R Notebook

Libraries

```
library(readxl)
library(data.table)
library(lubridate) #Date Time Object
```

```
##
## Attaching package: 'lubridate'
```

```
## The following objects are masked from 'package:data.table':
##
##     hour, isoweek, mday, minute, month, quarter, second, wday,
##     week, yday, year
```

```
## The following object is masked from 'package:base':
##
##     date
```

```
library(ggplot2)
library(stringr) #Regular expression string splitting
library(rio)
library(MASS)
library(tigris)
```

```
## As of version 0.5.1, tigris does not cache downloaded data by default. To enable caching of data, se
```

```
##
## Attaching package: 'tigris'
```

```
## The following object is masked from 'package:graphics':
##
##     plot
```

```
library(sp)
library(broom)
library(dplyr)
```

```
## ------------------------------------------------------------------------------
```

```
## data.table + dplyr code now lives in dtplyr.
## Please library(dtplyr)!
```

```
## ------------------------------------------------------------------------------
```

```
##
## Attaching package: 'dplyr'
```

```
## The following object is masked from 'package:MASS':
##
##     select
```

```
## The following objects are masked from 'package:lubridate':
##
##     intersect, setdiff, union
```

```
## The following objects are masked from 'package:data.table':
##
```

```
##     between, first, last
## The following objects are masked from 'package:stats':
##
##     filter, lag
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(rgdal)
```

```
## rgdal: version: 1.2-13, (SVN revision 686)
##  Geospatial Data Abstraction Library extensions to R successfully loaded
##  Loaded GDAL runtime: GDAL 2.1.2, released 2016/10/24
##  Path to GDAL shared files: /Library/Frameworks/R.framework/Versions/3.3/Resources/library/sf/gdal
##  Loaded PROJ.4 runtime: Rel. 4.9.1, 04 March 2015, [PJ_VERSION: 491]
##  Path to PROJ.4 shared files: /Library/Frameworks/R.framework/Versions/3.3/Resources/library/sf/proj
##  Linking to sp version: 1.2-5
```

```
library(stargazer)
```

```
##
## Please cite as:
##  Hlavac, Marek (2015). stargazer: Well-Formatted Regression and Summary Statistics Tables.
##  R package version 5.2. http://CRAN.R-project.org/package=stargazer
```

T21 External Raw Data

```
T21External <- import("../../Data/MACTownHealth21_v6.csv")
View(T21External)
```

Remove the 2 cases and remove the years before 2009 and removing 2018 (Data needs to be reformatted again)

```
T21ExtSub <- T21External[T21External$Event != 2, ]
T21ExtSub <- T21ExtSub[-1]
T21ExtSub <- subset(T21ExtSub, Year > 2010)
T21ExtSub <- subset(T21ExtSub, Year < 2018)
View(T21ExtSub)

write.csv(T21ExtSub, "../../Data/T21ExtSub.csv")
View(T21ExtSub)
```

All T21 Imports

```
AllT21TopRatios <- import("../../Data/AllT21TopRatios.csv")
View(AllT21TopRatios)
AllT21TopRatios <- AllT21TopRatios[-1]
View(AllT21TopRatios)
```

Merging the data file

```
T21Merged <- T21ExtSub %>% left_join(AllT21TopRatios, by = c("Name" ="CityTown", "Year" = "year"))
```

```
names(T21Merged) <- gsub(" ","",names(T21Merged))
```

```
str(T21Merged)
```

```
## 'data.frame':    2220 obs. of  42 variables:
##  $ Year                          : int  2011 2012 2013 2014 2015 2016 2017 2011 2012 2013 ...
##  $ Name                          : chr  "Abington" "Abington" "Abington" "Abington" ...
##  $ Event                         : chr  "0" "0" "0" "0" ...
##  $ Type                          : chr  "Town" "Town" "Town" "Town" ...
##  $ County                        : chr  "Plymouth" "Plymouth" "Plymouth" "Plymouth" ...
##  $ Formofgovernment              : chr  "Open town meeting" "Open town meeting" "Open town meeting" "
##  $ Population                    : chr  "15985.0" "15985.0" "15985.0" "15985.0" ...
##  $ FormYear                      : int  1712 1712 1712 1712 1712 1712 1712 1735 1735 1735 ...
##  $ T21Date                       : chr  NA NA NA NA ...
##  $ FIPS                          : chr  "25023" "25023" "25023" "25023" ...
##  $ State                         : chr  "Massachusetts" "Massachusetts" "Massachusetts" "Massachuset
##  $ No_PrematureDeaths            : int  4693 4692 4854 4854 4881 4926 5179 11206 10885 10899 ...
##  $ YearsofPotentialLifeLostRate  : num  5872 5676 5721 5721 5466 ...
##  $ Percent_FairOrPoor            : num  10.6 10.2 11 11 11 ...
##  $ PhysicallyUnhealthyDays       : num  3.1 3 3.2 3.1 3.1 ...
##  $ MentallyUnhealthyDays         : num  3.6 3.6 3.8 3.6 3.6 ...
##  $ PrecentLBW                    : num  7.76 7.8 7.7 7.7 7.49 ...
##  $ PercentSmokers                : num  20.9 19.9 19.1 17.8 17.8 ...
##  $ PercentObese                  : num  22.1 23.1 23.1 24.7 25.8 26.4 27.8 22.1 23.1 23.1 ...
##  $ FoodEnvironmentIndex          : num  NA NA NA 8.76 8.6 ...
##  $ PercentPhysicallyInactive     : num  NA 21 21 23.1 23.9 24.7 22.9 NA 21.1 21.1 ...
##  $ PercentWithAccessToExercise   : num  NA NA NA 80.2 87.4 ...
##  $ PercentExcessiveDrinking      : num  20.7 21.1 20.8 21.1 21.1 ...
##  $ NumberChlamydiaCases          : int  NA 1160 1414 1398 1502 1490 1176 NA 2562 2778 ...
##  $ ChlamydiaRate                 : num  NA 236 286 281 300 ...
##  $ TeenBirths                    : int  2194 2213 2102 2017 1935 1794 1613 4133 4157 4086 ...
##  $ TeenBirthRate                 : num  18.9 19.5 17.7 17 16.3 ...
##  $ PercentUninsured              : num  8.3 4.4 4.7 4.29 3.89 ...
##  $ NumberMedicareEnrollees       : int  100457 53185 54940 58071 61488 64212 66023 264855 133633 135
##  $ PercentDiabetics              : int  1769 5117 5350 5871 6281 6734 6928 4000 10819 11042 ...
##  $ PercentReceivingHbA1c         : num  86.3 86.5 88.4 88.6 89.9 ...
##  $ PercentUnemployed             : num  8.8 9.1 7.8 6.91 7.17 ...
##  $ PercentChildreninPoverty      : num  8.3 10.9 10.8 10.7 11.2 11.5 12.4 8.3 8.5 9.3 ...
##  $ PercentSingle-ParentHouseholds: num  24.1 24.7 25 25.5 25.2 ...
##  $ NumberViolentCrimes           : num  6299 6280 2179 2161 2024 ...
##  $ ViolentCrimeRate              : num  428 435 443 448 432 ...
##  $ T21year                       : int  NA NA NA NA NA NA NA 2015 2015 2015 ...
##  $ MAYOR                         : chr  NA NA NA NA ...
##  $ LENGTHOFTERM                  : chr  NA NA NA NA ...
##  $ NEXTELECTION                  : chr  NA NA NA NA ...
##  $ mayor_ind                     : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ NearbyT21Ratio                : num  0 0 0 0 0 0.5 0.5 0 0 0 ...
```

```
T21Merged$Population <- as.numeric (T21Merged$Population)
```

```
## Warning: NAs introduced by coercion
```

```
write.csv(T21Merged, "../../Data/T21Mergedsecondattempt.csv")
```

```
demo <- import("../../Data/T21Mergedsecondattempt.csv")
View(demo)
dim(demo)
```

```
## [1] 2220   43
```

```
T21Merged$Event <- as.numeric(T21Merged$Event)
```

Logit Model

```
logitmodel <- glm(Event ~  Population + NearbyT21Ratio + PhysicallyUnhealthyDays + MentallyUnhealthyDays

summary(logitmodel)
```

```
##
## Call:
## glm(formula = Event ~ Population + NearbyT21Ratio + PhysicallyUnhealthyDays +
##     MentallyUnhealthyDays + PercentSmokers + PercentExcessiveDrinking +
##     TeenBirthRate + ChlamydiaRate + PercentUnemployed + ViolentCrimeRate,
##     family = binomial(link = "logit"), data = T21Merged)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.5982  -0.3089  -0.2340  -0.1982   2.8605
##
## Coefficients:
##                              Estimate Std. Error z value Pr(>|z|)
## (Intercept)                -1.791e+00  1.859e+00  -0.963   0.3353
## Population                  5.869e-06  5.342e-06   1.099   0.2719
## NearbyT21Ratio              4.642e+00  5.010e-01   9.266   <2e-16 ***
## PhysicallyUnhealthyDays     1.830e-01  4.245e-01   0.431   0.6664
## MentallyUnhealthyDays      -1.730e-01  3.689e-01  -0.469   0.6390
## PercentSmokers              1.729e-03  9.345e-02   0.018   0.9852
## PercentExcessiveDrinking -1.311e-01  9.860e-02  -1.330   0.1835
## TeenBirthRate              -4.031e-02  2.937e-02  -1.373   0.1698
## ChlamydiaRate               2.983e-03  1.595e-03   1.871   0.0614 .
## PercentUnemployed           1.482e-01  1.494e-01   0.992   0.3211
## ViolentCrimeRate           -1.396e-03  1.613e-03  -0.866   0.3868
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 954.67  on 1837  degrees of freedom
## Residual deviance: 743.33  on 1827  degrees of freedom
##   (382 observations deleted due to missingness)
## AIC: 765.33
##
## Number of Fisher Scoring iterations: 6
```

Table format

```
stargazer(logitmodel, type = "text")
```

```
##
## ================================================
##                            Dependent variable:
##                        ------------------------
##                                   Event
## ------------------------------------------------
## Population                         0.00001
##                                   (0.00001)
```

```
##
## NearbyT21Ratio                    4.642***
##                                    (0.501)
##
## PhysicallyUnhealthyDays            0.183
##                                    (0.425)
##
## MentallyUnhealthyDays             -0.173
##                                    (0.369)
##
## PercentSmokers                     0.002
##                                    (0.093)
##
## PercentExcessiveDrinking          -0.131
##                                    (0.099)
##
## TeenBirthRate                     -0.040
##                                    (0.029)
##
## ChlamydiaRate                     0.003*
##                                    (0.002)
##
## PercentUnemployed                  0.148
##                                    (0.149)
##
## ViolentCrimeRate                  -0.001
##                                    (0.002)
##
## Constant                          -1.791
##                                    (1.859)
##
## ------------------------------------------------
## Observations                        1,838
## Log Likelihood                   -371.665
## Akaike Inf. Crit.                 765.330
## ================================================
## Note:                  *p<0.1; **p<0.05; ***p<0.01
```