

London Fire Station analysis in Excel

In this exercise we use Excel to explore a dataset with the aim of helping inform policy.

This exercise has been built and tested in Microsoft Excel 365. It is possible to complete this exercise in earlier versions of Excel, however please note that some of the instructions may not point exactly to the location of the required features.

In this exercise we are going to analyse the performance of the London Fire Service since the closure of a number of fire stations on 9 January 2014.

The dataset required is available from the London Data Store for reference.

<https://data.london.gov.uk/dataset/london-fire-brigade-incident-records>

A smaller, cut down version, as used in this exercise is also available at:

<http://training.theodi.org/csl-an2/LFB-Data.zip>

Note: Step 1 has already been completed on the dataset above.

Step 1 - Filter the data

The first step is to decide what time frame of data to analyse. The station closures took place on 9th January 2014. To compare performance before and after closures will require you to create two datasets representing these two time periods. Once done, proceed to step 2 with the dataset containing the incident records after the closure date.

Step 2 - Sanity check

Before we continue it would be worth checking that there are no remaining records for the closed fire stations.

For the purpose of this exercise we are looking at the response time of the first fire appliance only. To check that there are no records left for the 10 closed stations, apply a filter to the *FirstPumpArriving_DeployedFromStation* column.



In Excel the filter option (or icon) is available from the data toolbar or menu. To apply it, click any column heading and select filter.

With filters on, select the drop down in the *FirstPumpArriving_DeployedFromStation* column and ensure that none the following 10 stations are showing:

Belsize, Bow, Clerkenwell, Downham, Kingsland, Knightsbridge, Silvertown, Southwark, Westminster and Woolwich.

If any of these stations are present then we may either have an error in the data or the stations didn't fully close on the 9th January 2014. You may even need to consult a knowledgeable expert to find out the reasons for this.



For the purpose of this exercise, simply remove the extraneous rows from the data. Once done ensure you clear your filter to bring back all the data.

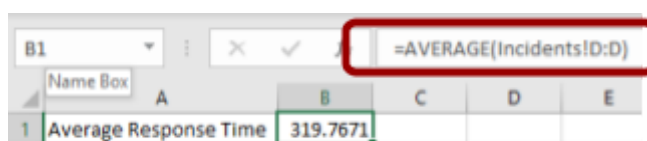
Step 3 - London performance

Before looking at the performance of each borough, let's look at London overall.

This can be done by generating the average first appliance response time.

To do this, **create a new sheet** with a row entitled "*First appliance average response time*". In the cell next to this use the average formula (=average) and highlight column "D" in the first sheet to average all of the data.

Your formula will read =average(Sheet1!D:D)



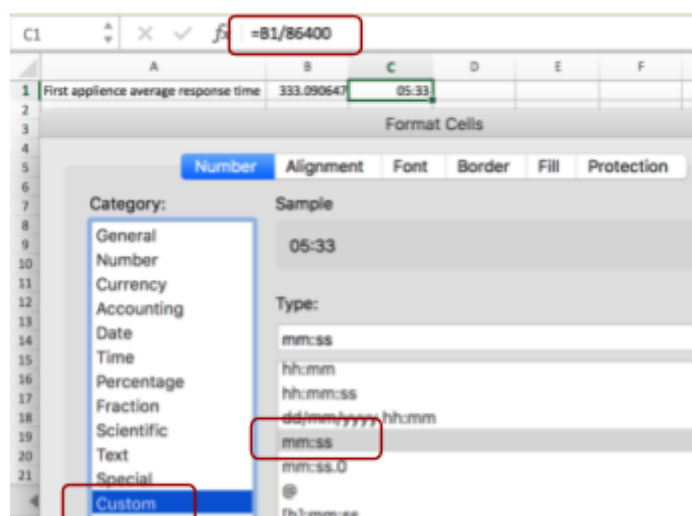
This will give you the average first appliance response time in seconds. To see this time in minutes:

- 1) You first need to convert the seconds to days by dividing the number by 86400 (the number of seconds in a day). Put this value in the next cell. (The Excel function that converts to minutes needs to know the number of days to perform the conversion.)
- 2) Format the cell (right click, format cells) with a custom format "*mm:ss*" to show the number of minutes and seconds.

Before the closures the average response time across London was 5 minutes and 34 seconds (according to the analysis available from <http://london-fire.labs.theodi.org/explore/>).

We can see that overall it appears to have improved by a second.

In this exercise we are not going to test for significance of this result or others.

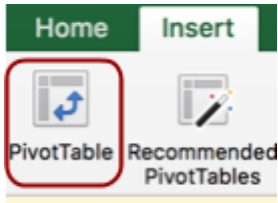


Step 4 - Performance by Borough

In this part of the exercise we are going to look at average time for the first appliance arriving on a per borough basis.

To do this we are going to use pivot tables in Excel. Pivot tables are normally available from the insert menu or panel in excel.

Before you create a pivot table, ensure you are on the full data sheet and that the data does not have any filters applied.



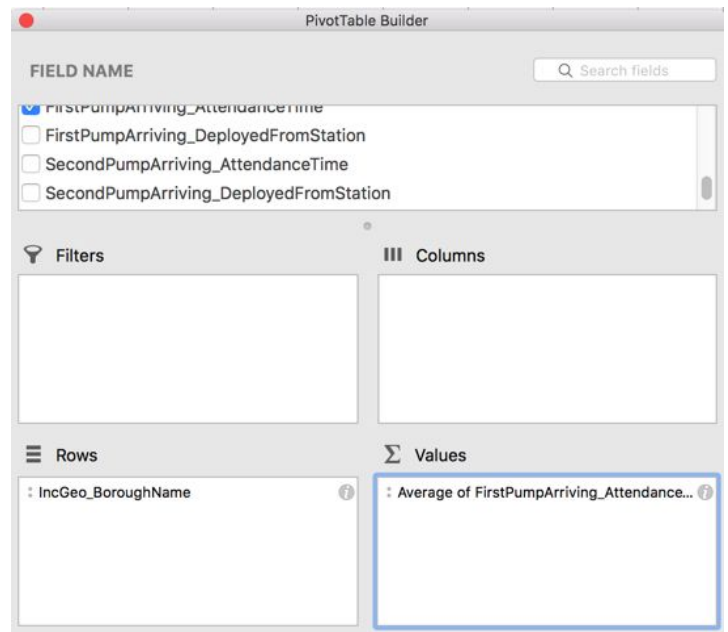
Before you insert a pivot table, select cell A1.

Now find your pivot table insertion tool and click to create a new pivot table. The default options in your pivot table will include all of the data and will create the result in a new sheet. Click OK.

On the next screen (or variant thereof) you will be able to construct your pivot table by dragging the column titles into the correct places.

To create a pivot table we want each row to represent a borough (*IncGeo_BoroughName*) and the values to be the **average** of the attendance time for the first appliance (*FirstPumpArriving_AttendanceTime*). Note that the default function on values is to count them, you can change this by clicking the *i* icon in Excel.

This will create a new pivot table showing the average attendance time in seconds. At this stage we are not going to convert this into minutes.



Step 5 - Comparing performance

In order to compare the performance to that of before the closures your could choose to download all the data prior to 7th January 2014 and follow the same steps as above to calculate the averages from before the closures.

In order to speed up the process the following file can be downloaded that contains the resultant data, per borough, in seconds.

<http://bit.ly/LFAbefore>

Add this to your table as a new column so performance can be compared.

The result may look like the following:

3	Row Labels	Average of FirstPumpArriving_AttendanceTime	Before closures
4	Barking And dagenham	354.0702273	347.5796946
5	Barnet	353.288791	382.4238233
6	Bexley	348.6156102	355.7133082
7	Brent	346.4291832	358.9557595
8	Bromley	360.9261916	383.9837057
9	Camden	299.269524	284.0886178
10	City Of london	303.0353895	290.1475347

There are many ways in which the performance of the fire service can be compared. You could subtract the old average from the new one as shown below.

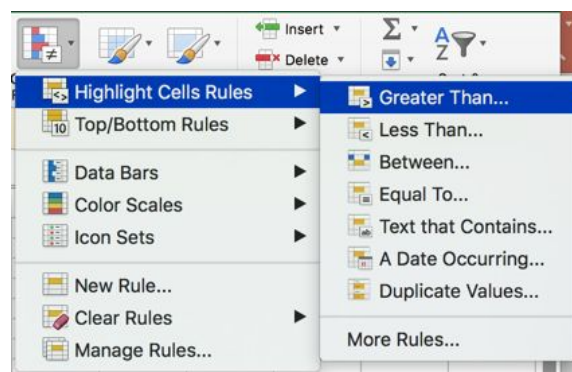
Row Labels	Average of FirstPumpArriving_AttendanceTime	Before closures	
Barking And dagenham	354.0702273	347.5796946	=B4-C4
Barnet	353.288791	382.4238233	
Bexley	348.6156102	355.7133082	

A result greater than 0 represents a degrade in performance.

You can then use conditional formatting to shade cells in red and green based upon the performance. In Excel, conditional formatting is available from the home tab bar. On the differences column you can use the less than and greater than template rules to shade values red or green based upon if they are greater than or less than 0.

The target response time for London is 6 minutes (360 seconds). So you could also choose to compare the performance of each borough to the 6 minute target. Thus you could also evaluate:

1. Which boroughs' performance are now better than the target (after closures)
2. Which are now worse as a result of the closures



Step 6 - Visualising performance

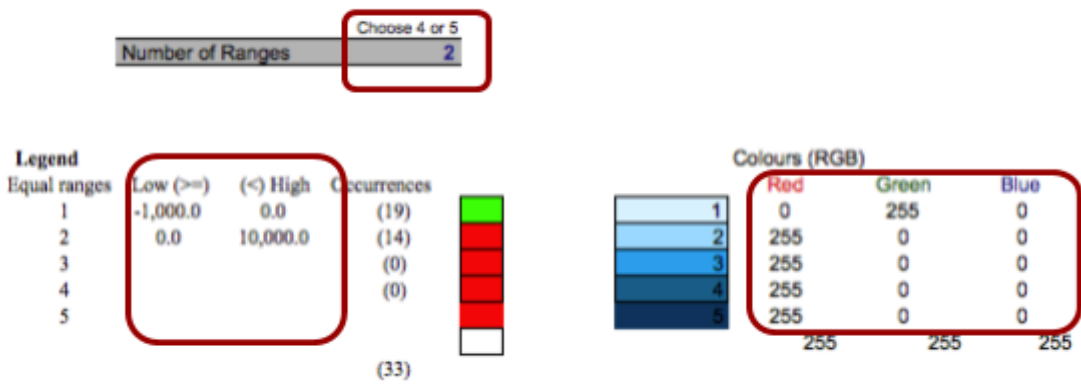
With a simple analysis completed we can now make a simple data visualisation using a great template available from the London Data Store.

<https://data.london.gov.uk/dataset/excel-mapping-template-for-london-boroughs-and-wards>

Download and open the standard borough thematic map, making sure you enable macros on opening.

You can then populate the value column with your time differences from step 5.

Much like conditional formatting you can then update the key on the right to shade the different boroughs based upon performance. In the example below we are simply using a binary scale to shade boroughs green where there has been an improvement and red where there has been a deterioration in service.



Once you have updated all the relevant data, click refresh map to see your output.