Homework 4

Jiahui Guan 995973611
Yepei Zhang 912465089

1. We create 4 tables for NHTS Data files, 1 table for EIA data, and use all the records. Although the process is time-consuming, it will not miss any data. The relations between the NHTS Data files are shown as the following diagram. We can apply Referential Integrity for the future update. The left side of the arrow is contains "referenced attributes" and right side of the arrow contains the "referencing attributes", which are the common variables of two files. Then, set the referential integrity options following Cascade Policy and Set-Null Policy so that when we update the referencing attribute, the referenced attribute will be update at the same time. However, this method is "one-way update". That is, updating the information of left side table will not effect the right side table. Moreover, when we update NHTS records, we can apply "TRANSACTIONS" to make sure the data update at the same time. There is no constrain for updating EIA Data.
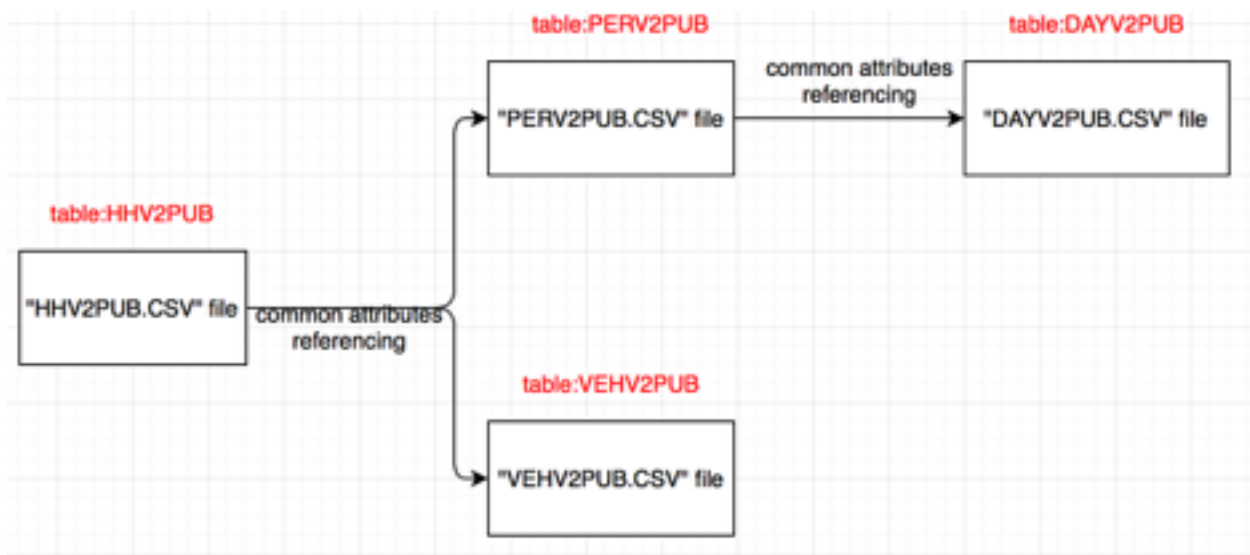
> For NHTS Data:
> DAYV2PUB(_, _, ….,_)
> FERV2PUB(_, _, ….,_)
> VEHV2PUB(_, _, ….,_)
> HHV2PUB(_, _, ….,_)
>
> For EIA Data:
> EIA (MSN, YYYYMM, Value, Column_Order, Description, Unit)

The common attributes of PERV2PUB and DAYV2PUB (49 attributes):

HOUSEID, PERSONID, FRSTHM, OUTOFTWN, HH_HISP, HH_RACE, DRIVER, R_SEX, WORKER, DRVRCNT, HHFAMINC, HHSIZE, HHVEHCNT, NUMADLT, FLAG100, LIF_CYC, CDIVMSAR, CENSUS_D, CENSUS_R, EDUC, HHRESP, HHSTATE, HHSTFIPS, MSACAT, MSASIZE, PRMACT, PROXY, R_AGE, RAIL, TRAVDAY, URBAN, URBANSIZE, URBRUR, USEINTST, USEPUBTR,WRKCOUNT,TDAYDATE,HOMEOWN,HOMETYPE,HBHUR, HTRESDN, HTHTNRNT, HTPPOPDN, HTEEMPDN, HBRESDN, HBHTNRNT, HBPPOPDN, HH_CBSA, HHC_MSA

The common attributes of PERV2PUB and HHV2PUB(37 attributes):
HOUSEID, HH_HISP, HH_RACE, DRVRCNT, HHFAMINC, HHSIZE, HHVEHCNT, NUMADLT, FLAG100, LIF_CYC, CDIVMSAR, CENSUS_D, CENSUS_R, HHRESP, HHSTATE, HHSTFIPS, MSACAT, MSASIZE, RAIL, TRAVDAY, URBAN, URBANSIZE, URBRUR, WRKCOUNT, TDAYDATE, HOMEOWN, HOMETYPE, HBHUR, HTRESDN, HTHTNRNT, HTPPOPDN, HTEEMPDN, HBRESDN, HBHTNRNT, HBPPOPDN, HH_CBSA, HHC_MSA

The common attributes of VEHV2PUB and HHV2PUB (37 attributes):

HOUSEID, WTHHFIN, DRVRCNT, HHFAMINC, HHSIZE, HHVEHCNT, NUMADLT, FLAG100, CDIVMSAR, CENSUS_D, CENSUS_R, HHSTATE, HHSTFIPS, MSACAT, MSASIZE, RAIL, TRAVDAY, URBAN, URBANSIZE, URBRUR, WRKCOUNT, TDAYDATE, HH_HISP, HH_RACE, HOMEOWN, HOMETYPE, LIF_CYC, HBHUR, HTRESDN, HTHTNRNT, HTPPOPDN, HTEEMPDN, HBRESDN, HBHTNRNT, HBPPOPDN, HH_CBSA, HHC_MSA

2.
We choose Python 3.5.1 to load all data into Postgresql, the codes are shown file: Codes.

3.
(a)

< 5	< 10	< 15	< 20	< 25	< 30	< 35	< 40	< 45	< 50
12.64%	23.21%	35.30%	43.83%	52.92%	58.91%	65.23%	69.50%	73.89%	76.82%
< 55	< 60	< 65	< 70	< 75	< 80	< 85	< 90	< 95	< 100
80.00%	82.03%	84.34%	85.82%	87.40%	88.53%	89.70%	90.54%	91.43%	92.02%

(b)

< 5	< 10	< 15	< 20	< 25	< 30	< 35	< 40	< 45	< 50
26.3980	26.3055	26.2870	26.2878	26.3006	26.3353	26.3447	26.3693	26.3847	26.3915
< 55	< 60	< 65	< 70	< 75	< 80	< 85	< 90	< 95	< 100
26.3957	26.4026	26.4096	26.4066	26.4159	26.4119	26.4067	26.4025	26.3969	26.4009

(c)

200803	200804	200805	200806	200807	200808	200809
51.1112%	56.72750%	55.66593%	59.58241%	61.26975%	61.54795%	59.5485%
200810	200811	200812	200901	200902	200903	200904
56.02017%	60.7717%	60.3811%	58.5087%	65.0865%	60.6595%	67.1607%

(d) order by test date from 200803 to 200904:

When X is 20 miles:
0.18975593238505675
0.21218732104295
0.1892405208603691
0.2064626028961217
0.15779288929389523
0.15776322765271739
0.19374300865905422
0.1737326588333338
0.1905798199388375
0.16940229401591902
0.18763800676048664
0.27800084211887105
0.2066829441097728
0.2272572044124278

When X is 40 miles:
0.22756891262223186
0.24708192948316943
0.22676730251320804
0.2426474831718979
0.1935546557909155
0.1934148043464966
0.22478971800265338
0.20800556927079855
0.22292348363996167
0.20426477653929942
0.2226766834532904
0.3126272330551132
0.24834393174742814
0.26963230755964085

When X is 60 miles:
0.24120708050862663
0.26050433491776215
0.2418596457818169
0.25853743967798487
0.2094484594498215
0.2091491772158931
0.23723839970324537
0.22208836071653573
0.2371782488510729
0.2182523775573295
0.23671189233150222
0.32600210627170445
0.26599867663472554
0.2880116969685234