

PL 9239: Assignment 5 Presentation

Deadline: Tuesday May 12th

The Data

For this last assignment, we will be analyzing real data that I took from the website of the Welsh government (link). The data collection is called WIMD, which stands for Welsh Index of Multiple Deprivation. It contains a series of data that are all measured at the lower layer super output areas, also called LSOA. LSOAs have an average population of 1500 people or 650 households, which means that we end up with a geographically very fine grained view of Wales: geographical units with the same population size, but of course different geographical sizes. Each of the data points we observe represents the average of a variable among this amount of people. As you can see on the home page, the data is published every five years. We work with the latest release from 2019. In general, this data covers a wide range of variables that are all super interesting—so go wild and explore if you want!

For this exercise, we will focus on a small subset of the variables and cover the relationship between educational achievement and distance from secondary schools. We will also add some control variables. Overall, we are covering the following variables:

Outcome Variable

- `adults.no.qualifications.percent` is the percentage of adults aged 25-64 with no qualifications.

Health Related Variables

- `chronic.condition.rate` is the GP recorded patients with a chronic condition per 100.
- `long.term.illness.rate` is the rate of patients with a limiting long term illness per 100.

Socio-Economic Variables

- `income.deprivation` measures the people in income deprivation in percent.

Urbanity

- `urban` is a dummy variable that equals 1 for Cardiff and Swansea, and is 0 for all other parts of Wales.

Travel Time to School

- `public.travel.prim.school` measures the average public return travel time to a primary school (minutes).
- `public.travel.sec.school` measures the average public return travel time to a secondary school (minutes).

Analyse the Relationship Between Share of Adults Without Qualification and Travel Time to Secondary School

This is your research question: *How does the average travel time to the secondary school affect the share of adults without qualification?*

In analyzing the data, you should of course study the relation between these two variables. In addition, make use of statistical control to also account for the influence of confounders, i.e. other variables that are related to the outcome, but might ‘confound’ the actual relation between the distance to the secondary school and educational achievement. The exercise will guide you through the analysis step by step.

First, get an intuition for the data and describe it.

- How are the values spread?
- What are typical values?
- Plotting can always be a good idea.
- Always remember that it can be very helpful to study bivariate relationships, too—in particular those with the outcome variable (share of adults without qualification) and the key explanatory variable we are interested in (travel time to secondary school)!

Then build your analysis step by step with the help of multivariate regression.

- Regress the travel times for primary and secondary school on the share of adults without qualification. What is the result? What do you conclude for the relationship between those variables in substantive terms? What in terms of their statistical significance?
- Add the variable that controls for urbanity. What is the effect for that variable? What happens to the relation between travel time to schools and educational achievement?
- Then also add the socio-economic variable on income deprivation as control. What do you conclude for the socio-economic variable? What do you conclude for the travel time to school in the light of adding this new control?
- Now add the two health variables as control. What do you conclude for the health variables? What do you conclude for the distance variables in the light of adding these controls?
- In sum, what do you ultimately conclude for the relationship between distance from secondary school and the education achievement in an area? What is the story that the data is telling?

Presentation of Your Findings

This time, for the evaluation you will not submit a lab sheet. You will submit a power point presentation that contains your results. Think of this presentation as the story that you want to tell—and try to really tell your findings as an interesting tale on the basis of the data. What is the relationship between educational achievement and distance to secondary schools in Wales? What does the data teach us?

You should be able to give this presentation as a briefing to policymakers. For that, you will rarely have more than 10 minutes, which is why you will submit a PowerPoint presentation that has five slides. Please structure the PowerPoint as follows.

- *Slide #1: Motivation.* Please describe the research question and its relevance.
- *Slide #2: Data Description.* Please describe the data and its most relevant features. The audience typically has heard of the source—the WIMD data—but they need to be introduced to how the values of the data look like. You cannot tell everything there is to know about the values in the data set in one slide, so try to balance on the one hand completeness with on the other hand the focus on the key aspects of your presentation.
- *Slide #3 & #4: Analysis.* Show how you analyze the data to come to your conclusion. You do not have to provide a full account of what you did. It is much more relevant to focus on key aspects for a potential listener: Imagine what a listener needs to know so that she can follow you how you are getting to your conclusion.
- *Slide #5: Conclusion.* What is your main finding? And what policy implications arise from your finding?

Special Prize for the Winning Team: Present your Findings with the Government

As previously mentioned: The winning team will have the opportunity to present the findings with the Welsh Government. So for those who are eager to go all in here, feel free to do so.

Of course, before you will actually present it, I will help you polish things (if necessary) and we will have test-runs of the presentations.

Marking Criteria

In line with existing marking criteria in our Department, the main criterium of marking your presentation is how convincing your argumentation is.

Criteria

- *Technical Evaluation:* How good is your analysis? Does it conclusively analyse the relationship between educational achievement and distance from secondary schools?
- *Presentation:* How well is the power point presented, both in words but also in its visual appeal? Is it easy to follow the main story you telling?
- *Focus:* Is there a good balance between focussing on the necessary facts, while at the same time containing all key information?
- *Argument:* How convincing is your overall argument?

Marks

- 90+: Outstanding work that would be publishable
- 80+: Excellent work with new insights
- 70+: Very good work that shows confident understanding
- 60+: Good work that indicates good argumentation
- 50+: Competent work that may have certain lapses
- 40+: Adequate to pass, but underdeveloped
- 30+: Attempt to engage, but confused and/or inappropriate argumentation
- 20+: Very poor engagement with serious deficiencies
- 10+: Minimal engagement
- 0: No work

Submission

For your submission, please submit your presentation to learning central. Make sure that you include the student IDs of the other students you are collaborating with in your cover sheet—but make sure to **clearly state your own id**.

Lastly: In case there is anything unclear about this assignment, please do get in touch.