

Create Hive database and Hive tables

We can create a database and tables in Hive after creating datewise_bookings table. This operation was already written inside **spark_local_flatten_datewise_aggregates.py**

```
# create Hive database
spark.sql("create database if not exists capstone_project")

# write dataframes to Hive
bookings.write.mode("overwrite").saveAsTable("capstone_project.bookings")
clicking_stream.write.mode("overwrite").saveAsTable("capstone_project.clicking_stream")
datewise_bookings.write.format("orc").mode("overwrite").saveAsTable("capstone_project.datewise_bookings")
print("Successfully loaded dataframes to Hive")
```

Make sure that we have pyspark==2.3.0 installed and run this command below. If already run this file in step 3, don't run it again.

```
python spark_local_flatten_datewise_aggregates.py
```

The first line is to create a database named capstone_project in Hive. And the rest are used to write tables to the database.

To check if the database and tables are created successfully, use this command in Hive

```
>> hive
>> use capstone_project;
>> show tables;
>> select * from bookings limit 20;
>> select * from clicking_stream limit 20;
>> select * from datewise_bookings limit 20;
```

Screenshots of command process

```
root@ip-10-0-0-174:~/capstoi x + v
[root@ip-10-0-0-174 capstone_project]# hadoop fs -ls /user/hive/warehouse/
Found 1 items
drwxrwxrwx - root ec2-user 0 2022-04-19 21:23 /user/hive/warehouse/capstone_project.db
[root@ip-10-0-0-174 capstone_project]# hive
OpenJDK 64-Bit Server VM warning: ignoring option MaxPermSize=512M; support was removed in 8.0
OpenJDK 64-Bit Server VM warning: ignoring option MaxPermSize=512M; support was removed in 8.0

Logging initialized using configuration in jar:file:/opt/cloudera/parcels/CDH-5.15.1-1.cdh5.15.1.p0.4/jars/hive-common-1.1.0-cdh5.15.1.jar!/hive-log4j.properties
WARNING: Hive CLI is deprecated and migration to Beeline is recommended.
hive> use capstone_project;
OK
Time taken: 1.826 seconds
hive> show tables;
OK
bookings
clicking_stream
datewise_bookings
Time taken: 0.216 seconds, Fetched: 3 row(s)
hive> |
```

```
hive> select * from bookings limit 20;
OK
BK8968087150 51811359 15055660 2.2.14 Android -49.431965 103.917854 -58.804386 146.47737 2020-06-23 19:33:10 2020
-06-06 09:02:10 534 83 INR black 054-38-4479 4 3 3
BK629851904 31663218 60872180 3.4.1 iOS -83.54084 175.80086 86.20705 128.36723 2020-05-23 12:22:04 2020
-08-09 19:02:56 126 67 INR lime 796-39-6801 3 2 4
BK1797410350 86869399 94276051 4.1.36 iOS -67.89307 55.234127 -51.1079 -31.07475 2020-05-19 14:14:32 2020
-08-23 18:38:39 297 63 INR olive 748-73-1579 1 3 3
BK5788246325 58230837 45457227 2.4.27 Android 13.707887 2 113.49995 54.38129 -18.437752 2020-03-24 01:30:15 2020
-05-10 11:16:45 932 32 INR white 558-80-6346 3 2 2
BK0342703255 84232510 86494681 4.1.34 Android -6.091461 -114.64979 22.84495 70.137825 2020-08-03 19:10:52 2020
-03-24 08:25:40 260 7 INR blue 068-72-1637 3 3 3
BK6015582453 11981042 35862658 2.4.39 iOS -18.910034 -70.1931 -10.182921 173.87721 2020-07-17 05:33:48 2020
-04-30 04:54:27 907 53 INR purple 102-10-5639 3 2 3
BK4529355854 60071878 78022360 2.1.9 iOS 1.215274 -56.014904 35.152878 104.324905 2020-01-02 01:48:40 2020
-02-16 04:28:55 547 17 INR teal 866-83-4349 2 3 4
BK9720088219 14327312 94427067 3.1.2 Android -55.482224 173.36226 65.01212 51.39075 2020-04-10 15:11:07 2020
-01-20 21:17:42 259 33 INR maroon 572-73-6526 3 3 2
BK7157532607 46407210 43160003 1.3.4 Android 46.005844 -16.826145 7.6126013 -156.42857 2020-06-09 05:56:31 2020
-03-19 01:53:16 787 21 INR olive 667-23-5880 2 2 3
BK5014871433 65861573 64708618 1.3.28 iOS -29.565327 64.84371 84.06811 -49.820835 2020-08-14 20:43:42 2020
-06-03 09:39:59 586 5 INR fuchsia 255-52-5654 5 5 1
BK9051488736 37721758 27297770 2.3.13 Android 61.936462 83.2497 0.0281895 115.4691 2020-04-07 04:27:59 2020-09-29 1
0:51:41 912 80 INR aqua 739-09-9569 2 1 2
BK243762319 62552969 45877457 3.3.9 iOS -62.651516 -139.15402 28.029999 -62.8556 2020-07-01 00:36:05 2020
-09-30 17:40:23 821 23 INR black 590-44-6613 2 3 4
BK4683595168 56801961 53401707 4.2.34 iOS -5.860265 -100.00404 25.01659 70.47136 2020-05-03 10:17:56 2020
-06-08 09:11:27 71 10 INR fuchsia 454-04-0608 5 2 3
BK9783284253 66909721 40509554 2.2.22 Android 36.191315 5.686264 88.908396 36.5806 2020-03-05 16:02:01 2020-05-29 1
3:36:15 26 81 INR black 600-17-7043 3 1 3
BK2880021380 50163555 34005420 3.4.23 Android -83.06599 106.26869 8.830086 74.87235 2020-01-15 02:00:07 2020
-05-12 21:53:04 571 99 INR navy 506-09-4981 1 5 3
BK4537426043 91111754 59250769 1.2.19 iOS -43.118843 -99.93572 3.7026224 -46.828716 2020-04-28 05:18:34 2020
-02-12 11:31:40 650 81 INR white 362-35-8054 5 5 2
BK998130731 67875357 14562526 3.1.15 Android -10.861959 -111.98985 57.23312 95.469986 2020-01-25 01:37:22 2020
-04-28 09:42:00 590 3 INR teal 359-51-9362 1 1 4
BK5645323730 18442993 84939946 3.1.29 Android -81.47224 -88.404915 12.690818 -140.99768 2020-09-24 08:18:31 2020
-07-16 05:12:24 515 1 INR blue 024-35-8771 1 3 4
BK6163608413 36591778 11946210 4.2.38 Android 60.20364 120.9885 32.103264 -50.551888 2020-07-26 06:12:56 2020
-04-23 06:57:20 810 58 INR silver 833-16-1378 3 5 1
```

```

hive> select * from clicking_stream limit 20;
OK
26564820      3.2.35  Android 16.4454865      99.902065      de545711-3914-4450-8c11-b17b8dabb5e1      fcba68aa-1231-11eb-adc1-0242ac120002      No      YesN
o      Yes      2020-09-14 09:59:07
31906387      2.4.7   iOS      -64.813749      -133.52704      de545711-3914-4450-8c11-b17b8dabb5e1      a95dd57b-779f-49db-819d-b6960483e554      No      No Y
es      Yes      2020-05-16 16:30:21
25713677      3.4.12  Android 89.943435      127.313415      b328829e-17ae-11eb-adc1-0242ac120002      fcba68aa-1231-11eb-adc1-0242ac120002      No      No Y
es      No      2020-02-09 00:52:13
83474293      3.1.8   Android -69.93907      -36.45167      e7bc5fb2-1231-11eb-adc1-0242ac120002      e1e99492-17ae-11eb-adc1-0242ac120002      Yes      No Y
es      No      2020-06-17 10:42:50
63727807      2.2.9   iOS      64.082108      -81.822078      e7bc5fb2-1231-11eb-adc1-0242ac120002      fcba68aa-1231-11eb-adc1-0242ac120002      No      YesY
es      Yes      2020-07-06 02:51:53
73737907      4.3.19  Android -18.850508      -116.358375      b328829e-17ae-11eb-adc1-0242ac120002      e1e99492-17ae-11eb-adc1-0242ac120002      No      YesN
o      Yes      2020-04-26 06:18:16
36927433      3.2.26  iOS      -84.6857245      -146.507678      de545711-3914-4450-8c11-b17b8dabb5e1      a95dd57b-779f-49db-819d-b6960483e554      Yes      YesN
o      Yes      2020-02-06 10:21:18
12691783      3.3.11  Android 54.3852925      -37.411814      de545711-3914-4450-8c11-b17b8dabb5e1      e1e99492-17ae-11eb-adc1-0242ac120002      Yes      YesN
o      No      2020-08-08 04:23:56
22635021      4.4.36  iOS      -31.8055      150.65565      e7bc5fb2-1231-11eb-adc1-0242ac120002      a95dd57b-779f-49db-819d-b6960483e554      No      No N
o      No      2020-08-02 00:33:50
23593546      1.2.16  Android 8.0918475      -83.929878      de545711-3914-4450-8c11-b17b8dabb5e1      e1e99492-17ae-11eb-adc1-0242ac120002      Yes      No Y
es      No      2020-07-23 23:59:10
16929816      1.2.25  Android -34.8576385      -136.639835      e7bc5fb2-1231-11eb-adc1-0242ac120002      fcba68aa-1231-11eb-adc1-0242ac120002      No      No N
o      Yes      2020-05-13 16:35:17
41929865      1.1.14  iOS      -34.3613155      -137.467833      b328829e-17ae-11eb-adc1-0242ac120002      a95dd57b-779f-49db-819d-b6960483e554      No      YesN
o      No      2020-04-13 07:58:26
68940320      1.1.17  Android 81.738239      -104.207763      de545711-3914-4450-8c11-b17b8dabb5e1      fcba68aa-1231-11eb-adc1-0242ac120002      Yes      No Y
es      No      2020-10-14 00:39:05
95485928      1.2.18  Android 44.185627      78.125488      e7bc5fb2-1231-11eb-adc1-0242ac120002      e1e99492-17ae-11eb-adc1-0242ac120002      No      No N
o      No      2020-07-03 19:15:34
56235860      1.4.16  Android -61.178407      8.534126      de545711-3914-4450-8c11-b17b8dabb5e1      a95dd57b-779f-49db-819d-b6960483e554      Yes      YesY
es      Yes      2020-06-02 21:32:44
81758845      4.4.30  Android 78.904984      167.315596      e7bc5fb2-1231-11eb-adc1-0242ac120002      a95dd57b-779f-49db-819d-b6960483e554      Yes      YesN
o      No      2020-04-17 19:52:56
46822067      1.2.26  Android 45.8993985      135.895118      de545711-3914-4450-8c11-b17b8dabb5e1      fcba68aa-1231-11eb-adc1-0242ac120002      Yes      YesY
es      Yes      2020-08-09 12:02:09
37623067      3.2.27  iOS      -74.1131505      -170.715421      b328829e-17ae-11eb-adc1-0242ac120002      a95dd57b-779f-49db-819d-b6960483e554      No      YesN
o      No      2020-01-20 00:06:06
14571693      3.4.21  iOS      -28.8329055      161.132369      e7bc5fb2-1231-11eb-adc1-0242ac120002      fcba68aa-1231-11eb-adc1-0242ac120002      No      YesN
o      No      2020-08-16 11:20:40

```

```

hive> select * from datewise_bookings limit 20;
OK
2020-01-01      1
2020-01-02      3
2020-01-03      2
2020-01-04      2
2020-01-05      2
2020-01-06      3
2020-01-07      2
2020-01-08      4
2020-01-09      2
2020-01-10      2
2020-01-11      3
2020-01-12      3
2020-01-14      2
2020-01-15      5
2020-01-16      3
2020-01-17      4
2020-01-18      4
2020-01-20      4
2020-01-21      1
2020-01-23      4
Time taken: 0.073 seconds, Fetched: 20 row(s)

```