

External Table Creation Guide

Please follow these simple steps to create external table on top of HDFS parquet data.

1. Run below script

```
/opt/jobs/CEM/hdfs_platform/monitoring/scripts/manual/create_external_table.sh
```

2. Select Storage Layer

Display all available storage layers from hdfs. Choose appropriate one by entering number.

```
Available Storage Layers:
1. items
2. ldr
3. qos
4. sdm
5. stg
6. test_migration
7. usm
Enter your choice (1-7): 4
```

3. Select File Directory

Choose folder containing the parquet files you want to register as an external table.

```
Selected Storage Layer: sdm
Available File Directories:
1. items
2. cgr_ntwrk_stats
3. dly_call_traffic_fct
4. dly_cemfixed_total_traffic_fct
5. dly_cstmr_ntwrk_fct
6. dly_data_throwpt_dtl_fct
7. dly_data_throwpt_fct
8. dly_fup_type_info
9. dly_iot_svcs_usg_pdp_fct
10. dly_iot_svcs_usg_tcp_fct
11. dly_roamers_fct
12. dly_sms_fct
13. dly_video_streaming_fct
14. dly_vo_wifi_fct
15. dly_voip_call_fct
16. dly_volte_addl_info
17. dly_volte_call_fct
18. dly_web_brwsng_fct
19. dly_web_brwsng_fct_new
20. dly_wifi_offloading_fct
21. dt=2025-09-17
22. qos_dly_web_brwsng_fct
23. test1
Enter your choice (1-23): 12
```

4. Choose Hive Database

Lists all hive databases. Pick your target schema for the table.

```
Selected File Directory: dly_sms_fct
Available Databases:
1. bi_onetime_team
2. ccex_reserved
3. ccex_rpt
4. default
5. dp_dfl_cem_edl
6. information_schema
7. prod_cem_ldr
8. prod_cem_sdm
9. prod_cem_sdm_arch
10. prod_cem_stg
11. prod_usm
12. prod_usm_view
13. qos
14. sys
15. test
16. tibco_test
Enter your choice (1-16): 15
```

5. Enter Table Name

Provide a name for new table.

```
Selected Database: test
Enter Table Name: sms_fct_demo
```

6. Schema Auto-Detection

Tool will do the following steps automatically:

- Find latest parquet file
- Extract schema and partition columns
- Generate external table DDL

7. Table Created & Repaired Successfully

External table has been created and partitions successfully repaired. If you wanna create another table then press **Y** otherwise **N**.

```
=====
[✓] Selected Storage Layer: sdm
[✓] Selected File Directory: dly_sms_fct
[✓] Selected Database: test
[✓] Table Name: sms_fct_demo
[✓] HDFS Location: /user/cxpprod/prod/ccex//sdm/dly_sms_fct
[✓] Status: Created! & Repaired
=====
Do you wanna create another External Table (Y/N): █
```

8. Exit the Process

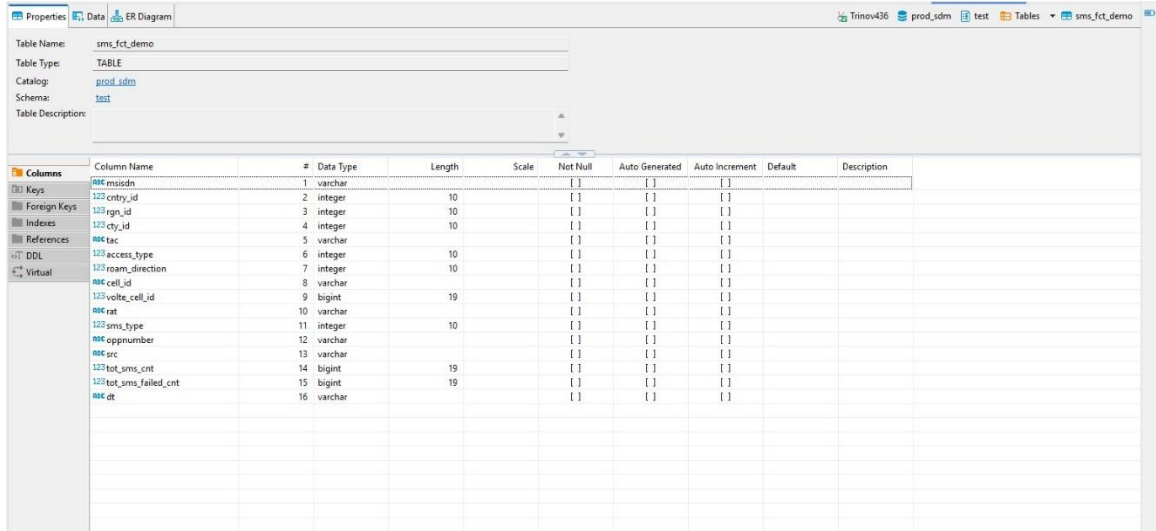
If you wish to exit external table creation process, simply press **N**

```
-----
List of tables created:
1. prod_sdm.test.sms_fct_demo

Thanks for Using!
[stccemprod@cemsem-wk07.stc.corp manual]$ █
```

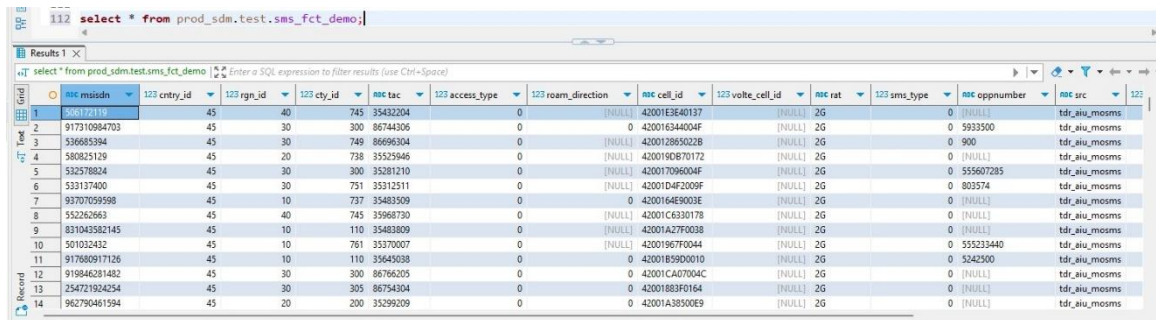
Tool Outputs

As shown in the following screenshots, table has been created and partitions have been successfully repaired.



The screenshot displays the 'Properties' tab of a database tool. The table 'sms_fct_demo' is selected, showing its metadata: Table Type is 'TABLE', Catalog is 'prod_sdm', Schema is 'test', and Table Description is empty. Below this, the 'Columns' tab is active, showing a list of 16 columns with their data types, lengths, scales, and other attributes. The columns are: nrc_msisdn (varchar, 10), nrc_cntry_id (integer, 10), nrc_rgn_id (integer, 10), nrc_cty_id (integer, 10), nrc_tac (varchar, 10), nrc_access_type (integer, 10), nrc_roam_direction (integer, 10), nrc_cell_id (varchar, 19), nrc_volte_cell_id (varchar, 19), nrc_rat (integer, 10), nrc_sms_type (integer, 10), nrc_opnnumber (varchar, 10), nrc_src (varchar, 10), nrc_tot_sms_cnt (bigint, 19), nrc_tot_sms_failed_cnt (bigint, 19), and nrc_dt (varchar, 10). The table is partitioned by nrc_dt.

Column Name	#	Data Type	Length	Scale	Not Null	Auto Generated	Auto Increment	Default	Description
nrc_msisdn	1	varchar	10		[]	[]	[]		
nrc_cntry_id	2	integer	10		[]	[]	[]		
nrc_rgn_id	3	integer	10		[]	[]	[]		
nrc_cty_id	4	integer	10		[]	[]	[]		
nrc_tac	5	varchar	10		[]	[]	[]		
nrc_access_type	6	integer	10		[]	[]	[]		
nrc_roam_direction	7	integer	10		[]	[]	[]		
nrc_cell_id	8	varchar	19		[]	[]	[]		
nrc_volte_cell_id	9	varchar	19		[]	[]	[]		
nrc_rat	10	integer	10		[]	[]	[]		
nrc_sms_type	11	integer	10		[]	[]	[]		
nrc_opnnumber	12	varchar	10		[]	[]	[]		
nrc_src	13	varchar	10		[]	[]	[]		
nrc_tot_sms_cnt	14	bigint	19		[]	[]	[]		
nrc_tot_sms_failed_cnt	15	bigint	19		[]	[]	[]		
nrc_dt	16	varchar	10		[]	[]	[]		



The screenshot displays the 'Results' tab of the same database tool. A query 'select * from prod_sdm.test.sms_fct_demo;' has been executed, resulting in 14 rows of data. The columns are the same as in the first screenshot. The data is as follows:

nrc_msisdn	nrc_cntry_id	nrc_rgn_id	nrc_cty_id	nrc_tac	nrc_access_type	nrc_roam_direction	nrc_cell_id	nrc_volte_cell_id	nrc_rat	nrc_sms_type	nrc_opnnumber	nrc_src
906172116	45	40	745	35432204	0	[NULL]	42001E3E40137	[NULL]	2G	0	[NULL]	tdr_eiu_mosms
917310994703	45	30	300	86744306	0	0	420016344004F	[NULL]	2G	0	5933500	tdr_eiu_mosms
536685394	45	30	749	86696304	0	0	420012985022B	[NULL]	2G	0	900	tdr_eiu_mosms
548025129	45	20	738	35525946	0	[NULL]	4200140870172	[NULL]	2G	0	[NULL]	tdr_eiu_mosms
532578824	45	30	300	35381210	0	[NULL]	420017096004F	[NULL]	2G	0	555607285	tdr_eiu_mosms
533117400	45	30	751	35312511	0	[NULL]	42001D4F2009F	[NULL]	2G	0	803574	tdr_eiu_mosms
93707059598	45	10	737	35483509	0	0	4200164E9003E	[NULL]	2G	0	[NULL]	tdr_eiu_mosms
552262663	45	40	745	35968730	0	[NULL]	42001C6330178	[NULL]	2G	0	[NULL]	tdr_eiu_mosms
831043582145	45	10	110	35483809	0	[NULL]	42001A27F0038	[NULL]	2G	0	[NULL]	tdr_eiu_mosms
501032432	45	10	761	35370007	0	[NULL]	42001967F0044	[NULL]	2G	0	555233440	tdr_eiu_mosms
917680917126	45	10	110	35645038	0	0	4200185900010	[NULL]	2G	0	5242500	tdr_eiu_mosms
919846281482	45	30	300	86766205	0	0	42001CA07004C	[NULL]	2G	0	[NULL]	tdr_eiu_mosms
254721924254	45	30	305	86754304	0	0	42001883F0164	[NULL]	2G	0	[NULL]	tdr_eiu_mosms
962790461594	45	20	200	35299209	0	0	42001A38500E9	[NULL]	2G	0	[NULL]	tdr_eiu_mosms