

The Scenario

This data comes from the London Fire Brigade. A panel of Fire service managers and local politicians want to better understand some particular aspects of the costs and response times associated with incidents during the time period in the data.

They have asked that you provide a report containing specific outputs and answering their specific questions.

The Structure

This should have two sections. The first section should include the code that you produce to perform all stages of the data analysis. It should read in the datafile provided, check the integrity of data, calculate all of the statistics requested by the company, and produce all of the figures/visualisations requested.

The purpose of this first section is that it could be shared with someone else who is not familiar with the data or project, but is an expert in R and statistics. They should be able to see how you achieved your results, and if necessary, easily modify the code to work with new or updated data. Therefore, the appropriateness of your comments and data documentation will also be assessed.

The second section should be a polished and professional report presenting and interpreting the findings for the panel of fire service managers and politicians. You should assume that the panel have a basic knowledge of what data is recorded about fire response and what the different variables mean, but are not experts in statistics or R. Therefore, you should communicate the results clearly and accurately, without relying on technical terms. The visualisations and figures in this section should be as clear as possible and of publication standard that could be presented to journalists or other professionals.

The Request

The company have requested the following:

The costs of responding to fires and false alarms

The panel would like to know the sum of all costs associated with responding to fires during the time period, and (separately) the sum of all costs associated with responding to false alarms. They would also like to know the average (mean) cost of responding to a fire, and of responding to a false alarm.

The distribution of response times

The panel ask that you provide a visualisation showing response times (as indicated in the 'FirstPumpArriving_AttendanceTime' variable) to all incidents where there was indeed a response.

In addition, they would also like for you to provide an additional version of the plot that allows for them to easily compare the distribution of response times for the three different types of incident ("Fire", "False Alarm" and "Special Service", as indicated by the 'IncidentGroup' variable).

Summary of special service response times

For this part you should use only special service incidents.

The panel ask that you provide the values below, separately for each type of special service (as indicated by the variable 'SpecialServiceType'):

A count of the number of each type of incident
The mean response time for each type of incident
The 10th percentile of response times for each type of incident
The 90th percentile of response times for each type of incident

A t-test comparing Ealing and Greenwich

Use a t-test to compare the mean response times in Ealing to those in Greenwich.

You should report the results using Null Hypothesis Significance Testing, and the estimation approach.