**Cassandra**

Cassandra is defined as an open-source NoSQL data storage system that leverages a distributed architecture to enable high availability, scalability, and reliability.

Apache Cassandra, a distributed database management system, is built to manage a large amount of data over several cloud data centers.

Go through the below training:

<https://dxc.udemy.com/course/learn-cassandra-from-scratch/>

**Keyspace**:

A keyspace is a data container in Cassandra, like a database in relational database management systems (RDMBS). A cluster has one keyspace per application, as many as needed, depending on requirements and system usage. Keyspaces are entirely separate entities, and the data they contain is unrelated to each other.

**Replication Strategy:**

When defining a keyspace, the replication strategy specifies the nodes where replicas will be placed. By using multiple nodes to place replicas, you achieve fault tolerance, high availability, and reliability.

There are two possible strategies:

**Simple Strategy**: - Use this strategy for test and development environments, and if you do not intend to deploy a cluster to more than one data center. The replication factor applies to the whole cluster. The partitioner decides where to put the first replica on a node. Then, other replicas are distributed clockwise on the next nodes irrespective of data center or location.

**Network Topology Strategy**: - This strategy is suitable when you need to deploy your cluster to multiple data centers. However, you can use it even with a single data center so you can expand later. Network Topology Strategy works for both production and development. It tends to place replicas on nodes that are not in the same rack to avoid issues when one rack goes down. Each data center can have a separate replication factor by using this option.

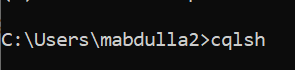
**Start the Cassandra Server**:

Open command prompt and execute the batch file “Cassandra.bat”. (Before starting the server setup the system environment for Cassandra server).



**Open the CQL prompt**:

Open command prompt and execute the command cqlsh.



**Create the keyspace and tables**.

**Create Keyspace**:

create keyspace firstkeyspace with replication = {'class': 'SimpleStrategy', ’replication\_factor’: '1' };

**Create table**:

create table employee(

emp\_id int,

emp\_name text,

emp\_age int,

PRIMARY KEY((emp\_id)));

**Index**:

An index provides a means to access data in Cassandra using attributes other than the partition key

**Create index**:

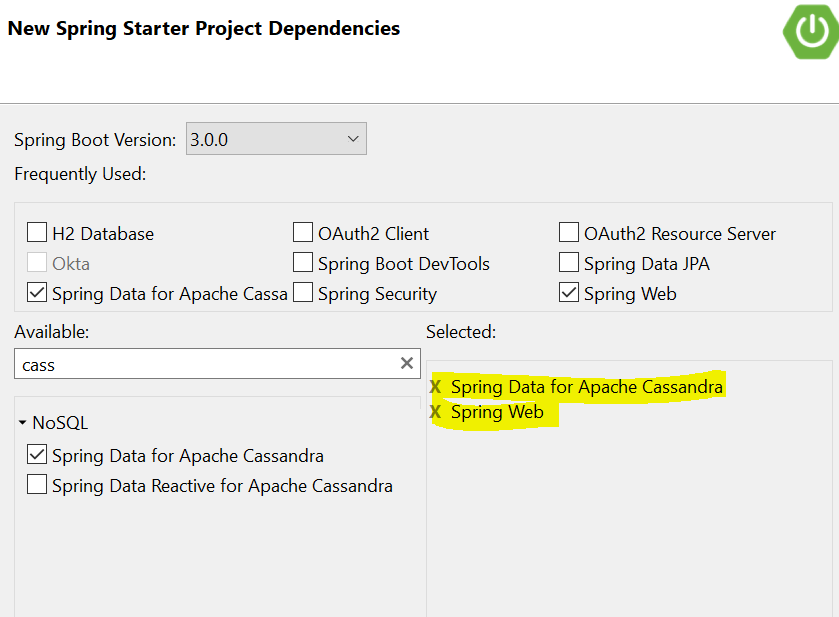
CREATE INDEX index\_name ON table\_name (entries (<map column name>));//for Map

CREATE INDEX index\_name ON table\_name (values (<list/set column name>));//for list/set

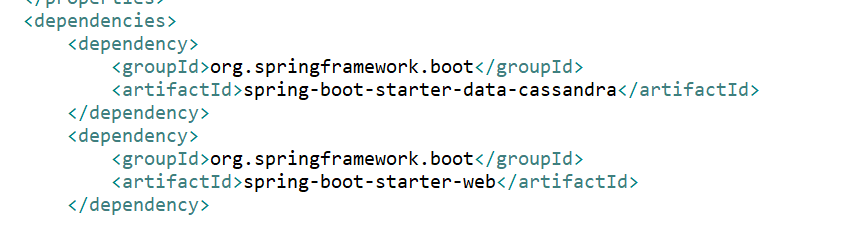
CREATE INDEX index\_name ON table\_name (<column name>);//for column

**Create spring boot application with Cassandra**

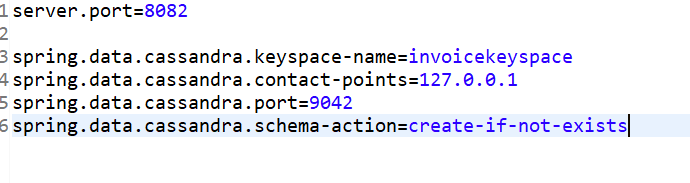
Create new spring maven project:



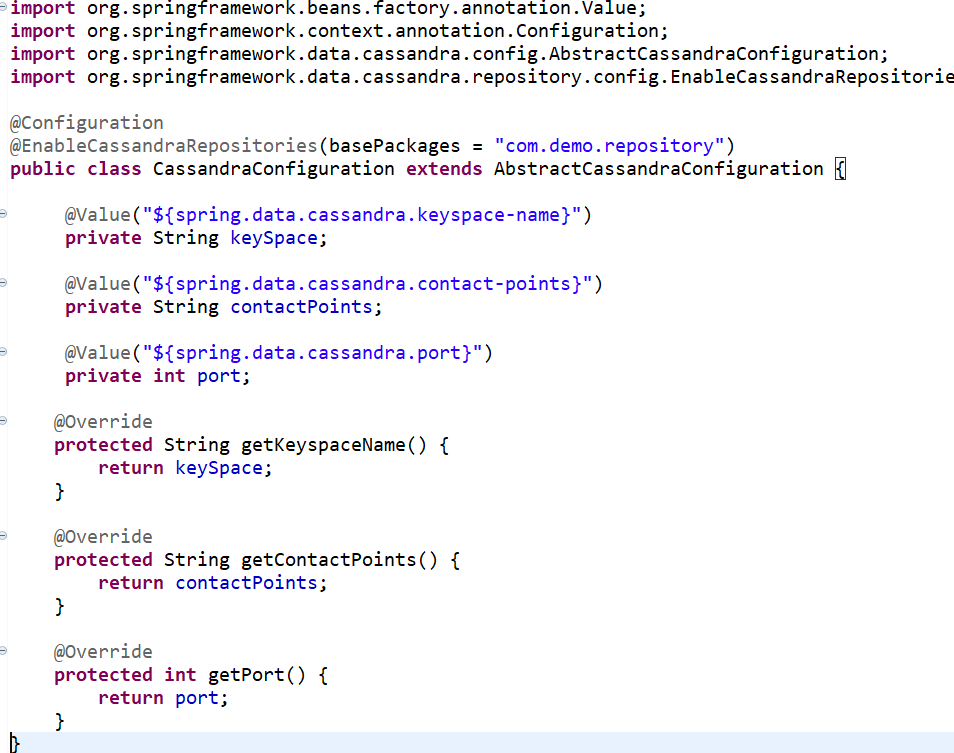
Verify below dependencies after creating the project:



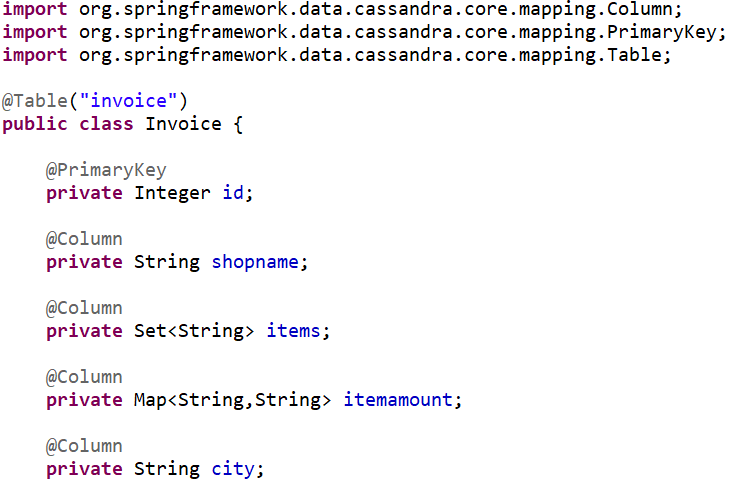
Set the property file like below:



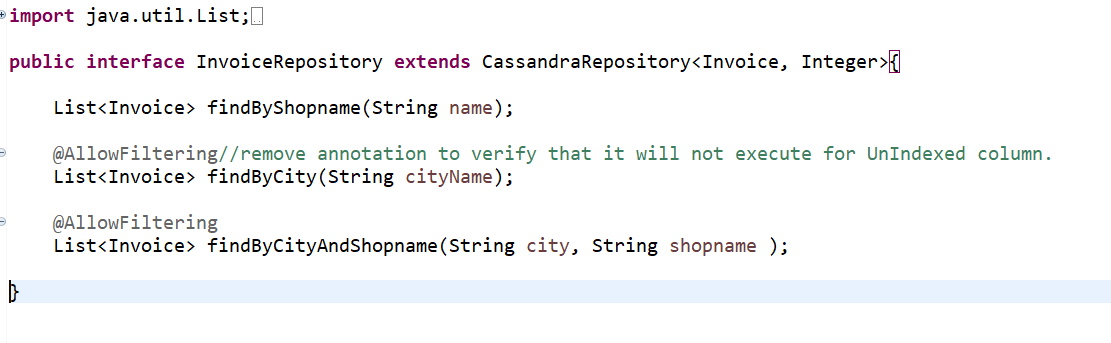
Create configuration class:



Create model class and add some property and apply getter/setter:



Create repository class:



And call the operation from end point/controller:

