Workflow COVID19 project

Glenn Magerman

Version: March 30, 2020

Abstract

This document describes the workflow for our umbrella COVID 19 project. It first describes the idea and focus of the project. It then describes the general workflow and tools.

1 General structure

The umbrella project is called "Learning from the curve". There is a website and a GitHub page. It is a collaboration that started between several researchers at ECARES, ULB and we invite other contributors to collaborate.

Learning from the curve focuses on the economic aspects of COVID. In particular, we want to provide:

1. **Context:** regularly updated numbers on various aspects of the outbreak (disease, economics, social related).
2. **Insights:** understand the possible impact of COVID on the economic tissue of countries, regions and demographics.
3. **Expertise:** contribute as economists in several fields (health economics, econometrics, macro, political economy, networks, . . .)

Learning from the curve is an umbrella for different projects. We invite researchers to contribute to projects in the following ways:

1. **Short run:** Write individual or collaborative blog posts related to COVID and economics. These can be one-shot or regularly updated. Writing style is punchy, not academics.
2. **Medium run:** Collaborate on structural content of the umbrella project. This can be in the form of providing new data, adding features to existing projects, new correlations, new variables etc.
3. **Long run:** Collect information and knowledge to set up long term academic research projects with new data and possibly new collaborations.

2 Workflow

Note: The following notes describe our general pipeline. Please do not let this hold you back to produce output and collaborate. If you want to submit output in the form of an article only, this is absolutely perfect.

For people that want to share their code/data and collaborate, here is the work flow:

1. **Create a project:**

   (a) Find a set of people to collaborate or work individually on a topic that crosses COVID with economics.

   (b) Ask Fabrizio Leone to join the Learning from the curve organization on GitHub so you can create a repo for your project. Each project is a separate repo.

2. **Collect data:**

   (a) Several datasets on the topic are made available here, including global COVID case data, population statistics, and national accounts for EU28. More datasets will be added as we proceed. Links to other potential datasets are available [HERE].

   (b) Run the script to get the latest data, and collect other data yourself. Save the data locally to work on with your team.

   (c) If you use some data that is not on the repo, and that can be of interest to other projects as well, ask Fabrizio Leone to add them to the GitHub repo. You can also contribute by cloning the repo, adding datasets/features, and pushing it to GitHub again.

   (d) This repo is maintained by Fabrizio Leone. Please contact him if you want to contribute or have questions.

3. **Code and analysis:**

   (a) Code up your analysis, possibly in Python (to keep our learning curve for the Python-do-group). But let the framework not hinder your workflow. Feel free to work in Stata, R, or anything that keeps your pace high.

4. **Share code and output:**

   (a) Push your code to GitHub once you are happy and have a stable release.

   (b) Any potential collaborator can contribute to any released project by opening issues to ask for additional features, functionalities, or clone & push the repo to contribute directly.

   (c) Create a GitHub project on your project for task management: next steps to do, what is done, etc. This also invites other people to collaborate on your project and keep moving when stuck. There is also a project for the website itself.

5. **Publish article/blog:**

   (a) When you are done, open an issue under the project repo to Glenn Magerman and Federico Gallina. We will proofread and format for "publication".

   (b) For articles, you can send the output in Markdown. Writing in Markdown is surprisingly easy: iPython notebooks, Python IDEs, Stata or R.

   (c) If you are most comfortable with Word or LateX, feel free to send the output in those formats. We will translate these to Markdown for release on the site.

   (d) All blogs/articles will be published on website. People can subscribe on the RSS feed, and we seek to reach out to organizations/outlets to take us up too.