|  |
| --- |
| Tutorial 12(Week starting on 6-jun-2022) |

|  |
| --- |
| **Objectives**   * Learning how to present data properly * Working with geo-referenced data   **Working materials**   * **Schwabish, J. A. [2014]**. An economist's guide to visualizing data. *Journal of Economic Perspectives*, *28*(1), 209-34. * **Tufte, E. R. [1997]**. “Visual explanations: images and quantities, evidence and narrative” *Ch 2* Ed. Graphics Press. * **QGIS** * You will need **database “**Datos ejercitación 12” which can be found on the virtual campus. |

**Exercise guide**

The exercises marked with an asterisk (\*) are compulsory and must be submitted by email **before 12.00 on Wednesday 15th June**. [[1]](#footnote-1) To the mark obtained on the evaluation of this assignment,**0.05 points will be subtracted for each minute late**. Consult the course program with regards to the formalities of the presentation.

1. (\*) Explain briefly the cases of the discovery of the transmission of cholera and the launch of the Challenger space shuttle. According to Tufte [1997], in what lies the success and failure of each of the cases?
2. (\*) Write the model that John Snow had in mind when he proposed his hypothesis. Recall that the simplest model is , where is the explanatory/independent variable, is the explained/dependent variable, and is the error term. How would you interpret in this model?
3. (\*) Explain in no more than one paragraph why this model might have endogeneity problems. Do not give a general answer by naming the theory, but apply what you know to this specific case.
4. What are the principles of "graphical excellence" by Tufte [2007]? What are the three basic principles presented by Schwabish [2014] for graphical representation of data?
5. (\*) Take two graphs from any two papers seen in the lectures or tutorials and propose possible improvements, following Tufte's and Schwabish's principles. Include the original graph in your answer, but you do not need to make an "improved" one. Recommendations should be textual.

**Exercises in QGIS**

1. Add the image "OSMap\_Grayscale.tif" via Layer / Add Layer /Add Raster Layer. Then add the information about "deaths" and about "water pumps". To do this go to Layer / Add Layer / Add Vector Layer. This will take you to a new window. Under "Source": find the file "Pumps" and "Cholera\_Deaths". Expand the markers corresponding to the "Pumps". Change the format of the markers corresponding to the deaths. What can we observe at a glance? Try saving these changes as a single image.
2. Create a heat map for the cholera deaths. Set 50 meters as Radius. Modify the colors so that your heat map is more intuitive. To do this go to "Processing" -> "Toolbox" -> "Interpolation" -> "Heatmap". Improve the style of the graphic in "Properties" by right clicking on the "Heatmap" layer. Choose "Singleband pseudocolor".
3. Finally, let's see what the current location of the deaths in London today would look like. To do this, go to "Pluggins" -> "Manage and install pluggins" -> find and install "QuickMapServices". A world icon will be added. Click on the icon and look for the OSM option. Choose OSM Standard.

**Some useful links:**

* <https://mappinggis.com/2018/03/como-crear-mapas-de-calor-en-qgis-3/>
* <https://mappinggis.com/2016/09/plugin-quickmapservices-capas-base-de-google-landsat-openstreetmap-para-qgis/>

1. In the case of cholera, it was a disease believed to be transmitted through the air instead of through water. Proving reality was successful thanks to John Snow plotting all cases in the city geographically and highlighting their closeness to infected water pump. Its success was due to it clearly suggesting the desired cause and effect relation and by including counterfactual proof by showing those near uninfected pumps weren’t contaminated. In the case of the space shuttle, it was launched in temperatures too low to be safe, but which fact wasn’t effectively communicated graphically. The failure lay in displaying too much information and not doing so in such an order that the relation between low temperatures and accidents was emphasized.
2. We though the model John Snow probably had in mind was something in the region of: the chance of getting infected by cholera as a function of one’s distance to an infected water pump. In this case, x would be the vector distance to the pump and the explained variable

Endogeneity means there might be an omitted variable affecting both relevant terms. In this case, we though age could be a confounder since, according to the national institute for rare diseases children are “Children are more susceptible to cholera than adults” and there was a school near the pump, so being school-aged would explain proximity too. This might be addressed using a dummy variable controlling for age.

There probably are other more relevant endogeneity problems which we couldn’t think of.

2. This looks pretty cluttered, could have just been points

Gráfico, Gráfico de dispersión

Descripción generada automáticamente

There really is no point in having all the text of the country code if it cant even be read because its superimposed, maybe highlight the relevant ones. (Acemoglu, colonial origins)



1. Check the course’s syllabus on presentation [↑](#footnote-ref-1)