databricks04 - Timestamp Functions



Create Tables

Run the cell below to create tables for the questions in this notebook.

%run ../Utilities/04-CreateTables

Declared the following table:

• timetable1

Declared the following table:

• timetable2

Question 1: Extract Year and Month

Summary

Extract the year and month from the Timestamp field in the table timetable1 and store records with only the 12th month into q1Results table.

Steps to complete

Write a SQL query that achieves the following:

- Extracts the year and month from the Timestamp column from the table timetable1
 - Timestamp is an integer representing seconds since midnight on January 1st, 1970 (e.g. 1519344286). You must cast it as a timestamp to to extract years and months.
- Filters records to include only month 12
- Stores the resulting records in a temporary view named q1Results with the following schema.

column	type
Date	timestamp
Year	integer
Month	integer

A properly completed solution should produce a DataFrame similar to this sample output:

Data	Vaar	Manth
Date	Year	Month

SELECT * **FROM** timetable1

	index	Name	ID 🛕	Password
1	0	Non	4895554448746211000	s4ZMhBTflBw7
2	1	Laborum	8151144905769343000	vzZjMwK7us26

3	2	Laborum Tempor	5045821528940880000	MHv3FWPwuDw3
4	3	Ullamco Commodo	6878394900174072000	v4FGotjBrabl
5	4	Veniam In	8025583264729804000	rKYgKzhtxmdz
6	5	Dolor Excepteur	7990147143746850000	3gszsHrM6068
7	6	Enim Consectetur	9086843171616272000	5jSJPrQdXMsn

Truncated results, showing first 1000 rows.

CREATE OR REPLACE TEMPORARY VIEW q1Results AS **SELECT**

```
FROM_UNIXTIME(Timestamp) Date,
month(FROM_UNIXTIME(Timestamp)) Month,
year(FROM_UNIXTIME(Timestamp)) Year
FROM
timetable1
```

SELECT * **FROM** q1Results

OK

	Date	Month	Year 🔺
1	2015-06-08 16:22:55	6	2015
2	2015-03-26 20:44:38	3	2015
3	2016-11-30 22:10:12	11	2016
4	2003-09-03 09:12:14	9	2003
5	2021-08-15 13:27:25	8	2021
6	2017-03-19 21:10:25	3	2017
7	2017-08-10 14:18:40	8	2017

Truncated results, showing first 1000 rows.

Question 2: Extract year, month, and day of year

Summary

Extract the year, month and dayofyear from the Timestamp field in the table timetable2 and return records for only the 4th month.

Steps to complete

Write a SQL query that achieves the following:

- Create Year, Month and DayOfYear columns from the Timestamp column in the timetable2 table
 - Timestamp is an integer representing seconds since midnight on January 1st, 1970 (e.g. 1519344286). You must cast it as a timestamp to to extract years, months, and day of year.
- Filter records to include only month 4
- Stores the records in a temporary table named q2Results with the following schema:

column	type
Date	timestamp
Year	integer
Month	integer
DayOfYear	integer

• A properly completed solution should produce a DataFrame similar to this sample output:

Date	Year	Month	DayOfYear
2002-04-22 06:41:39	2002	4	112
2012-04-01 05:00:06	2012	4	92

CREATE OR REPLACE TEMPORARY VIEW q2Results AS **SELECT**

```
FROM_UNIXTIME(Timestamp) Date,
month(FROM_UNIXTIME(Timestamp)) Month,
year(FROM_UNIXTIME(Timestamp)) Year,
dayofyear(FROM_UNIXTIME(Timestamp)) DayOfYear
--date_format(split(from_unixtime(Timestamp), ' ')[0], 'EEEE') DayOfYear
FROM
timetable1
SELECT
FROM
q2Results
```

	Date	Month 🔺	Year 🔺	DayOfYear 🔺
1	2015-06-08 16:22:55	6	2015	159
2	2015-03-26 20:44:38	3	2015	85
3	2016-11-30 22:10:12	11	2016	335
4	2003-09-03 09:12:14	9	2003	246
5	2021-08-15 13:27:25	8	2021	227
6	2017-03-19 21:10:25	3	2017	78
7	2017-08-10 14:18:40	8	2017	222

Truncated results, showing first 1000 rows.

© 2020 Databricks, Inc. All rights reserved.

Apache, Apache Spark, Spark and the Spark logo are trademarks of the Apache Software Foundation (http://www.apache.org/).