

Kunal R. Agarwal

18  
201700110

Date: \_\_/\_\_/\_\_

### BDA Assignment NO 3

#### JAQL Query Language

Q1) What is jaql? Where does it fit? Explain its format  
JAQL is a functional data processing and query language most commonly used for JSON query processing on big data. JAQL supports lazy evaluation, so expressions are only materialized when needed.

Working together with a workflow orchestrator, JAQL is used to exchange data between storage, processing and analytics job. It also provides link to external data and services, including relational databases and machine learning data.

JAQL data model is based on JavaScript Object Notation or JSON. Data structure of JSON on which JAQL operates is built on two structures:-

- A collection of name/value pairs such as an object, record, struct, dictionary, hash table, keyed list.
- An ordered list of values such as an array, vector, list, or sequence.

#### JAQL Functions

→ Hive

→ Pig

→

Q.2

Let "animals.txt" be

cow

dog

bee

cow

1. `jaql > myfile = read (lines ("animals.txt"));`2. `jaql > myfile;`Output

[

"cow",

"dog",

"bee",

"cow",

]

3. `jaql > myfile -> group by animal = $ into {animal, total : count ($)};`Output

[

{ "animal": "bee",

"total": 1 },

{

"animal": "cow",

"total": 2 },

{

"animal": "dog",

"total": 1 }

]



4. `myfile = read (lines ("animals.txt"));`

`myfile`

→ `group by animal = $ into { animal, total: count($) }`

→ `write (del ("animal-count.csv"));`

Output

bee, 1

cow, 2

dog, 1

Q.3 Write different type of data types in Jaql

A.3. Jaql data type is usually inferred by how a value is provided.

#### Standard JSON types

- 1 null - null
- 2 boolean - true, false
- 3 string - "hi"
- 4 long - 10
- 5 double - 10.2, 10d, 10e-2
- 6 array - [1, 2, 3]
- 7 record - {a: 1, b: 2}

#### Jaql Extension

1. decimal - 3.5m, 10e-2m
2. binary - hex("FOA0DEE")
3. date
4. schema - schema [long\*]
5. function - fn(a, b) a+b
6. comparator
7. regex - regex("[a-c].\*")

Q.4 List all the Jaql Operators

A) Expand :-

Flattens nested arrays.



201700110

B. Filter -

Similar to a WHERE clause, filtering array data

C. GROUP -

Groups one or more arrays based upon key values, applying an aggregate expression

D. Sort -

Sorts the contents of an array

E. top -

Returns the first N element of an array

F. transform -

Transforms elements of the input array to produce a modified output array.

Q.5 Explain in detail I/O in Jaql.

A.5 I/O in Jaql is performed through I/O adapters. Adapters return a description of how to access and process a data source. It handles processing of data in its stored format.

Once created the I/O adapter is passed into an I/O function to perform the actual operation.

Some Operations that can be performed on I/O adapters

- read()
- local Read()
- write()