Instructions

**For Data Science Quest I**

# Resources

* [GitHub Repo](https://github.com/minghao51/iris_similar) (where all the scripts for the app, python, Dockerfile, Slides, Docs can be found)
* [Google Colab](https://colab.research.google.com/drive/1n3ghm1LyYxwMFhtdcpi8lEa2yOnPGmOX?usp=sharing) (used for the development of python scripts, require google chrome)
* [Deployed Web App](https://irissimilar-jdc7q5tvtq-as.a.run.app/) (the web is deployed through via docker on GCP cloud run)

# Context

Please refer to the files within Slides and Docs folder for information on the background.

# Prerequisite

* [Docker](https://docs.docker.com/desktop/)
* Linux/PowerShell (for commands)
* [Git](https://gitforwindows.org/)
* Chrome/Browser

# To build and run the codes

Through PowerShell/Git/Docker

1. **Extract all the necessary files.***git clone* [*https://github.com/minghao51/iris\_similar.git*](https://github.com/minghao51/iris_similar.git)  
   One can also just extract the zip file (attachment in email), download from google drive or [Github Repo](https://github.com/minghao51/iris_similar.git).
2. **Navigate to the extracted folder and compile the docker***docker build --tag iris-app:1.1 .*
3. **Docker Run***docker run --rm -p 8050:8050 -it iris-app:1.1*
4. **On a local browser (preferably chrome), insert the URL**[*http://localhost:8050/*](http://localhost:8050/)  
   The dash app website formatting may be off on other browsers.

**To trigger the app/run data quality check manually within docker**

*docker run --rm -p 8050:8050 -it --entrypoint /bin/sh iris-app:1.1*

Web app  
*python app.py*

Data quality check  
*python test\_iris\_similar.py  
cat data\_quality.output*

# Open-sourced tools used

* Dash (for the react based web components, and Plotly charts)
* Python
* Docker (To construct the environment, deployment)