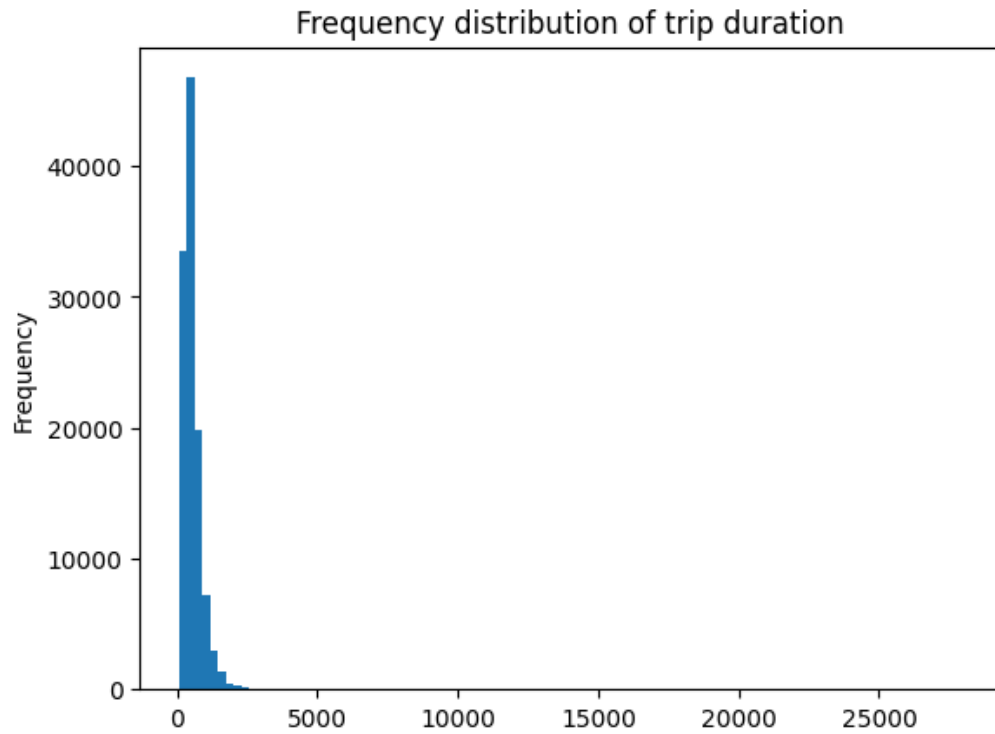


QUESTION

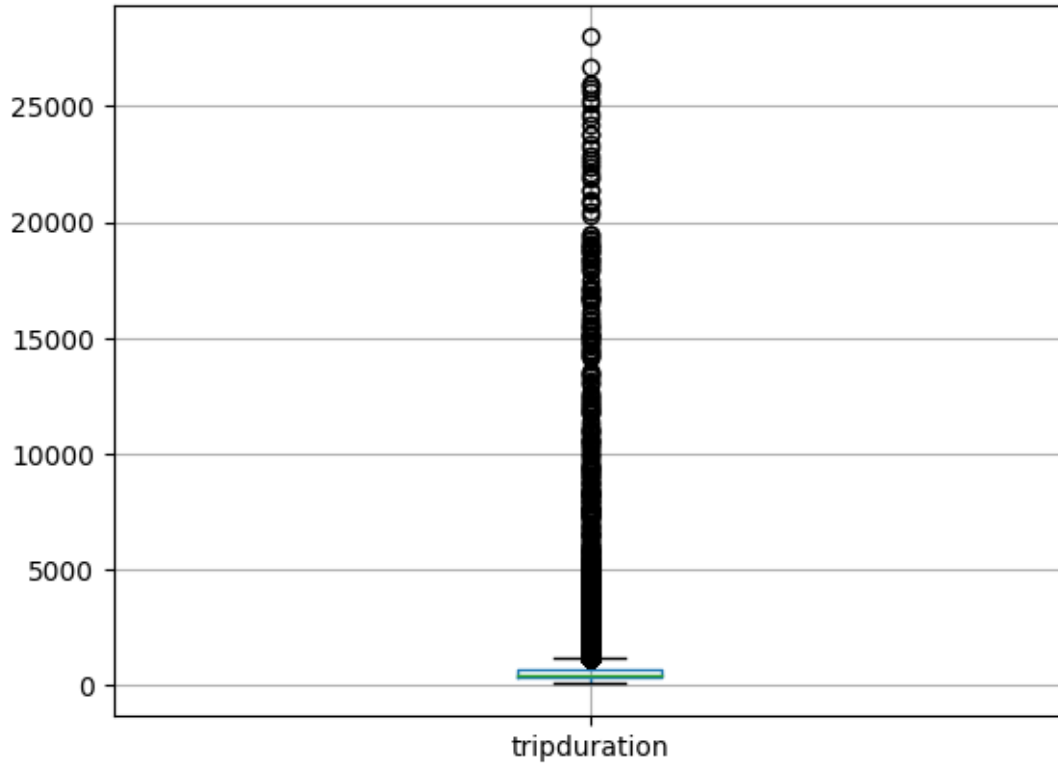
1. Find the mean, median, and mode of the trip duration of gender type male.
2. By looking at the numbers obtained earlier, in your opinion is the distribution symmetric or skewed? If skewed, then is it positively skewed or negatively skewed?
3. Plot a frequency distribution of trip duration for trips availed by gender type male. Does it validate your inference as you did so in the previous question?
4. Plot a box plot of the trip duration of trips taken by males. Do you think any outliers exist?
5. Apply the formula in Listing 1-19 to determine the percentage of observations for which outliers exists.
6. Perform the treatment of outliers by incorporating one of the methods we discussed earlier for the treatment of outliers (lihat penjelasan di halaman 28). Untuk STB Ganjil gunakan metode "replace outliers with mean", STB Genap gunakan metode "transform the outlier values to upper boundary atau to lower boundary).

ANSWER

1. Berdasarkan gambar dibawah ini nilai mean 563.402797, nilai median 458.451500 dan nilai mode 466.174000
2. Distribusi akan miring dari kiri ke kanan yg artinya positif karena nilai Mean < Median < Mode.
3. Iya, plot distribusi miring dari kiri ke kanan.



4. Berdasarkan box plot berikut dapat dilihat kalau terdapat banyak outliers dari data.



5. Berdasarkan perhitungan nilai proporsi pada data sebagai outliers sebesar 112845.030104 persen
6. Dengan mengganti outliers dengan mean maka boxplot-nya lebih bersih dari outliers dan distribusi data lebih meluas.

