

Homework:
Percentiles & Quantiles

MATH 150

Due: Feb 7, 2024

Sepehr Akbari

Problem 1

The first two problems refer to the erykah data set, available on Moodle. Compute the five-number summary and inter-quartile range of the tempo variable. You should only need one R function for each.

Answer

Using R:

```
library(readxl)
erykah <- read_excel("Documents/LFC/MATH 150/DSs/erykah.xlsx")
View(erykah)

fivenum(erykah$tempo) # 0.000 82.485 92.109 129.868 215.079
IQR(erykah$tempo) # 47.383
```

Problem 2

What is the 40th percentile of the duration variable? How long is this in minutes?

Answer

Using R:

```
quantile(erykah$duration_ms, 0.4) # 40th quantile = 248309
quantile(erykah$duration_ms, 0.4) / 60000 # approx. 4.138483 minutes
```

Problem 3

The remaining problems refer to the following data, which represents the ages of 18 customers at a restaurant.

```
49 58 61 39 55 57 53 50 64
42 45 57 45 51 30 37 44 49
```

Which ages are below the 30th percentile?

Answer

Using R:

```
data <- c(49,58,61,39,55,57,53,50,64,42,45,57,45,51,30,37,44,49)

data[data < quantile(data, 0.3)] # 39 42 30 37 44
```

Problem 4

Which ages are above the 60th percentile?

Answer

Using R:

```
data[data > quantile(data, 0.6)] # 58 61 55 57 53 64 57
```