Tech Module Exercises: 0 / 65

Hive and Pig

* Mini Projects

|  |  |  |
| --- | --- | --- |
| S.No. | Mini-project Description | Topics Covered |
| 1 | Pig Mini project | Pig |
| 2 | Database NYSE in hive | Hive |

* + In this TECH Module, you are expected to complete the below Mini-Project(s)
* Topics to Learn

To complete the above project, you will need to lean the below technical topics-

* + **HIVE - Introduction**

Learning Material for **HIVE - Introduction**

Below is the learning material that you are expected to read along with completion of the hands-on assignments. The material is mentioned is the order in which it should be read.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| No | Material Title | Material Location | Type of Material | Classification |
| 1. | Hive - Introduction | <https://www.tutorialspoint.com/hive/hive_introduction.htm> | Web | Mandatory |
| 2. | Hive – Installation | <https://www.tutorialspoint.com/hive/hive_installation.htm> | Web | Mandatory |
| 3. | Hive – Data Types | <https://www.tutorialspoint.com/hive/hive_data_types.htm> | Web | Mandatory |

* + - Hands-on Assignments for **HIVE - Introduction**
      * No Hands-on Assignments for this topic
  + **HIVE – Creating and Managing Database**

Learning Material for **HIVE – Creating and Managing Database**

Below is the learning material that you are expected to read along with completion of the hands-on assignments. The material is mentioned is the order in which it should be read.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| No | Material Title | Material Location | Type of Material | Classification |
| 1. | Hive – Create Database | <https://www.tutorialspoint.com/hive/hive_create_database.htm> | Web | Mandatory |
| 2. | Hive – Create Database | <https://www.youtube.com/watch?v=kEUYpCmW2qo&list=PL310uEg7PPwlC74uTYw8d0BjPNxKBQHDn> | Video | Suggestive |
| 3. | Hive – Drop Database | <https://www.tutorialspoint.com/hive/hive_drop_database.htm> | Web | Mandatory |
| 4. | Hive – Create Table | <https://www.tutorialspoint.com/hive/hive_create_table.htm> | Web | Mandatory |
| 5. | Hive – Create Table | <https://www.youtube.com/watch?v=UWDjlJwFcS0&list=PL310uEg7PPwlC74uTYw8d0BjPNxKBQHDn&index=3> | Video | Suggestive |
| 6. | Hive – Alter Table | <https://www.tutorialspoint.com/hive/hive_alter_table.htm> | Web | Mandatory |
| 7. | Hive – Alter Table | <https://www.youtube.com/watch?v=bvBQ1NQu56Q&list=PL310uEg7PPwlC74uTYw8d0BjPNxKBQHDn&index=10> | Video | Suggestive |
| 8. | Hive – Alter Table | <https://www.youtube.com/watch?v=iynMeumb4mg&list=PL310uEg7PPwlC74uTYw8d0BjPNxKBQHDn&index=9> | Video | Suggestive |
| 9. | Hive – Drop Table | <https://www.tutorialspoint.com/hive/hive_drop_table.htm> | Web | Mandatory |

* + - Hands-on Assignments for **HIVE – Creating and Managing Database**

Complete the below hands-on assignments before proceeding with the next Topic

|  |  |  |  |
| --- | --- | --- | --- |
| No. | Hands-on Assignment | Topic Covered | Status |
| 1. | Create a database named empdata . | HIVE - Creating and Managing Database |  |
| 2. | Create an employee table in the above created database ‘empdata’ with the below set of data  empid, age, salary, department  1001,25,23000,D1  1004,27,27000,D3  1002,23,21000,D2  1003,28,33000,D2  1006,31,29000,D3  1007,41,38000,D4  1009,27,21000,D3 | HIVE - Creating and Managing Database |  |
| 3. | Create a departments table in the above database ‘empdata’ with the below set of data deptid,deptname,deptloc  D1,IT,BDC  D2,HR,CDC  D3,Finance,HDC  D4,Sales,KDC | HIVE - Creating and Managing Database |  |
| 4. | Create another database named empdata 1 | HIVE - Creating and Managing Database |  |
| 5. | Create a table named order to store the below data. Save the data into a file named orders.txt  Order.txt  OrderID, CustID, OrderDate  101,C001,10/12/2015  102,C002,11/10/2015  103,C003,15/12/2015 | HIVE - Creating and Managing Database |  |
| 6. | Create a table named order\_detail to store the below data. Save the data into a file named Order\_details.txt  Order\_details.txt  OrderID,ItemID,Qty,Rate  101,Item1,10,2500  101,Item2,2,500  102,Item3,5,400  102,Item1,5,500  103,Item1,10,500 | HIVE - Creating and Managing Database |  |
| 7. | Create a Database  named hdp1 . | HIVE - Creating and Managing Database |  |
| 8. | Create a Database hdp2, verify if it is already existing. | HIVE - Creating and Managing Database |  |
| 9. | List out all the existing databases in Hive | HIVE - Creating and Managing Database |  |
| 10. | Drop the database hdp1 and hdp2 after verifying if they are existing. | HIVE - Creating and Managing Database |  |

* + **HIVE - Partitioning**

Learning Material for **HIVE - Partitioning**

Below is the learning material that you are expected to read along with completion of the hands-on assignments. The material is mentioned is the order in which it should be read.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| No | Material Title | Material Location | Type of Material | Classification |
| 1. | Hive - Partitioning | <https://www.tutorialspoint.com/hive/hive_partitioning.htm> | Web | Mandatory |
| 2. | Hive - Partitioning | <https://www.youtube.com/watch?v=OyiGcAIQf_Q&list=PL310uEg7PPwlC74uTYw8d0BjPNxKBQHDn&index=8> | Video | Suggestive |

* + - Hands-on Assignments for **HIVE - Partitioning**

Complete the below hands-on assignments before proceeding with the next Topic

|  |  |  |  |
| --- | --- | --- | --- |
| No. | Hands-on Assignment | Topic Covered | Status |
| 1. | Create a table P1 in the default database. | HIVE - Partitioning |  |
| 2. | Add a partition pt1  to the table P1 by  specifying an year. | HIVE - Partitioning |  |
| 3. | Load the values into the table P1 using above partition. | HIVE - Partitioning |  |
| 4. | Add another partition pt2 to another year for above table. | HIVE - Partitioning |  |
| 5. | Change the partition year for pt2. | HIVE - Partitioning |  |
| 6. | Drop the partition after verifying its existence. | HIVE - Partitioning |  |

* + **Hive – Built-in Operators and Functions**

Learning Material for **Hive – Built-in Operators and Functions**

Below is the learning material that you are expected to read along with completion of the hands-on assignments. The material is mentioned is the order in which it should be read.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| No | Material Title | Material Location | Type of Material | Classification |
| 1. | Hive – Built-in Operators | <https://www.tutorialspoint.com/hive/hive_built_in_operators.htm> | Web | Mandatory |
| 2. | Hive – Built-in Operators | <https://www.tutorialspoint.com/hive/hive_built_in_functions.htm> | Web | Mandatory |

* + - Hands-on Assignments for **Hive – Built-in Operators and Functions**

Complete the below hands-on assignments before proceeding with the next Topic

|  |  |  |  |
| --- | --- | --- | --- |
| No. | Hands-on Assignment | Topic Covered | Status |
| 1. | List out all employees with band as D1 and earning a salary less than 30000. | Hive - Built-in Operators and Functions |  |
| 2. | List out all employees who are not in band D2. | Hive - Built-in Operators and Functions |  |
| 3. | List how many employees are there in band D2. | Hive - Built-in Operators and Functions |  |
| 4. | List out the min, max and average salary of each band. | Hive - Built-in Operators and Functions |  |

* + **Hive – View and Indexes**

Learning Material for **Hive – View and Indexes**

Below is the learning material that you are expected to read along with completion of the hands-on assignments. The material is mentioned is the order in which it should be read.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| No | Material Title | Material Location | Type of Material | Classification |
| 1. | Hive – View and Indexes | <https://www.tutorialspoint.com/hive/hive_views_and_indexes.htm> | Web | Mandatory |
| 2. | Hive – View and Indexes | <https://www.youtube.com/watch?v=Jr5aYLIMi0U&list=PL310uEg7PPwlC74uTYw8d0BjPNxKBQHDn&index=22> | Video | Suggestive |

* + - Hands-on Assignments for **Hive – View and Indexes**

Complete the below hands-on assignments before proceeding with the next Topic

|  |  |  |  |
| --- | --- | --- | --- |
| No. | Hands-on Assignment | Topic Covered | Status |
| 1. | Create a table with below data set  ID,Name, Age,Address,Salary  1,Ross,32,Ahmedabad,2000  2,Heeln,25,Delhi,1500  3,Chandler,23,Kota,2000  4,Monika,25,Mumbai,6500  5,Mike,27,Bhopal,8500  6,Phoebe,22,MP, 4500  7,Joey,24,Indore,10000 | Hive - View and Indexes |  |
| 2. | Create a view V1 with name, annual salary and age. | Hive - View and Indexes |  |
| 3. | Create a view V2 with id, name,age,salary . | Hive - View and Indexes |  |
| 4. | Load some data into the above view v1 and v2  and verify the accessibility | Hive - View and Indexes |  |
| 5. | Create an index on id of above data. | Hive - View and Indexes |  |
| 6. | Drop the index created on id | Hive - View and Indexes |  |

* + **HiveQL – Basics of SELECT**

Learning Material for **HiveQL – Basics of SELECT**

Below is the learning material that you are expected to read along with completion of the hands-on assignments. The material is mentioned is the order in which it should be read.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| No | Material Title | Material Location | Type of Material | Classification |
| 1. | HiveQL – Select-Where | <https://www.tutorialspoint.com/hive/hiveql_select_where.htm> | Web | Mandatory |
| 2. | HiveQL – Select-Order By | <https://www.tutorialspoint.com/hive/hiveql_select_order_by.htm> | Web | Mandatory |
| 3. | HiveQL – Select-Order By | <https://www.youtube.com/watch?v=nIYW1mN_Hl4&list=PL310uEg7PPwlC74uTYw8d0BjPNxKBQHDn&index=4> | Video | Suggestive |
| 4. | HiveQL – Select-Order By | <https://www.youtube.com/watch?v=23wAWS-Z29M&list=PL310uEg7PPwlC74uTYw8d0BjPNxKBQHDn&index=11> | Video | Suggestive |
| 5. | HiveQL – Aggregations | <https://www.youtube.com/watch?v=kBsxXUfxcqk&list=PL310uEg7PPwlC74uTYw8d0BjPNxKBQHDn&index=2> | Video | Suggestive |
| 6. | HiveQL – Select-Group By | <https://www.tutorialspoint.com/hive/hiveql_group_by.htm> | Web | Mandatory |

* + - Hands-on Assignments for **HiveQL – Basics of SELECT**

Complete the below hands-on assignments before proceeding with the next Topic

|  |  |  |  |
| --- | --- | --- | --- |
| No. | Hands-on Assignment | Topic Covered | Status |
| 1. | Write a query to list details of all employees in HDC having salary >= 25000 | HiveQL - Basics of SELECT |  |
| 2. | Write a query to list the sum of salaries based on deptname | HiveQL - Basics of SELECT |  |

* + **HiveQL – Basics of JOINS**

Learning Material for **HiveQL – Basics of JOINS**

Below is the learning material that you are expected to read along with completion of the hands-on assignments. The material is mentioned is the order in which it should be read.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| No | Material Title | Material Location | Type of Material | Classification |
| 7 | HiveQL – Select-Joins | <https://www.tutorialspoint.com/hive/hiveql_joins.htm> | Web | Mandatory |
| 8 | HiveQL – Select-Joins | <https://www.youtube.com/watch?v=5GAsfUCwQrY&list=PL310uEg7PPwlC74uTYw8d0BjPNxKBQHDn&index=5> | Video | Suggestive |
| 9 | HiveQL – Select-Outer Joins | <https://www.youtube.com/watch?v=OnKBnd_bOPI&list=PL310uEg7PPwlC74uTYw8d0BjPNxKBQHDn&index=6> | Video | Suggestive |

* + - Hands-on Assignments for **HiveQL – Basics of JOINS**

Complete the below hands-on assignments before proceeding with the next Topic

|  |  |  |  |
| --- | --- | --- | --- |
| No. | Hands-on Assignment | Topic Covered | Status |
| 1. | Write a query to display empdetails with the department details | HiveQL - Basics of JOINS |  |
| 2. | Write a query to display ename who are working in Finance department | HiveQL - Basics of JOINS |  |
| 3. | Write a query to find Total Amount earned for each order | HiveQL - Basics of JOINS |  |
| 4. | Write a query to find Total Tax Amount for each order (Tax Amount = total amount \* 10%) | HiveQL - Basics of JOINS |  |

* + **PIG – Introductions and Architecture**

Learning Material for **PIG – Introductions and Architecture**

Below is the learning material that you are expected to read along with completion of the hands-on assignments. The material is mentioned is the order in which it should be read.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| No | Material Title | Material Location | Type of Material | Classification |
| 1 | Apache Pig - Overview | <https://www.tutorialspoint.com/apache_pig/apache_pig_overview.htm> | Web | Mandatory |
| 2 | Apache Pig – Introduction | <https://www.youtube.com/watch?v=qr_awo5vz0g> | Video | Suggestive |
| 3 | Apache Pig - Architecture | <https://www.tutorialspoint.com/apache_pig/apache_pig_architecture.htm> | Web | Mandatory |
| 4 | Apache Pig - Architecture | <https://www.youtube.com/watch?v=60xGpKuxG5M> | Video | Suggestive |

* + - Hands-on Assignments for **PIG – Introductions and Architecture**
      * No Hands-on Assignments for this topic
  + **PIG – Installation**

Learning Material for **PIG – Installation**

Below is the learning material that you are expected to read along with completion of the hands-on assignments. The material is mentioned is the order in which it should be read.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| No | Material Title | Material Location | Type of Material | Classification |
| 1 | Apache Pig - Installation | <https://www.tutorialspoint.com/apache_pig/apache_pig_installation.htm> | Web | Mandatory |
| 2 | Apache Pig - Installation | <https://www.youtube.com/watch?v=J9l_TP8Wwpg> | Video | Suggestive |

* + - Hands-on Assignments for **PIG – Installation**
      * No Hands-on Assignments for this topic
  + **PIG – Grunt Shell**

Learning Material for **PIG – Grunt Shell**

Below is the learning material that you are expected to read along with completion of the hands-on assignments. The material is mentioned is the order in which it should be read.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| No | Material Title | Material Location | Type of Material | Classification |
| 1 | Apache Pig - Execution | <https://www.tutorialspoint.com/apache_pig/apache_pig_execution.htm> | Web | Mandatory |
| 2 | Apache Pig – Grunt Shell | <https://www.tutorialspoint.com/apache_pig/apache_pig_grunt_shell.htm> | Web | Mandatory |
| 3 | Apache Pig – Grunt Shell | <https://www.youtube.com/watch?v=gRNHYO1nPsE> | Video | Suggestive |
| 4 | Pig Latin - Basics | <https://www.tutorialspoint.com/apache_pig/pig_latin_basics.htm> | Web | Mandatory |

* + - Hands-on Assignments for **PIG – Grunt Shell**

Complete the below hands-on assignments before proceeding with the next Topic

|  |  |  |  |
| --- | --- | --- | --- |
| No. | Hands-on Assignment | Topic Covered | Status |
| 1. | Go to Grunt Shell and write a shell command to list out all the contents in present directory | PIG - Grunt Shell |  |
| 2. | Use the fs command to see the content in the present directory | PIG - Grunt Shell |  |
| 3. | Use clear command to clear the screen | PIG - Grunt Shell |  |
| 4. | Use the history command to see all the commands that you have used till now | PIG - Grunt Shell |  |

* + **PIG – Working with data**

Learning Material for **PIG – Working with data**

Below is the learning material that you are expected to read along with completion of the hands-on assignments. The material is mentioned is the order in which it should be read.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| No | Material Title | Material Location | Type of Material | Classification |
| 1 | Apache Pig – Reading Data | <https://www.tutorialspoint.com/apache_pig/apache_pig_reading_data.htm> | Web | Mandatory |
| 2 | Apache Pig – Reading Data / loading data | <https://www.youtube.com/watch?v=wX65bpf11_w> | Video | Suggestive |
| 3 | Apache Pig – Storing Data | <https://www.tutorialspoint.com/apache_pig/apache_pig_storing_data.htm> | Web | Mandatory |
| 4 | Apache Pig – Diagnostic Operators | <https://www.tutorialspoint.com/apache_pig/apache_pig_diagnostic_operators.htm> | Web | Mandatory |
| 5 | Apache Pig – Describe Operators | <https://www.tutorialspoint.com/apache_pig/apache_pig_describe_operator.htm> | Web | Mandatory |
| 6 | Apache Pig – Explain Operators | <https://www.tutorialspoint.com/apache_pig/apache_pig_explain_operator.htm> | Web | Mandatory |
| 7 | Apache Pig – Illustrate Operators | <https://www.tutorialspoint.com/apache_pig/apache_pig_illustrate_operator.htm> | Web | Mandatory |
| 8 | Apache Pig – Describe / explain illustrate | <https://www.youtube.com/watch?v=vUctEFt2nlw> | Video | Suggestive |

* + - Hands-on Assignments for **PIG – Working with data**

Complete the below hands-on assignments before proceeding with the next Topic

|  |  |  |  |
| --- | --- | --- | --- |
| No. | Hands-on Assignment | Topic Covered | Status |
| 1. | Create a Directory in HDFS using MKDIR as Pig\_data | PIG - Working with data |  |
| 2. | Place the below data in files in local directory as Order.txt  101,C001,10/12/2015  102,C002,11/10/2015  103,C003,15/12/2015  104,C001,15/12/2015 | PIG - Working with data |  |
| 3. | Place the below data in files in local directory as Order\_detalis.txt  101,Item1,10,2500  101,Item2,2,800  102,Item3,5,400  102,Item1,5,500  103,Item1,10,500  104,Item3,3,400 | PIG - Working with data |  |
| 4. | Use the Order.txt by using copyFromLocal command to move the file to the HDFS directory Pig\_data | PIG - Working with data |  |
| 5. | Use the Order\_details.txt to move the file to HDFS directory by using put command | PIG - Working with data |  |
| 6. | Verify the files in Pig\_data | PIG - Working with data |  |
| 7. | create a table with name as Order with below data  OrderID, CustID, OrderDate  101,C001,10/12/2015  102,C002,11/10/2015  103,C003,15/12/2015  104,C001,15/12/2015 | PIG - Working with data |  |
| 8. | Use the data below for the OrderDetals table  OrderID,ItemID,Qty,Rate  101,Item1,10,2500  101,Item2,2,800  102,Item3,5,400  102,Item1,5,500  103,Item1,10,500  104,Item3,3,400 | PIG - Working with data |  |
| 9. | Create relation called Order and load the data from order.txt by using LOAD command | PIG - Working with data |  |
| 10. | Create relation called Order\_details and load the data from order\_details.txt by using LOAD command | PIG - Working with data |  |
| 11. | Store both the relations by using store in the HDFS directory /pig\_Mydata/ | PIG - Working with data |  |
| 12. | Verify the data in the /pig\_Mydata/ folder | PIG - Working with data |  |
| 13. | Use the Dump operator with the relations order and order\_Details and observe the MapReduce Job | PIG - Working with data |  |
| 14. | Describe the relation order and order\_details and check the outputs | PIG - Working with data |  |
| 15. | Use the explain command on the order relations to display the logical, physical, and MapReduce execution plans of a relation. | PIG - Working with data |  |
| 16. | Use illustrate operator on Order relation to get you the step-by-step execution of a sequence of statements. | PIG - Working with data |  |

* + **PIG – Group and Cogroup**

Learning Material for **PIG – Group and Cogroup**

Below is the learning material that you are expected to read along with completion of the hands-on assignments. The material is mentioned is the order in which it should be read.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| No | Material Title | Material Location | Type of Material | Classification |
| 1. | Apache Pig – Group Operator | <https://www.tutorialspoint.com/apache_pig/apache_pig_group_operator.htm> | Web | Mandatory |
| 2. | Apache Pig –Cogroup Operator | <https://www.tutorialspoint.com/apache_pig/apache_pig_cogroup_operator.htm> | Web | Mandatory |
| 3. | Apache Pig – Group / cogroup Operator | <https://www.youtube.com/watch?v=xB4jkp9kolM> | Video | Suggestive |

* + - Hands-on Assignments for **PIG – Group and Cogroup**

Complete the below hands-on assignments before proceeding with the next Topic

|  |  |  |  |
| --- | --- | --- | --- |
| No. | Hands-on Assignment | Topic Covered | Status |
| 1. | Create Group\_Item on order\_details relation by itemID | PIG - Group and Cogroup |  |
| 2. | Verify the Group\_Item by using Dump | PIG - Group and Cogroup |  |
| 3. | Create multiple group as Group\_Item\_qty on order\_details relation by iteamID and qty | PIG - Group and Cogroup |  |
| 4. | Use cogroup on the two relations order and order\_details created by orderDate and ItemId | PIG - Group and Cogroup |  |

* + **PIG – Join Union Split**

Learning Material for **PIG – Join Union Split**

Below is the learning material that you are expected to read along with completion of the hands-on assignments. The material is mentioned is the order in which it should be read.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| No | Material Title | Material Location | Type of Material | Classification |
| 1. | Apache Pig – Join Operator | <https://www.tutorialspoint.com/apache_pig/apache_pig_join_operator.htm> | Web | Mandatory |
| 2. | Apache Pig – Cross Operator | <https://www.tutorialspoint.com/apache_pig/apache_pig_cross_operator.htm> | Web | Mandatory |
| 3. | Apache Pig – Join Operator | <https://www.youtube.com/watch?v=R9QuCowBYxA> | Video | Suggestive |
| 4. | Apache Pig – Union Operator | <https://www.tutorialspoint.com/apache_pig/apache_pig_union_operator.htm> | Web | Mandatory |
| 5. | Apache Pig – Split Operator | <https://www.tutorialspoint.com/apache_pig/apache_pig_split_operator.htm> | Web | Mandatory |
| 6. | Apache Pig – cross / Split / Union | <https://www.youtube.com/watch?v=IThQBAp0ZW0> | Video | Suggestive |

* + - Hands-on Assignments for **PIG – Join Union Split**

Complete the below hands-on assignments before proceeding with the next Topic

|  |  |  |  |
| --- | --- | --- | --- |
| No. | Hands-on Assignment | Topic Covered | Status |
| 1. | Create relation using Join with columns as custid,itemid and qty by using orderid as a common column | PIG - Join Union Split |  |
| 2. | Create relation using left outer Join with columns as custid,itemid and qty by using orderid as a common column | PIG - Join Union Split |  |
| 3. | Create relation using right outer Join with columns as custid,itemid and qty by using orderid as a common column | PIG - Join Union Split |  |
| 4. | Use the union operator on the Order and orderdetails relation and verify the output | PIG - Join Union Split |  |
| 5. | Split the data into two relations based on the rate:  Rates more than 500 in order1 and remaining moved to order2 relations | PIG - Join Union Split |  |

* + **PIG – Filter Foreach Order By**

Learning Material for **PIG – Filter Foreach Order By**

Below is the learning material that you are expected to read along with completion of the hands-on assignments. The material is mentioned is the order in which it should be read.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| No | Material Title | Material Location | Type of Material | Classification |
| 1 | Apache Pig – Filter Operator | <https://www.tutorialspoint.com/apache_pig/apache_pig_filter_operator.htm> | Web | Mandatory |
| 2 | Apache Pig – Filter Operator | <https://www.youtube.com/watch?v=_iTp1W-H5XU> | Video | Suggestive |
| 3 | Apache Pig – Distinct Operator | <https://www.tutorialspoint.com/apache_pig/apache_pig_distinct_operator.htm> | Web | Mandatory |
| 4 | Apache Pig – Foreach Operator | <https://www.tutorialspoint.com/apache_pig/apache_pig_foreach_operator.htm> | Web | Mandatory |
| 5 | Apache Pig – Order By | <https://www.tutorialspoint.com/apache_pig/apache_pig_order_by.htm> | Web | Mandatory |

* + - Hands-on Assignments for **PIG – Filter Foreach Order By**

Complete the below hands-on assignments before proceeding with the next Topic

|  |  |  |  |
| --- | --- | --- | --- |
| No. | Hands-on Assignment | Topic Covered | Status |
| 1. | Create relation as Qty\_above\_5 . Filter the data from orderdetails dump the qty value above 5 to it. | PIG - Filter Foreach Order By |  |
| 2. | Use foreach on orderdetails and display all the data | PIG - Filter Foreach Order By |  |

* + **PIG – Other functions**

Learning Material for **PIG – Other functions**

Below is the learning material that you are expected to read along with completion of the hands-on assignments. The material is mentioned is the order in which it should be read.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| No | Material Title | Material Location | Type of Material | Classification |
| 1 | Apache Pig – Limit Operator | <https://www.tutorialspoint.com/apache_pig/apache_pig_limit_operator.htm> | Web | Mandatory |
| 2 | Apache Pig – Eval Functions | <https://www.tutorialspoint.com/apache_pig/apache_pig_eval_functions.htm> | Web | Mandatory |
| 3 | Apache Pig – Load & Store Functions | <https://www.tutorialspoint.com/apache_pig/apache_pig_load_store_functions.htm> | Web | Mandatory |
| 4 | Apache Pig Load dump sort Limit | <https://www.youtube.com/watch?v=9hHb3TmYHYg&t=21s> | Video | Suggestive |
| 5 | Apache Pig – Bag & Tuple Functions | <https://www.tutorialspoint.com/apache_pig/apache_pig_bag_tuple_functions.htm> | Web | Mandatory |
| 6 | Apache Pig – Bag & Tuple Functions | <https://www.youtube.com/watch?v=CoG6cANYU50> | Video | Suggestive |
| 7 | Apache Pig – String Functions | <https://www.tutorialspoint.com/apache_pig/apache_pig_string_functions.htm> | Web | Mandatory |
| 8 | Apache Pig – Date-time Functions | <https://www.tutorialspoint.com/apache_pig/apache_pig_date_time_functions.htm> | Web | Mandatory |
| 9 | Apache Pig – Math Functions | <https://www.tutorialspoint.com/apache_pig/apache_pig_math_functions.htm> | Web | Mandatory |
| 10 | Apache Pig – Math Functions and string functions | <https://www.youtube.com/watch?v=JdP93O104HY&t=10s> | Video | Suggestive |
| 11 | Apache Pig – User Defined Functions | <https://www.tutorialspoint.com/apache_pig/apache_pig_user_defined_functions.htm> | Web | Mandatory |
| 12 | Apache Pig – Running Scripts | <https://www.tutorialspoint.com/apache_pig/apache_pig_running_scripts.htm> | Web | Mandatory |

* + - Hands-on Assignments for **PIG – Other functions**

Complete the below hands-on assignments before proceeding with the next Topic

|  |  |  |  |
| --- | --- | --- | --- |
| No. | Hands-on Assignment | Topic Covered | Status |
| 1. | Using Foreach display all the itemid in upper case | PIG - Other functions |  |
| 2. | Using foreach display all the month part from the order relation of orderdate | PIG - Other functions |  |

* Learning Outcomes

Upon completion of this TECH Module, you should be able to:

* + Introduction to Hive
  + How to Installation
  + The Data Types in Hive
  + How to Create Database
  + How to Drop Database
  + How to Create Table
  + How to Alter Table
  + How to Drop Table
  + Partitioning in Hive
  + Use of Built-in Operators
  + Use of Built-in Functions
  + View and Indexes in Hive
  + How to use Select-Where
  + How to use Select-Order By
  + How to use Select-Order By
  + How to use Select-Order By
  + Aggregations in Hive
  + How to use Select-Group By
  + How to use Select-Joins
  + How to use Select-Joins
  + How to use Select-Outer Joins
  + Apache Pig - Overview
  + Understand the Architecture
  + How to install Pig and execution
  + Apache Pig - Execution
  + What is Grunt Shell and it’s usage
  + Pig Latin - Basics
  + Reading Data and Storing data in Pig
  + Describe all the types of Operator
  + Use of Foreach Operator
  + Understand Apache Pig – Limit Operator
  + Understand Apache Pig – Eval Functions
  + Understand Apache Pig – Load & Store Functions
  + Understand Apache Pig – Bag & Tuple Functions
  + Understand Apache Pig – String Functions
  + Understand Apache Pig – Date-time Functions
  + Understand Apache Pig – Math Functions
  + Understand Apache Pig – User Defined Functions
  + Understand Apache Pig – Running Scripts
* Sample Questions

After gaining knowledge of the above module, below are the possible interview questions that you should be able to confidently answer.

|  |  |
| --- | --- |
| No. | Questions |
| 1 | What are the different types of tables available in HIve? |
| 2 | Is Hive suitable to be used for OLTP systems? Why? |
| 3 | Can a table be renamed in Hive? |
| 4 | What is a metastore in Hive? |
| 5 | What is the need for custom Serde? |
| 6 | Why do we need Hive? |
| 7 | What is the default location where hive stores table data? |
| 8 | What are the three different modes in which hive can be run? |
| 9 | Is there a date data type in Hive? |
| 10 | What are collection data types in Hive? |
| 11 | Can we run unix shell commands from hive? Give example. |
| 12 | What is a Hive variable? What for we use it? |
| 13 | Can hive queries be executed from script files? How? |
| 14 | What are the default record and field delimiter used for hive text files? |
| 15 | What do you mean by schema on read? |
| 16 | What does the “USE” command in hive do? |
| 17 | How do you check if a particular partition exists? |
| 18 | How can Hive avoid mapreduce? |
| 19 | Can we LOAD data into a view? |
| 20 | What makes easier to program in Apache Pig than Hadoop MapReduce? |
| 21 | What are the modes of Execution for Apache Pig |
| 22 | Compare Apache Pig and SQL. |
| 23 | Explain the need for MapReduce while programming in Apache Pig. |
| 24 | What is the usage of foreach operation in Pig scripts? |
| 25 | How do users interact with the shell in Apache Pig? |
| 26 | What is illustrate used for in Apache Pig? |
| 27 | Differentiate between the logical and physical plan of an Apache Pig script |
| 28 | Explain the difference between COUNT\_STAR and COUNT functions in Apache Pig? |
| 29 | What are the commonalities between Pig and Hive? |
| 30 | Explain about the scalar datatypes in Apache Pig. |
| 31 | How do users interact with HDFS in Apache Pig ? |
| 32 | What is the use of having Filters in Apache Pig ? |
| 33 | What is a UDF in Pig? |
| 34 | Can you join multiple fields in Apache Pig Scripts ? |
| 35 | Does Pig support multi-line commands? |