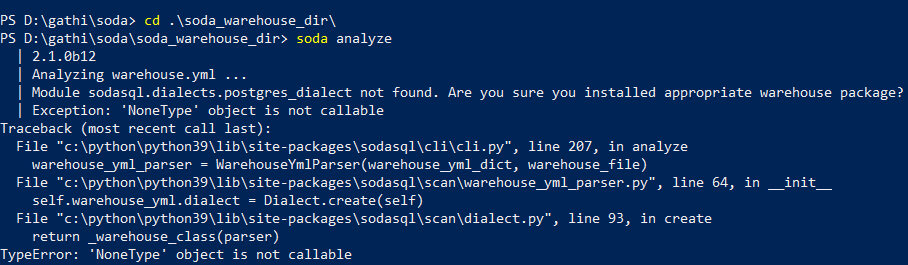
Soda First Take

Programmatic access: Soda SQL

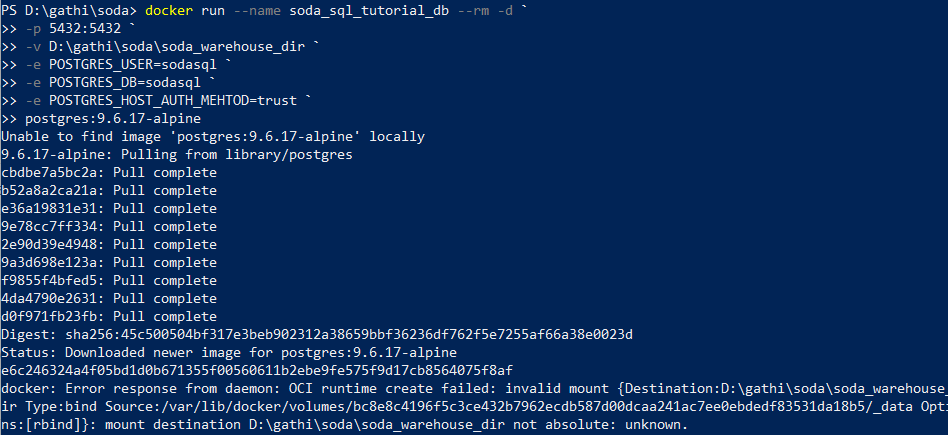
Web-app access: Soda Cloud

Soda SQL:

1. Connect data source
   1. Seems to be only able to connect with database (not a single file)
   2. Users need to edit warehouse.yml and env\_vars.yml manually (one is database access details like port, one is credentials)
      1. Different yaml files for different schema databases
2. Must do a first scan (similar to our data registration)
   1. Scan yaml needs to define table, metrics to be used, test to be operated or optionally the columns in the table that will be tested
      1. Different scan yaml files for different datasets within the same database
3. Then we can apply default metrics or user-defined metrics to columns. (user-defined metrics are more like SQL queries). It also needs to submit a scan yaml file to do the jobs.

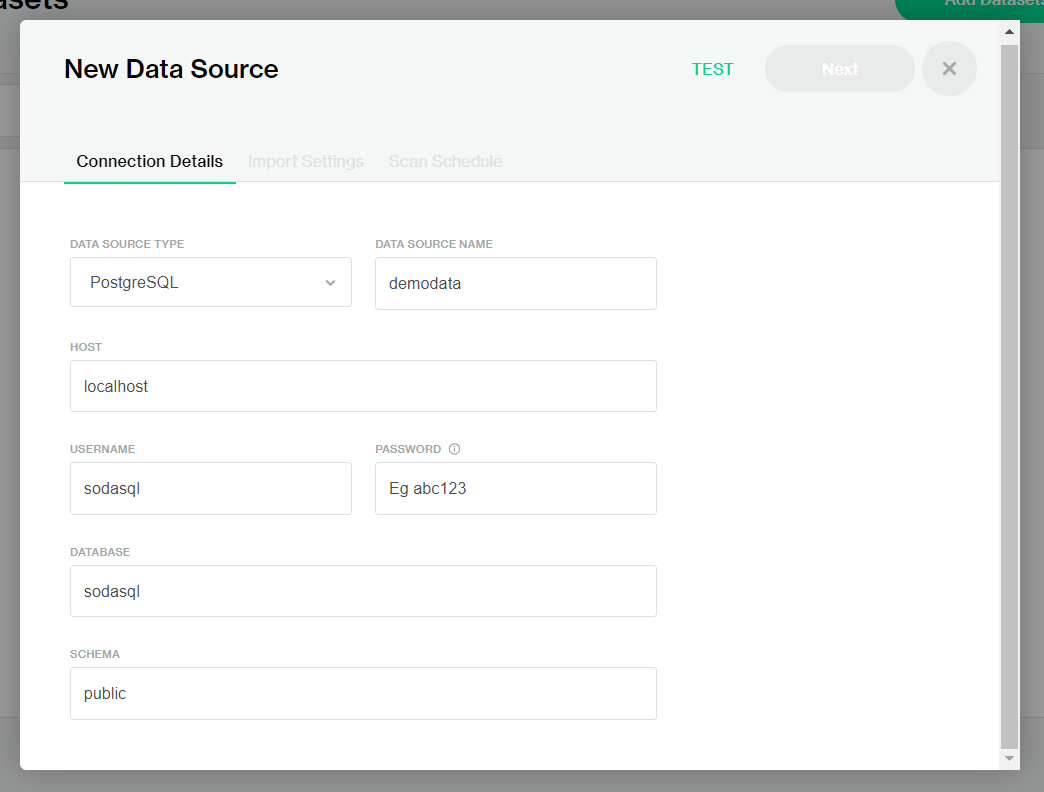
I tried to followed the instructions to use soda SQL, but I ran into errors. I got a couple warnings when I did “pip install soda-sql-postgresql” but it still went through. When I tried to use “soda create” to create the data warehouse yaml file. It failed. So was the env\_var yaml file. Therefore, I manually created them. Then run “soda analyze” but it still couldn’t go through. I think it has to do with the yaml file parser. 

After failed to use my own dataset. I tried to use their tutorial dataset which was containerized. And it still wouldn’t work for me. No containers were running with no errors.





Soda Cloud:

1. They have a clean UI. However, I was not able to connect to their tutorial database neither.
2. 
3. According to the documentation, users can edit data source, edit tests, and define metrics when creating new monitors.

Metrics:

I can find their files and view them in Pycharm or Notepad++ but I didn’t understand them. The hierarchical architecture about different modules is a bit complex for me. It would take me a while to understand them.

