# open git bash 1.

ssh asekows@129.150.69.91

hdfs dfs -mkdir tmp/drugs

hdfs dfs -chmod -R o+w tmp/

# open git bash 2.

ssh asekows@129.150.69.91

beeline

!connect jdbc:hive2://bigdai-nov-bdcsce-1:2181,bigdai-nov-bdcsce-2:2181,bigdai-nov-bdcsce-3:2181/;serviceDiscoveryMode=zooKeeper;zooKeeperNamespace=hiveserver2?tez.queue.name=interactive bdcsce\_admin

use asekows;

CREATE TABLE drugs\_year AS SELECT reporter\_state, drug\_name, quantity, REVERSE(SUBSTR(REVERSE(transaction\_date), 0, 4)) as tdate FROM arcos\_all\_washpost1;

CREATE TABLE IF NOT EXISTS drugs\_state\_download ROW FORMAT DELIMITED FIELDS TERMINATED BY "," STORED AS TEXTFILE LOCATION "/user/asekows/tmp/drugs" AS SELECT reporter\_state, drug\_name, tdate, SUM (quantity) FROM drugs\_year GROUP BY reporter\_state, drug\_name, tdate;

# in git bash 1:

hdfs dfs -get /user/asekows/tmp/drugs/00000\*\_0

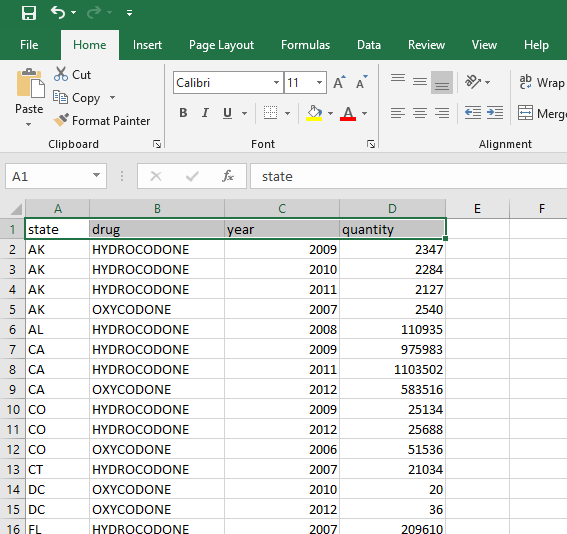
cat 000000\_0 000001\_0 000002\_0 000003\_0 000004\_0 000005\_0 000006\_0 000007\_0 > drugs\_out.csv

# open git bash 3:

scp [asekows@129.150.69.91:/home/asekows/drugs\_out.csv drugs\_out.csv](mailto:asekows@129.150.69.91:/home/asekows/drugs_out.csv%20drugs_out.csv)

# Open downloaded file in Excel

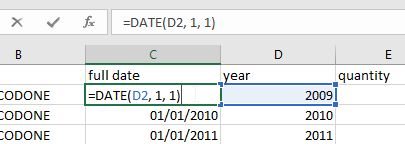
Add column headings



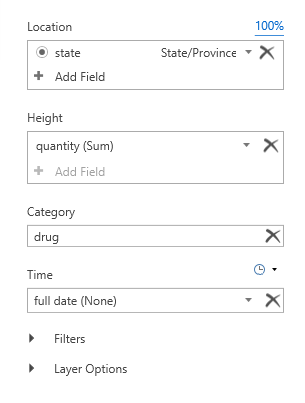
Save as xlsx

Add a new column “full date”. We need to create a formatted date column that we can use for timeline.

Use following formula for the “full date” column:



Create 3d Map and use:



# Result

