

# SI100B Bonus Assignment (Fall, 2021)

## Before you start...

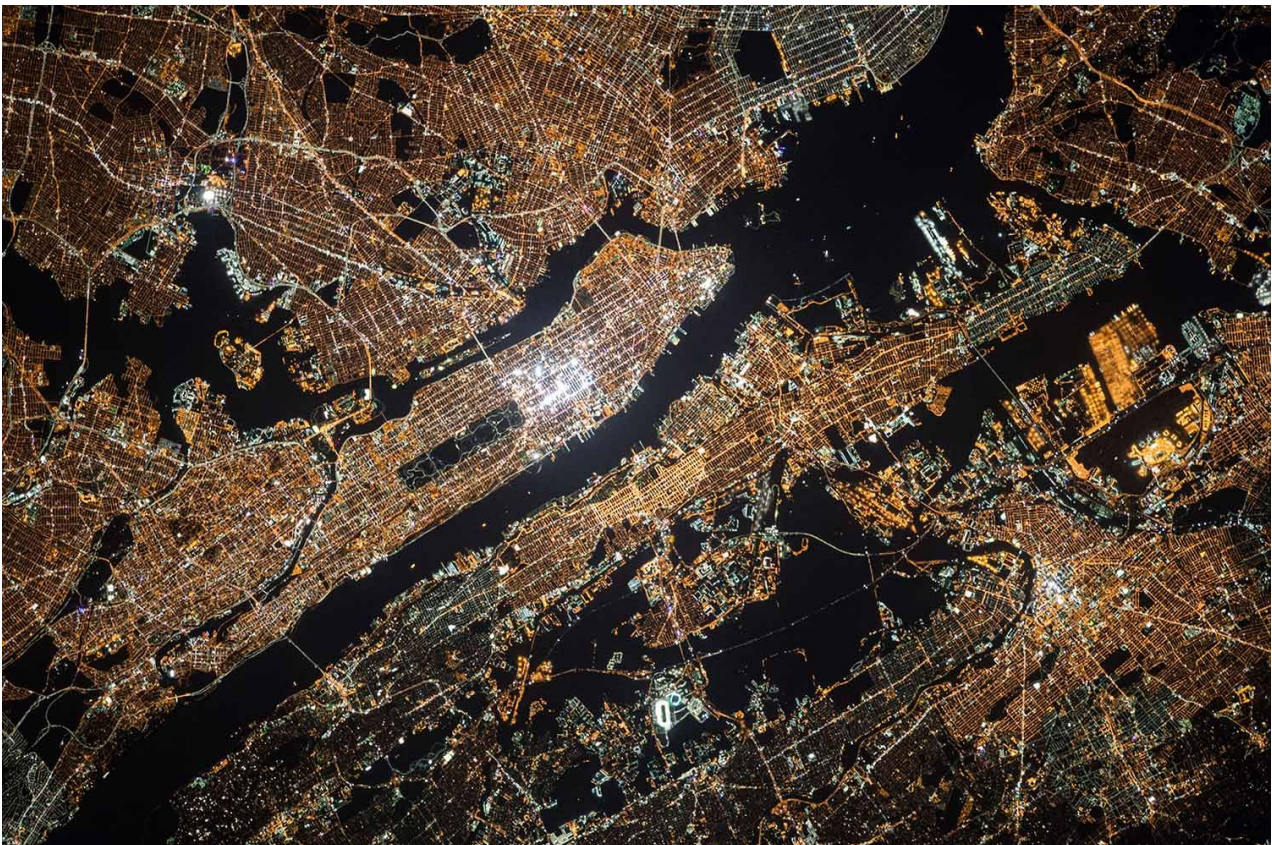
The bonus assignment is not for every student, and **it's OK for you to skip this work safely**. To get into it, you are supposed to have the following accomplishments beforehand:

- Rich experience in Python programming.
- Basic front-end knowledge, like javascript (for web crawler in Python).

We suggest that if you are innocent in programming before this course, you may not have the adequate ability for this assignment. However, for those who are with a related basis, you are encouraged to have a try. Working as a team of 1-3 students is allowed. At the end of the semester, we will evaluate and grade your results.

Besides, it is quite an open assignment, which means **you will be highly free to develop it with your ideas**. You can always feel free to add features and change the presentation using your creativity. By the way, this task is meaningful for the current time, and your program may even work in real life to help people in need. Ready for it? Let's get started then.

---

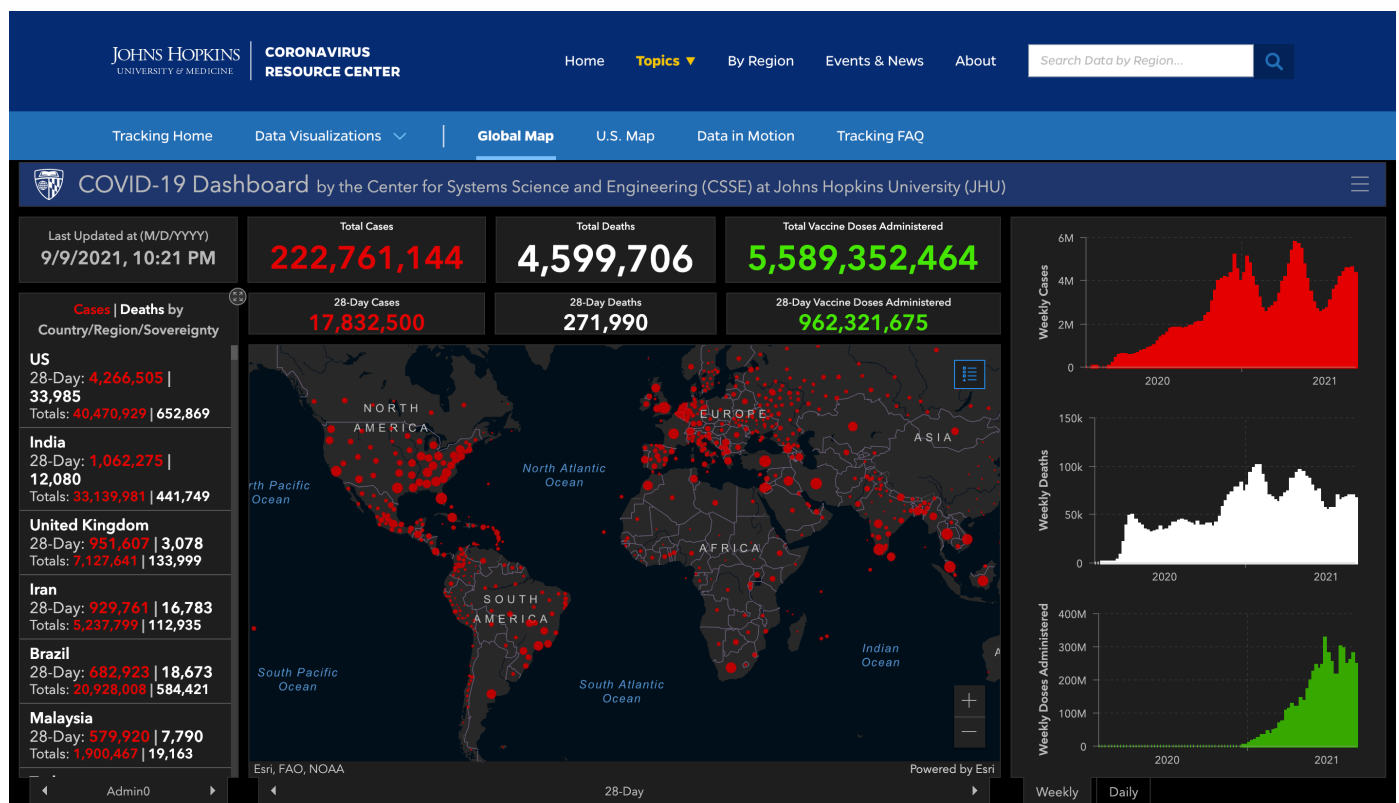


# COVID-19 Dashboard

The [COVID-19 pandemic](#), also known as the coronavirus pandemic, is an ongoing global pandemic of coronavirus disease 2019 (COVID-19) caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The novel virus was first identified in Wuhan, China, in December 2019, and it spread to other parts of mainland China and around the world. The World Health Organization (WHO) declared a Public Health Emergency of International Concern on 30 January 2020, and a pandemic on 11 March 2020. Since 2021, variants of the virus have emerged and become dominant in many countries, with the Delta, Alpha and Beta variants being the most virulent. As of 9 September 2021, more than 222 million cases and 4.59 million deaths have been confirmed, making it one of the deadliest pandemics in history. [1]

The COVID-19 pandemic has had a severe collision on humans in the past period. With the vaccination and drug development, we are back to normal comparatively, despite uncertainty brought by the spread of variants of the virus. To be sure, its impact on us will continue for a long time still.

To increase transparency and help people stay informed with the latest status, international organizations, governments, and institutions have established various information release platforms. Among the well-known sources are [WHO](#), [Johns Hopkins University](#), and [CDCs](#), etc. The function layouts of these websites do not look the same, but most of them contain a **dashboard**. You can see statistical charts and maps in this dashboard. The statistical charts reflect the trend of the epidemic index over time, and the map works as another visualization to show them over regions.



COVID-19 Dashboard by the CSSE at Johns Hopkins University

All you need to do is build a dashboard like those and meet the following requirements:

- Your program should **crawl** the latest data about COVID-19 on the Internet and use it to do **further analysis**.
- Your data should have reliable and publicly open sources (for example, JHU).
- You should use cube data (i.e., multidimensional data), which may contain: population density, age pyramid, vaccination status, protection rate, virus transmission, variants, new confirmed, number of cures,

hospitalizations, mortality, government policies, etc. You are supposed to decide which indices to choose.

- You can choose your target as the whole world or a country (for the richness and transparency of data, we recommend the United States). If you choose the world as your target, your largest basic unit is each country (may not include all countries); Or if you prefer to do with a country, your model should be accurate to the state or province.
- The **data visualization** module. It includes at least the two modules mentioned above: **epidemic map** and **statistical charts**. You are free to design and add new elements for sure.
- **Prediction module**. Use your data to predict the trend in the epidemic situation in various regions. You should at least give the epidemic trend of a week, and the performance should be better than the **moving average** predictor. You can freely use formulas in epidemics; Or use machine learning methods by **scikit-learn** and **PyTorch** to build the predictor. Your results should also show on the visualization modules.
- Use python as the principal development language. If you have a wealth of front-end knowledge, it's a good idea to present your program as a web page or application (desktop or mobile), which will improve your score in this project. If not, you may use platforms such as **Jupyter Notebook** to make it work manually.
- Any **python library** is allowed to use at will. Also, you can freely **use any code** on the Internet, as long as specifying the reference in standardized manners and paying attention to **license** issues. However, including your original ideas is required. **Originality and innovation** will significantly affect your final grades. If your program is almost the same as on the Internet, your score will be invalid.

Your final score contains two parts, **70% for the primary score and 30% for the evaluation score**; If your program meets all the above requirements, you will get full marks in the first item. The second item depends on the aesthetics of your **UI, usability, prediction accuracy, feature-richness**, etc.

You can spend the whole semester working on this assignment. Notice that each team carries 1-3 students, and each person should have a **clear division of work with an equally contributed load**. We will verify the contribution from each of you in the final grading check. We regard **sharing code or pseudo-code between teams as academic misconduct**, but bouncing ideas is allowed. We encourage you to share ideas and materials publicly on **Piazza**, which can benefit more students and reduce the risk of unconsciously violating the rules of academic integrity. Before the end of the semester, **do not disclose the source code** in any way. If you use online code hosting services (such as **GitHub**), temporarily set the repository as private and only authorize your team members to access it. Please do not break the above rules. Otherwise, your score will be invalid after all, and there might be other serious consequences.

Many great inventions in the world were born in course projects at college. We sincerely hope to see your proud achievements. Please give full play to your learning ability and creativity. Enjoy it!

## Reference

[1] [https://en.wikipedia.org/wiki/COVID-19\\_pandemic](https://en.wikipedia.org/wiki/COVID-19_pandemic)

[2] <https://covid19.who.int/>

[3] <https://coronavirus.jhu.edu/map.html>

[4] <https://covid.cdc.gov/covid-data-tracker>

---

Author: Yucong Chen, please contact [chenyc@shanghaitech.edu.cn](mailto:chenyc@shanghaitech.edu.cn) for any questions.