Question 5: How many orders are not rated? [1 mark]

To determine the number of orders that are not rated in your dataset, we can use the **value\_counts()** method in pandas to count the frequency of each unique value in the 'rating' column. By default, unrated or null ratings might be recorded as a specific value or as NaNs (null values).

However, since the dataset doesn't have null values as we checked earlier, we'll need to identify if there's a specific marker (like 0 or a negative number) used for unrated orders. Assuming the absence of a rating is indicated by a rating value of '0', I'll write the function to count how many orders have a '0' rating.

Let's proceed with the coding of this function.

I've prepared the function to count the number of unrated orders, assuming '0' as the indicator for unrated orders. Here’s the Python code for that function:

The output from the function indicates that there are \*\*736 unrated orders\*\* in the dataset. The ratings distribution also shows the count for other ratings:

- 588 orders with a rating of 5

- 386 orders with a rating of 4

- 188 orders with a rating of 3

It seems that the dataset uses "Not given" to mark unrated orders. This was important to adjust in the function's assumption if we're to automate or standardize this process further. If you need adjustments or further analysis, feel free to let me know!

**df['rating'].value\_counts() ## Complete the code**

**Question 6:** Explore all the variables and provide observations on their distributions. (Generally, histograms, boxplots, countplots, etc. are used for univariate exploration.) [9 marks]

import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

def plot\_distributions(file\_path):

# Load the dataset

data = pd.read\_csv(file\_path)

# Setting up the figure and axes for the plots

fig, axes = plt.subplots(nrows=3, ncols=2, figsize=(16, 18))

# Histograms for numeric variables

sns.histograph(data['cost\_of\_the\_order'], bins=30, kde=True, ax=axes[0, 0])

axes[0, 0].set\_title('Histogram of Cost of the Order')

sns.histograph(data['food\_preparation\_time'], bins=15, kde=True, ax=axes[0, 1])

axes[0, 1].set\_title('Histogram of Food Preparation Time')

sns.histograph(data['delivery\_time'], bins=15, kde=True, ax=axes[1, 0])

axes[1, 0].set\_title('Histogram of Delivery Time')

# Boxplots for numeric variables

sns.boxplot(y=data['cost\_of\_the\_order'], ax=axes[1, 1])

axes[1, 1].set\_title('Boxplot of Cost of the Order')

sns.boxplot(y=data['food\_preparation\_time'], ax=axes[2, 0])

axes[2, 0].set\_title('Boxplot of Food Preparation Time')

sns.boxplot(y=data['delivery\_time'], ax=axes[2, 1])

axes[2, 1].set\_title('Boxplot of Delivery Time')

plt.tight\_layout()

plt.show()

# Example usage, replace 'your\_dataset.csv' with the path to your dataset

plot\_distributions("path\_to\_your\_dataset.csv")