

BUAN 6346 Big Data

Homework 7

Partition Data in Impala or Hive

Step1-3:

The screenshot shows the Impala Query Editor interface. On the left, the 'DATABASE' dropdown is set to 'default', and a list of tables is visible: accounts, accounts_avro, accounts_by_areacode, customerservicerep, device, and webpage. The main query editor area contains the following SQL code:

```
1 CREATE EXTERNAL TABLE accounts_by_areacode (  
2 acct_num INT,  
3 first_name STRING,  
4 last_name STRING,  
5 phone_number STRING )  
6 PARTITIONED BY (areacode STRING)  
7 ROW FORMAT DELIMITED  
8 FIELDS TERMINATED BY ','  
9 LOCATION '/loudacre/accounts_by_areacode';  
10  
11  
12 SELECT acct_num, first_name, last_name,  
13 phone_number, SUBSTR(phone_number,1,3) AS areacode  
14 FROM accounts_avro;  
15  
16 Insert into accounts by areacode PARTITION(areacode) SELECT acct_num, first_name, last_name,  
17 phone_number, SUBSTR(phone_number,1,3) AS areacode  
18 FROM accounts_avro;  
19
```

Below the query editor, there are buttons for 'Execute', 'Save as...', 'Explain', and 'New query'. The 'Results' tab is selected, showing the message: 'The operation has no results.'

Step4: SELECT * FROM accounts_by_areacode LIMIT 10

The screenshot shows the Impala Query Editor interface with the same database and table list. The main query editor area contains the following SQL code:

```
1 SELECT * FROM accounts_by_areacode LIMIT 10
```

Below the query editor, there are buttons for 'Execute', 'Save as...', 'Explain', and 'New query'. The 'Results' tab is selected, displaying the following data:

	acct_num	first_name	last_name	phone_number	areacode
0	64915	Thomas	Egan	7759708972	775
1	64916	Mary	Walker	7758009605	775
2	64922	Lisa	Ray	7752225702	775
3	64998	Darla	Hammons	7754380170	775
4	65006	Jason	Fallon	7756034547	775
5	65022	Cassie	Devlin	7751365937	775
6	65031	Tamika	Coleman	7756074591	775
7	65041	Natalia	Barrett	7756572963	775
8	65118	Thelma	Laroche	7758320564	775
9	65120	Scott	Shultis	7759372692	775

Step5:

File Browser

Actions

Move to trash

Upload

New

	Name	Size	User	Group	Permissions	Date
	.		training	supergroup	drwxrwxrwx	June 27, 2020 11:24 AM
	..		impala	supergroup	drwxrwxrwx	June 27, 2020 11:37 AM
	_impala_insert_staging		impala	supergroup	drwxrwxrwx	June 27, 2020 11:37 AM
	areacode=209		impala	supergroup	drwxrwxrwx	June 27, 2020 11:37 AM
	areacode=213		impala	supergroup	drwxrwxrwx	June 27, 2020 11:37 AM
	areacode=310		impala	supergroup	drwxrwxrwx	June 27, 2020 11:37 AM
	areacode=408		impala	supergroup	drwxrwxrwx	June 27, 2020 11:37 AM
	areacode=415		impala	supergroup	drwxrwxrwx	June 27, 2020 11:37 AM
	areacode=424		impala	supergroup	drwxrwxrwx	June 27, 2020 11:37 AM
	areacode=503		impala	supergroup	drwxrwxrwx	June 27, 2020 11:37 AM
	areacode=510		impala	supergroup	drwxrwxrwx	June 27, 2020 11:37 AM
	areacode=530		impala	supergroup	drwxrwxrwx	June 27, 2020 11:37 AM
	areacode=541		impala	supergroup	drwxrwxrwx	June 27, 2020 11:37 AM
	areacode=559		impala	supergroup	drwxrwxrwx	June 27, 2020 11:37 AM
	areacode=562		impala	supergroup	drwxrwxrwx	June 27, 2020 11:37 AM

Home

/ loudacre / accounts_by_areacode / areacode=209 / d241ec54a4535e99-6b063307e127f893_293771212_data.0.

Page 1 of 26

Previous

Next

```

64906, Teresa, Gomez, 2090238951
65026, Stan, Myers, 2090051921
65046, Cecil, Williams, 2090918849
65057, Shawn, Davis, 2093582566
65067, Timothy, Okeefe, 2094851742
65102, Edith, Fitzgerald, 2090709340
65147, Eddie, Singh, 2095010447
65181, Kenny, Palacios, 2094731413
65226, Britt, Richards, 2091494429
65301, Gerald, Cohen, 2094117753
65307, Kenneth, Rogers, 2098601991
65318, Michael, Long, 2094906167
65390, Robert, Bartels, 2098051404
65429, Kelvin, Rowe, 2095625371
65471, William, Benton, 2096044868
65514, Martin, Martin, 2092251868
65516, James, Watson, 2096404880
65642, Judith, Goad, 2097876624
65680, Jack, Shirley, 2092610569
65752, Debbie, Day, 2099193260
65762, Claudine, Exley, 2096450671
65795, Willard, Meaney, 2090863109
65826, Lori, Walls, 2097908538
65831, Randall, Vazquez, 2097078563
65854, Kenneth, Haire, 2091746081

```

ACTIONS

View as binary

Edit file

Download

View file location

Refresh

INFO

Last modified

June 27, 2020 11:37 a.m.

User

impala

Group

supergroup

Size

100.3 KB

Mode

100666

2