1.DISTRIBUTION OF TARGET VARIABLE
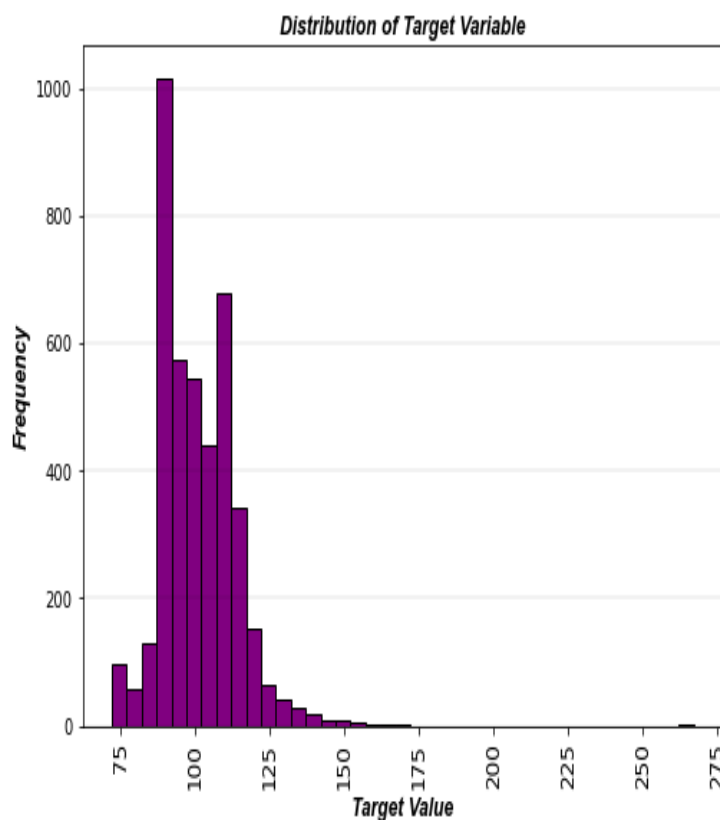
DEFINE THE RANGE OF VALUES FOR THE HISTOGRAM BINS

```python
In [192]: # Define the range of values for the histogram bins
          bin_width = 5
          bins = np.arange(train_df['y'].min(), train_df['y'].max() + bin_width, bin_width)

          # Create a histogram of the target variable with a bin width of 5
          font_style = {'family': 'Arial', 'size': 12, 'weight': 'bold', 'style': 'italic'}
          plt.figure(figsize=(8,6))
          plt.grid(axis='y',ls='solid',color ='k',lw=0.2,alpha=0.5)
          plt.hist(train_df['y'], bins=bins, color='purple', edgecolor='black')

          # Add labels and a title
          plt.xlabel('Target Value',fontdict=font_style)
          plt.ylabel('Frequency',fontdict=font_style)
          plt.title('Distribution of Target Variable',fontdict=font_style)
          plt.xticks(rotation='vertical', ha='center',size=12)

          # Display the chart
          plt.show()
```


Distribution of Target Variable

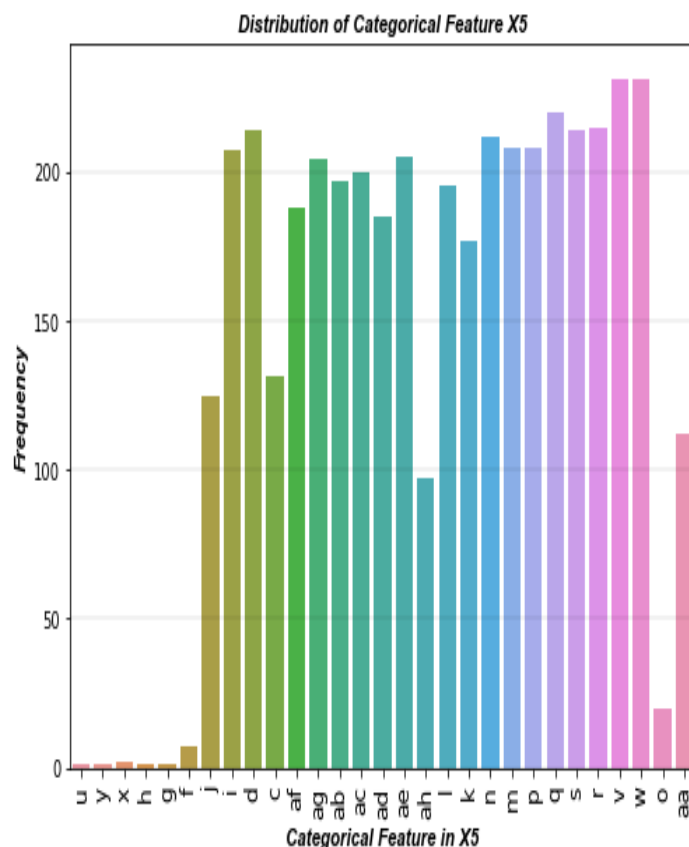## 2. DISTRIBUTION OF CATEGORICAL FEATURE X5

DEFINE THE RANGE OF VALUES THE HISTOGRAM BINS

```python
In [145]: # Define the range of values for the histogram bins
          bin_width = 5
          bins = np.arange(train_df['y'].min(), train_df['y'].max() + bin_width, bin_width)

          # Create a histogram of the target variable with a bin width of 5
          font_style = {'family': 'Arial', 'size': 11.5, 'weight': 'bold', 'style': 'italic'}
          plt.figure(figsize=(8,6))
          plt.grid(axis='y',ls='solid',color ='k',lw=0.2,alpha=0.5)
          sns.countplot(x='X5', data=train_df)

          # Add labels and a title
          plt.xlabel('Categorical Feature in X5',fontdict=font_style)
          plt.ylabel('Frequency',fontdict=font_style)
          plt.title('Distribution of Categorical Feature X5',fontdict=font_style)
          plt.xticks(rotation='vertical', ha='center',size=12)

          # Display the chart
          plt.show()
```

3.DISTRIBUTION OF PREDICTED TARGET VARIABLE FROM GIVEN TEST DATA

DEFINE THE RANGE OF VALUES FOR THE HISTOGRAM BINS

In [191]:
```python
# Define the range of values for the histogram bins
bin_width = 1
bins = np.arange(submission['y'].min(), submission['y'].max() + bin_width, bin_width)

# Create a histogram of the target variable with a bin width of 5
font_style = {'family': 'Arial', 'size': 12, 'weight': 'bold', 'style': 'italic'}
plt.figure(figsize=(8,6))
plt.grid(axis='y',ls='solid',color ='k',lw=0.2,alpha=0.5)
plt.hist(submission['y'], bins=bins, color='purple', edgecolor='black')

# Add labels and a title
plt.xlabel('Target Value',fontdict=font_style)
plt.ylabel('Frequency',fontdict=font_style)
plt.title('Distribution of Target Variable',fontdict=font_style)
plt.xticks(rotation='vertical', ha='center',size=12)

# Display the chart
plt.show()
```