

# **Introduction to Statistics**

**Exercises: Part 1**

Sophie Lee

## Exercise 1: Missing data

A mobile phone app collects user information. Some of this data are missing as users chose to opt out of location-based services.

Give one scenario where the missing data would lead to **biased** analysis results.

Give another scenario where the missing data leads to a reduced sample but could still be used to produce **unbiased** results.

## Exercise 2: Summary statistics

### Question 1

The following line graphs are taken from the Criminal Courts statistics quarterly report, showing the average number of days from offense to completion for defendants at the Crown Court. Which graph is the most appropriate to display the average time and why?

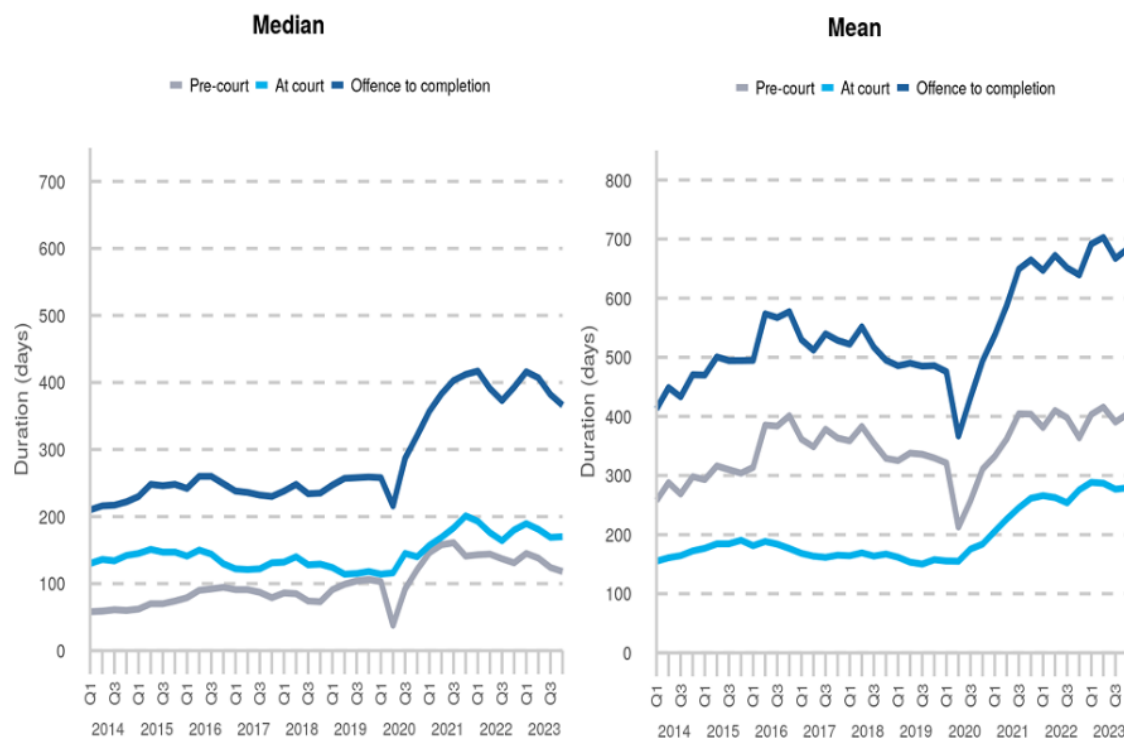


Figure 1: Average number of days from offence to completion for defendants dealt with at the Crown Court, Q1 2014 – Q4 2023

## Question 2

The mean and standard deviation of waiting times in weeks for hearings in the Crown Court between 2020 and 2023 are given in the table below. Using this information, what can you tell about the distribution of these times?

Year	Mean wait (weeks)	SD wait (weeks)
2020	13.5	10.2
2021	19.1	15.6
2022	20.9	13.9
2023	22.8	14.7