Gurteg Singh
ECN 145
PS3

1.) Log fiile, etc.

2.)

	hour	dow	speed	volume	density	latitude	longitude
count	16688.000000	16688.000000	16688.000000	16688.000000	16688.000000	16688.000000	16688.000000
mean	11.581675	3.032838	26.309771	1136.211170	52.813201	41.873955	-87.722030
std	6.914101	1.984992	6.311214	526.586662	43.620730	0.000185	0.000958
min	0.000000	0.000000	1.601355	12.000000	0.305129	41.873770	-87.722990
25%	6.000000	1.000000	26.128127	696.000000	23.755864	41.873770	-87.722990
50%	12.000000	3.000000	28.028109	1272.000000	47.454847	41.873800	-87.721120
75%	18.000000	5.000000	29.575019	1536.000000	60.606186	41.874140	-87.721120
max	23.000000	6.000000	94.285750	2436.000000	329.720700	41.874150	-87.719710

In this dataset, we are given 16,688 observations.

3.) Given as tables: Speed, Density, and Volume. (Weekday West, Weekday East, Weekend East, Weekend West) ~ respectively.

	volume	density	speed
	mean	mean	mean
hour			
0	717.120000	24.047256	29.908310
1	381.418327	12.639181	30.213591
2	271.210762	8.842141	30.755575
3	202.272000	6.634341	30.626803
4	263.285714	8.430507	31.493004
5	467.376000	14.886739	31.471577
6	1020.286853	32.875283	31.044235
7	1600.685259	60.568092	28.316333
8	1277.264822	107.209929	18.654843
9	1214.476190	49.449569	28.135879
10	1279.344000	45.777560	28.612860
11	1372.695652	48.877130	28.136247
12	1434.877470	50.925178	28.262708
13	1497.168000	55.702042	27.602356
14	1599.325301	73.146679	24.708708
15	1482.741036	123.238704	15.713682
16	1287.571429	167.808139	9.196917
17	1198.904762	178.725376	8.459885
18	1326.142857	154.332831	10.939984
19	1397.312253	92.347280	20.346458
20	1312.932806	47.611531	27.846385
21	1214.857143	43.970925	28.033382
22	1264.771654	47.181551	27.609065
23	1053.741176	39.906308	28.218666

	volume	density	speed
	mean	mean	mean
hour			
0	495.130435	17.153190	28.910556
1	271.039370	9.347138	29.090033
2	184.143498	6.310785	29.319726
3	144.430279	4.895416	29.730965
4	196.299213	6.534243	30.527254
5	534.000000	17.999198	29.781358
6	1513.328063	54.752082	27.789674
7	1981.565217	86.175328	23.514348
8	1958.164706	110.360446	18.476587
9	1817.196850	106.280942	18.315053
10	1639.952381	72.580408	23.880693
11	1538.086957	61.243160	25.650249
12	1475.102362	57.011662	26.219475
13	1452.000000	56.318273	26.296000
14	1489.632000	57.385507	26.213834
15	1532.964427	58.280191	26.457837
16	1603.446640	65.085751	25.299045
17	1651.636364	79.289826	23.116072
18	1643.525692	82.345025	22.336191
19	1575.685039	64.896967	25.422794
20	1340.031496	49.473470	27.214400
21	1148.632411	41.674090	27.605565
22	1073.882353	39.401210	27.662940
23	860.671875	30.610824	28.286806

	volume	density	speed
	mean	mean	mean
hour			
0	921.840000	32.721071	28.213558
1	590.326531	20.496044	28.856581
2	383.181818	13.177590	29.186067
3	271.733333	9.215081	29.554105
4	183.191489	6.215328	29.884411
5	251.733333	8.483898	29.821774
6	492.000000	16.960507	29.544001
7	819.625000	27.851923	29.684216
8	986.081633	35.590759	28.393577
9	1218.432990	43.629275	28.065836
10	1432.000000	52.349415	27.451965
11	1528.653061	56.510472	27.120492
12	1513.375000	56.607873	26.831362
13	1520.484848	57.135107	26.720161
14	1531.393939	57.865126	26.654797
15	1515.835052	56.952842	26.741823
16	1508.121212	57.543911	26.415657
17	1524.000000	58.821121	26.287674
18	1514.181818	58.039507	26.240440
19	1469.510204	55.565252	26.640942
20	1325.696970	50.199897	26.934701
21	1183.272727	43.495306	27.325851
22	1164.600000	42.825339	27.416261
23	1041.720000	37.726018	27.996343

	volume	density	speed
	mean	mean	mean
hour			
0	1112.969697	39.300787	28.495246
1	797.443299	27.082247	29.544135
2	642.344828	21.318199	30.222795
3	526.516854	17.131918	30.841095
4	427.612903	13.811754	31.114703
5	354.606742	11.412742	31.243478
6	412.926316	13.417084	31.617629
7	592.673684	18.635785	31.908576
8	804.371134	25.935401	31.097374
9	1004.125000	32.712094	30.782448
10	1202.693878	42.424051	29.821434
11	1390.886598	47.538181	29.331545
12	1472.336842	51.871012	28.740248
13	1514.081633	57.370592	27.594508
14	1535.877551	64.175290	26.342339
15	1493.625000	77.197857	23.522992
16	1487.632653	87.483523	21.888223
17	1417.090909	76.803490	23.495805
18	1395.428571	66.415304	24.978515
19	1308.123711	51.044121	27.225582
20	1246.391753	44.296957	28.233427
21	1204.701031	43.993335	27.870778
22	1199.142857	43.339606	27.970266
23	1111.346939	42.573730	27.779128

Here we can see that the main traffic slow downs occur in the morning (rush hour), slightly past noon, and in the evening (people are getting back from work, etc.)

4.)
For the weekend tables (subsequent third and fourth table), we see a subtle decrease in speeds during the afternoon/evening timeframe.

5.)

```
modelW = ols('speed ~ C(hour)', nonwW).fit()
print(modelW.summary())
```

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                  speed   R-squared:                       0.670
Model:                            OLS   Adj. R-squared:                  0.669
Method:                 Least Squares   F-statistic:                     528.6
Date:                Thu, 12 Nov 2020   Prob (F-statistic):               0.00
Time:                        15:46:18   Log-Likelihood:                -18170.
No. Observations:                6012   AIC:                         3.639e+04
Df Residuals:                    5988   BIC:                         3.655e+04
Df Model:                          23
Covariance Type:            nonrobust
=================================================================================
                    coef    std err          t      P>|t|      [0.025      0.975]
---------------------------------------------------------------------------------
Intercept        29.9083      0.315     94.972      0.000      29.291      30.526
C(hour)[T.1]      0.3053      0.445      0.686      0.493      -0.567       1.177
C(hour)[T.2]      0.8473      0.459      1.847      0.065      -0.052       1.746
C(hour)[T.3]      0.7185      0.445      1.613      0.107      -0.155       1.592
C(hour)[T.4]      1.5847      0.444      3.565      0.000       0.713       2.456
C(hour)[T.5]      1.5633      0.445      3.510      0.000       0.690       2.436
C(hour)[T.6]      1.1359      0.445      2.553      0.011       0.264       2.008
C(hour)[T.7]     -1.5920      0.445     -3.578      0.000      -2.464      -0.720
C(hour)[T.8]    -11.2535      0.444    -25.344      0.000     -12.124     -10.383
C(hour)[T.9]     -1.7724      0.444     -3.988      0.000      -2.644      -0.901
C(hour)[T.10]    -1.2954      0.445     -2.909      0.004      -2.169      -0.422
C(hour)[T.11]    -1.7721      0.444     -3.991      0.000      -2.643      -0.902
C(hour)[T.12]    -1.6456      0.444     -3.706      0.000      -2.516      -0.775
C(hour)[T.13]    -2.3060      0.445     -5.178      0.000      -3.179      -1.433
C(hour)[T.14]    -5.1996      0.446    -11.663      0.000      -6.074      -4.326
C(hour)[T.15]   -14.1946      0.445    -31.904      0.000     -15.067     -13.322
C(hour)[T.16]   -20.7114      0.444    -46.598      0.000     -21.583     -19.840
C(hour)[T.17]   -21.4484      0.444    -48.256      0.000     -22.320     -20.577
C(hour)[T.18]   -18.9683      0.444    -42.676      0.000     -19.840     -18.097
C(hour)[T.19]    -9.5619      0.444    -21.534      0.000     -10.432      -8.691
C(hour)[T.20]    -2.0619      0.444     -4.644      0.000      -2.932      -1.191
C(hour)[T.21]    -1.8749      0.444     -4.218      0.000      -2.746      -1.004
C(hour)[T.22]    -2.2992      0.444     -5.183      0.000      -3.169      -1.430
C(hour)[T.23]    -1.6896      0.443     -3.813      0.000      -2.558      -0.821
==============================================================================
Omnibus:                      734.794   Durbin-Watson:                   1.249
Prob(Omnibus):                  0.000   Jarque-Bera (JB):             5603.110
Skew:                          -0.327   Prob(JB):                         0.00
Kurtosis:                       7.684   Cond. No.                         25.0
==============================================================================
```

Our model accounts for about 67% of the data's variation. Holding our 'hour' variable as discrete, and hour = 24 (midnight) as our base value, we can see the effects each other hour has in regards to our dependent speed variable. For example, _Ihour17_ (5pm) has a coefficient of about -24.448 which means, all other variables held constant (and there are none), at 5 pm, our speed will be 24.4 miles per hour *less* than the intercept value (midnight). This equates to us having a speed near 8.5 miles per hour.

6.) Since there is only one independent variable in our regression, there is definitely omitted variable bias, which may result from weather conditions, and/or traffic accidents, and - much more common today, Covid-19! These externalities can lessen or make greater, our coefficients at certain times, or increase depending on the situation.

7.) Utilizing our regression in part 5, we can see that our speed is going to be near 8.5 miles per hour. Plugging this in to the formula, we get a Marginal Private Cost of about $2.50 miles per hour.

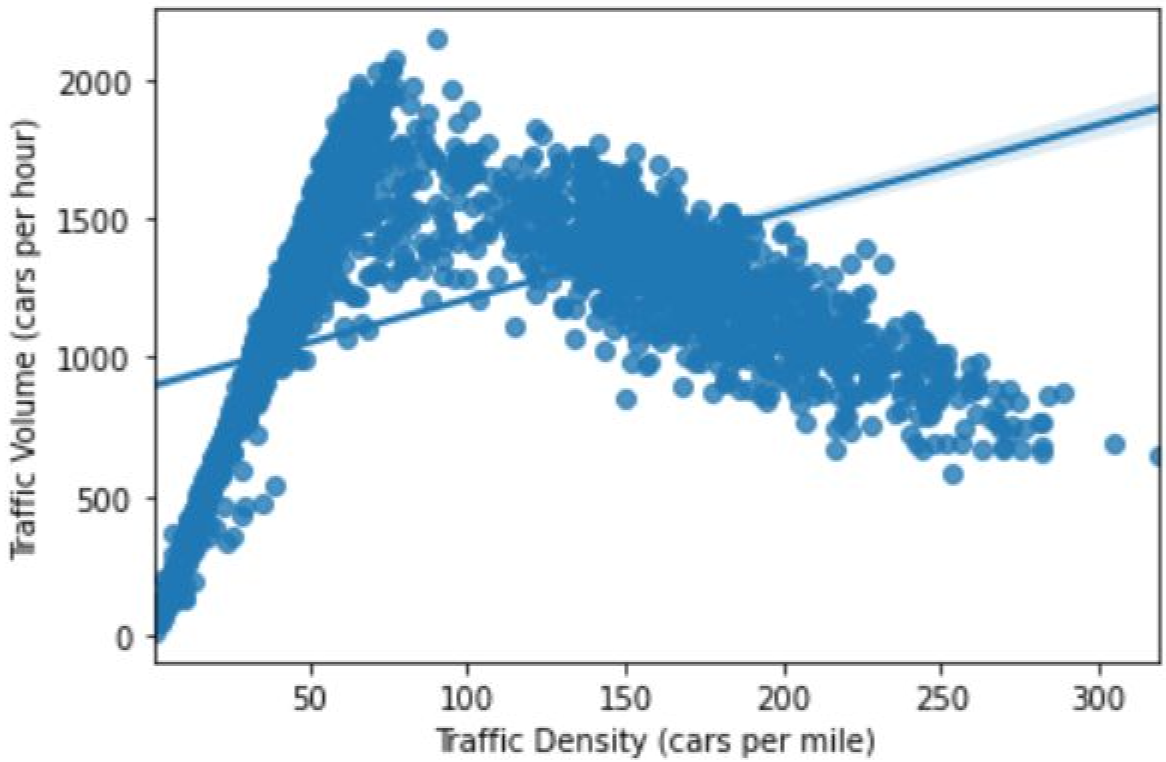
8.)
By calculating average speed, volume and densities per hour for weekends, and weekdays, for both directions, we can divide the volume averages by the speed averages, and find our density values.

9.)

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                  speed   R-squared:                       0.932
Model:                            OLS   Adj. R-squared:                  0.932
Method:                 Least Squares   F-statistic:                     3428.
Date:                Thu, 12 Nov 2020   Prob (F-statistic):               0.00
Time:                        15:46:19   Log-Likelihood:                -13414.
No. Observations:                6012   AIC:                         2.688e+04
Df Residuals:                    5987   BIC:                         2.705e+04
Df Model:                          24
Covariance Type:            nonrobust
=================================================================================
                    coef    std err          t      P>|t|      [0.025      0.975]
---------------------------------------------------------------------------------
Intercept        33.3029      0.145    230.440      0.000      33.020      33.586
C(hour)[T.1]     -1.3051      0.202     -6.461      0.000      -1.701      -0.909
C(hour)[T.2]     -1.2991      0.208     -6.233      0.000      -1.708      -0.891
C(hour)[T.3]     -1.7396      0.203     -8.587      0.000      -2.137      -1.342
C(hour)[T.4]     -0.6198      0.202     -3.068      0.002      -1.016      -0.224
C(hour)[T.5]      0.2701      0.202      1.337      0.181      -0.126       0.666
C(hour)[T.6]      2.3821      0.202     11.799      0.000       1.986       2.778
C(hour)[T.7]      3.5634      0.205     17.420      0.000       3.162       3.964
C(hour)[T.8]      0.4860      0.216      2.254      0.024       0.063       0.909
C(hour)[T.9]      1.8134      0.203      8.937      0.000       1.416       2.211
C(hour)[T.10]     1.7720      0.203      8.732      0.000       1.374       2.170
C(hour)[T.11]     1.7330      0.203      8.552      0.000       1.336       2.130
C(hour)[T.12]     2.1485      0.203     10.591      0.000       1.751       2.546
C(hour)[T.13]     2.1625      0.204     10.598      0.000       1.762       2.563
C(hour)[T.14]     1.7314      0.207      8.356      0.000       1.325       2.138
C(hour)[T.15]    -0.1925      0.222     -0.868      0.385      -0.627       0.242
C(hour)[T.16]    -0.4178      0.242     -1.729      0.084      -0.892       0.056
C(hour)[T.17]     0.3863      0.247      1.561      0.119      -0.099       0.871
C(hour)[T.18]    -0.5769      0.235     -2.455      0.014      -1.038      -0.116
C(hour)[T.19]     0.0795      0.211      0.377      0.706      -0.334       0.493
C(hour)[T.20]     1.2645      0.203      6.244      0.000       0.867       1.661
C(hour)[T.21]     0.9375      0.202      4.633      0.000       0.541       1.334
C(hour)[T.22]     0.9664      0.202      4.778      0.000       0.570       1.363
C(hour)[T.23]     0.5491      0.201      2.725      0.006       0.154       0.944
density          -0.1412      0.001   -152.120      0.000      -0.143      -0.139
==============================================================================
Omnibus:                     1679.323   Durbin-Watson:                   1.463
Prob(Omnibus):                  0.000   Jarque-Bera (JB):            30282.721
Skew:                          -0.869   Prob(JB):                         0.00
Kurtosis:                      13.857   Cond. No.                     2.16e+03
==============================================================================
```

Now, our model accounts for about 93.2% of our data’s variation. The coefficient of hour density variable tells us that for each additional car added per mile on our road, we can expect an about -0.141 mile per hour *decrease*. However, in actuality, this coefficient is probably going to shift down -in magnitude- as most negative externalities follow a marginally diminishing return function. Our model now can account for both, the time of day, as well as driving factors such as traffic ‘density’, hence our lower variable bias, and increase in R-square.

10.)



Our critical density point occurs between any real number in the set (75,100) and our volume reaches any real number in the set (2000,2200).

11.)
Using our Weekday West table, we can see that our volume is around 178 cars, and our price of time is given at $20 an hour. Our average speed is about 8.5, and density of around -0.141, leading to our Marginal External Cost of Congestion of about $6.97 per mile.

12.)
Using the same table (the first one), we can see that our 6 am coefficient is about 31.472. With this in mind, using the same values for our constants in Question #11, we can estimate a Marginal External Cost of Congestion near $0.089 per mile.