

A dark blue vertical bar runs down the left side of the slide. A blue arrow points to the right from this bar, containing the text 'Big Data Lab'.

Big Data Lab

Pig Programs

Several thin, dark blue curved lines originate from the bottom left corner and sweep upwards and to the right, creating a sense of motion or data flow.

1. Write a program to filter data by city.

pig_script_filter.pig

```
student_details = LOAD '/home/msrit/Downloads/pig/test/Filter/student_details.txt'
USING PigStorage(',') as (id:int, firstname:chararray, lastname:chararray, age:int,
phone:chararray, city:chararray);

filter_data = FILTER student_details BY city == 'Chennai';

Dump filter_data;
```

2. Write a program to group students by age.

pig_script_grouping.pig

```
student = LOAD '/home/msrit/Downloads/pig/test/Grouping/student_details.txt' USING
PigStorage(',') as (id:int, firstname:chararray, lastname:chararray, age:int,
phone:chararray, city:chararray);

group_data = GROUP student by age;

Dump group_data;
```

3. Write a program to join 2 tables by id and order by customer_id

pig_script_joining.pig

```
customers = LOAD '/home/msrit/Downloads/pig/test/Joining/customer.txt' USING
PigStorage(',') as (id:int, name:chararray, age:int, address:chararray,
salary:int);

orders = LOAD '/home/msrit/Downloads/pig/test/Joining/order.txt' USING
PigStorage(',') as (oid:int, date:chararray, customer_id:int, amount:int);

join_result = JOIN customers BY id, orders BY customer_id;

Dump join_result
```

4. Write a program to sort student by age in descending order and limit the results to a maximum of 4 records.

pig_script_sorting.pig

```
student = LOAD '/home/msrit/Downloads/pig/test/Sorting/student_details.txt' USING
PigStorage(',') as (id:int, firstname:chararray, lastname:chararray, age:int,
phone:chararray, city:chararray);

student_order = ORDER student BY age DESC;

student_limit = LIMIT student_order 4;

Dump student_limit;
```

5. Write a program to create a union between 2 tables.

pig_script_union.pig

```
cust1 = LOAD '/home/msrit/Downloads/pig/test/Union/customer1.txt' USING
PigStorage(',') as (id:int, firstname:chararray, lastname:chararray,
phone:chararray, city:chararray);

cust2 = LOAD '/home/msrit/Downloads/pig/test/Union/customer2.txt' USING
PigStorage(',') as (id:int, firstname:chararray, lastname:chararray,
phone:chararray, city:chararray);

cust = UNION cust1, cust2;

Dump cust
```