# **INTERPRETATION OF THE BOX PLOT**

The box plot generated is intended to visualize the distribution of the 'total\_bill' column with respect to different days of the week, grouped by whether the customer is a smoker or not.

- The x-axis represents the days of the week (Thur, Fri, Sat, Sun).

- The y-axis represents the total bill amount.

- The boxes in the plot represent the interquartile range (IQR), which spans from the 25th percentile (Q1) to the 75th percentile (Q3) of the data.

- The horizontal line within each box represents the median (50th percentile) of the data.

- The "whiskers" extend from the boxes to the minimum and maximum values within a certain range, usually defined as 1.5 times the IQR.

- Any data points outside the whiskers are considered outliers and are plotted individually as points.

# **REPORT**

The box plot you've generated provides insights into the distribution of total bill amounts for different days of the week, considering whether the customers are smokers or not. Here are some observations:

## **1. MEDIAN TOTAL BILL AMOUNTS:**

- On all days, the median total bill amount for non-smokers is generally lower than that for smokers.

## **2. INTERQUARTILE RANGE (IQR):**

- The IQR indicates the range in which the middle 50% of the data falls. For both smokers and non-smokers, the IQR tends to be wider on Saturdays and Sundays, suggesting more variability in the total bill amounts on weekends.

## **3. OUTLIERS:**

- There are several outlier data points on Fridays, Saturdays, and Sundays, especially for non-smokers. These outliers might represent large bills for special occasions or other factors.

## **4. COMPARISON BY SMOKING STATUS:**

- For non-smokers, the median total bill amount is relatively consistent across all days of the week, with Sunday showing a slightly higher median.

- For smokers, the median total bill amount appears to be higher on Fridays, followed by Saturdays and Sundays.

## **5. VARIABILITY:**

- Saturdays and Sundays exhibit a wider range of total bill amounts compared to Thursdays and Fridays, regardless of smoking status.

## **6. OBSERVATION ON DAYS:**

- Thursdays have relatively lower variability in total bill amounts, possibly due to being weekdays and associated with lunchtime.

# **CONCLUSION:**

The box plot illustrates that total bill amounts tend to be higher for smokers, and there is also variability based on the day of the week. Fridays, Saturdays, and Sundays generally have higher median total bill amounts, potentially indicating that customers tend to spend more during the weekend. The presence of outliers suggests that there are certain instances where customers spend significantly more, possibly indicating special events or larger group gatherings.

