

## What about the rest of the country? COVID-19 cases in the UK

*“An analysis is only as good as the data”.*

I hear this sentence very often in the office. Let me begin this post with my latest version of this mantra:

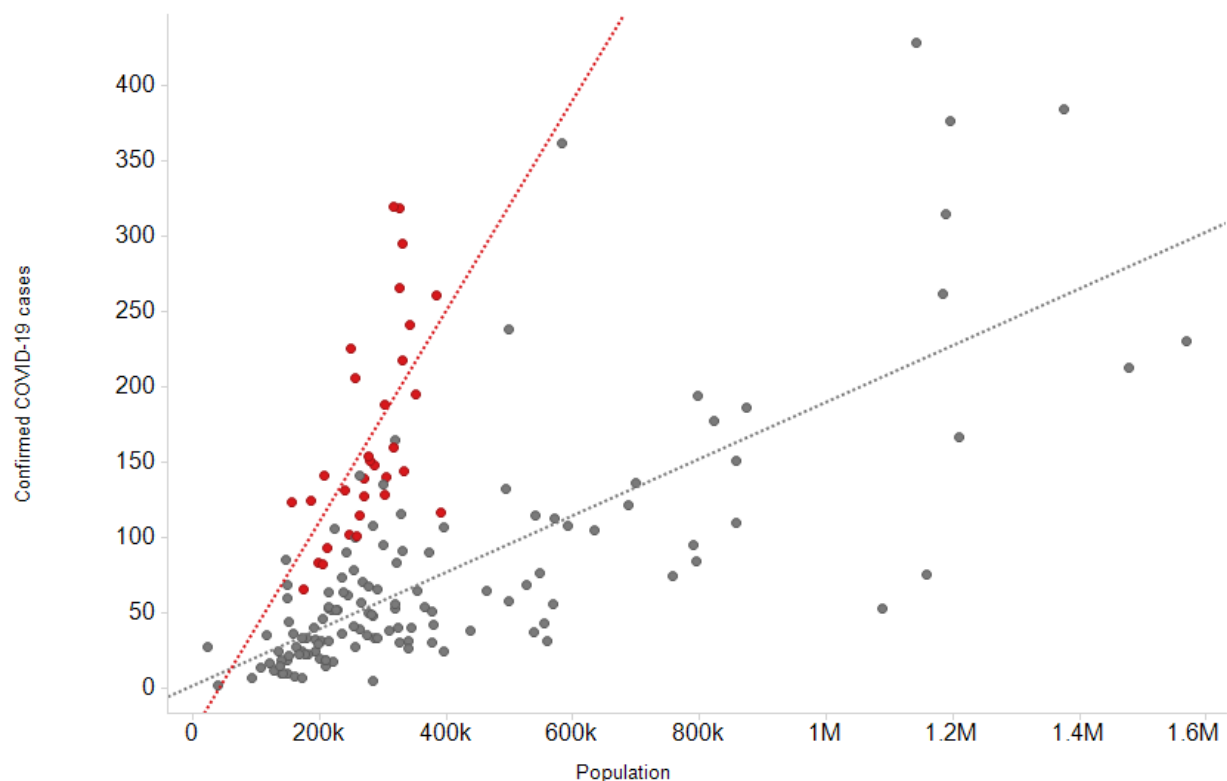
*“An analysis is only as good as the data available”.*

For the UK, the only available corona virus data that can be broken down geographically is the number of confirmed cases. This figure is most probably much smaller than the real number of cases as people experiencing mild symptoms are not being tested. Not to mention people who have no symptoms at all...so let's assume that the numbers we are getting are somewhat proportional to the real number of cases and use this as our starting point (warning: this also assumes that all geographical areas are testing at the same rate...which might or might not be true!).

Ok, now that the disclaimer is out of the way, let's dig in. Last week, I plotted [a chart](#) of the spread of corona virus in London which highlighted how some of the wealthiest boroughs have the most number of cases.

This week, I'm looking at the UK overall. In the following chart, each dot is an UTLA (or district, roughly speaking): the UTLAs in red are in London, the rest of the UK is plotted in grey. As you can see, London is doing pretty badly - even the linear fit (dashed lines) has a much more dramatic gradient!

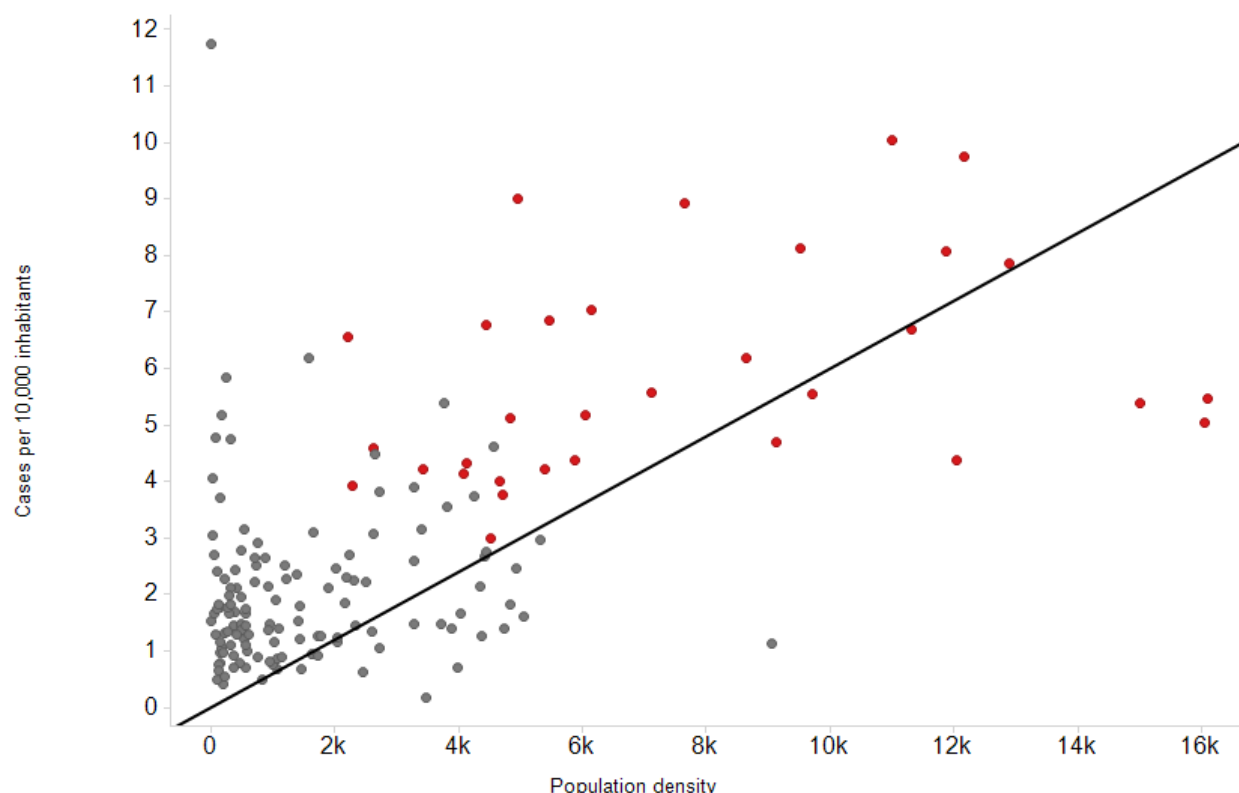
UK UTLAs (Last updated: March 29th)



But there are some areas outside London that seems to be behaving like the capital - some of these are large cities such as Sheffield and Liverpool. This might indicate that population density is perhaps a better parameter for studying the virus spread.

This second graph is showing the number of cases per 10,000 inhabitants on the vertical axis (you can think of this as a "squashed" version of the graph above - now, instead of two variables, we just have one) and population density (inhabitants per squared kilometre) on the horizontal axis. You know the drill, London is still in red...the rest of the UK is still in grey.

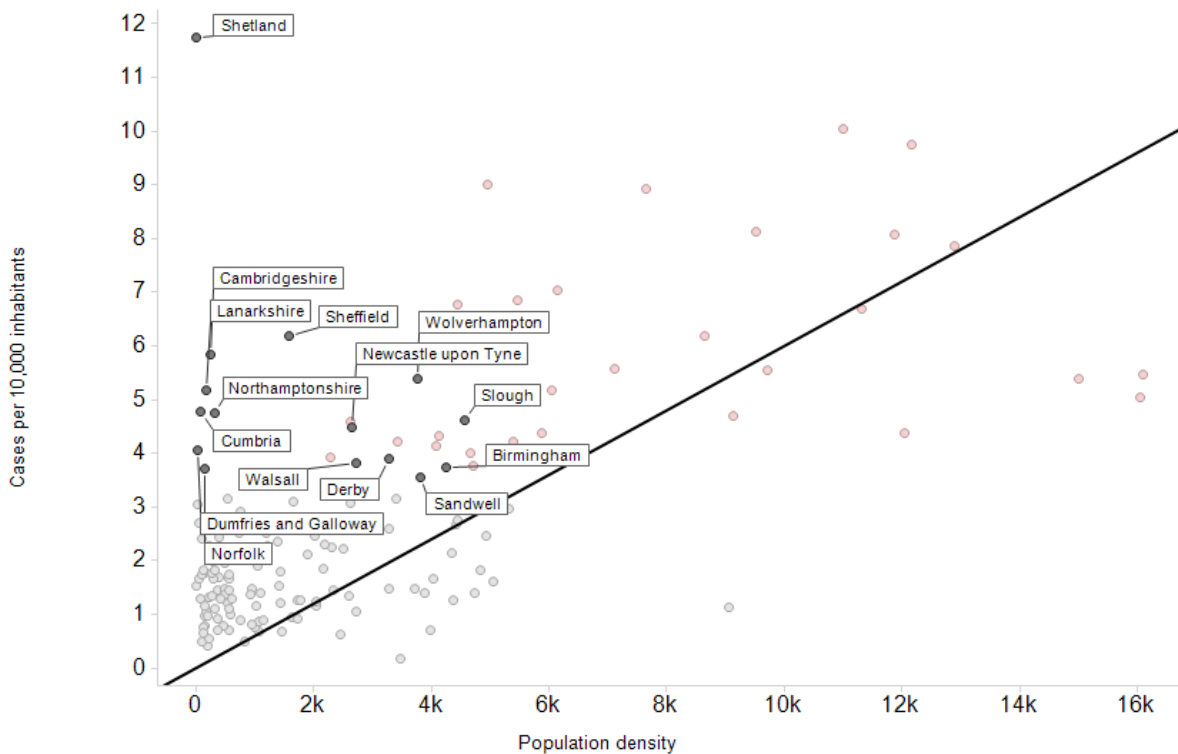
UK UTLAs (Last updated: March 22nd)



Now, this is one of those cases when a linear fit (the black line) is just there to guide the eye. It's not a good fit. But it's a pretty decent indication that population density is indeed somewhat related to the number of cases...which is not an unreasonable claim as the virus spreads proportionally to the number of people met every day by an infected person.

Onto the exceptions now (don't we love exceptions when we're trying to draw a straight line?): this next chart is the same as the one above - I'm just showing the names for some of the areas.

UK UTLAs (Last updated: March 22nd)



What is the deal with Shetland? Well, I checked and re-checked the data to make sure it wasn't some kind of typo. No typos. Part of the explanation could be that Shetland Islands density (15 inhabitants/km<sup>2</sup>) is somewhat misleading as the territory is very vast and the population is concentrated in the towns (Lerwick has a density of about 2,000 inhabitants/km<sup>2</sup>). However, this doesn't really fully explain what's going on: if we move the Shetland dot to 2k inhabitants/km<sup>2</sup> we still don't get it to fit the rest of the population...anyone has any guesses? Are the Shetland Islands testing more?

This is all for this weekend. Over and out.

---

## Data sources

- Number of confirmed cases: [The Guardian](#) publishes every day an extract in tabular form. This tends to cover more areas than the [GOV.UK](#) dashboard.
- Population: mostly [Wikipedia](#). I had to do some manual insert for some of the UTLAs that didn't match up so not all data is from the same year but bear with me.
- Population density: mostly [Wikipedia](#). Again, I had to do some googling. You might be surprised to know that population density can be measured in many ways (inhabitants per KM<sup>2</sup>, per hectare, per square mile...) which is of course a delight to work with.

## Acknowledgements

Credit goes to Irene Stratton and Fran Begley for the helpful discussions and suggestions.

### **Disclaimer**

This disclaimer informs readers that the views, thoughts, and opinions expressed in the text belong solely to the author, and not necessarily to the author's employer, organization, committee or other group or individual.

Olimpia Onelli, March 2020