1. Restaurant X dataset head:

drive_times_x dataframe

	drive_times_x
0	174.0
1	105.0
2	233.0
3	167.0
4	154.0

2. Restaurant Y dataset head:

drive_times_y dataframe

	drive_times_y
0	240.0
1	225.0
2	130.0
3	226.0
4	217.0

3. Summary Statistics for Restaurant X and Y

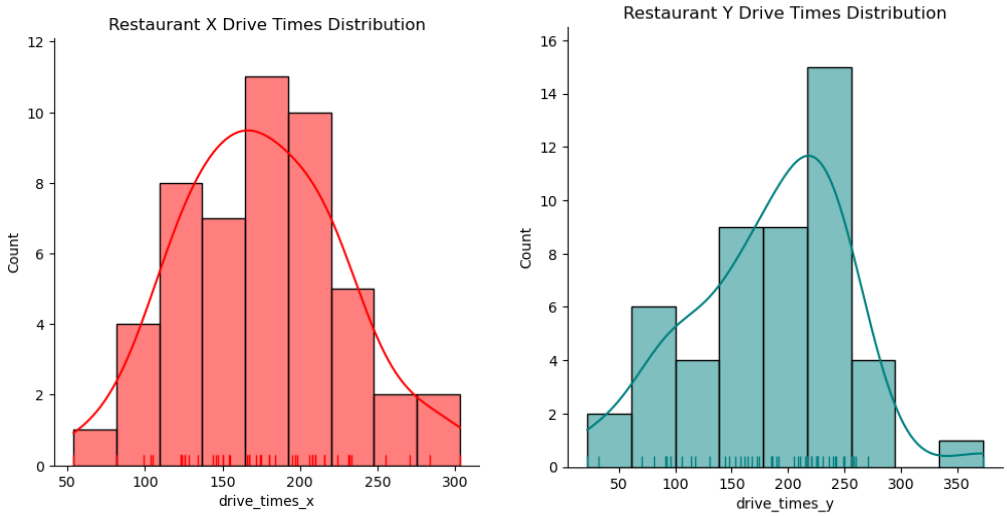
Statistics	**Restaurant X**	**Restaurant Y**
Mean	174.640	183.180
Median	173	191
stDev	52.084	68.351
Q1	136.500	145
Q3	209.500	226
Max	303	373
Min	54	22

- Restaurant X exhibits a data range spanning 249 units between its maximum and minimum values, with an interquartile range of 73. The standard deviation is 52.084. Notably, the mean of 174.64 closely aligns with the median of 173, indicating a limited presence of outliers that could significantly impact the mean. Restaurant Y drive-times data displays a substantial range of 351 units between its maximum and minimum values, accompanied by an interquartile range of 81. The standard deviation is 68.351. However, the mean of 183.18 significantly

differs from the median of 191, indicating the likely presence of outliers that are exerting an influence on the mean.

1.
   - Based on the provided statistics, it appears that the distribution for Restaurant X is approximately normal and exhibits a relatively symmetrical pattern, with the majority of data points clustered around the mean.
     In contrast, when examining the statistics for Restaurant Y, it becomes evident that the distribution is more dispersed and contains outliers that could potentially distort the accurate representation of the mean drive time.
- Both the histograms and the boxplots appear to support the above conclusions.

4. Here are both side-by-side box plots and duo histograms. I've included both types as box plots provide a clearer view of outliers and ranges, while histograms effectively illustrate the contrast between Restaurant X and Restaurant Y



- When examining the boxplots, the presence of outliers in Restaurant Y becomes more evident. Additionally, the boxplot shows a notably extended lower whisker, indicating increased variability in the data points below the first quartile (Q1). In the case of Restaurant Y, outliers are visually marked by Python using the threshold values as whisker end markers.
- The boxplot for Restaurant X exhibits a balanced distribution, utilizing both the minimum and maximum values as whisker end markers.
- Both graphs validate the description of the distributions. While Restaurant X and Restaurant Y share similarities, such as mostly adhering to a normal distribution with a mere 8-unit difference in their interquartile ranges (IQR), they also exhibit differences. Notably, Restaurant Y contains two outliers, identified through the calculation of the IQR and the detection of data values deviating 1.5 times above or below the IQR. Furthermore, Restaurant X appears to display a more evenly

distributed pattern with a narrower spread compared to Restaurant Y

5. 99% Confidence intervals for both Restaurant x and Restaurant y.

99% CI for Mean Drive Times	
Restaurant x	(155.67, 193.61)
Restaurant y	(158.28, 208.08)

   - I chose METHOD 3 to calculate the confidence intervals because there was sufficient data to use a normal distribution and the Scipy code is more efficient and requires less lines of code than METHOD 1.
   - The 99% confidence intervals for Restaurant X and Restaurant Y exhibit an overlap within the range of (158.28, 193.61). This overlap indicates that there is no statistically significant difference between the two confidence intervals.

6. 90% Confidence intervals for both Restaurant x and Restaurant y.
   - The 90% confidence intervals are narrower and provide a lower level of confidence compared to the wider 99% confidence intervals, which offer a higher level of confidence but cover a broader range of values.

7. The distributions of mean drive-through times for Restaurant X and Restaurant Y exhibit similarities. While there is an absolute difference of 8.54 between their respective mean drive times, there is a much larger absolute difference of 102.0 in the range of drive times between the two restaurants. This broader range in Restaurant Y's drive times can be attributed to the presence of two outliers and a larger standard deviation in the dataset.

   However, it's noteworthy that there is an overlap in both the 90% and 99% confidence intervals, suggesting that no statistically significant difference exists between the drive times of Restaurant X and Restaurant Y