# Project Module – 3: NoSQL Cassandra Data Modeling

Port the Web Service API from Project 1 to [ScyllaDB](https://www.scylladb.com/), a wide-column NoSQL database compatible with [Cassandra](http://cassandra.apache.org/).

### Test Environment

#### Installing and Configuring Docker

The easiest way to install and run ScyllaDB is using [Docker](https://www.scylladb.com/download/#docker). Use the following commands to install and configure Docker on the Tuffix VM:

$ **sudo apt install --yes docker.io**

$ **sudo usermod -aG docker $USER**

Log out and back into the VM to pick up changes before running the docker command below.

#### Installing ScyllaDB

To start a single instance of ScyllaDB, use the following command (entered all on one line):

$ **docker run --name scylla -d scylladb/scylla --smp 1 --memory 1G --overprovisioned 1 --developer-mode 1 --experimental 1**

Wait a few moments, then check that ScyllaDB is up with

$ **docker exec -it scylla nodetool status**

If this command fails, you can check for errors in the ScyllaDB logs by running

$ **docker logs scylla**

##### Memory Allocation

Note that the --memory 1G switch is a minimum value for starting ScyllaDB on a Tuffix VM with 3G of RAM. If you run ScyllaDB on another platform, you may need to adjust that figure. You may also wish to increase it if your host has more RAM available.

Leaving off the --memory switch entirely is not recommended unless you are on hardware dedicated exclusively to running ScyllaDB.

#### Managing ScyllaDB

Once ScyllaDB is up, you can execute CQL commands using

$ **docker exec -it scylla cqlsh**

If you need to stop ScyllaDB, use

$ **docker stop scylla**

and restart with

$ **docker start scylla**

You can remove ScyllaDB completely and start over with

$ **docker rm -f scylla && docker rmi scylladb/scylla**

#### Python Cassandra Driver

Install the DataStax [Python Cassandra Driver](http://datastax.github.io/python-driver/), either by running

$ **sudo apt install --yes python3-cassandra**

or following the [instructions](http://datastax.github.io/python-driver/installation.html) provided.

Verify that you can connect to ScyllaDB from Python by [creating a Cluster and retrieving a Session](http://datastax.github.io/python-driver/getting_started.html). Use the IP address shown by the nodetool command you ran when installing ScyllaDB.

#### Cassandra Flask Extension

You may wish to install and use the [Flask-Cassandra](https://github.com/TerbiumLabs/flask-cassandra) extension to configure your Cassandra connection, but this is not a requirement.

### Cassandra Data Model

Create a single keyspace and at least one column family (table). You may wish to create a separate column family to store login information. Keep in mind that Cassandra supports [collections](https://cassandra.apache.org/doc/latest/cql/types.html#collections) as a column type.

Don’t forget the advice from Chapter 8 of the textbook: identify the access patterns for each API call, and create additional indexes for any queries not based on primary key. See [Cassandra Succinctly](https://www.syncfusion.com/ebooks/cassandra) for more information and examples.

### Database Population

You will need to update your init\_db Flask CLI command to connect to Cassandra and use CQL.