

#Assignment3 Data Aggregration

Name : Prerana Dhakal

Sd id: c0935990

Video Game data

12...

default

Browse

Select or search database/schema

1 CREATE EXTERNAL TABLE sales
2
3 (
4 Name string,
5 Platform string,
6 Year_of_Release int,
7 Genre string,
8 Publisher string,
9 NA_Sales float,
10 EU_Sales float,
11 JP_Sales float,
12 Other_Sales float,
13 Global_Sales float,
14 Critic_Score int,
15)
16)

Execute Save As Insert UDF Visual Explain

RESULTS LOG VISUAL EXPLAIN TEZ UI

Filter columns

sales.name	sales.platform	sales.year_of_release	sales.genre	sales.publisher	sales.na_sales	sales.eu_sales	sales.jp_sales	sa
Wii Sports	Wii	2006	Sports	Nintendo	41.36	28.96	3.77	8.4
Super Mario Bros.	NES	1985	Platform	Nintendo	29.08	3.58	6.81	0.1
Mario Kart Wii	Wii	2008	Racing	Nintendo	15.68	12.76	3.79	3.1

default

Tables(7)

Search Tables

hivesampletable

sales

sales_genre

shooter_games

shooter_games_sales_count

sports_games

sports_games_sales_count

DATABASE

default

Browse

Select or search database/schema

1 CREATE EXTERNAL TABLE sales_genre
2
3 (
4 Genre string,
5 Global_sales float,
6 Critic_Score int
7)

Execute Save As Insert UDF Visual Explain

RESULTS LOG VISUAL EXPLAIN TEZ UI

Filter columns

default

Tables(7)

Search Tables

hivesampletable

sales

sales_genre

shooter_games

shooter_games_sales_count

sports_games

sports_games_sales_count

Licensed under the Apache License, Version 2.0.
See third party tools/resources that Ambari uses and their respective authors

Select or search database/schema default Browse

```
1 SELECT * FROM sales_genre LIMIT 10;
```

Execute Save As Insert UDF Visual Explain

RESULTS LOG VISUAL EXPLAIN TEZ UI

Filter columns x

sales_genre.genre	sales_genre.global_sales	sales_genre.critic_score
Action	11.0	98
Action	11.0	98
Sports	5.0	98

default Tables(8)

Search Tables 🔍

- employee
- hivesampletable
- sales
- sales_genre
- shooter_games
- shooter_games_sales_count
- sports_games**
- sports_games_sales_count

Answer 1

QUERY JOB TABLES SAVED QUERIES UDFs SETTINGS NOTIFICATIONS

Worksheet1 * x Worksheet2 * x Worksheet3 * x **Worksheet4 * x** Worksheet5 * x Worksheet6 * x Worksheet7 * x Worksheet8 * x

Worksheet9 * x Worksheet10 * x Worksheet11 * x Worksheet12 * x hr 1 * x Worksheet14 * x Worksheet15 * x +

DATABASE default Browse

```
1
2 CREATE EXTERNAL TABLE sports_games (
3   sports_critic_score DECIMAL(5, 2)
4 );
5
6
7 CREATE EXTERNAL TABLE shooter_games (
8   shooter_critic_score DECIMAL(5, 2)
9 );
10
11
```

Execute Save As Insert UDF Visual Explain

default Tables(8)

Search Tables 🔍

- employee
- hivesampletable
- sales
- sales_genre
- shooter_games
- shooter_games_sales_count
- sports_games**

```
1 SELECT
2   AVG(critic_score) AS sports_critic_score
3 FROM sales_genre
4 WHERE genre LIKE "%Sports%"
```

Execute Save As Insert UDF Visual Explain

RESULTS LOG VISUAL EXPLAIN TEZ UI

Filter columns x

sports_critic_score
71.96311818943839

default Tables(8)

Search Tables 🔍

- employee
- hivesampletable
- sales
- sales_genre
- shooter_games
- shooter_games_sales_count
- sports_games**
- sports_games_sales_count

default

Select or search database/schema

1 SELECT
2 AVG(critic_score) AS shooter_critic_score
3 FROM sales_genre
4 WHERE genre LIKE "XshooterX"

default ✓ Tables(8)

Search Tables

- employee
- hivesampletable
- sales
- sales_genre
- shooter_games
- shooter_games_sales_count
- sports_games
- sports_games_sales_count

Execute Save As Insert UDF Visual Explain

RESULTS LOG VISUAL EXPLAIN TEZ UI

Filter columns

shooter_critic_score

70.16932907348243

QN 2

Worksheet1 * Worksheet2 * Worksheets * Worksheet4 * Worksheets * Worksheet6 * Worksheet7 * Worksheets *

Worksheet9 * Worksheet10 * Worksheet11 * Worksheet12 *

DATABASE

Select or search database/schema

default

1 SELECT
2 AVG(global_sales) AS average_global_sales,
3 MIN(global_sales) AS min_global_sales,
4 MAX(global_sales) AS max_global_sales
5 FROM sales_genre;
6

default ✓ Tables(7)

Search Tables

- hivesampletable
- sales
- sales_genre
- shooter_games
- shooter_games_sales_count
- sports_games
- sports_games_sales_count

Execute Save As Insert UDF Visual Explain

RESULTS LOG VISUAL EXPLAIN TEZ UI

Filter columns

average_global_sales	min_global_sales	max_global_sales
0.6133925268220496	0.0	83.0

DATABASE
Select or search database/schema

default

Browse

```
1 SELECT
2   AVG(global_sales) AS average_global_sales,
3   MIN(global_sales) AS min_global_sales,
4   MAX(global_sales) AS max_global_sales
5 FROM sales_genre
6 WHERE genre LIKE "%shooter%"
```

default Tables(7)
Search Tables
hivesampletable
sales
sales_genre
shooter_games
shooter_games_sales_count
sports_games
sports_games_sales_count

Execute Save As Insert UDF Visual Explain

RESULTS LOG VISUAL EXPLAIN TEZ UI

Filter columns

average_global_sales	min_global_sales	max_global_sales
0.8178913738019169	0.0	15.0

All Bookmarks

DATABASE
Select or search database/schema

default

Browse

```
1 SELECT
2   AVG(global_sales) AS average_global_sales,
3   MIN(global_sales) AS min_global_sales,
4   MAX(global_sales) AS max_global_sales
5 FROM sales_genre
6 WHERE genre LIKE "%Sports%"
```

default Tables(7)
Search Tables
hivesampletable
sales
sales_genre
shooter_games
shooter_games_sales_count
sports_games
sports_games_sales_count

Execute Save As Insert UDF Visual Explain

RESULTS LOG VISUAL EXPLAIN TEZ UI

Filter columns

average_global_sales	min_global_sales	max_global_sales
0.6906957250628667	0.0	83.0

QN 3

Ambari

Dashboard

Services

DFS

ARN

MapReduce2

ez

ive

ig

qoop

ozie

ooKeeper

Ambari Metrics

WebHcat

osts

Alerts

Cluster Admin

1 SELECT global_sales, COUNT(global_sales) AS count

2 FROM sales_genre

3 GROUP BY global_sales;

Execute

Save As

Insert UDF

Visual Explain

RESULTS

LOG

VISUAL EXPLAIN

TEZ UI

Filter columns

global_sales

count

3.0

168

5.0

54

9.0

8

10.0

6

11.0

7

default

Tables(7)

Search Tables

hivesampletable

sales

sales_genre

shooter_games

shooter_games_sales_count

sports_games

sports_games_sales_count

Worksheet9

Worksheet10

Worksheet11

Worksheet12

DATABASE

default

Browse

1 CREATE TABLE sports_games_sales_count AS

2 SELECT

3 global_sales AS global_sales,

4 COUNT(*) AS count

5 FROM sales_genre

6 WHERE genre LIKE "%Sports%"

7 GROUP BY global_sales;

8

Execute

Save As

Insert UDF

Visual Explain

RESULTS

LOG

VISUAL EXPLAIN

TEZ UI

default

Tables(7)

Search Tables

hivesampletable

sales

sales_genre

shooter_games

shooter_games_sales_count

sports_games

sports_games_sales_count

No Results available

le Nand...

Select or search database/schema default Browse

```
1 SELECT * FROM sports_games_sales_count LIMIT 10;
```

default Tables(7)

Search Tables

- hivesampletable
- sales
- sales_genre
- shooter_games
- shooter_games_sales_count
- sports_games
- sports_games_sales_count

Execute Save As Insert UDF Visual Explain

RESULTS LOG VISUAL EXPLAIN TEZ UI

Filter columns x

sports_games_sales_count.global_sales	sports_games_sales_count.count
0.0	757
1.0	306
2.0	69

1 CREATE TABLE shooter_games_sales_count AS
2 SELECT
3 global_sales AS global_sales,
4 COUNT(*) AS count
5 FROM sales_genre
6 WHERE genre LIKE '%shooter%'
7 GROUP BY global_sales;
8 SELECT * FROM shooter_games_sales_count LIMIT 10;

default Tables(7)

Search Tables

- employee
- hivesampletable
- sales
- sales_genre
- shooter_games
- sports_games
- sports_games_sales_count

Execute Save As Insert UDF Visual Explain

RESULTS LOG VISUAL EXPLAIN TEZ UI

Filter columns x

shooter_games_sales_count.global_sales	shooter_games_sales_count.count
3.0	30
5.0	5
9.0	2
10.0	3
11.0	1

Human resource data

Nand...

All Bookmarks

Hive View 2.0

margie

admin

HIVE

NEW JOB

NEW TABLE

QUERY

JOBS

TABLES

SAVED QUERIES

UDFs

SETTINGS

NOTIFICATIONS

Worksheet1

Worksheet2

Worksheet3

Worksheet4

Worksheet5

Worksheet6

+

DATABASE

default

Browse

1 CREATE EXTERNAL TABLE employee

2 {

3 {

4 Age int,

5 Attrition string,

6 BusinessTravel string,

7 DailyRate int,

8 Department string,

9 DistanceFromHome int,

10 Education int,

11 EducationField string,

12 EmployeeCount int,

13 EmployeeNumber int,

14 EnvironmentSatisfaction int,

Execute

Save As

Insert UDF

Visual Explain

default

Tables(7)

Search Tables

employee

employee_1

employee_sales

employee_sales1

hivesampletable

salespeople_attrition

salespeople_non_attrition

RESULTS

LOG

VISUAL EXPLAIN

TEZ UI

Filter columns

employee.age

employee.attrition

employee.businesstravel

employee.dailyrate

employee.department

employee.distancefromhome

41	Yes	Travel_Rarely	1102	Sales	1
49	No	Travel_Frequently	279	Research & Development	8
37	Yes	Travel_Rarely	1373	Research & Development	2
33	No	Travel_Frequently	1392	Research & Development	3
27	No	Travel_Rarely	591	Research & Development	2
32	No	Travel_Frequently	1005	Research & Development	2
59	No	Travel_Rarely	1324	Research & Development	3
30	No	Travel_Rarely	1358	Research & Development	24
38	No	Travel_Frequently	216	Research & Development	23
36	No	Travel_Rarely	1299	Research & Development	27

Worksheet1

Worksheet2

Worksheet3

Worksheet4

Worksheet5

Worksheet6

+

DATABASE

default

Browse

1 CREATE EXTERNAL TABLE employee_1 AS

2 SELECT * FROM employee

3 WHERE department LIKE "%Sales%"

4

Execute

Save As

Insert UDF

Visual Explain

RESULTS

LOG

VISUAL EXPLAIN

TEZ UI

Filter columns

1 SELECT * FROM employee_1 LIMIT 10;

Execute

Save As

Insert UDF

Visual Explain

RESULTS

LOG

VISUAL EXPLAIN

TEZ UI

Filter columns

x

employee_1.age	employee_1.attrition	employee_1.businesstravel	employee_1.dailyrate	employee_1.department	employee_1.distance
41	Yes	Travel_Rarely	1102	Sales	1
53	No	Travel_Rarely	1219	Sales	2
36	Yes	Travel_Rarely	1218	Sales	9
42	No	Travel_Rarely	691	Sales	8

default

Tables(7)

Search Tables

employee

employee_1

employee_sales

employee_sales1

hivesampletable

salespeople_attrition

salespeople_non_attrition

QUERY

JOBS

TABLES

SAVED QUERIES

UDFs

SETTINGS

NOTIFICATIONS

Worksheet1

Worksheet2

Worksheet2 *

Worksheet3 *

Worksheet4 *

Worksheet5 *

Worksheet6 *

+

DATABASE

Select or search database/schema

default

Browse

1 CREATE EXTERNAL TABLE salespeople_attrition

2

3 (

4 attrition string

5);

6

7 CREATE EXTERNAL TABLE salespeople_non_attrition

8

9 (

10 non_attrition string

11);

Execute

Save As

Insert UDF

Visual Explain

default

Tables(7)

Search Tables

employee

employee_1

employee_sales

employee_sales1

hivesampletable

salespeople_attrition

salespeople_non_attrition

1 SELECT attrition, COUNT(attrition) AS attrition_count

2 FROM employee_1

3 WHERE attrition = 'Yes'

4 GROUP BY attrition;

5

Execute

Save As

Insert UDF

Visual Explain

RESULTS

LOG

VISUAL EXPLAIN

TEZ UI

Filter columns

x

attrition	attrition_count
Yes	92

default

Tables(7)

Search Tables

employee

employee_1

employee_sales

employee_sales1

hivesampletable

salespeople_attrition

salespeople_non_attrition

```
1 SELECT attrition, COUNT(attrition) AS non_attrition_count
2 FROM employee_1
3 WHERE attrition = 'No'
4 GROUP BY attrition;
```

default Tables(7)

Search Tables

employee

employee_1

employee_sales

employee_sales1

hivesampletable

salespeople_attrition

salespeople_non_attrition

Execute Save As Insert UDF Visual Explain

RESULTS LOG VISUAL EXPLAIN TEZ UI

Filter columns

attrition non_attrition_count

No 354

QN 2

```
1 SELECT
2 AVG(monthlyincome) AS average_monthly_income,
3 MIN(monthlyincome) AS min_monthly_income,
4 MAX(monthlyincome) AS max_monthly_income
5 FROM employee_1;
6
```

default Tables(7)

Search Tables

employee

employee_1

employee_sales

employee_sales1

hivesampletable

salespeople_attrition

salespeople_non_attrition

Execute Save As Insert UDF Visual Explain

RESULTS LOG VISUAL EXPLAIN TEZ UI

Filter columns

average_monthly_income min_monthly_income max_monthly_income

6959.17264573991 1052 19847

Licensed under the Apache License, Version 2.0

See third-party tools/resources that Ambari uses and their respective authors

12°C

10:40 PM

11/6/2023

DATABASE

Select or search database/schema

default

Browse

```

1 SELECT
2   AVG(monthlyincome) AS average_monthly_income,
3   MIN(monthlyincome) AS min_monthly_income,
4   MAX(monthlyincome) AS max_monthly_income
5 FROM employee_1
6 WHERE attrition LIKE 'No';

```

Execute

Save As

Insert UDF

Visual Explain

RESULTS

LOG

VISUAL EXPLAIN

TEZ UI

Filter columns

average_monthly_income	min_monthly_income	max_monthly_income
7232.24011299435	1052	19847

default

Tables(7)

Search Tables

employee

employee_1

employee_sales

employee_sales1

hivesampletable

salespeople_attrition

salespeople_non_attrition

default

Tables(8)

Search Tables

employee

employee_1

hivesampletable

salespeople_attrition

salespeople_non_attrition

salespeople_with_attrition_income_count

salespeople_with_attrition_income_count1

salespeople_without_attrition_income_count

default

Tables(8)

Search Tables

employee

employee_1

hivesampletable

salespeople_attrition

salespeople_non_attrition

salespeople_with_attrition_income_count

salespeople_with_attrition_income_count1

salespeople_without_attrition_income_count

DATABASE

Select or search database/schema

default

Browse

```

1 SELECT
2   AVG(monthlyincome) AS average_monthly_income,
3   MIN(monthlyincome) AS min_monthly_income,
4   MAX(monthlyincome) AS max_monthly_income
5 FROM employee_1
6 WHERE attrition LIKE 'Yes';
7

```

Execute

Save As

Insert UDF

Visual Explain

RESULTS

LOG

VISUAL EXPLAIN

TEZ UI

Filter columns

average_monthly_income	min_monthly_income	max_monthly_income
5908.45652173913	1081	19845

QN 3

QUERYJOBSTABLESSAVED QUERIESUDFSSettingsNOTIFICATIONS

Worksheet1 x Worksheet2 x Worksheet3 x Worksheet4 x Worksheet5 x Worksheet7 x Worksheet8 x Worksheet9 x

Worksheet1 * x Worksheet2 * x +

DATABASE

Select or search database/schema

default

Browse

1 CREATE TABLE salespeople_with_attrition_income_count1 AS

2 SELECT

3 monthlyincome AS monthlyincome,

4 COUNT(*) AS count

5 FROM employee_1

6 WHERE attrition = 'Yes'

7 GROUP BY monthlyincome;

8

Execute

Save As

Insert UDF

Visual Explain

RESULTS

LOG

VISUAL EXPLAIN

TEZ UI

default

Tables(7)

Search Tables

employee

employee_1

hivesampletable

salespeople_attrition

salespeople_non_attrition

salespeople_with_attrition_income_count

salespeople_with_attrition_income_count1

Select or search database/schema

default

Browse

1 SELECT * FROM salespeople_with_attrition_income_count1 LIMIT 10;

Execute

Save As

Insert UDF

Visual Explain

RESULTS

LOG

VISUAL EXPLAIN

TEZ UI

Filter columns x

salespeople_with_attrition_income_count1.monthlyincome salespeople_with_attrition_income_count1.count

1081 1

1091 1

1118 1

Select or search database/schema

default

Browse

```
1 CREATE TABLE salespeople_without_attrition_income_count AS
2 SELECT
3   monthlyincome AS monthlyincome,
4   COUNT(*) AS count
5 FROM employee_1
6 WHERE attrition = 'No'
7 GROUP BY monthlyincome;
8 SELECT * FROM salespeople_without_attrition_income_count LIMIT 10;
9
```

Execute Save As Insert UDF Visual Explain

RESULTS LOG VISUAL EXPLAIN TEZ UI

Filter columns

salespeople_without_attrition_income_count.monthlyincome

salespeople_without_attrition_income_count.count

1052	1
2014	1
2096	1
2207	1

default

Tables(7)

Search Tables

employee

employee_1

hivesampletable

salespeople_attrition

salespeople_non_attrition

salespeople_with_attrition_income_count

salespeople_with_attrition_income_count1