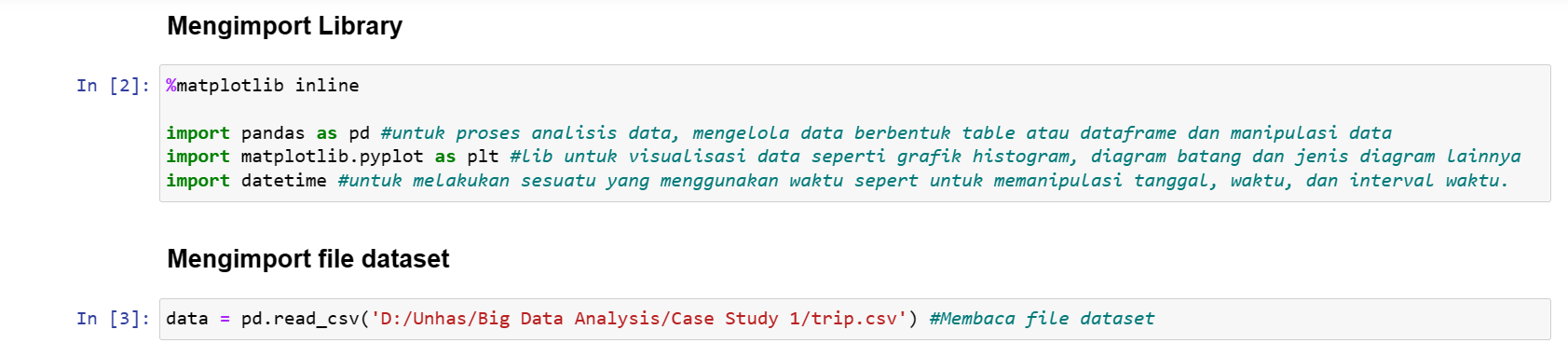
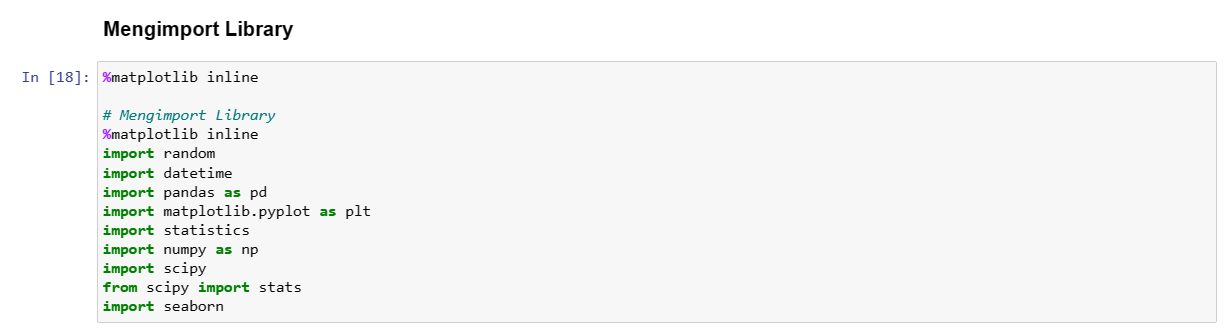
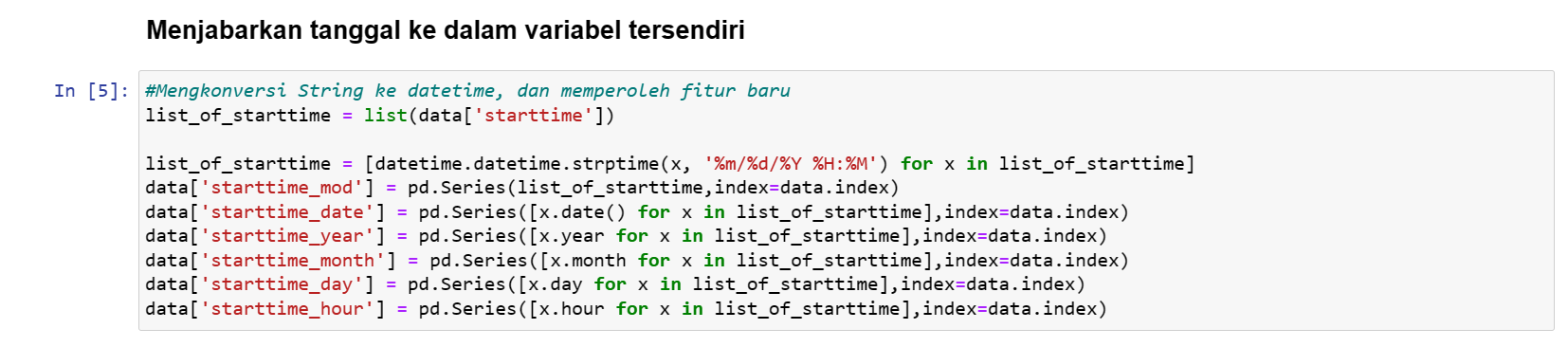
**Nama : Emil Hardiansyah**

**NIM : D082221005**

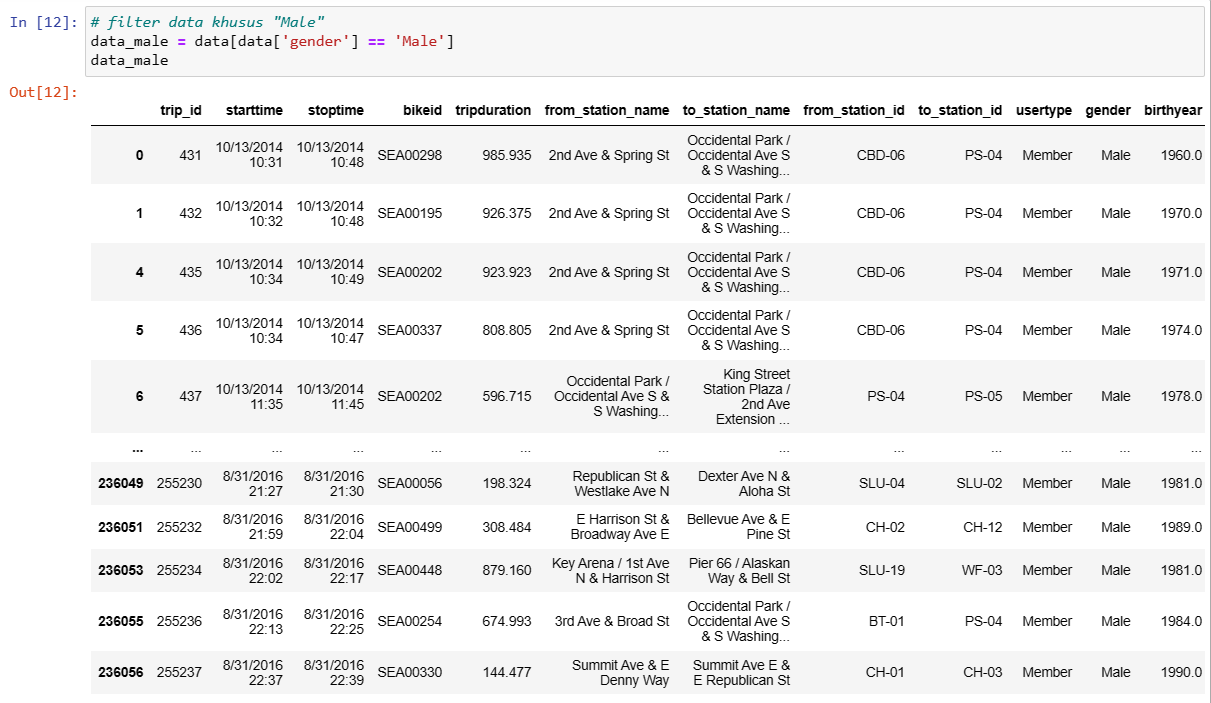
**Tugas 2 Case Study – Cycle Sharing Scheme**

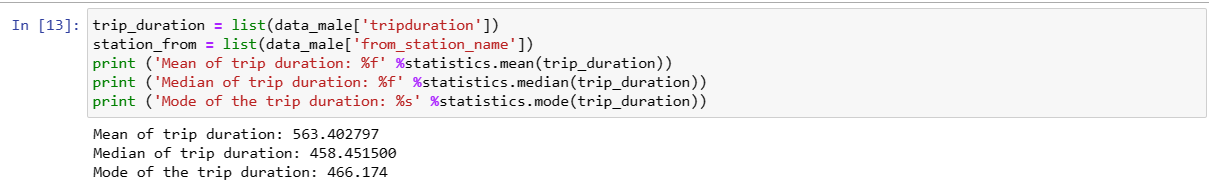




1. Find the mean, median, and mode of the trip duration of gender type male.

**Answer:**

****

****

Berdasarkan gambar diatas nilai mean 563.402797, nilai median 458.451500 dan nilai mode 466.174000

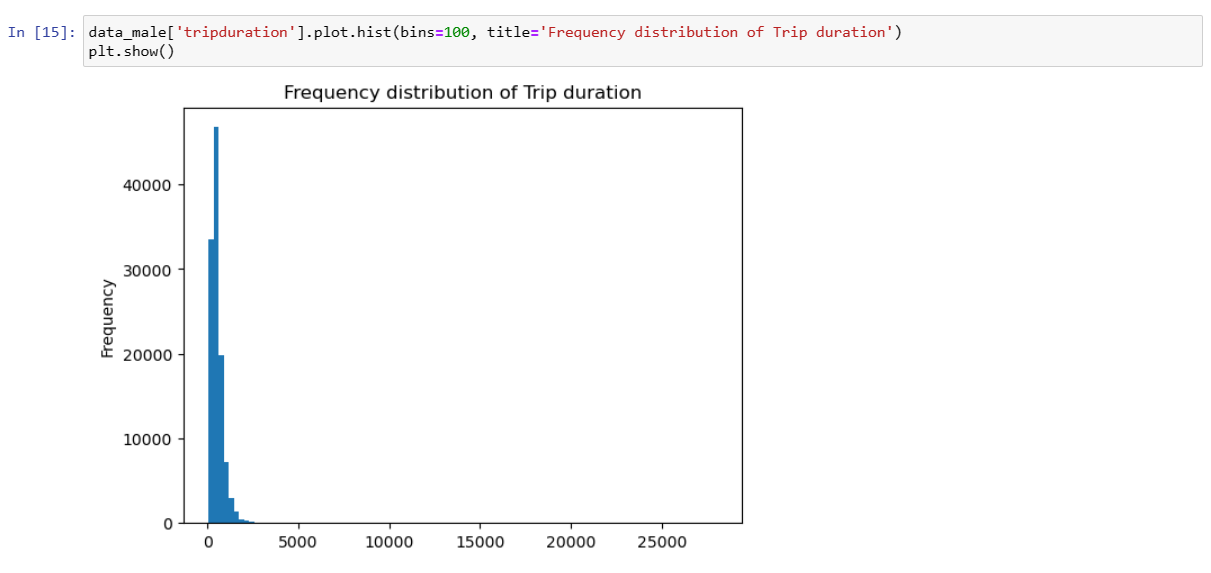
1. By looking at the numbers obtained earlier, in your opinion is the distribution symmetric or skewed? If skewed, then is is it positively skewed or negatively skewed?

**Answer:**

Berdasarkan rata-rata yang lebih besar dibandingkan nilai tengah atau Mean > Median ada kemungkinan distribusinya tidak simetris karena jarak antara mean dan median jauh, dan distribusinya kemungkinan miring. Untuk kemiringan distribusi tidak bisa ditebak, karena positively skewed dan negatively skewed sama-sama bisa terjadi karena mean > median begitupun sebaliknya.

1. 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?

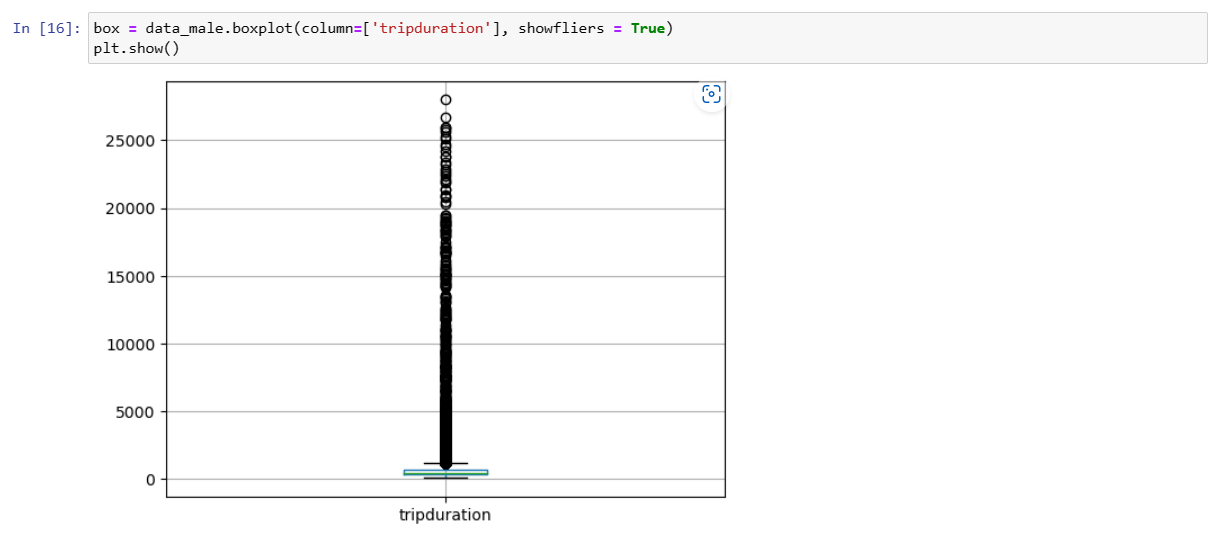
**Answer:**



Iya, plot distribusi miring dari kiri ke kanan.

1. Plot a box plot of the trip duration of trips taken by males. Do you think any outliers exist?

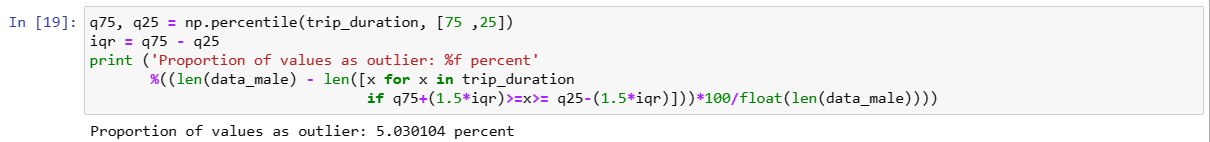
**Answer:**



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

1. Apply the formula in Listing 1-19 to determine the percentage of observations for which outliers exists.

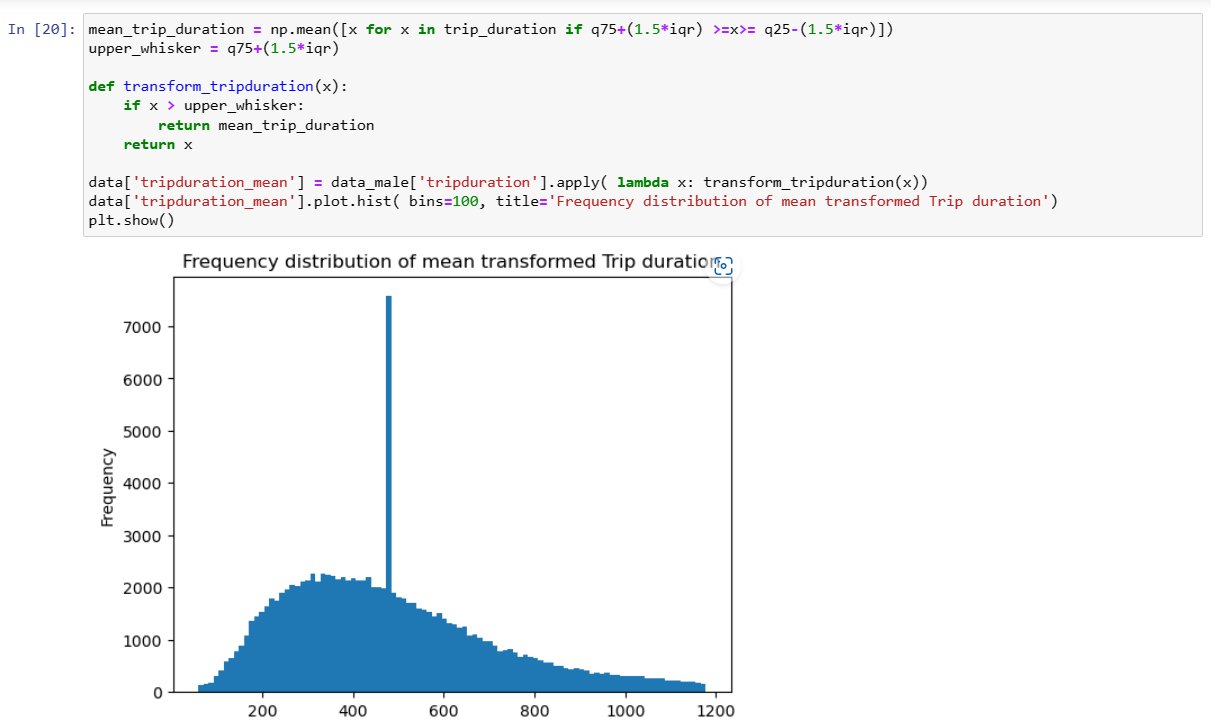
**Answer:**



Berdasarkan perhitungan nilai proporsi pada data sebagai outliers sebesar 5.030104 persen.

1. 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:**



Dengan mengganti outliers dengan mean maka boxplot-nya lebih bersih dari outliers dan distribusi data lebih meluas.