Level 2: Full C3 AI COVID-19 Data Lake Access

Full access to the C3 AI Convid-19 Data Lake offers all stored data while allowing the researcher to use whatever analysis framework they choose with their own compute resources. This level offers the fastest startup time while ensuring access to all data. Once you learn how to query data in C3 AI Suite, that data can be streamed to your compute resources where you can use your language and tools of choice.

Pros:

- Full access to C3 AI Convid-19 Data Lake.
- Ability to use any language and tools of your choosing.

Cons:

- C3 AI environment access required.
- All data must be streamed to your local compute resources.

Getting Started

Please see the DTI Readiness checklist. If you pass the checklist, you're ready to start learning!

Training Curriculum

Once you have completed the DTI Readiness checklist, you're ready to start learning.

These hands on resources give you an in-depth guide on how to start using the C3 AI Suite immediately followed by explicit examples showing how to use C3 AI documentation, Types, Timeseries, and Metrics. Each of these examples is grounded in the C3 AI Covid-19 Data Lake. Additionally, links to appropriate official C3 AI documentation and videos are provided throughout the quickstart guide.

- DTI C3 AI Suite Quickstart Guide
  - This guide gives a quick hands-on introduction to how researchers can fetch and analyze data from the C3 AI Covid-19 Data Lake.
- Example Workflows
  - Cases example (C3 connector)
  - Vaccine Listing example (C3 connector)
  - Epidemiology example (C3 connector)
  - Phylogenetic Tree (C3 connector)

These C3 AI created materials give more information:

- C3 AI Getting Started Guide
  - If you encounter any authentication problem, go to https://developer.c3.ai/ to login your C3 AI account.
  - Not all sections of the **Getting Started Guide** are necessary. The recommended sections are:
    - Introduction to C3 AI Suite
    - Architecture
      - Basics
      - Understanding Clusters, Tenants, and Tags
      - Model-driven Architecture
        - Types
        - Fields and Methods
        - Type Inheritance
    - Time Series Data Section
    - Metrics Section

Additional Resources

Some users may find these additional resources helpful at this stage:

- c3python module for easy python interactivity
- DTI Guide: Utilizing C3 AI Documentation

Challenge Problems

This set of problems exercises the concepts defined under the 'Level 2' header. You should be able to solve these problems by fetching types, executing existing Metrics, and creating new Metrics. These may require processing outside of the C3 AI platform, for instance to produce plots.

- What is the age distribution of coronavirus patients? Does this vary by location?
- How does likelihood of death change with age and sex?
- Do coronavirus case reporters agree with each other on how many new cases a given US county or state have? How often do they disagree?
- Is there a seasonality to case reporting data? What period is it?