

# CISC 886

# Cloud and Big Data

Lab 4

Shahenda Youssef

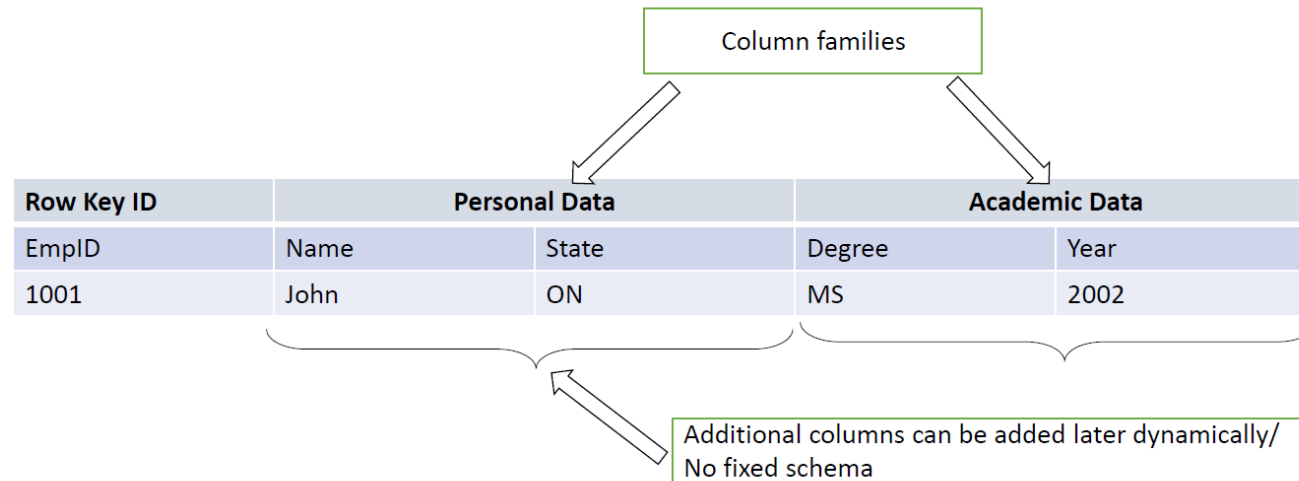
# MySQL Commands

<https://www.javatpoint.com/mysql-commands-cheat-sheet>

- Connect with MySQL Server with a username and passwords  
\$mysql -u root -p
- \$show databases;
- \$use retail\_db;
- \$show tables; (categories, customers, departments, ...)
- \$desc customers;

# Hbase Commands

<https://www.tutorialspoint.com/hbase/index.htm>



- Run Hbase `$hbase shell`
- Create Table `$create '<table name>','<column family>'`  
`$create 'emp', 'Personal Data', 'Academic Data'`
- Verify whether the table is created `$list`
- Show description of a table `$describe 'tablename'`

# Hbase Commands

- Show the records
  - `$scan 'tablename'`
  - `$scan 'emp', {LIMIT => 5}`
- Inserting Data
  - `$put 'tablename', 'row1', 'colfamily:key', 'value'`
  - `$put 'emp','1', 'Personal Data:name', 'Ali'`
- Reading Data
  - `$get 'table name','row1'`
  - `$get 'emp', '1'`
- Fetch the content of a column family `$get 'tablename ', 'rowkey ', 'column'`
- Fetch the content of a cell: `get 'tablename ', 'rowkey ', 'columnfamily:key'`

# Sqoop Commands

[https://sqoop.apache.org/docs/1.4.6/SqoopUserGuide.html#\\_syntax](https://sqoop.apache.org/docs/1.4.6/SqoopUserGuide.html#_syntax)

- Import data from external datastores into HDFS:

```
$sqoop import --connect jdbc:mysql://192.168.88.137/retail_db --table departments --username root --password bigdata --target-dir /user/osboxes/data1
```

- ```
$sqoop import --connect jdbc:mysql://192.168.88.137/retail_db --table departments --username root --password bigdata --columns "department_id, department_name" --where "department_id > 5" -m 1 --target-dir /user/osboxes/stocks_selective
```

# Sqoop Commands

- Importing the data from sqoop to Hbase:
  - Move the data from the MySQL table to HBase with respect to the primary key and the **individual** column families.

```
$sqoop import --connect jdbc:mysql://192.168.88.137/retail_db --  
username root --password bigdata --table Customers --columns "col1,  
col2" --hbase-table emp --column-family Personal Data --hbase-row-key  
emp_id -m 1
```

# Lab 4

This lab requires that you use MySQL, Sqoop and HBase. The details are provided in the following steps:

1. Create a Customers table in Hbase with two column families namely CustomerInfo and ContactInfo with row key as customer\_id
2. From the given MYSQL database named 'retail\_db' within VM, import the data of Customers table into the Hbase Customers table. Make sure all data of MySQL Customers table is imported. Use Sqoop to complete this import job. Note that:
  - a) Hint: The sqoop import command may include:
    - hbase-table tablename
    - column-family columnfamilyname
    - hbase-row-key 'rowkeyname'
  - b) CustomerInfo column family may contain the following keys (columns) and associated value: customer\_fname, customer\_lname, customer\_email, customer\_password
  - c) ContactInfo column family may contain the following keys and associated values: customer\_street, customer\_city, customer\_state, customer\_zip

# Lab 4

3. Verify that all data has been imported from the MySQL customers table.
4. List top 10 records of NY and VA state from the Hbase table.
5. Write report explaining all the above steps and show the screenshot.