

In [13]:

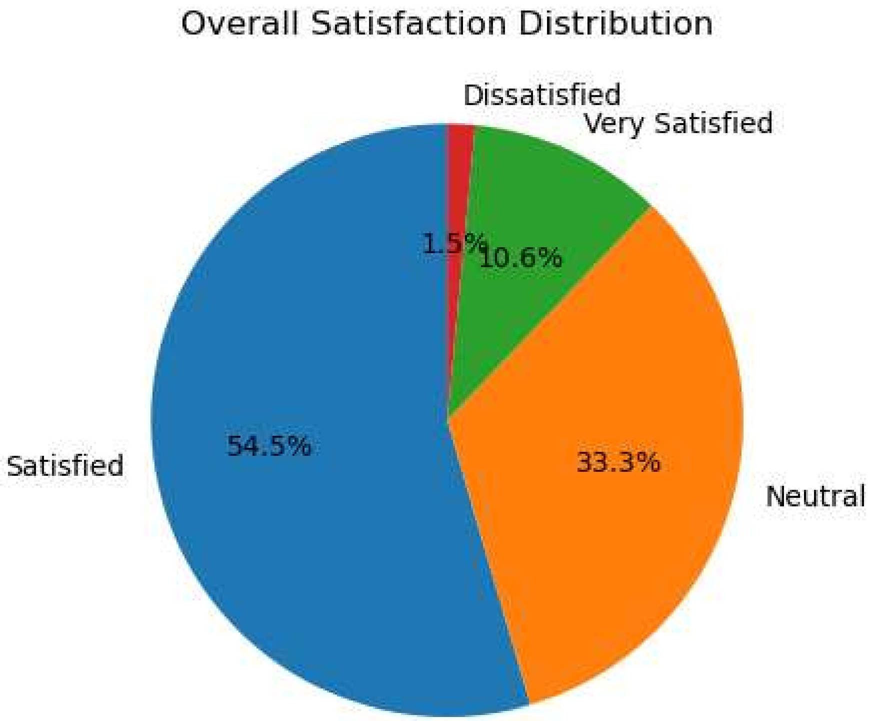
```python
import matplotlib.pyplot as plt
import seaborn as sns

sns.countplot(x='Preferred Airline', data=data)
plt.title('Preferred Airline Distribution')
plt.show()
```



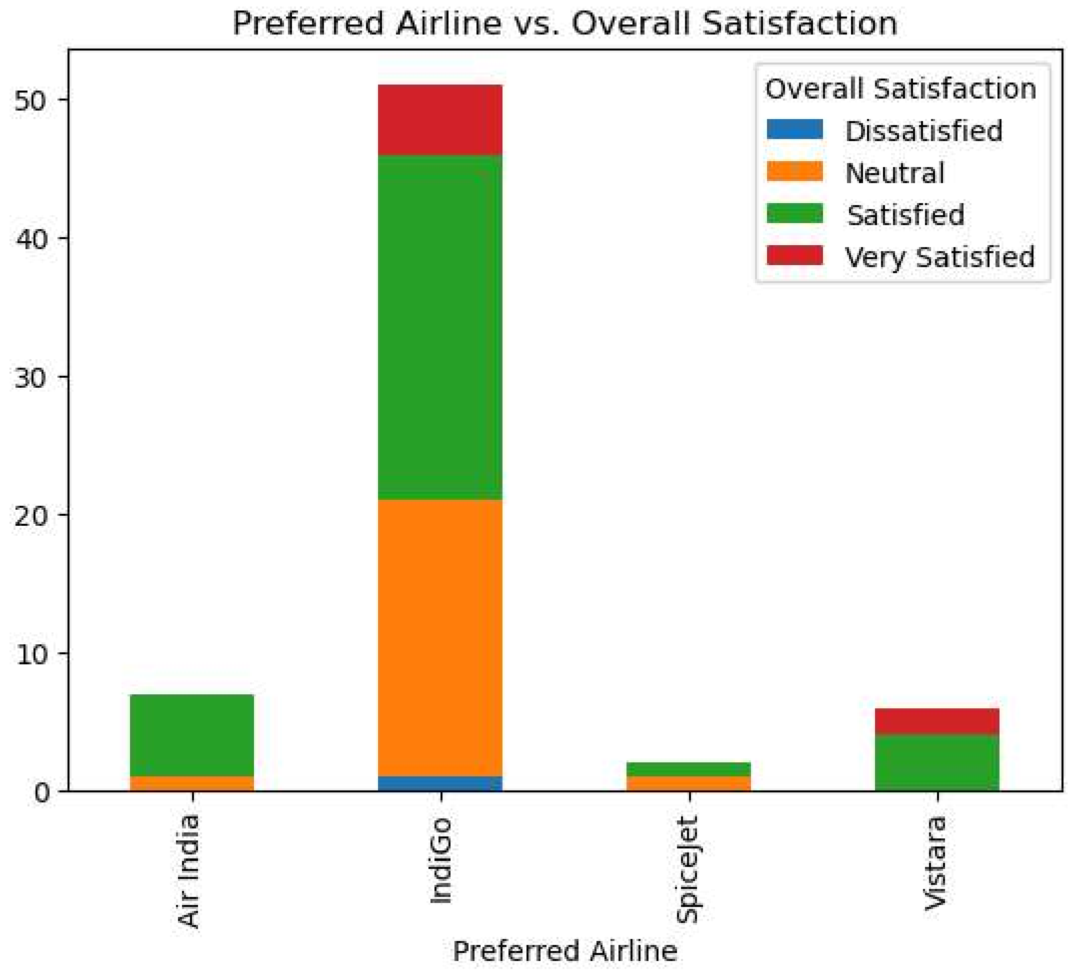
In [14]:

```
data['Overall Satisfaction'].value_counts().plot.pie(autopct='%1.1f%%', startangle=90)
plt.title('Overall Satisfaction Distribution')
plt.ylabel('')
plt.show()
```



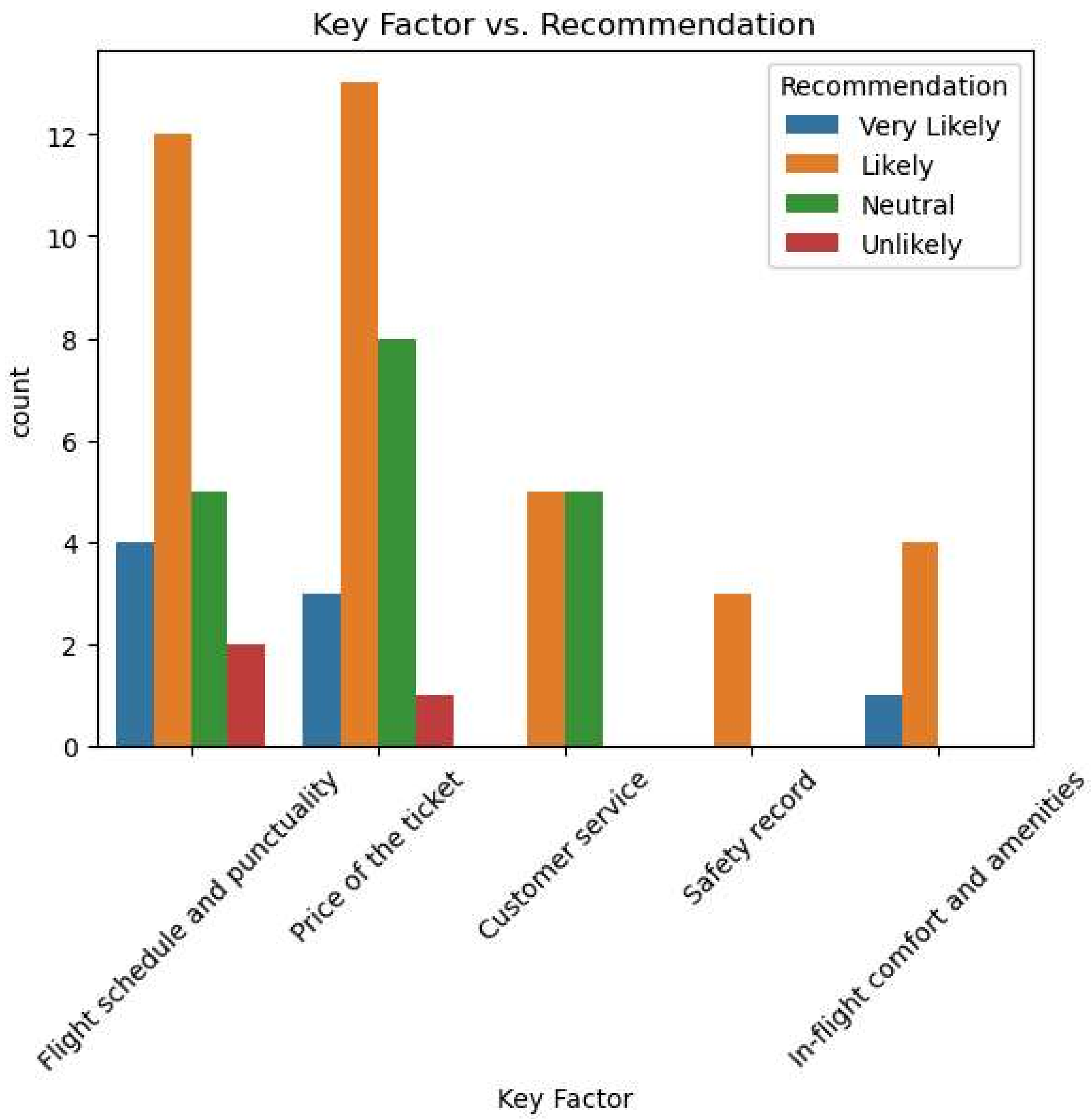
In [15]:

```
pd.crosstab(data['Preferred Airline'], data['Overall Satisfaction']).plot(kind='bar',
plt.title('Preferred Airline vs. Overall Satisfaction')
plt.show()
```



In [16]:

```
sns.countplot(x='Key Factor', hue='Recommendation', data=data)
plt.title('Key Factor vs. Recommendation')
plt.xticks(rotation=45)
plt.show()
```


Key Factor vs. Recommendation
Recommendation
Very Likely
Likely
Neutral
Unlikely
count
0
2
4
6
8
10
12
Flight schedule and punctuality
Price of the ticket
Customer service
Safety record
In-flight comfort and amenities
Key Factor