

HW11+_additional_codes_output

Environment: Dumbo- Hive

beeline

!connect jdbc:hive2://babar.es.its.nyu.edu:10000/

use *your_netid*;

In order to run the codes below, there should be 8 tables created in your database
show tables;

should show tables like below

| tab_name |
|--------------|
| sw_cc |
| sw_cs |
| sweden_groc |
| sweden_parks |
| sweden_res |
| taiwan_groc |
| taiwan_parks |
| taiwan_res |
| tw_cc |
| tw_cs |

ALL mobility data in ONE TABLE

select sg.date, sg.val as sweden_grocery, tg.val as taiwan_grocery,
sr.val as sweden_residential, tr.val as taiwan_residential,
sp.val as sweden_outdoor, tp.val as taiwan_outdoor
FROM sweden_groc sg

INNER JOIN taiwan_groc tg on tg.date=sg.date

INNER JOIN sweden_res sr on sr.date=sg.date

INNER JOIN taiwan_res tr on tr.date=sg.date

INNER JOIN sweden_parks sp on sp.date=sg.date

INNER JOIN taiwan_parks tp on tp.date=sg.date;

Sample output: (very long)

| sg.date | sweden_grocery | taiwan_grocery | sweden_residential | taiwan_residential | sweden_outdoor | taiwan_outdoor |
|------------|----------------------|----------------------|--------------------|--------------------|--------------------|---------------------|
| 2020-02-17 | -3.3329999446868896 | -5.0 | 1.3329999446868896 | 3.3329999446868896 | -10.0 | -12.333000183105469 |
| 2020-02-18 | -2.5 | -4.25 | 1.25 | 3.75 | -4.25 | -10.0 |
| 2020-02-19 | -1.7999999523162842 | -4.199999809265137 | 1.2000000476837158 | 3.7999999523162842 | 1.7999999523162842 | -6.400000095367432 |
| 2020-02-20 | -1.3329999446868896 | -4.333000183105469 | 1.1670000553131104 | 3.3329999446868896 | 1.3329999446868896 | -6.0 |
| 2020-02-21 | -1.143000066757202 | -3.7139999866485596 | 1.143000066757202 | 3.571000099182129 | 4.285999774932861 | -4.285999774932861 |
| 2020-02-22 | -1.2860000133514404 | -3.0 | 1.2860000133514404 | 4.14300012588501 | 2.428999900817871 | -3.571000099182129 |
| 2020-02-23 | -0.28600001335144043 | -1.2860000133514404 | 1.0 | 3.7139999866485596 | 10.0 | 2.8570001125335693 |
| 2020-02-24 | 0.14300000667572021 | -0.7139999866485596 | 0.8569999933242798 | 3.428999900817871 | 13.857000350952148 | 5.714000225067139 |
| 2020-02-25 | 1.4290000200271606 | -0.42899999022483826 | 0.8569999933242798 | 3.0 | 13.142999649047852 | 6.714000225067139 |
| 2020-02-26 | 1.5709999799728394 | -0.28600001335144043 | 0.8569999933242798 | 2.7139999866485596 | 10.0 | 6.0 |
| 2020-02-27 | 2.0 | 0.28600001335144043 | 0.8569999933242798 | 2.1429998874664307 | 11.571000099182129 | 5.285999774932861 |
| 2020-02-28 | 2.2860000133514404 | 1.8569999933242798 | 0.8569999933242798 | 3.1429998874664307 | 10.71399974822998 | 13.71399974822998 |
| 2020-02-29 | 3.0 | 1.8569999933242798 | 0.7139999866485596 | 3.1429998874664307 | 13.28600025177002 | 17.714000701904297 |
| 2020-03-01 | 2.7139999866485596 | 1.5709999799728394 | 0.8569999933242798 | 3.0 | 9.0 | 17.285999298095703 |

BROAD COMPARISON - avg value for 2020 per monthly - all outdoor, residential, grocery data

select MONTH(TO_DATE(FROM_UNIXTIME(UNIX_TIMESTAMP(sg.date,
'yyyy-MM-dd')))) as month,

```

avg(sg.val) as grocery_SW , avg(tg.val) as grocery_TG,
avg(sr.val) as res_SW , avg(tr.val) as res_TG,
avg(sp.val) as outdoor_SW , avg(tp.val) as outdoor_TG
FROM sweden_groc sg
INNER JOIN taiwan_groc tg on tg.date=sg.date
INNER JOIN sweden_res sr on sr.date=sg.date
INNER JOIN taiwan_res tr on tr.date=sg.date
INNER JOIN sweden_parks sp on sp.date=sg.date
INNER JOIN taiwan_parks tp on tp.date=sg.date
GROUP BY MONTH(TO_DATE(FROM_UNIXTIME(UNIX_TIMESTAMP(sg.date,
'yyyy-MM-dd'))));

```

| month | grocery_sw | grocery_tg | res_sw | res_tg | outdoor_sw | outdoor_tg |
|-------|----------------------|---------------------|--------------------|---------------------|--------------------|---------------------|
| 2 | -0.09630768115703876 | -1.7855384601996496 | 1.0290769246908336 | 3.3243076617901144 | 6.012999974764311 | 1.1853077411651611 |
| 3 | -0.06912902478248842 | 1.8202258079282698 | 4.958516147828871 | 2.2580967526282034 | 16.728129156174198 | -2.1106129461719143 |
| 4 | -5.15723333756129 | 0.895199986298879 | 9.785633325576782 | 4.32850002447764 | 63.461966323852536 | -6.157133344809214 |
| 5 | -1.6083548338182512 | 2.8663548433011576 | 7.672838672514884 | 2.6634638811812864 | 82.02306464410597 | -7.834032231761563 |
| 6 | 4.4952666640281675 | 6.385633345444997 | 3.68566672514884 | 1.642833361943563 | 187.45240020751953 | -0.9809667110443115 |
| 7 | 3.7188387224751134 | 8.88016128540039 | 5.967774175828593 | 1.5347096881558817 | 233.78338524603075 | 3.5022580729376886 |
| 8 | 0.8064193485244628 | 6.359451624654954 | 2.8294516032741917 | 1.6083548338182512 | 228.3594215146957 | 1.7327741730597712 |
| 9 | -2.747666676839193 | 2.542833352088928 | 2.5713666796684267 | 0.2198333167235056 | 94.45243326822917 | -4.2333666751782095 |
| 10 | -2.2395806552902346 | 5.133580619289029 | 4.013806443060598 | 0.6820322582798619 | 43.61290347191595 | 1.5529999963698848 |
| 11 | -6.059499993920326 | 2.9107499917348227 | 7.791666686534882 | 0.17845832804838815 | 19.87504158417384 | -5.732083360354106 |

GROCERY DATA AVG -PER MONTH COMPARISON (TW vs SW)

```

select
MONTH(TO_DATE(FROM_UNIXTIME(UNIX_TIMESTAMP(sg.date, 'yyyy-MM-dd')))) as
month,
avg(sg.val) as average_groc_valS,
avg(tg.val) as average_groc_valT
from sweden_groc sg
inner join taiwan_groc tg on tg.date=sg.date
group by MONTH(TO_DATE(FROM_UNIXTIME(UNIX_TIMESTAMP(sg.date,
'yyyy-MM-dd'))));

```

| month | average_groc_val_s | average_groc_val_t |
|-------|----------------------|---------------------|
| 2 | -0.09630768115703876 | -1.7855384601996496 |
| 3 | -0.06912902478248842 | 1.8202258079282698 |
| 4 | -5.15723333756129 | 0.895199986298879 |
| 5 | -1.6083548338182512 | 2.8663548433011576 |
| 6 | 4.4952666640281675 | 6.385633345444997 |
| 7 | 3.7188387224751134 | 8.88016128540039 |
| 8 | 0.8064193485244628 | 6.359451624654954 |
| 9 | -2.747666676839193 | 2.542833352088928 |
| 10 | -2.2395806552902346 | 5.133580619289029 |
| 11 | -6.059499993920326 | 2.9107499917348227 |

OUTDOOR(PARK) DATA -AVG PER MONTH COMPARISON (TW vs SW)

```
select
```

```

MONTH(TO_DATE(FROM_UNIXTIME(UNIX_TIMESTAMP(sg.date, 'yyyy-MM-dd')))) as
month,
avg(sg.val) as average_parks_val_S,
avg(tg.val) as average_parks_val_T
from sweden_parks sg
inner join taiwan_parks tg on tg.date=sg.date
group by MONTH(TO_DATE(FROM_UNIXTIME(UNIX_TIMESTAMP(sg.date,
'yyyy-MM-dd'))));

```

| month | average_parks_val_s | average_parks_val_t |
|-------|---------------------|---------------------|
| 2 | 6.012999974764311 | 1.1853077411651611 |
| 3 | 16.728129156174198 | -2.1106129461719143 |
| 4 | 63.461966323852536 | -6.157133344809214 |
| 5 | 82.02306464410597 | -7.834032231761563 |
| 6 | 187.45240020751953 | -0.9809667110443115 |
| 7 | 233.78338524603075 | 3.5022580729376886 |
| 8 | 228.3594215146957 | 1.7327741730597712 |
| 9 | 94.45243326822917 | -4.2333666751782095 |
| 10 | 43.61290347191595 | 1.5529999963698848 |
| 11 | 19.87504158417384 | -5.732083360354106 |

RESIDENTIAL DATA -AVG PER MONTH COMPARISON (TW vs SW)

```

select
MONTH(TO_DATE(FROM_UNIXTIME(UNIX_TIMESTAMP(sg.date, 'yyyy-MM-dd')))) as
month,
avg(sg.val) as average_res_val_S,
avg(tg.val) as average_res_val_T
from sweden_res sg
inner join taiwan_parks tg on tg.date=sg.date
group by MONTH(TO_DATE(FROM_UNIXTIME(UNIX_TIMESTAMP(sg.date,
'yyyy-MM-dd'))));

```

| month | average_res_val_s | average_res_val_t |
|-------|--------------------|---------------------|
| 2 | 1.0290769246908336 | 1.1853077411651611 |
| 3 | 4.958516147828871 | -2.1106129461719143 |
| 4 | 9.785633325576782 | -6.157133344809214 |
| 5 | 7.672838672514884 | -7.834032231761563 |
| 6 | 3.6856666723887126 | -0.9809667110443115 |
| 7 | 5.967774175828503 | 3.5022580729376886 |
| 8 | 2.8294516032741917 | 1.7327741730597712 |
| 9 | 2.5713666796684267 | -4.2333666751782095 |
| 10 | 4.013806443060598 | 1.5529999963698848 |
| 11 | 7.791666686534882 | -5.732083360354106 |

BROAD_COMPARISON- ALL IN ONE TABLE -monthly SUM value- all outdoor, residential, grocery data

```

select MONTH(TO_DATE(FROM_UNIXTIME(UNIX_TIMESTAMP(sg.date,
'yyyy-MM-dd')))) as month,
sum(sg.val) as grocery_SW , sum(tg.val) as grocery_TG,

```

```

sum(sr.val) as res_SW , sum(tr.val) as res_TG,
sum(sp.val) as outdoor_SW ,sum(tp.val) as outdoor_TG
FROM sweden_groc sg
INNER JOIN taiwan_groc tg on tg.date=sg.date
INNER JOIN sweden_res sr on sr.date=sg.date
INNER JOIN taiwan_res tr on tr.date=sg.date
INNER JOIN sweden_parks sp on sp.date=sg.date
INNER JOIN taiwan_parks tp on tp.date=sg.date
group by MONTH(TO_DATE(FROM_UNIXTIME(UNIX_TIMESTAMP(sg.date,
'yyyy-MM-dd'))));

```

| month | grocery_sw | grocery_tg | res_sw | res_tg | outdoor_sw | outdoor_tg |
|-------|---------------------|--------------------|--------------------|--------------------|--------------------|---------------------|
| 2 | -1.251999855041504 | -23.21199982595444 | 13.378000020980835 | 43.215999603271484 | 78.16899967193604 | 15.409000635147095 |
| 3 | -2.142999768257141 | 56.42700004577637 | 153.714000582695 | 70.0009993314743 | 518.5720038414001 | -65.42900133132935 |
| 4 | -154.71700012683868 | 26.85599958896637 | 293.56899976730347 | 129.85500073432922 | 1903.8589897155762 | -184.71400034427643 |
| 5 | -49.858999848365784 | 88.85700014233589 | 237.85799884796143 | 82.56800031661987 | 2542.715003967285 | -242.85499918460846 |
| 6 | 134.85799992084503 | 191.56900036334991 | 110.57000017166138 | 49.28500008583069 | 5623.572006225586 | -29.429001331329346 |
| 7 | 115.28400039672852 | 275.2849998474121 | 185.0009994506836 | 47.576000332832336 | 7247.284942626953 | 108.5700026106834 |
| 8 | 24.998999804258347 | 197.1430003643036 | 87.71299970149994 | 49.858999848365784 | 7079.142066955566 | 53.715999364852905 |
| 9 | -82.43000030517578 | 76.28500056266785 | 77.1410003900528 | 6.570999950170517 | 2833.572998046875 | -127.0010002553463 |
| 10 | -69.427000031399727 | 159.1409991979599 | 124.42799973487854 | 21.14300000667572 | 1352.0000076293945 | 48.14299988746643 |
| 11 | -145.42799985408783 | 69.85799980163574 | 187.00000047683716 | 4.282999873161316 | 477.0009980201721 | -137.57000064849854 |

10 rows selected (50.824 seconds)

FINDING monthly MAX values-all outdoor, residential, grocery data (sweden vs taiwan)

```

select MONTH(TO_DATE(FROM_UNIXTIME(UNIX_TIMESTAMP(sg.date,
'yyyy-MM-dd')))) as month,
max(sg.val) as sweden_grocery_max,
max(tg.val) as taiwan_grocery_max,
max(sr.val) as sweden_resident_max,
max(tr.val) as taiwan_resident_max,
max(sp.val) as sweden_outdoor_max,
max(tp.val) as taiwan_outdoor_max
FROM sweden_groc sg
INNER JOIN taiwan_groc tg on tg.date=sg.date
INNER JOIN sweden_res sr on sr.date=sg.date
INNER JOIN taiwan_res tr on tr.date=sg.date
INNER JOIN sweden_parks sp on sp.date=sg.date
INNER JOIN taiwan_parks tp on tp.date=sg.date
group by MONTH(TO_DATE(FROM_UNIXTIME(UNIX_TIMESTAMP(sg.date,
'yyyy-MM-dd'))));

```

| month | sweden_grocery_max | taiwan_grocery_max | sweden_resident_max | taiwan_resident_max | sweden_outdoor_max | taiwan_outdoor_max |
|-------|---------------------|--------------------|---------------------|---------------------|--------------------|---------------------|
| 2 | 3.0 | 1.8569999933242798 | 1.3329999446868896 | 4.14300012588501 | 13.857000350952148 | 17.714000701904297 |
| 3 | 4.85699987411499 | 5.85699987411499 | 9.571000099182129 | 3.1429998874664307 | 38.42900085449219 | 17.285999298095703 |
| 4 | 0.8569999933242798 | 2.428999900817871 | 12.857000350952148 | 7.428999900817871 | 97.0 | 4.428999900817871 |
| 5 | 5.428999900817871 | 6.85699987411499 | 9.71399974822998 | 4.14300012588501 | 140.85699462890625 | 9.0 |
| 6 | 11.571000099182129 | 10.28600025177002 | 5.0 | 3.7139999866485596 | 268.2860107421875 | 19.14299964904785 |
| 7 | 5.85699987411499 | 12.0 | 7.14300012588501 | 4.0 | 319.4289855957031 | 15.71399974822998 |
| 8 | 3.428999900817871 | 8.428999900817871 | 5.85699987411499 | 2.571000099182129 | 304.5710144042969 | 7.714000225067139 |
| 9 | -1.0 | 7.0 | 3.2860000133514404 | 2.2860000133514404 | 110.0 | 0.14300000667572021 |
| 10 | 1.0 | 9.571000099182129 | 5.85699987411499 | 2.2860000133514404 | 65.28600311279297 | 15.857000350952148 |
| 11 | -1.4290000200271606 | 4.14300012588501 | 9.571000099182129 | 0.8569999933242798 | 36.2859992980957 | 2.2860000133514404 |

FINDING monthly MIN values-all outdoor, residential, grocery data (sweden vs taiwan)

```
select MONTH(TO_DATE(FROM_UNIXTIME(UNIX_TIMESTAMP(sg.date,
'yyyy-MM-dd')))) as month,
min(sg.val) as sweden_grocery_min,
min(tg.val) as taiwan_grocery_min,
min(sr.val) as sweden_resident_min,
min(tr.val) as taiwan_resident_min,
min(sp.val) as sweden_outdoor_min,
min(tp.val) as taiwan_outdoor_min
FROM sweden_groc sg
INNER JOIN taiwan_groc tg on tg.date=sg.date
INNER JOIN sweden_res sr on sr.date=sg.date
INNER JOIN taiwan_res tr on tr.date=sg.date
INNER JOIN sweden_parks sp on sp.date=sg.date
INNER JOIN taiwan_parks tp on tp.date=sg.date
group by MONTH(TO_DATE(FROM_UNIXTIME(UNIX_TIMESTAMP(sg.date,
'yyyy-MM-dd'))));
```

| month | sweden_grocery_min | taiwan_grocery_min | sweden_resident_min | taiwan_resident_min | sweden_outdoor_min | taiwan_outdoor_min |
|-------|---------------------|---------------------|---------------------|----------------------|--------------------|---------------------|
| 2 | -3.3329999446868896 | -5.0 | 0.7139999866485596 | 2.1429998874664307 | -10.0 | -12.333000183105469 |
| 3 | -5.571000099182129 | -0.7139999866485596 | 0.8569999933242798 | 1.4290000200271606 | 5.14300012588501 | -10.571000099182129 |
| 4 | -12.28600025177002 | -0.7139999866485596 | 7.571000099182129 | 2.8570001125335693 | 22.429000854492188 | -14.142999649047852 |
| 5 | -6.571000099182129 | -3.2860000133514404 | 4.285999774932861 | 1.4290000200271606 | 44.143001556396484 | -25.714000701904297 |
| 6 | -1.5709999799728394 | 1.8569999933242798 | 2.2860000133514404 | 0.14300000667572021 | 116.42900085449219 | -12.142999649047852 |
| 7 | 2.571000099182129 | 4.285999774932861 | 3.2860000133514404 | 1.1430000066757202 | 133.0 | -6.14300012588501 |
| 8 | -1.0 | 3.0 | 1.2860000133514404 | 0.8569999933242798 | 105.57099914550781 | -6.571000099182129 |
| 9 | -4.14300012588501 | -0.5709999799728394 | 1.7139999866485596 | -0.7139999866485596 | 66.85700225830078 | -12.857000350952148 |
| 10 | -4.0 | 1.7139999866485596 | 2.8570001125335693 | -0.42899999022483826 | 22.0 | -9.0 |
| 11 | -8.428999900817871 | 1.7139999866485596 | 5.85699987411499 | -0.42899999022483826 | 4.14300012588501 | -13.428999900817871 |

STANDARD DEVIATION - standard deviation- for all outdoor, residential, grocery data

select STDDEV(groc) as standard_deviation from sweden_groc;

```
+-----+
| standard_deviation |
+-----+
| 3.9069031426705645 |
+-----+
1 row selected (30.846 seconds)
```

select STDDEV(res) as standard_deviation from sweden_res;

```
+-----+
| standard_deviation |
+-----+
| 2.9586431071308072 |
+-----+
1 row selected (30.588 seconds)
```

select STDDEV(parks) as standard_deviation from sweden_parks;

```

+-----+
| standard_deviation |
+-----+
| 90.2448010600201   |
+-----+
1 row selected (29.017 seconds)

```

select STDDEV(groc) as standard_deviation from taiwan_groc;

```

+-----+
| standard_deviation |
+-----+
| 3.358274984703127  |
+-----+
1 row selected (29.102 seconds)

```

select STDDEV(res) as standard_deviation from taiwan_res;

```

+-----+
| standard_deviation |
+-----+
| 1.4776547816898955 |
+-----+
1 row selected (29.455 seconds)

```

select STDDEV(parks) as standard_deviation from taiwan_parks;

```

+-----+
| standard_deviation |
+-----+
| 7.814206553780813  |
+-----+
1 row selected (28.972 seconds)

```