

CSEE5590-0001/490-0003: Big Data Programming

Lesson Plan #7

ICP Feedback and Submission Link :

<https://forms.gle/xMAmr3zATrtMG5cX7>

Lesson Title: *Cassandra*

Lesson Description: *Independent Column based No SQL Tool - Cassandra*

Lesson Overview:

- Cassandra Introduction
- Strengths
- Architecture
- Data Types
- Commands

Cheat Sheet: <https://umkc.box.com/s/xw305tyjz44wdo237l1g zr2fzvso2wwm>

Installation guidelines:

<https://umkc.box.com/s/6xc5q85r7aht8fx8wadb03se4l8d79o4>

In Class Exercise

Data:

<https://umkc.box.com/s/zm2e2py1hifcw9pckrn83ofhke1prve0>

Cassandra (CQL Command Line)

Write a document and include your CQL screen shots for the commands below. Use the datasets provided to you in class (employees data) and create a KEYSPACE with the name “test” or use appropriate KEYSPACE according to your choice. You have the flexibility to change data according to your requirements.

To create KEYSPACE :

```
CREATE KEYSPACE KEYSPACENAME WITH REPLICATION = {'class' : 'SimpleStrategy',  
'replication_factor' : 3 }
```

Two key parameters to take note of is the class and the replication factor. The class defines the replication strategy i.e. algorithm used to determine the placement of the replicas. The replication factor determines the number of replicas.

DESC KEYSPACES

Execute the above command to check if keyspace is created or not.

USE KEYSPACENAME

Next, let's connect to the created keyspace with the help the of the USE command. The USE command enables switching context between keyspaces.

1. List the empID,ename,jobtitle,and hiredate of employee from the employee table.
2. List the name,salary of the employees who are clerks.
3. List the name,job,salary of every employee joined on 'february 18,2000',
4. List name and annual salary of all the employees.
5. Display employees' names, salary and manager values of those employees whose salary is 45000 from EMP table using SELECT statement.

Bonus

1. Write a query that displays last names with first letter capitalized and all other letters in lowercase and length of lastnames whose lastnames start with J, G or S.
2. Import any data from the given data set and apply any commands like above.

ICP Submission Guidelines:

1. ICP Submission is in pairs of three/four students.
2. Once completed, must be presented to TA or Instructor before the completion of the class
3. Submission after the deadline is considered as late submission. (Check the late submission policy in the syllabus)
4. ICP Code with brief explanation should be pushed to GitHub.
5. Submit your screenshots as well to GitHub and documentation. The screenshot should have both the code and the output.
6. Submit a demo video 2-3 min showing your assignment with a voice over explaining your work if you are unable to complete ICP within the deadline due to genuine reason.
7. Provide the video submission link through the GitHub and submission form <https://forms.gle/xMAmr3zATrtMG5cX7>

Cheating, plagiarism, disruptive behavior and other forms of unacceptable conduct are subject to strong sanctions in accordance with university policy. See detailed description of university policy at the following URL:
<https://catalog.umkc.edu/special-notices/academic-honesty/>