Midterm Assessment 2026

POL272 Quantitative Methods for Social Science Research

Dr Sebastian Koehler
2026-02-13

Instructions

- The coursework is due on **10 March 2026 at 13:00**. Please follow all designated SPIR submission guidelines for online submission as detailed on the POL272 QMPlus page. Standard late submission penalties apply. **Please add the latest SPIR coverpage and an AI declaration.**
- This is an individual assessed piece of coursework (worth 50% of your final module mark) for the POL272 module; collaboration and/or discussion of the coursework with anyone is strictly prohibited. The rules for plagiarism apply and any cases of suspected plagiarism of published work or the work of classmates will be taken seriously.
- The course is in the amber category. The regulations of the SPIR AI Policy for the amber category apply to this assessment. You can find the policy on QMplus. *You must add an AI declaration to the submission*. Failure to add the declaration will incur penalties in line with SPIR regulations.
- Along with the coursework itself, the dataset for the coursework can be loaded using:

dataset <- read.csv("https://raw.githubusercontent.com/QMUL-SPIR/Public_files/master/datasets/Covid2020.csv")

Alternatively, you can find the dataset on QMPlus, or on the POL272 website

- Coursework must be submitted via the appropriate link on the assessment tab of the QMPlus module page.
- The format of this submission is a TikTok-style video, alongside a text submission. Please see below for more details on the video instructions.

- The text submission consists of a script of the video (i.e., the "scenes", tables, and words spoken or written in the video in the same order as they appear) and the R code used to produce the results.
- The word count for the written part of the assessment is **2000 words.** This does not include the appendix (i.e., the R code), or any words or numbers contained in plots and tables. You **WILL NOT** be penalised for writing less than 2,000 words. But you **WILL** be penalised if you write more than 2,000 words.
- Please submit your written work in a single document. Create an appendix section at the end which contains all the R code needed to reproduce your results (you do not need to include the code that failed to run, but just the cleaned-up version. Your code has to work when we run it). Failure to include the R code means that the coursework will be marked incomplete.
- You may assume the methods you have used (e.g. descriptive statistics, plots, difference-in-means estimators, linear regression etc) are understood by the tutors and do not need definitions, but you do need to explain why you have used these methods.
- Round all numbers to two digits after the decimal point.

Structure

Your main submission will consist on a 3-minute Tiktok-style video in which you will address the main requirements of the assessment. **Note: there is no need to upload this to TikTok, or for you to have a Tiktok account, the reference is only in relation to the format.** The main parts of the video should be:

1. A short introduction explaining the puzzle and why this is interesting, relevant and important.
2. A brief discussion of the theory, particularly the references presented below. Here you should explain why the variables of interest are likely or unlikely to be associated with one another, and what you might expect to find.
3. The results, where you include your statistics, tables and graphs (see below). This is the section where you interpret your results.
4. A conclusion section where you discuss your findings and propose other potential explanations for countries' performances during the pandemic (again, see below).

Dataset

Covid2020.csv

This dataset is a combination of data from various different sources, including from the data published in the [Our World in Data project] (https://ourworldindata.org) in 2020, the [Quality of Government Institute] (https://www.qogdata.pol.gu.se/) and Wikipedia. The following list of variables is your codebook:

Variable	Description
iso_code	ISO codes for country names.
continent	Continent of the country.
location	Country name.
date	Date of the report for Covid-19 related figures.
total_cases	Total (cumulative) number of confirmed cases of Covid-19.
total_deaths	Total (cumulative) number of Covid-19 related deaths.
total_cases_per_million	Total number of confirmed cases of Covid-19 per million inhabitants.
total_deaths_per_million	Total number of Covid-19 related deaths per million inhabitants.
total_tests	Total cumulative number of Covid-19 tests.
total_tests_per_thousand	Total cumulative number of Covid-19 tests per thousand inhabitants.
tests_per_case	Ratio of Covid-19 tests per confirmed case.
positive_rate	Percentage of Covid-19 tests that gave a positive outcome.
population	Estimated population size.
population_density	Estimated population per \(Km^2\)
median_age	Estimated median age in the country.
aged_65_older	Estimated share of the population who is aged 65 or older.
extreme_poverty	Estimated hare of the population living in extreme poverty.

Variable	Description
life_expectancy	Estimated life expectancy at birth.
human_development_index	UNDP Human Development Index.
vdem_polyarchy	Level of Liberal Democracy, where lower values represent lower levels of liberal democracy and higher values represent higher levels of democracy (V-Dem Institute).
female_HoG	Female Head of Government. It is coded 0 if the HoG is male, 1 if the HoG is female (Wikipedia).

Project

Motivation

Imagine you are a data journalist specialising in political and international issues. You have been hired to produce TikTok videos for a mainstream British newspaper that is looking to appeal to younger audiences. As such, you are requested to produce a 3-minute TikTok-style video analysing the political and social drivers for success/failure in tacking the Covid-19 pandemic in 2020, **particularly in terms of the number of deaths per million inhabitants.** The editor is interested in understanding whether:

1. Countries led by women are more successful than those led by men.
2. Richer countries do better or worse than poorer countries.
3. More democratic countries are more successful than authoritarian countries.

Specific tasks

In order to answer the 3 questions posited by the editor, you should first examine the relevant variables individually (e.g. total_deaths_per_million, female_HoG, gdp_per_capita and vdem_polyarchy). After you do that, you should examine the relationship between the total Covid-19 related deaths per million share and each of the relevant variables that relate to the editor's 3 key areas of interest. In all cases, you should state your justification for the proposed relationship between these two variables (i.e., explain why they are likely or unlikely to be associated with one another, and what you expect this relationship to look like). For some ideas, you

might look at this paper or this paper). For the first issue the editor is interested in (1), concerning whether countries led by women are more successful than those led by men, you should calculate and interpret the difference-in-means estimator and formulate a statement as to what you conclude from this, and what assumptions you are making (and whether these are reasonable).

For the second issue the editor is interested in (2), concerning whether richer countries do better or worse than poorer countries, you should produce a scatterplot (including a line of best fit) and find the correlation coefficient, and interpret your findings. For the third issue the editor is interested in (3), around whether more democratic countries are more successful than more authoritarian countries when it comes to Covid-19, you should answer this question by running a linear regression model and interpreting the results of this. Finally, you should use the difference-in-means estimator, your scatter plot (including the line of best fit), your interpretation of the correlation coefficient relating to this and the results from your regression model in your discussion in the video. You should explain whether there is evidence to support your theories and how they can be improved. Please include all of the cleaned R code used to conduct the analysis activities set out here in the appendix of your text submission, but do not include this in your Tiktok-style video. You need only include the results in this video (e.g., the difference-in-means estimator, the scatterplot (with line of best fit), the correlation coefficient and regression results).

Useful examples

Here are some useful examples of TikTok videos that use empirical information and/or quantitative methods to discuss political and social questions. You are welcome to use them as a reference for the style and organisation. You are also welcome to use any other format you consider useful:

1. https://www.tiktok.com/@charlottechaze/video/7250998634630679851?lang=en
2. https://www.tiktok.com/@arghavansallesmd-phd/video/7328865157222845742?lang=en
3. https://www.tiktok.com/@econwithsarah/video/7299568383475879211?lang=en
4. https://www.tiktok.com/@stumbling.through.oxford/video/7330948933897817376?lang=en

5. https://www.tiktok.com/@learningcast/video/7247854060500421934?lang=en
6. https://www.tiktok.com/@learningcast/video/7167010863784693035?lang=en
7. https://www.tiktok.com/@tldrnews/video/7283142663723896096?lang=en
8. https://www.tiktok.com/@matthew.rein/video/7063301422242139439?lang=en
9. https://www.tiktok.com/@susiemicolta/video/7280205738557852961?lang=en
10. https://www.tiktok.com/@verydata/video/7020761056066686213
11. https://www.tiktok.com/@bylinetimes/video/7139240008996769029
12. https://www.tiktok.com/@guardian/video/7153667120675441925
13. https://www.tiktok.com/@guardian/video/7150594136708582662
14. https://www.tiktok.com/@yahoouk/video/7155852506071928069
15. https://www.tiktok.com/@yahoouk/video/7145205780948602118