CKME 136 Initial Results

Step 1: Data Description

```
#Read data into a dataframe
data <- read.csv(file="https://data.ontario.ca/dataset/f4112442-bdc8-45d2-be3c-12efae72fb27/resource/455fd63b-603d-4608-8216-7d8647f43350/download/conposcovidloc.csv",header=T,sep=",",na.strings=c("","  ","NA"))

#summary statistics of dataset
summary(data)
```

```
##      Row_ID      Accurate_Episode_Date  Case_Reported_Date
##  Min.   :    1   2020-04-17:  685      2020-04-17:  753
##  1st Qu.: 8988   2020-04-13:  655      2020-04-15:  614
##  Median :17975   2020-04-15:  642      2020-04-20:  611
##  Mean   :17975   2020-04-14:  618      2020-04-13:  609
##  3rd Qu.:26961   2020-04-16:  605      2020-05-29:  609
##  Max.   :35948   (Other)   :32741      2020-04-18:  603
##                  NA's      :    2      (Other)   :32149
##   Test_Reported_Date    Specimen_Date     Age_Group        Client_Gender
##  2020-04-17:  690    2020-04-13:  744   50s    :5811   FEMALE     :19157
##  2020-04-20:  611    2020-04-15:  734   20s    :5514   MALE       :16509
##  2020-04-18:  610    2020-04-17:  691   40s    :5116   OTHER      :    9
##  2020-04-13:  606    2020-04-14:  689   30s    :5026   TRANSGENDER:    8
##  2020-04-15:  592    2020-04-16:  667   60s    :4091   UNKNOWN    :  265
##  (Other)   :32418    (Other)   :32120   80s    :3475
##  NA's      :  421    NA's      :  303   (Other):6915
##       Case_AcquisitionInfo         Outcome1     Outbreak_Related
##  CC              :10839     Fatal       : 2689   Yes :14267
##  No Epi-link     : 6687     Not Resolved: 1833   NA's:21681
##  No Info-Missing: 1122      Resolved    :31426
##  No Info-Unk     : 1280
##  OB              :14259
##  Travel          : 1761
##
##                               Reporting_PHU
##  Toronto Public Health              :13431
##  Peel Public Health                 : 5961
##  York Region Public Health Services: 3059
##  Ottawa Public Health               : 2117
##  Durham Region Health Department    : 1715
##  Windsor-Essex County Health Unit   : 1662
##  (Other)                            : 8003
##                      Reporting_PHU_Address   Reporting_PHU_City
##  277 Victoria Street, 5th Floor:13431     Toronto    :13431
##  7120 Hurontario Street        : 5961     Mississauga: 5961
```

```
##  17250 Yonge Street            : 3059     Newmarket  : 3059
##  100 Constellation Drive       : 2117     Ottawa     : 2117
##  605 Rossland Road East        : 1715     Whitby     : 1715
##  1005 Ouellette Avenue         : 1662     Windsor    : 1662
##  (Other)                       : 8003     (Other)    : 8003
##  Reporting_PHU_Postal_Code
##  M5B 1W2:13431
##  L5W 1N4: 5961
##  L3Y 6Z1: 3059
##  K2G 6J8: 2117
##  L1N 0B2: 1715
##  N9A 4J8: 1662
##  (Other): 8003
##                                                        Reporting_PHU_Website
##  www.toronto.ca/community-people/health-wellness-care/        :13431
##  www.peelregion.ca/health/                                    : 5961
##  www.york.ca/wps/portal/yorkhome/health/                      : 3059
##  www.ottawapublichealth.ca                                    : 2117
##  www.durham.ca/en/health-and-wellness/health-and-wellness.aspx: 1715
##  www.wechu.org                                                : 1662
##  (Other)                                                      : 8003
##  Reporting_PHU_Latitude Reporting_PHU_Longitude
##  Min.   :42.31          Min.   :-94.49
##  1st Qu.:43.65          1st Qu.:-79.71
##  Median :43.66          Median :-79.38
##  Mean   :43.74          Mean   :-79.52
##  3rd Qu.:43.66          3rd Qu.:-79.38
##  Max.   :49.77          Max.   :-74.74
##
```

```
str(data)
```

```
## 'data.frame':    35948 obs. of  17 variables:
##  $ Row_ID                  : int  1 2 3 4 5 6 7 8 9 10 ...
##  $ Accurate_Episode_Date   : Factor w/ 151 levels "2020-01-01","2020-01-1
0",..: 17 18 17 24 21 10 26 25 24 26 ...
##  $ Case_Reported_Date      : Factor w/ 134 levels "2020-01-23","2020-01-2
4",..: 8 8 8 8 9 9 10 10 12 12 ...
##  $ Test_Reported_Date      : Factor w/ 133 levels "2020-01-27","2020-02-0
3",..: 9 8 8 8 8 8 11 11 12 12 ...
##  $ Specimen_Date           : Factor w/ 138 levels "2020-01-23","2020-01-2
4",..: 12 11 11 10 12 12 14 15 15 15 ...
##  $ Age_Group               : Factor w/ 10 levels "<20","20s","30s",..: 7
6 7 5 6 7 6 5 6 4 ...
##  $ Client_Gender           : Factor w/ 5 levels "FEMALE","MALE",..: 1 1 1
2 2 1 2 1 2 2 ...
##  $ Case_AcquisitionInfo    : Factor w/ 6 levels "CC","No Epi-link",..: 6
6 6 6 6 6 6 6 6 6 ...
##  $ Outcome1                : Factor w/ 3 levels "Fatal","Not Resolved",..
: 3 3 3 3 3 3 3 3 3 3 ...
```

```
##  $ Outbreak_Related         : Factor w/ 1 level "Yes": NA NA NA NA NA NA N
A NA NA NA ...
##  $ Reporting_PHU            : Factor w/ 34 levels "Algoma Public Health Un
it",..: 34 34 34 31 31 31 31 24 21 31 ...
##  $ Reporting_PHU_Address    : Factor w/ 34 levels "100 Constellation Drive
",..: 14 14 14 23 23 23 23 33 32 23 ...
##  $ Reporting_PHU_City       : Factor w/ 34 levels "Barrie","Belleville",..
: 14 14 14 31 31 31 31 32 12 31 ...
##  $ Reporting_PHU_Postal_Code: Factor w/ 34 levels "K2G 6J8","K6J 5T1",..:
11 11 11 16 16 16 16 18 13 16 ...
##  $ Reporting_PHU_Website    : Factor w/ 34 levels "www.algomapublichealth.
com",..: 34 34 34 31 31 31 31 26 20 31 ...
##  $ Reporting_PHU_Latitude   : num  44 44 44 43.7 43.7 ...
##  $ Reporting_PHU_Longitude  : num  -79.5 -79.5 -79.5 -79.4 -79.4 ...
```

Step 1: Data Preparation

```
#Find missing values in data
sapply(data, function(x) sum(is.na(x)))
```

```
##                     Row_ID     Accurate_Episode_Date
##                          0                         2
##        Case_Reported_Date        Test_Reported_Date
##                          0                       421
##             Specimen_Date                 Age_Group
##                        303                         0
##             Client_Gender      Case_AcquisitionInfo
##                          0                         0
##                  Outcome1          Outbreak_Related
##                          0                     21681
##             Reporting_PHU     Reporting_PHU_Address
##                          0                         0
##        Reporting_PHU_City Reporting_PHU_Postal_Code
##                          0                         0
##     Reporting_PHU_Website    Reporting_PHU_Latitude
##                          0                         0
##   Reporting_PHU_Longitude
##                          0
```

```
sapply(data, function(x) length(unique(x)))
```

```
##                     Row_ID     Accurate_Episode_Date
##                      35948                       152
##        Case_Reported_Date        Test_Reported_Date
##                       134                       134
##             Specimen_Date                 Age_Group
##                       139                        10
##             Client_Gender      Case_AcquisitionInfo
##                         5                         6
##                  Outcome1          Outbreak_Related
```

```
##                          3                          2
##             Reporting_PHU     Reporting_PHU_Address
##                         34                         34
##        Reporting_PHU_City Reporting_PHU_Postal_Code
##                         34                         34
##     Reporting_PHU_Website    Reporting_PHU_Latitude
##                         34                         34
##   Reporting_PHU_Longitude
##                         34
```

```
#install.packages("Amelia")
library(Amelia)
```

```
## Warning: package 'Amelia' was built under R version 3.5.3
```

```
## Loading required package: Rcpp
```

```
## Warning: package 'Rcpp' was built under R version 3.5.3
```

```
## ##
## ## Amelia II: Multiple Imputation
## ## (Version 1.7.6, built: 2019-11-24)
## ## Copyright (C) 2005-2020 James Honaker, Gary King and Matthew Blackwell
## ## Refer to http://gking.harvard.edu/amelia/ for more information
## ##
```

```
missmap(data, main = "Missing values vs observed")
#4% of data is missing

#Remove Not Resolved rows in Outcome1
dataclean <- droplevels(data[!data$Outcome1 == 'Not Resolved',])

#Outbreak Related has "missing values" but should be No
sum(is.na(dataclean$Outbreak_Related) == TRUE)
```

```
## [1] 20230
```

```
length(dataclean$Outbreak_Related)
```

```
## [1] 34115
```

```
#Replace NA with No
dataclean$Outbreak_Related <- factor(dataclean$Outbreak_Related, exclude = NU
LL,
               levels = c("Yes", NA),
               labels = c("Yes", "No"))
table(dataclean$Outbreak_Related, useNA = "always")
```

```
##
##   Yes    No  <NA>
## 13885 20230     0
```

```r
str(dataclean$Outbreak_Related)
```

```
##  Factor w/ 2 levels "Yes","No": 2 2 2 2 2 2 2 2 2 2 ...
```

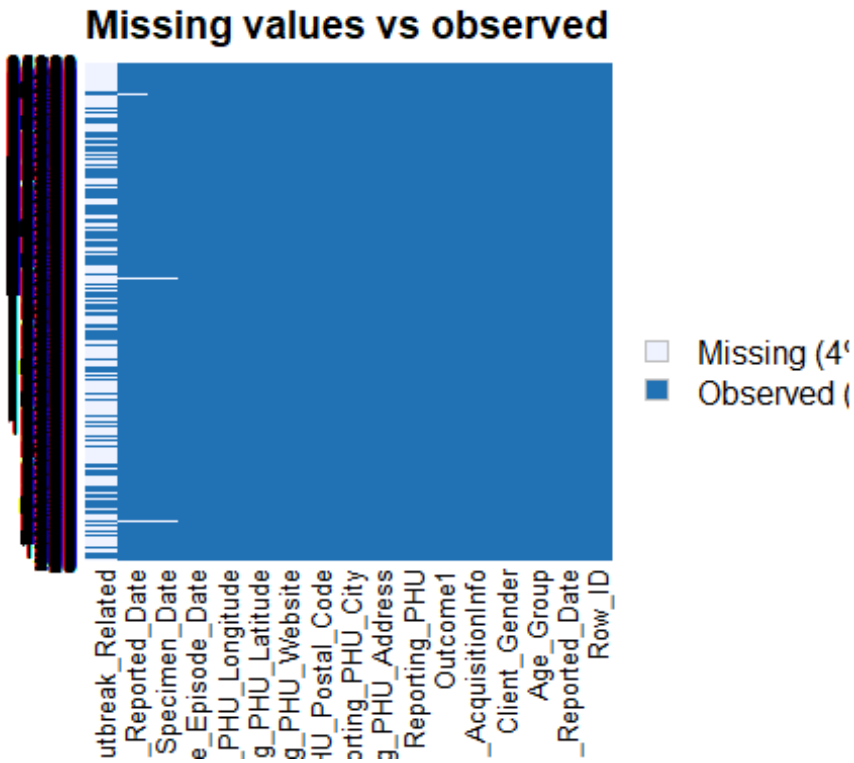
```r
#Remove repetitive variables
dataclean$Reporting_PHU <- NULL
dataclean$Reporting_PHU_Address <- NULL
dataclean$Reporting_PHU_Postal_Code <- NULL
dataclean$Reporting_PHU_Website <- NULL

#summary statistics of clean dataset
summary(dataclean)
```

```
##      Row_ID      Accurate_Episode_Date  Case_Reported_Date
##  Min.   :    1   2020-04-17:  681      2020-04-17:  752
##  1st Qu.: 8554   2020-04-13:  652      2020-04-15:  612
##  Median :17139   2020-04-15:  640      2020-05-29:  609
##  Mean   :17550   2020-04-14:  614      2020-04-20:  608
##  3rd Qu.:26743   2020-04-16:  604      2020-04-13:  607
##  Max.   :35948   (Other)   :30922      2020-04-18:  599
##                  NA's      :    2      (Other)   :30328
##   Test_Reported_Date    Specimen_Date     Age_Group        Client_Gender
##  2020-04-17:  687    2020-04-13:  741   50s    :5546   FEMALE     :18364
##  2020-04-20:  610    2020-04-15:  730   20s    :5062   MALE       :15484
##  2020-04-13:  605    2020-04-17:  687   40s    :4876   OTHER      :    8
##  2020-04-18:  604    2020-04-14:  686   30s    :4705   TRANSGENDER:    8
##  2020-04-16:  590    2020-04-16:  666   60s    :3914   UNKNOWN    :  251
##  (Other)   :30612    (Other)   :30313   80s    :3404
##  NA's      :  407    NA's      :  292   (Other):6608
##       Case_AcquisitionInfo     Outcome1     Outbreak_Related
##  CC              :10261     Fatal   : 2689   Yes:13885
##  No Epi-link     : 6295     Resolved:31426   No :20230
##  No Info-Missing:  748
##  No Info-Unk     : 1211
##  OB              :13877
##  Travel          : 1723
##
##    Reporting_PHU_City Reporting_PHU_Latitude Reporting_PHU_Longitude
##  Toronto    :12861    Min.   :42.31          Min.   :-94.49
##  Mississauga: 5556    1st Qu.:43.65          1st Qu.:-79.71
##  Newmarket  : 2887    Median :43.66          Median :-79.38
##  Ottawa     : 2070    Mean   :43.76          Mean   :-79.49
##  Whitby     : 1683    3rd Qu.:43.90          3rd Qu.:-79.38
##  Windsor    : 1352    Max.   :49.77          Max.   :-74.74
##  (Other)    : 7706
```

```r
str(dataclean)
```

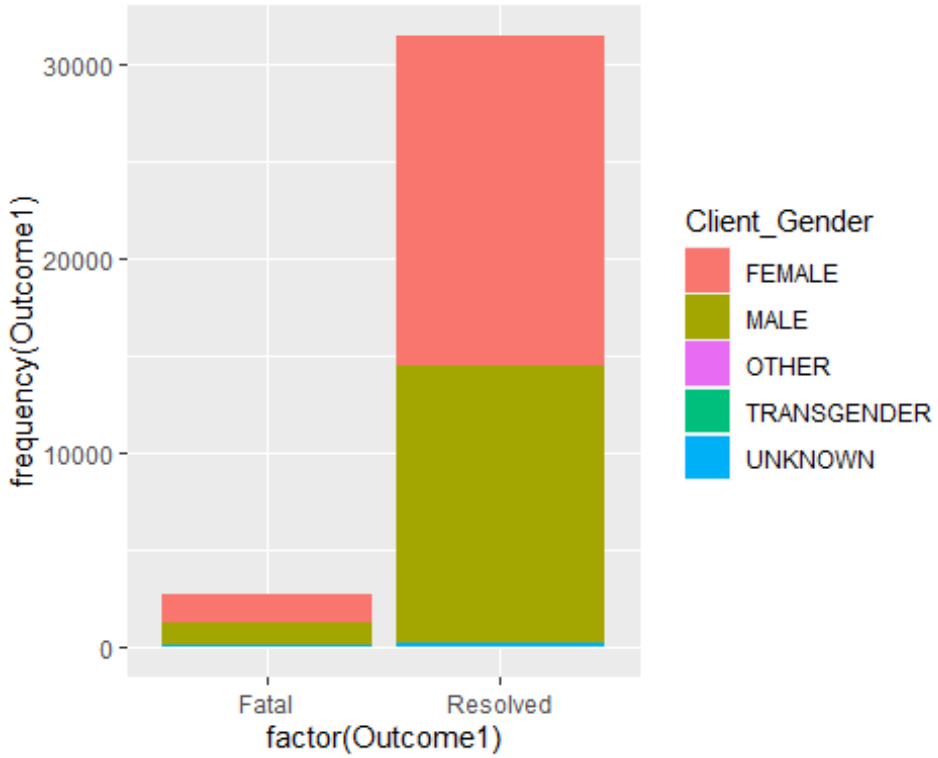
```
## 'data.frame':    34115 obs. of  13 variables:
##  $ Row_ID                 : int  1 2 3 4 5 6 7 8 9 10 ...
##  $ Accurate_Episode_Date  : Factor w/ 145 levels "2020-01-01","2020-01-10"
```

```
,..: 17 18 17 24 21 10 26 25 24 26 ...
##  $ Case_Reported_Date     : Factor w/ 133 levels "2020-01-23","2020-01-24"
,..: 8 8 8 8 9 9 10 10 12 12 ...
##  $ Test_Reported_Date     : Factor w/ 132 levels "2020-01-27","2020-02-03"
,..: 9 8 8 8 8 8 11 11 12 12 ...
##  $ Specimen_Date          : Factor w/ 136 levels "2020-01-23","2020-01-24"
,..: 12 11 11 10 12 12 14 15 15 15 ...
##  $ Age_Group              : Factor w/ 10 levels "<20","20s","30s",..: 7 6
7 5 6 7 6 5 6 4 ...
##  $ Client_Gender          : Factor w/ 5 levels "FEMALE","MALE",..: 1 1 1 2
2 1 2 1 2 2 ...
##  $ Case_AcquisitionInfo   : Factor w/ 6 levels "CC","No Epi-link",..: 6 6
6 6 6 6 6 6 6 ...
##  $ Outcome1               : Factor w/ 2 levels "Fatal","Resolved": 2 2 2 2
2 2 2 2 2 2 ...
##  $ Outbreak_Related       : Factor w/ 2 levels "Yes","No": 2 2 2 2 2 2 2 2
2 2 ...
##  $ Reporting_PHU_City     : Factor w/ 34 levels "Barrie","Belleville",..:
14 14 14 31 31 31 31 32 12 31 ...
##  $ Reporting_PHU_Latitude : num  44 44 44 43.7 43.7 ...
##  $ Reporting_PHU_Longitude: num  -79.5 -79.5 -79.5 -79.4 -79.4 ...

#bar charts
library(ggplot2)
```

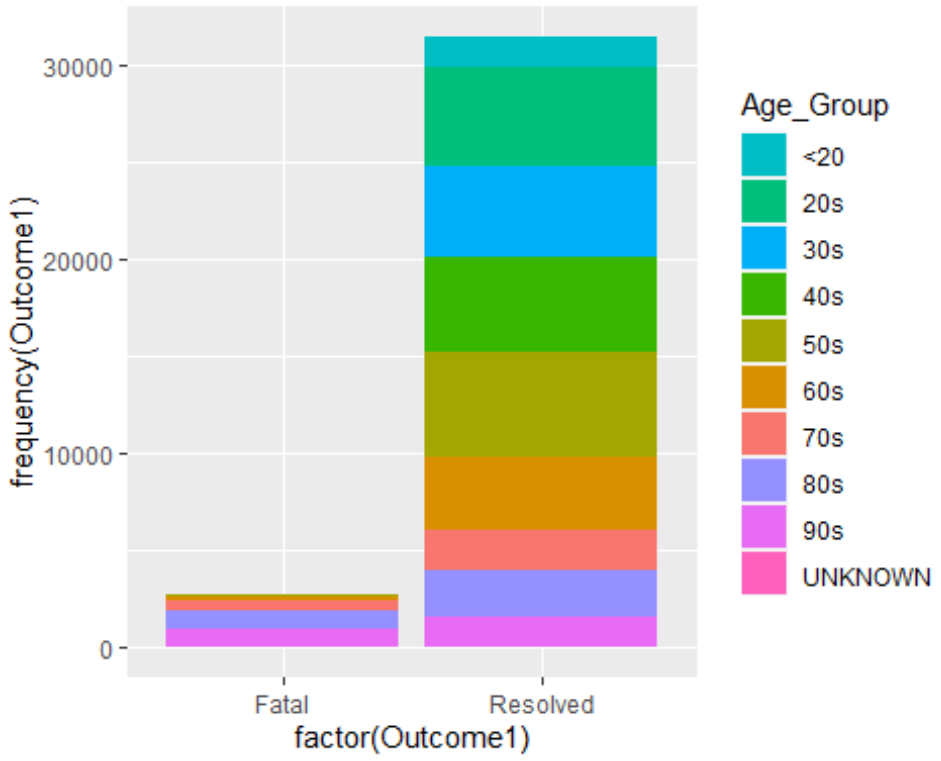



```
#Access ggplot-colors
gg_color_hue <- function(n) {
```

```
  hues = seq(15, 375, length=n+1)
  hcl(h=hues, l=65, c=100)[1:n]
}

#create custom palette for Client Gender
mycols <- gg_color_hue(length(unique(dataclean$Client_Gender)))
names(mycols) <- unique(dataclean$Client_Gender)

#stacked bar chart for Client Genders in Outcome
ggplot(dataclean, aes(x = factor(Outcome1), y = frequency(Outcome1), fill=Cli
ent_Gender)) + geom_bar(stat = 'identity') + scale_fill_manual(values = mycol
s)
```

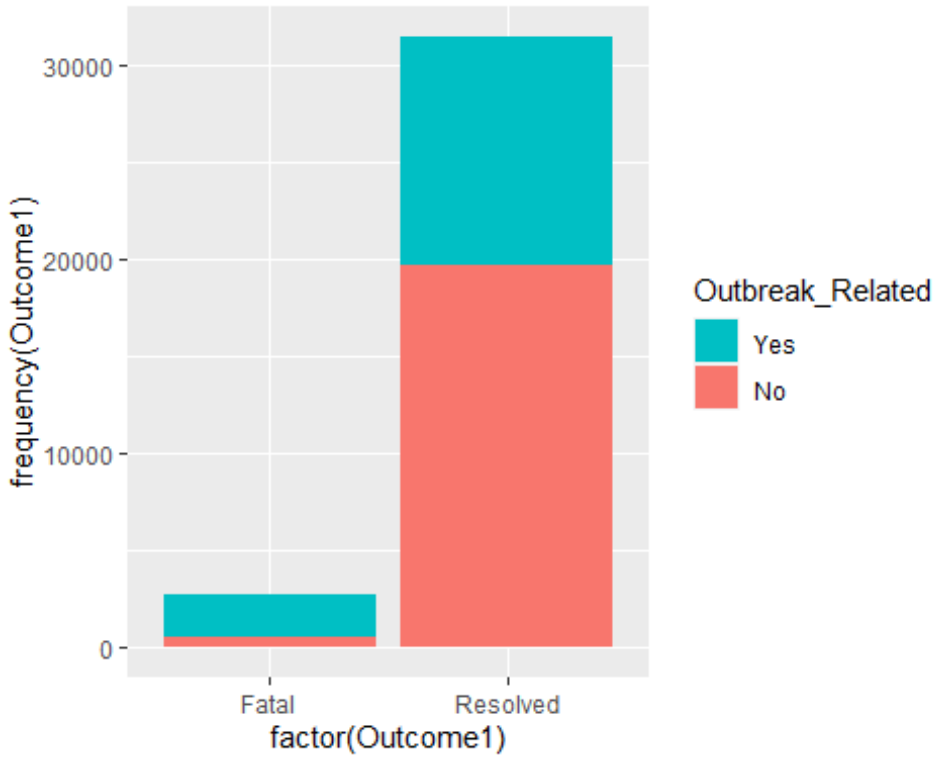



```
#Age Group stacked bar chart
mycols <- gg_color_hue(length(unique(dataclean$Age_Group)))
names(mycols) <- unique(dataclean$Age_Group)
ggplot(dataclean, aes(x = factor(Outcome1), y = frequency(Outcome1), fill=Age
_Group)) + geom_bar(stat = 'identity') + scale_fill_manual(values = mycols)
```

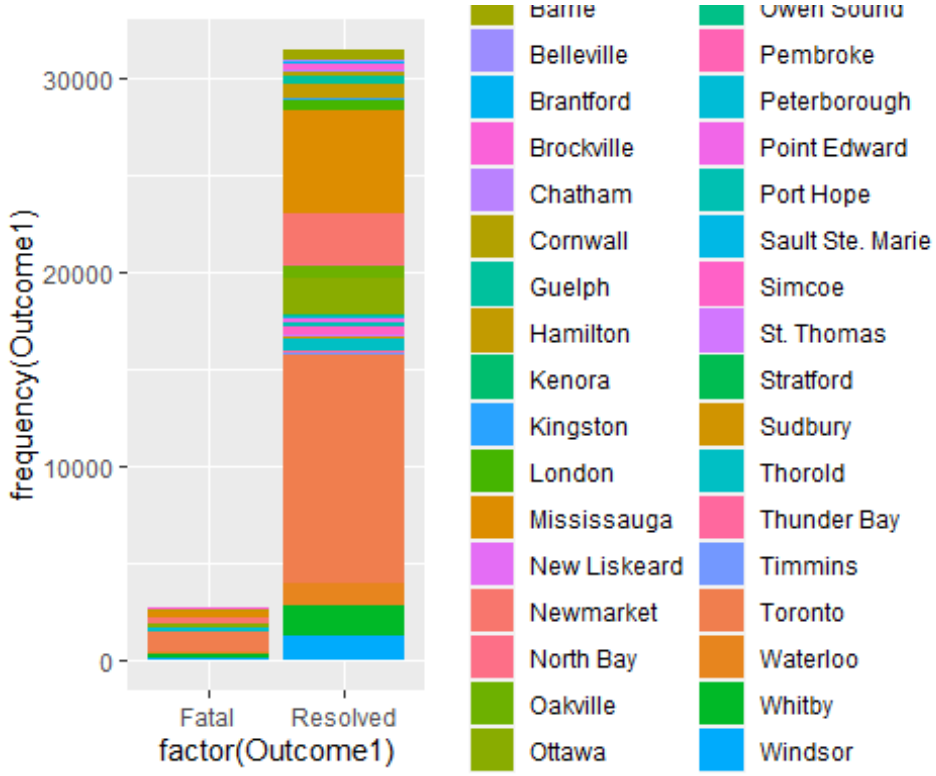



```
#Case Acquisition Info stacked bar chart
mycols <- gg_color_hue(length(unique(dataclean$Case_AcquisitionInfo)))
names(mycols) <- unique(dataclean$Case_AcquisitionInfo)
ggplot(dataclean, aes(x = factor(Outcome1), y = frequency(Outcome1), fill=Cas
e_AcquisitionInfo)) + geom_bar(stat = 'identity') + scale_fill_manual(values
= mycols)
```

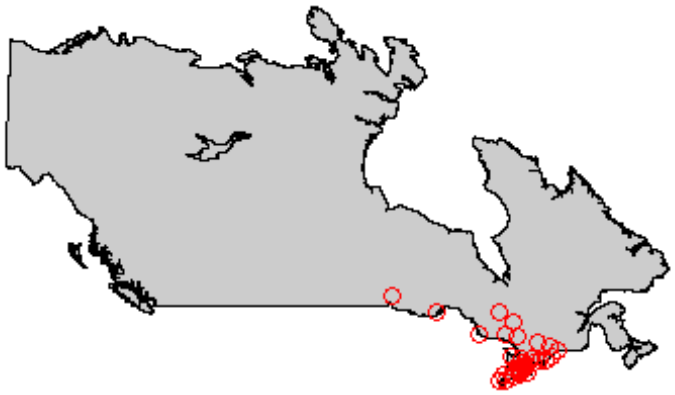

```
#Outbreak Related stacked bar chart
mycols <- gg_color_hue(length(unique(dataclean$Outbreak_Related)))
names(mycols) <- unique(dataclean$Outbreak_Related)
ggplot(dataclean, aes(x = factor(Outcome1), y = frequency(Outcome1), fill=Out
break_Related)) + geom_bar(stat = 'identity') + scale_fill_manual(values = my
cols)
```




```
#Reporting City stacked bar chart
mycols <- gg_color_hue(length(unique(dataclean$Reporting_PHU_City)))
names(mycols) <- unique(dataclean$Reporting_PHU_City)
ggplot(dataclean, aes(x = factor(Outcome1), y = frequency(Outcome1), fill=Rep
orting_PHU_City)) + geom_bar(stat = 'identity') + scale_fill_manual(values =
mycols)
```




```
#install.packages("maps")
#install.packages("mapdata")
#install.packages("mapproj")
library(maps)

## Warning: package 'maps' was built under R version 3.5.3

library(mapdata)

## Warning: package 'mapdata' was built under R version 3.5.3

library(mapproj)

## Warning: package 'mapproj' was built under R version 3.5.3

map(database = "worldHires", "Canada", xlim=c(-140,-110),ylim=c(48,64), col="
grey80", fill=TRUE, projection="gilbert", orientation= c(90,0,225))
lon <- c(dataclean$Reporting_PHU_Longitude)
lat <- c(dataclean$Reporting_PHU_Latitude)
coord <- mapproject(lon, lat, proj="gilbert", orientation=c(90, 0, 225))  #co
nvert points to projected lat/long
points(coord, pch=21, cex=1.2, col="red")
```



Step 3:

Experimental Design

```
#Split data into training (10%) and test (90%) sets
#createDataPartition function does stratified random sampling
set.seed(100)
#install.packages("caret")
library(caret)

## Warning: package 'caret' was built under R version 3.5.3

## Loading required package: lattice

trainingRows <- createDataPartition(dataclean$Outcome1, p = 0.9, list = FALSE
)
training <- dataclean[trainingRows,]
test  <- dataclean[-trainingRows,]

#see if proportion of training and test set are the same
prop.table(table(training$Outcome1))

##
##      Fatal   Resolved
## 0.07884709 0.92115291

prop.table(table(test$Outcome1))
```

```
##
##      Fatal   Resolved
## 0.07859238 0.92140762

#proportions are the same
```

Step 4: Modeling - (1) Logistic Regression

```
#install.packages("caret")
library(caret)
#Fit training data into Multinomial Logistic Regression Model
model.LogReg <- glm(Outcome1 ~ Client_Gender + Age_Group + Case_AcquisitionIn
fo + Outbreak_Related + Reporting_PHU_City, family = binomial(link = "logit")
, data=training)
summary (model.LogReg)

##
## Call:
## glm(formula = Outcome1 ~ Client_Gender + Age_Group + Case_AcquisitionInfo
+
##     Outbreak_Related + Reporting_PHU_City, family = binomial(link = "logit
"),
##     data = training)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -3.8747   0.0395   0.0951   0.2964   1.4507
##
## Coefficients:
##                                      Estimate Std. Error z value Pr(>|z|)
## (Intercept)                          20.04917  812.78069   0.025  0.98032
## Client_GenderMALE                    -0.60264    0.05090 -11.839  < 2e-16
## Client_GenderOTHER                   13.33216  789.16302   0.017  0.98652
## Client_GenderTRANSGENDER             -1.79175    1.45689  -1.230  0.21875
## Client_GenderUNKNOWN                 -0.22006    0.19943  -1.103  0.26983
## Age_Group20s                          0.25025    1.15557   0.217  0.82855
## Age_Group30s                         -0.43440    1.08130  -0.402  0.68788
## Age_Group40s                         -1.67264    1.02477  -1.632  0.10263
## Age_Group50s                         -2.86320    1.00774  -2.841  0.00449
## Age_Group60s                         -4.26561    1.00360  -4.250 2.13e-05
## Age_Group70s                         -5.44600    1.00281  -5.431 5.61e-08
## Age_Group80s                         -5.90095    1.00275  -5.885 3.99e-09
## Age_Group90s                         -6.31217    1.00326  -6.292 3.14e-10
## Age_GroupUNKNOWN                      9.77257  842.70109   0.012  0.99075
## Case_AcquisitionInfoNo Epi-link      -0.50266    0.11588  -4.338 1.44e-05
## Case_AcquisitionInfoNo Info-Missing  -0.41185    0.26470  -1.556  0.11972
## Case_AcquisitionInfoNo Info-Unk      -1.16888    0.16715  -6.993 2.69e-12
## Case_AcquisitionInfoOB              -13.03734  812.78005  -0.016  0.98720
## Case_AcquisitionInfoTravel           -0.03869    0.17578  -0.220  0.82578
## Outbreak_RelatedNo                  -12.13594  812.78004  -0.015  0.98809
```

```
## Reporting_PHU_CityBelleville          -0.41673    0.59789  -0.697  0.48580
## Reporting_PHU_CityBrantford           -0.07830    0.55340  -0.141  0.88748
## Reporting_PHU_CityBrockville           0.06791    0.27211   0.250  0.80293
## Reporting_PHU_CityChatham              0.44607    1.05613   0.422  0.67276
## Reporting_PHU_CityCornwall             0.13252    0.41912   0.316  0.75187
## Reporting_PHU_CityGuelph              -0.10067    0.29708  -0.339  0.73472
## Reporting_PHU_CityHamilton             0.36784    0.27998   1.314  0.18890
## Reporting_PHU_CityKenora              12.35282  359.99805   0.034  0.97263
## Reporting_PHU_CityKingston            12.79529  271.79503   0.047  0.96245
## Reporting_PHU_CityLondon              -0.28220    0.27365  -1.031  0.30243
## Reporting_PHU_CityMississauga         -0.15482    0.22390  -0.691  0.48927
## Reporting_PHU_CityNew Liskeard        13.34994  499.24058   0.027  0.97867
## Reporting_PHU_CityNewmarket           -0.13194    0.22669  -0.582  0.56056
## Reporting_PHU_CityNorth Bay            0.22505    1.08321   0.208  0.83541
## Reporting_PHU_CityOakville             0.73051    0.31440   2.324  0.02015
## Reporting_PHU_CityOttawa              -0.38107    0.22761  -1.674  0.09408
## Reporting_PHU_CityOwen Sound          13.87333  214.86525   0.065  0.94852
## Reporting_PHU_CityPembroke             0.49911    1.10375   0.452  0.65113
## Reporting_PHU_CityPeterborough         0.85948    0.78402   1.096  0.27297
## Reporting_PHU_CityPoint Edward         0.12178    0.31859   0.382  0.70228
## Reporting_PHU_CityPort Hope            0.07806    0.35183   0.222  0.82441
## Reporting_PHU_CitySault Ste. Marie    13.64713  437.69633   0.031  0.97513
## Reporting_PHU_CitySimcoe              -0.25936    0.30760  -0.843  0.39914
## Reporting_PHU_CitySt. Thomas          -0.80742    0.54315  -1.487  0.13713
## Reporting_PHU_CityStratford           -0.39837    0.57863  -0.688  0.49115
## Reporting_PHU_CitySudbury              0.22039    0.79329   0.278  0.78115
## Reporting_PHU_CityThorold              0.11343    0.26115   0.434  0.66403
## Reporting_PHU_CityThunder Bay          0.60473    1.05884   0.571  0.56792
## Reporting_PHU_CityTimmins             -1.19619    0.53711  -2.227  0.02594
## Reporting_PHU_CityToronto              0.05799    0.21638   0.268  0.78871
## Reporting_PHU_CityWaterloo            -0.16747    0.24340  -0.688  0.49141
## Reporting_PHU_CityWhitby              -0.13470    0.23151  -0.582  0.56067
## Reporting_PHU_CityWindsor              0.01875    0.25843   0.073  0.94215
##
## (Intercept)
## Client_GenderMALE                    ***
## Client_GenderOTHER
## Client_GenderTRANSGENDER
## Client_GenderUNKNOWN
## Age_Group20s
## Age_Group30s
## Age_Group40s
## Age_Group50s                         **
## Age_Group60s                         ***
## Age_Group70s                         ***
## Age_Group80s                         ***
## Age_Group90s                         ***
## Age_GroupUNKNOWN
## Case_AcquisitionInfoNo Epi-link      ***
## Case_AcquisitionInfoNo Info-Missing
```

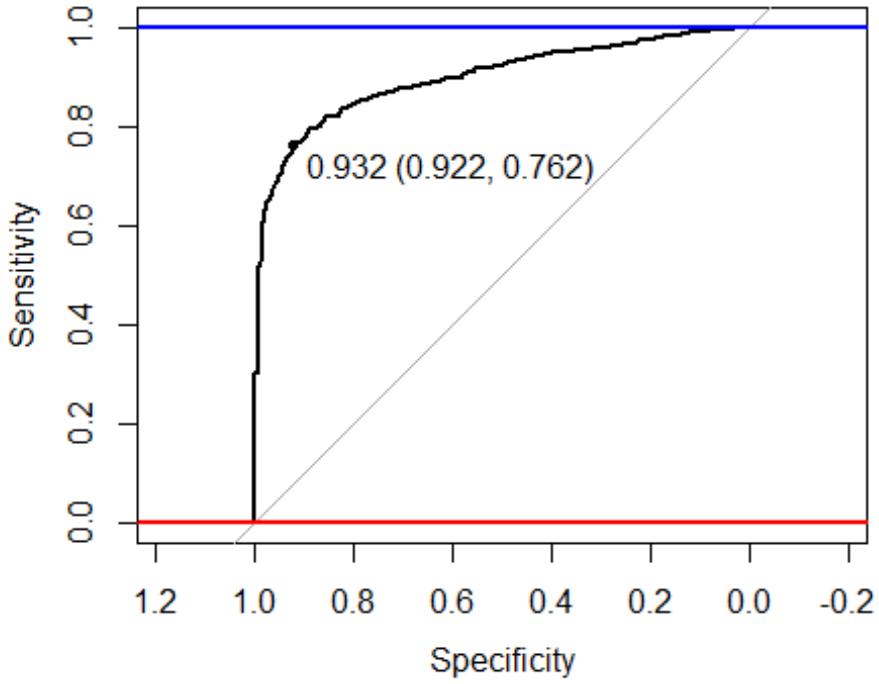
```
## Case_AcquisitionInfoNo Info-Unk     ***
## Case_AcquisitionInfoOB
## Case_AcquisitionInfoTravel
## Outbreak_RelatedNo
## Reporting_PHU_CityBelleville
## Reporting_PHU_CityBrantford
## Reporting_PHU_CityBrockville
## Reporting_PHU_CityChatham
## Reporting_PHU_CityCornwall
## Reporting_PHU_CityGuelph
## Reporting_PHU_CityHamilton
## Reporting_PHU_CityKenora
## Reporting_PHU_CityKingston
## Reporting_PHU_CityLondon
## Reporting_PHU_CityMississauga
## Reporting_PHU_CityNew Liskeard
## Reporting_PHU_CityNewmarket
## Reporting_PHU_CityNorth Bay
## Reporting_PHU_CityOakville          *
## Reporting_PHU_CityOttawa            .
## Reporting_PHU_CityOwen Sound
## Reporting_PHU_CityPembroke
## Reporting_PHU_CityPeterborough
## Reporting_PHU_CityPoint Edward
## Reporting_PHU_CityPort Hope
## Reporting_PHU_CitySault Ste. Marie
## Reporting_PHU_CitySimcoe
## Reporting_PHU_CitySt. Thomas
## Reporting_PHU_CityStratford
## Reporting_PHU_CitySudbury
## Reporting_PHU_CityThorold
## Reporting_PHU_CityThunder Bay
## Reporting_PHU_CityTimmins           *
## Reporting_PHU_CityToronto
## Reporting_PHU_CityWaterloo
## Reporting_PHU_CityWhitby
## Reporting_PHU_CityWindsor
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 16946  on 30704  degrees of freedom
## Residual deviance: 11217  on 30652  degrees of freedom
## AIC: 11323
##
## Number of Fisher Scoring iterations: 15

#ANOVA test
anova(model.LogReg, test="Chisq")
```

```
## Analysis of Deviance Table
##
## Model: binomial, link: logit
##
## Response: Outcome1
##
## Terms added sequentially (first to last)
##
##
##                      Df Deviance Resid. Df Resid. Dev  Pr(>Chi)
## NULL                                 30704      16946
## Client_Gender         4     20.2     30700      16926 0.0004466 ***
## Age_Group             9   5483.5     30691      11442 < 2.2e-16 ***
## Case_AcquisitionInfo  5    136.8     30686      11305 < 2.2e-16 ***
## Outbreak_Related      1      0.3     30685      11305 0.6115390
## Reporting_PHU_City   33     88.1     30652      11217 6.547e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#Predict on test data
test.probs <- predict(model.LogReg, test, type = "response")
pred.log <- rep("Fatal", length(test.probs))
pred.log[test.probs>=0.5] <- "Resolved"

#Confusion Matrix
confusionMatrix(factor(pred.log),factor(test$Outcome1))
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction Fatal Resolved
##   Fatal       16       15
##   Resolved   252     3127
##
##                Accuracy : 0.9217
##                  95% CI : (0.9122, 0.9305)
##     No Information Rate : 0.9214
##     P-Value [Acc > NIR] : 0.4909
##
##                   Kappa : 0.0922
##
##  Mcnemar's Test P-Value : <2e-16
##
##             Sensitivity : 0.059701
##             Specificity : 0.995226
##          Pos Pred Value : 0.516129
##          Neg Pred Value : 0.925422
##              Prevalence : 0.078592
##          Detection Rate : 0.004692
##    Detection Prevalence : 0.009091
```

```
##       Balanced Accuracy : 0.527464
##
##        'Positive' Class : Fatal
##

#ROC Curve
#install.packages("pROC")
library(pROC)

## Warning: package 'pROC' was built under R version 3.5.3

## Type 'citation("pROC")' for a citation.

##
## Attaching package: 'pROC'

## The following objects are masked from 'package:stats':
##
##     cov, smooth, var

roc.curve <- roc(test$Outcome1, test.probs)

## Setting levels: control = Fatal, case = Resolved

## Setting direction: controls < cases

print(roc.curve)

##
## Call:
## roc.default(response = test$Outcome1, predictor = test.probs)
##
## Data: test.probs in 268 controls (test$Outcome1 Fatal) < 3142 cases (test$
Outcome1 Resolved).
## Area under the curve: 0.898

plot(roc.curve, ylim=c(0,1), print.thres=TRUE)
abline(h=1,col='blue',lwd=2)
abline(h=0,col='red',lwd=2)
```



Step 4: Modeling - (2) Naive Bayes Classifier

```
#install.packages("e1071")
library(e1071)

## Warning: package 'e1071' was built under R version 3.5.3

#install.packages("gmodels")
library(gmodels)

## Warning: package 'gmodels' was built under R version 3.5.3

##
## Attaching package: 'gmodels'

## The following object is masked from 'package:pROC':
##
##     ci

#Building model on training set
NBC.classifier <- naiveBayes(training, training$Outcome1, laplace = 1)
NBC.classifier

##
## Naive Bayes Classifier for Discrete Predictors
##
## Call:
```

```
## naiveBayes.default(x = training, y = training$Outcome1, laplace = 1)
##
## A-priori probabilities:
## training$Outcome1
##      Fatal   Resolved
## 0.07884709 0.92115291
##
## Conditional probabilities:
##                   Row_ID
## training$Outcome1     [,1]     [,2]
##          Fatal    14407.57 10049.92
##          Resolved 17805.37 10401.55
##
##                   Accurate_Episode_Date
## training$Outcome1   2020-01-01   2020-01-10   2020-01-21   2020-01-22
##          Fatal    3.897116e-04 3.897116e-04 3.897116e-04 3.897116e-04
##          Resolved 1.055335e-04 7.035565e-05 7.035565e-05 7.035565e-05
##                   Accurate_Episode_Date
## training$Outcome1   2020-01-24   2020-02-01   2020-02-05   2020-02-07
##          Fatal    3.897116e-04 3.897116e-04 3.897116e-04 3.897116e-04
##          Resolved 7.035565e-05 7.035565e-05 7.035565e-05 7.035565e-05
##                   Accurate_Episode_Date
## training$Outcome1   2020-02-10   2020-02-14   2020-02-15   2020-02-16
##          Fatal    3.897116e-04 3.897116e-04 3.897116e-04 3.897116e-04
##          Resolved 1.055335e-04 7.035565e-05 3.517782e-05 7.035565e-05
##                   Accurate_Episode_Date
## training$Outcome1   2020-02-17   2020-02-19   2020-02-20   2020-02-21
##          Fatal    3.897116e-04 3.897116e-04 3.897116e-04 3.897116e-04
##          Resolved 3.517782e-05 7.035565e-05 1.407113e-04 7.035565e-05
##                   Accurate_Episode_Date
## training$Outcome1   2020-02-22   2020-02-23   2020-02-24   2020-02-25
##          Fatal    3.897116e-04 3.897116e-04 3.897116e-04 3.897116e-04
##          Resolved 1.758891e-04 1.758891e-04 1.758891e-04 1.407113e-04
##                   Accurate_Episode_Date
## training$Outcome1   2020-02-26   2020-02-27   2020-02-28   2020-02-29
##          Fatal    3.897116e-04 3.897116e-04 3.897116e-04 3.897116e-04
##          Resolved 1.055335e-04 7.035565e-05 2.462448e-04 2.110669e-04
##                   Accurate_Episode_Date
## training$Outcome1   2020-03-01   2020-03-02   2020-03-03   2020-03-04
##          Fatal    1.169135e-03 3.897116e-04 3.897116e-04 1.169135e-03
##          Resolved 5.628452e-04 5.276674e-04 5.276674e-04 3.517782e-04
##                   Accurate_Episode_Date
## training$Outcome1   2020-03-05   2020-03-06   2020-03-07   2020-03-08
##          Fatal    3.897116e-04 1.558846e-03 1.169135e-03 1.558846e-03
##          Resolved 8.442678e-04 1.055335e-03 9.146234e-04 1.583002e-03
##                   Accurate_Episode_Date
## training$Outcome1   2020-03-09   2020-03-10   2020-03-11   2020-03-12
##          Fatal    1.169135e-03 1.558846e-03 7.794232e-04 7.794232e-04
##          Resolved 2.075492e-03 3.201182e-03 3.306715e-03 4.150983e-03
##                   Accurate_Episode_Date
```

```
## training$Outcome1   2020-03-13   2020-03-14   2020-03-15   2020-03-16
##          Fatal    2.338270e-03 3.507405e-03 3.507405e-03 4.286828e-03
##          Resolved 4.713828e-03 4.186161e-03 5.698807e-03 8.020544e-03
##                   Accurate_Episode_Date
## training$Outcome1   2020-03-17   2020-03-18   2020-03-19   2020-03-20
##          Fatal    5.455963e-03 5.455963e-03 7.014809e-03 8.183944e-03
##          Resolved 8.301966e-03 8.126077e-03 8.090899e-03 1.030710e-02
##                   Accurate_Episode_Date
## training$Outcome1   2020-03-21   2020-03-22   2020-03-23   2020-03-24
##          Fatal    5.066251e-03 6.235386e-03 7.794232e-03 7.404521e-03
##          Resolved 8.337144e-03 7.809477e-03 9.849791e-03 7.985366e-03
##                   Accurate_Episode_Date
## training$Outcome1   2020-03-25   2020-03-26   2020-03-27   2020-03-28
##          Fatal    9.742790e-03 1.169135e-02 1.052221e-02 8.963367e-03
##          Resolved 8.618567e-03 7.739121e-03 9.111056e-03 8.829634e-03
##                   Accurate_Episode_Date
## training$Outcome1   2020-03-29   2020-03-30   2020-03-31   2020-04-01
##          Fatal    1.208106e-02 1.831645e-02 1.052221e-02 2.883866e-02
##          Resolved 7.528054e-03 1.086995e-02 9.216590e-03 1.276955e-02
##                   Accurate_Episode_Date
## training$Outcome1   2020-04-02   2020-04-03   2020-04-04   2020-04-05
##          Fatal    1.831645e-02 2.182385e-02 2.065472e-02 1.792673e-02
##          Resolved 9.814613e-03 1.217153e-02 1.129208e-02 9.322123e-03
##                   Accurate_Episode_Date
## training$Outcome1   2020-04-06   2020-04-07   2020-04-08   2020-04-09
##          Fatal    1.909587e-02 2.143414e-02 2.533125e-02 1.870616e-02
##          Resolved 1.213635e-02 1.224188e-02 1.143279e-02 1.255848e-02
##                   Accurate_Episode_Date
## training$Outcome1   2020-04-10   2020-04-11   2020-04-12   2020-04-13
##          Fatal    2.689010e-02 2.494154e-02 2.377241e-02 3.039751e-02
##          Resolved 1.495058e-02 1.526718e-02 1.350828e-02 1.801105e-02
##                   Accurate_Episode_Date
## training$Outcome1   2020-04-14   2020-04-15   2020-04-16   2020-04-17
##          Fatal    2.143414e-02 3.273578e-02 2.727981e-02 2.844895e-02
##          Resolved 1.765927e-02 1.755373e-02 1.649840e-02 1.913674e-02
##                   Accurate_Episode_Date
## training$Outcome1   2020-04-18   2020-04-19   2020-04-20   2020-04-21
##          Fatal    2.494154e-02 1.831645e-02 2.260327e-02 1.948558e-02
##          Resolved 1.607627e-02 9.990502e-03 1.519682e-02 1.224188e-02
##                   Accurate_Episode_Date
## training$Outcome1   2020-04-22   2020-04-23   2020-04-24   2020-04-25
##          Fatal    2.455183e-02 1.948558e-02 1.519875e-02 1.402962e-02
##          Resolved 1.291026e-02 1.238259e-02 1.153833e-02 1.252331e-02
##                   Accurate_Episode_Date
## training$Outcome1   2020-04-26   2020-04-27   2020-04-28   2020-04-29
##          Fatal    1.091193e-02 1.480904e-02 1.208106e-02 1.519875e-02
##          Resolved 9.955324e-03 1.371935e-02 1.157350e-02 1.108101e-02
##                   Accurate_Episode_Date
## training$Outcome1   2020-04-30   2020-05-01   2020-05-02   2020-05-03
##          Fatal    8.573655e-03 7.404521e-03 7.014809e-03 2.727981e-03
```

```
##          Resolved 1.217153e-02 1.343793e-02 9.533190e-03 8.126077e-03
##                   Accurate_Episode_Date
## training$Outcome1   2020-05-04   2020-05-05   2020-05-06   2020-05-07
##          Fatal    7.404521e-03 8.963367e-03 9.742790e-03 8.963367e-03
##          Resolved 9.462835e-03 9.884969e-03 7.739121e-03 9.814613e-03
##                   Accurate_Episode_Date
## training$Outcome1   2020-05-08   2020-05-09   2020-05-10   2020-05-11
##          Fatal    9.742790e-03 7.404521e-03 3.117693e-03 7.014809e-03
##          Resolved 1.027192e-02 8.513033e-03 8.055722e-03 1.132726e-02
##                   Accurate_Episode_Date
## training$Outcome1   2020-05-12   2020-05-13   2020-05-14   2020-05-15
##          Fatal    7.404521e-03 7.014809e-03 5.066251e-03 6.625097e-03
##          Resolved 1.009604e-02 9.111056e-03 1.132726e-02 1.410631e-02
##                   Accurate_Episode_Date
## training$Outcome1   2020-05-16   2020-05-17   2020-05-18   2020-05-19
##          Fatal    6.235386e-03 8.963367e-03 4.676539e-03 7.014809e-03
##          Resolved 9.603546e-03 9.251768e-03 1.129208e-02 1.164386e-02
##                   Accurate_Episode_Date
## training$Outcome1   2020-05-20   2020-05-21   2020-05-22   2020-05-23
##          Fatal    4.676539e-03 5.455963e-03 3.117693e-03 3.117693e-03
##          Resolved 1.210117e-02 1.065888e-02 8.266789e-03 8.899989e-03
##                   Accurate_Episode_Date
## training$Outcome1   2020-05-24   2020-05-25   2020-05-26   2020-05-27
##          Fatal    2.727981e-03 2.727981e-03 2.727981e-03 4.286828e-03
##          Resolved 8.548211e-03 1.143279e-02 9.568368e-03 8.829634e-03
##                   Accurate_Episode_Date
## training$Outcome1   2020-05-28   2020-05-29   2020-05-30   2020-05-31
##          Fatal    2.338270e-03 2.727981e-03 1.948558e-03 1.558846e-03
##          Resolved 7.809477e-03 1.111619e-02 8.864812e-03 6.648609e-03
##                   Accurate_Episode_Date
## training$Outcome1   2020-06-01   2020-06-02   2020-06-03   2020-06-04
##          Fatal    2.338270e-03 1.169135e-03 1.558846e-03 1.948558e-03
##          Resolved 1.122173e-02 7.492876e-03 7.211454e-03 6.965209e-03
##                   Accurate_Episode_Date
## training$Outcome1   2020-06-05   2020-06-06   2020-06-07   2020-06-08
##          Fatal    1.948558e-03 1.948558e-03 1.558846e-03 7.794232e-04
##          Resolved 6.683787e-03 5.769163e-03 4.960073e-03 6.507897e-03
##                   Accurate_Episode_Date
## training$Outcome1   2020-06-09   2020-06-10   2020-06-11   2020-06-12
##          Fatal    1.558846e-03 1.169135e-03 1.169135e-03 3.897116e-04
##          Resolved 5.522918e-03 6.543075e-03 5.487741e-03 5.733985e-03
##                   Accurate_Episode_Date
## training$Outcome1   2020-06-13   2020-06-14   2020-06-15   2020-06-16
##          Fatal    7.794232e-04 7.794232e-04 7.794232e-04 1.948558e-03
##          Resolved 3.588138e-03 4.573117e-03 6.296830e-03 5.135962e-03
##                   Accurate_Episode_Date
## training$Outcome1   2020-06-17   2020-06-18   2020-06-19   2020-06-20
##          Fatal    1.169135e-03 2.338270e-03 3.897116e-04 7.794232e-04
##          Resolved 5.382207e-03 4.678651e-03 5.769163e-03 4.889718e-03
##                   Accurate_Episode_Date
```

```
## training$Outcome1   2020-06-21   2020-06-22   2020-06-23   2020-06-24
##          Fatal    3.897116e-04 3.897116e-04 7.794232e-04 1.169135e-03
##          Resolved 4.221339e-04 4.924895e-04 7.035565e-05 7.035565e-05
##                   Accurate_Episode_Date
## training$Outcome1   2020-06-25   2020-06-27   2020-06-29   2020-06-30
##          Fatal    1.169135e-03 3.897116e-04 3.897116e-04 7.794232e-04
##          Resolved 1.758891e-04 1.407113e-04 1.055335e-04 7.035565e-05
##                   Accurate_Episode_Date
## training$Outcome1   2020-07-01
##          Fatal    3.897116e-04
##          Resolved 7.035565e-05
##
##                   Case_Reported_Date
## training$Outcome1   2020-01-23   2020-01-24   2020-02-21   2020-02-25
##          Fatal    3.915427e-04 3.915427e-04 3.915427e-04 3.915427e-04
##          Resolved 1.055706e-04 7.038041e-05 7.038041e-05 7.038041e-05
##                   Case_Reported_Date
## training$Outcome1   2020-02-26   2020-02-27   2020-02-28   2020-02-29
##          Fatal    3.915427e-04 3.915427e-04 3.915427e-04 3.915427e-04
##          Resolved 7.038041e-05 1.055706e-04 2.111412e-04 1.759510e-04
##                   Case_Reported_Date
## training$Outcome1   2020-03-01   2020-03-03   2020-03-04   2020-03-05
##          Fatal    3.915427e-04 3.915427e-04 3.915427e-04 3.915427e-04
##          Resolved 1.055706e-04 1.055706e-04 7.038041e-05 1.759510e-04
##                   Case_Reported_Date
## training$Outcome1   2020-03-06   2020-03-07   2020-03-08   2020-03-09
##          Fatal    3.915427e-04 3.915427e-04 7.830854e-04 3.915427e-04
##          Resolved 1.759510e-04 1.407608e-04 1.759510e-04 2.463314e-04
##                   Case_Reported_Date
## training$Outcome1   2020-03-10   2020-03-11   2020-03-12   2020-03-13
##          Fatal    3.915427e-04 3.915427e-04 7.830854e-04 7.830854e-04
##          Resolved 3.519020e-04 8.093747e-04 1.196467e-03 1.196467e-03
##                   Case_Reported_Date
## training$Outcome1   2020-03-14   2020-03-15   2020-03-16   2020-03-17
##          Fatal    3.915427e-04 3.915427e-04 7.830854e-04 1.174628e-03
##          Resolved 1.020516e-03 1.196467e-03 1.759510e-03 1.759510e-03
##                   Case_Reported_Date
## training$Outcome1   2020-03-18   2020-03-19   2020-03-20   2020-03-21
##          Fatal    1.174628e-03 1.174628e-03 1.566171e-03 3.523884e-03
##          Resolved 1.935461e-03 2.216983e-03 2.991167e-03 1.829891e-03
##                   Case_Reported_Date
## training$Outcome1   2020-03-22   2020-03-23   2020-03-24   2020-03-25
##          Fatal    1.957713e-03 3.523884e-03 2.740799e-03 4.306969e-03
##          Resolved 2.744836e-03 4.328395e-03 4.997009e-03 4.046873e-03
##                   Case_Reported_Date
## training$Outcome1   2020-03-26   2020-03-27   2020-03-28   2020-03-29
##          Fatal    4.306969e-03 6.656226e-03 7.439311e-03 1.018011e-02
##          Resolved 7.249182e-03 8.832741e-03 6.756519e-03 6.334237e-03
##                   Case_Reported_Date
## training$Outcome1   2020-03-30   2020-03-31   2020-04-01   2020-04-02
```

```
##          Fatal    7.830854e-03 1.213782e-02 1.213782e-02 1.801096e-02
##          Resolved 1.340747e-02 1.221100e-02 1.351304e-02 1.210543e-02
##                   Case_Reported_Date
## training$Outcome1   2020-04-03   2020-04-04   2020-04-05   2020-04-06
##          Fatal    2.975724e-02 1.292091e-02 1.566171e-02 2.897416e-02
##          Resolved 1.245733e-02 1.002921e-02 1.052187e-02 1.277404e-02
##                   Case_Reported_Date
## training$Outcome1   2020-04-07   2020-04-08   2020-04-09   2020-04-10
##          Fatal    2.310102e-02 2.388410e-02 1.683634e-02 2.114330e-02
##          Resolved 1.199986e-02 1.319633e-02 1.386494e-02 1.083858e-02
##                   Case_Reported_Date
## training$Outcome1   2020-04-11   2020-04-12   2020-04-13   2020-04-14
##          Fatal    2.349256e-02 2.819107e-02 2.584182e-02 2.858262e-02
##          Resolved 1.252771e-02 1.467431e-02 1.710244e-02 1.530774e-02
##                   Case_Reported_Date
## training$Outcome1   2020-04-15   2020-04-16   2020-04-17   2020-04-18
##          Fatal    3.210650e-02 3.093187e-02 3.719655e-02 2.740799e-02
##          Resolved 1.675054e-02 1.534293e-02 2.079741e-02 1.629306e-02
##                   Case_Reported_Date
## training$Outcome1   2020-04-19   2020-04-20   2020-04-21   2020-04-22
##          Fatal    1.996868e-02 1.918559e-02 2.270948e-02 2.701644e-02
##          Resolved 1.689130e-02 1.741915e-02 1.470950e-02 1.541331e-02
##                   Case_Reported_Date
## training$Outcome1   2020-04-23   2020-04-24   2020-04-25   2020-04-26
##          Fatal    2.310102e-02 1.527016e-02 2.114330e-02 1.487862e-02
##          Resolved 1.379456e-02 1.238695e-02 1.196467e-02 9.958827e-03
##                   Case_Reported_Date
## training$Outcome1   2020-04-27   2020-04-28   2020-04-29   2020-04-30
##          Fatal    1.722788e-02 1.527016e-02 1.487862e-02 1.370399e-02
##          Resolved 1.397051e-02 1.104972e-02 1.168315e-02 1.196467e-02
##                   Case_Reported_Date
## training$Outcome1   2020-05-01   2020-05-02   2020-05-03   2020-05-04
##          Fatal    1.409554e-02 8.222396e-03 9.005482e-03 5.873140e-03
##          Resolved 1.407608e-02 9.923637e-03 1.020516e-02 9.712496e-03
##                   Case_Reported_Date
## training$Outcome1   2020-05-05   2020-05-06   2020-05-07   2020-05-08
##          Fatal    9.005482e-03 1.174628e-02 7.830854e-03 8.222396e-03
##          Resolved 1.196467e-02 1.031073e-02 8.691980e-03 8.903121e-03
##                   Case_Reported_Date
## training$Outcome1   2020-05-09   2020-05-10   2020-05-11   2020-05-12
##          Fatal    8.613939e-03 9.397024e-03 6.264683e-03 9.005482e-03
##          Resolved 8.797551e-03 8.128937e-03 6.474997e-03 7.319562e-03
##                   Case_Reported_Date
## training$Outcome1   2020-05-13   2020-05-14   2020-05-15   2020-05-16
##          Fatal    8.222396e-03 7.439311e-03 8.222396e-03 6.264683e-03
##          Resolved 1.069782e-02 9.782876e-03 1.112010e-02 1.150720e-02
##                   Case_Reported_Date
## training$Outcome1   2020-05-17   2020-05-18   2020-05-19   2020-05-20
##          Fatal    8.222396e-03 7.830854e-03 4.698512e-03 7.830854e-03
##          Resolved 1.252771e-02 1.122567e-02 9.818067e-03 1.203505e-02
```

```
##                   Case_Reported_Date
## training$Outcome1   2020-05-21   2020-05-22   2020-05-23   2020-05-24
##          Fatal    4.698512e-03 7.047768e-03 6.656226e-03 4.698512e-03
##          Resolved 1.052187e-02 1.185910e-02 1.122567e-02 1.059225e-02
##                   Case_Reported_Date
## training$Outcome1   2020-05-25   2020-05-26   2020-05-27   2020-05-28
##          Fatal    5.873140e-03 2.740799e-03 2.740799e-03 4.306969e-03
##          Resolved 1.238695e-02 8.234508e-03 1.016997e-02 9.853257e-03
##                   Case_Reported_Date
## training$Outcome1   2020-05-29   2020-05-30   2020-05-31   2020-06-01
##          Fatal    5.481597e-03 3.132341e-03 3.915427e-03 2.349256e-03
##          Resolved 1.854524e-02 1.083858e-02 1.256290e-02 1.252771e-02
##                   Case_Reported_Date
## training$Outcome1   2020-06-02   2020-06-03   2020-06-04   2020-06-05
##          Fatal    3.915427e-04 2.740799e-03 1.957713e-03 1.957713e-03
##          Resolved 1.164796e-02 1.073301e-02 8.867931e-03 8.762361e-03
##                   Case_Reported_Date
## training$Outcome1   2020-06-06   2020-06-07   2020-06-08   2020-06-09
##          Fatal    1.957713e-03 3.915427e-04 1.566171e-03 2.349256e-03
##          Resolved 7.002850e-03 6.474997e-03 6.580568e-03 6.826899e-03
##                   Case_Reported_Date
## training$Outcome1   2020-06-10   2020-06-11   2020-06-12   2020-06-13
##          Fatal    1.566171e-03 2.349256e-03 7.830854e-04 1.174628e-03
##          Resolved 7.425133e-03 5.524862e-03 7.636274e-03 5.595242e-03
##                   Case_Reported_Date
## training$Outcome1   2020-06-14   2020-06-15   2020-06-16   2020-06-17
##          Fatal    3.915427e-04 7.830854e-04 1.174628e-03 1.566171e-03
##          Resolved 4.328395e-03 5.630432e-03 5.384101e-03 5.947144e-03
##                   Case_Reported_Date
## training$Outcome1   2020-06-18   2020-06-19   2020-06-20   2020-06-21
##          Fatal    2.740799e-03 1.174628e-03 1.957713e-03 1.566171e-03
##          Resolved 5.982335e-03 6.087905e-03 5.102579e-03 4.997009e-03
##                   Case_Reported_Date
## training$Outcome1   2020-06-22   2020-06-23   2020-06-24   2020-06-25
##          Fatal    7.830854e-04 7.830854e-04 7.830854e-04 1.174628e-03
##          Resolved 5.102579e-03 3.061548e-03 2.111412e-03 1.302038e-03
##                   Case_Reported_Date
## training$Outcome1   2020-06-26   2020-06-27   2020-06-28   2020-06-29
##          Fatal    1.174628e-03 3.915427e-04 3.915427e-04 3.915427e-04
##          Resolved 1.372418e-03 1.231657e-03 1.055706e-03 5.982335e-04
##                   Case_Reported_Date
## training$Outcome1   2020-06-30   2020-07-01   2020-07-02   2020-07-03
##          Fatal    7.830854e-04 3.915427e-04 1.174628e-03 3.915427e-04
##          Resolved 9.149453e-04 5.278530e-04 6.686139e-04 1.055706e-04
##                   Case_Reported_Date
## training$Outcome1   2020-07-04
##          Fatal    3.915427e-04
##          Resolved 3.167118e-04
##
##                   Test_Reported_Date
```

```
## training$Outcome1   2020-01-27   2020-02-03   2020-02-24   2020-02-25
##          Fatal    3.972984e-04 3.972984e-04 3.972984e-04 3.972984e-04
##          Resolved 1.068110e-04 7.120732e-05 7.120732e-05 7.120732e-05
##                   Test_Reported_Date
## training$Outcome1   2020-02-27   2020-02-28   2020-02-29   2020-03-01
##          Fatal    3.972984e-04 3.972984e-04 3.972984e-04 3.972984e-04
##          Resolved 1.068110e-04 1.780183e-04 7.120732e-05 2.848293e-04
##                   Test_Reported_Date
## training$Outcome1   2020-03-02   2020-03-03   2020-03-04   2020-03-05
##          Fatal    3.972984e-04 3.972984e-04 3.972984e-04 3.972984e-04
##          Resolved 1.068110e-04 7.120732e-05 1.068110e-04 1.780183e-04
##                   Test_Reported_Date
## training$Outcome1   2020-03-07   2020-03-08   2020-03-09   2020-03-10
##          Fatal    3.972984e-04 3.972984e-04 3.972984e-04 3.972984e-04
##          Resolved 1.068110e-04 1.424146e-04 1.780183e-04 1.780183e-04
##                   Test_Reported_Date
## training$Outcome1   2020-03-11   2020-03-12   2020-03-13   2020-03-14
##          Fatal    3.972984e-04 3.972984e-04 1.191895e-03 3.972984e-04
##          Resolved 3.916403e-04 7.476769e-04 8.900915e-04 1.459750e-03
##                   Test_Reported_Date
## training$Outcome1   2020-03-15   2020-03-16   2020-03-17   2020-03-18
##          Fatal    3.972984e-04 7.945967e-04 1.191895e-03 1.589193e-03
##          Resolved 8.900915e-04 9.256952e-04 1.424146e-03 1.637768e-03
##                   Test_Reported_Date
## training$Outcome1   2020-03-19   2020-03-20   2020-03-21   2020-03-22
##          Fatal    3.972984e-04 7.945967e-04 2.781089e-03 1.589193e-03
##          Resolved 1.637768e-03 2.492256e-03 1.388543e-03 2.314238e-03
##                   Test_Reported_Date
## training$Outcome1   2020-03-23   2020-03-24   2020-03-25   2020-03-26
##          Fatal    3.178387e-03 1.191895e-03 5.562177e-03 4.767580e-03
##          Resolved 4.343647e-03 4.806494e-03 4.735287e-03 7.298750e-03
##                   Test_Reported_Date
## training$Outcome1   2020-03-27   2020-03-28   2020-03-29   2020-03-30
##          Fatal    5.959476e-03 4.767580e-03 6.754072e-03 9.932459e-03
##          Resolved 8.829708e-03 6.195037e-03 7.049525e-03 1.395663e-02
##                   Test_Reported_Date
## training$Outcome1   2020-03-31   2020-04-01   2020-04-02   2020-04-03
##          Fatal    1.231625e-02 1.390544e-02 2.145411e-02 2.701629e-02
##          Resolved 1.167800e-02 1.203404e-02 1.246128e-02 1.249688e-02
##                   Test_Reported_Date
## training$Outcome1   2020-04-04   2020-04-05   2020-04-06   2020-04-07
##          Fatal    2.185141e-02 1.470004e-02 2.423520e-02 2.463250e-02
##          Resolved 1.096593e-02 9.434970e-03 1.171360e-02 1.278171e-02
##                   Test_Reported_Date
## training$Outcome1   2020-04-08   2020-04-09   2020-04-10   2020-04-11
##          Fatal    1.946762e-02 2.105681e-02 2.105681e-02 2.820818e-02
##          Resolved 1.381422e-02 1.377862e-02 1.100153e-02 1.281732e-02
##                   Test_Reported_Date
## training$Outcome1   2020-04-12   2020-04-13   2020-04-14   2020-04-15
##          Fatal    2.026222e-02 2.463250e-02 2.820818e-02 3.098927e-02
```

```
##          Resolved 1.388543e-02 1.755260e-02 1.577242e-02 1.655570e-02
##                   Test_Reported_Date
## training$Outcome1   2020-04-16   2020-04-17   2020-04-18   2020-04-19
##          Fatal    3.218117e-02 3.496226e-02 3.098927e-02 1.867302e-02
##          Resolved 1.591484e-02 1.929718e-02 1.659131e-02 1.538078e-02
##                   Test_Reported_Date
## training$Outcome1   2020-04-20   2020-04-21   2020-04-22   2020-04-23
##          Fatal    2.026222e-02 2.105681e-02 2.463250e-02 2.542710e-02
##          Resolved 1.758821e-02 1.513156e-02 1.473992e-02 1.456190e-02
##                   Test_Reported_Date
## training$Outcome1   2020-04-24   2020-04-25   2020-04-26   2020-04-27
##          Fatal    1.668653e-02 2.582439e-02 1.350814e-02 1.748113e-02
##          Resolved 1.274611e-02 1.441948e-02 9.541781e-03 1.345818e-02
##                   Test_Reported_Date
## training$Outcome1   2020-04-28   2020-04-29   2020-04-30   2020-05-01
##          Fatal    1.191895e-02 1.271355e-02 1.390544e-02 1.271355e-02
##          Resolved 9.862214e-03 1.121515e-02 1.182042e-02 1.434828e-02
##                   Test_Reported_Date
## training$Outcome1   2020-05-02   2020-05-03   2020-05-04   2020-05-05
##          Fatal    8.740564e-03 8.740564e-03 7.151371e-03 1.112435e-02
##          Resolved 9.969025e-03 1.046748e-02 1.032506e-02 1.146438e-02
##                   Test_Reported_Date
## training$Outcome1   2020-05-06   2020-05-07   2020-05-08   2020-05-09
##          Fatal    9.932459e-03 1.112435e-02 5.959476e-03 9.932459e-03
##          Resolved 1.018265e-02 7.975220e-03 9.577385e-03 8.188842e-03
##                   Test_Reported_Date
## training$Outcome1   2020-05-10   2020-05-11   2020-05-12   2020-05-13
##          Fatal    4.767580e-03 6.754072e-03 8.343266e-03 7.548669e-03
##          Resolved 8.260049e-03 6.622281e-03 7.939616e-03 1.068110e-02
##                   Test_Reported_Date
## training$Outcome1   2020-05-14   2020-05-15   2020-05-16   2020-05-17
##          Fatal    1.032976e-02 9.137863e-03 6.754072e-03 8.343266e-03
##          Resolved 1.007584e-02 1.117955e-02 1.157119e-02 1.306654e-02
##                   Test_Reported_Date
## training$Outcome1   2020-05-18   2020-05-19   2020-05-20   2020-05-21
##          Fatal    6.754072e-03 7.151371e-03 7.151371e-03 5.959476e-03
##          Resolved 1.103713e-02 9.399366e-03 1.206964e-02 1.093032e-02
##                   Test_Reported_Date
## training$Outcome1   2020-05-22   2020-05-23   2020-05-24   2020-05-25
##          Fatal    5.959476e-03 7.945967e-03 5.562177e-03 7.548669e-03
##          Resolved 1.185602e-02 1.199843e-02 1.036067e-02 1.580803e-02
##                   Test_Reported_Date
## training$Outcome1   2020-05-26   2020-05-27   2020-05-28   2020-05-29
##          Fatal    1.191895e-03 2.781089e-03 4.370282e-03 5.562177e-03
##          Resolved 7.654787e-03 9.506177e-03 1.025385e-02 1.751700e-02
##                   Test_Reported_Date
## training$Outcome1   2020-05-30   2020-05-31   2020-06-01   2020-06-02
##          Fatal    3.575685e-03 3.178387e-03 3.575685e-03 7.945967e-04
##          Resolved 1.288852e-02 1.107274e-02 1.128636e-02 1.203404e-02
##                   Test_Reported_Date
```

```
## training$Outcome1   2020-06-03   2020-06-04   2020-06-05   2020-06-06
##          Fatal    2.383790e-03 1.589193e-03 1.986492e-03 1.986492e-03
##          Resolved 1.185602e-02 8.865311e-03 8.794104e-03 6.978317e-03
##                   Test_Reported_Date
## training$Outcome1   2020-06-07   2020-06-08   2020-06-09   2020-06-10
##          Fatal    3.972984e-04 2.383790e-03 1.986492e-03 1.986492e-03
##          Resolved 6.551073e-03 7.085128e-03 7.120732e-03 7.654787e-03
##                   Test_Reported_Date
## training$Outcome1   2020-06-11   2020-06-12   2020-06-13   2020-06-14
##          Fatal    1.986492e-03 7.945967e-04 1.191895e-03 3.972984e-04
##          Resolved 5.803397e-03 7.263147e-03 5.589775e-03 4.770890e-03
##                   Test_Reported_Date
## training$Outcome1   2020-06-15   2020-06-16   2020-06-17   2020-06-18
##          Fatal    7.945967e-04 1.589193e-03 1.589193e-03 2.383790e-03
##          Resolved 5.589775e-03 5.589775e-03 5.554171e-03 5.910208e-03
##                   Test_Reported_Date
## training$Outcome1   2020-06-19   2020-06-20   2020-06-21   2020-06-22
##          Fatal    1.191895e-03 1.986492e-03 1.589193e-03 7.945967e-04
##          Resolved 6.408659e-03 4.735287e-03 5.162531e-03 5.340549e-03
##                   Test_Reported_Date
## training$Outcome1   2020-06-23   2020-06-24   2020-06-25   2020-06-26
##          Fatal    7.945967e-04 1.191895e-03 7.945967e-04 1.191895e-03
##          Resolved 3.097518e-03 2.065012e-03 1.139317e-03 1.530957e-03
##                   Test_Reported_Date
## training$Outcome1   2020-06-27   2020-06-28   2020-06-29   2020-06-30
##          Fatal    3.972984e-04 3.972984e-04 3.972984e-04 7.945967e-04
##          Resolved 1.139317e-03 1.068110e-03 7.476769e-04 8.544878e-04
##                   Test_Reported_Date
## training$Outcome1   2020-07-01   2020-07-02   2020-07-03   2020-07-04
##          Fatal    3.972984e-04 7.945967e-04 3.972984e-04 3.972984e-04
##          Resolved 5.696586e-04 6.764695e-04 1.068110e-04 3.560366e-04
##
##                   Specimen_Date
## training$Outcome1   2020-01-23   2020-01-24   2020-01-25   2020-02-20
##          Fatal    3.957262e-04 3.957262e-04 3.957262e-04 3.957262e-04
##          Resolved 7.095973e-05 7.095973e-05 7.095973e-05 7.095973e-05
##                   Specimen_Date
## training$Outcome1   2020-02-22   2020-02-23   2020-02-25   2020-02-26
##          Fatal    3.957262e-04 3.957262e-04 3.957262e-04 3.957262e-04
##          Resolved 1.419195e-04 7.095973e-05 7.095973e-05 7.095973e-05
##                   Specimen_Date
## training$Outcome1   2020-02-27   2020-02-28   2020-02-29   2020-03-01
##          Fatal    3.957262e-04 3.957262e-04 3.957262e-04 3.957262e-04
##          Resolved 1.773993e-04 1.773993e-04 1.773993e-04 1.419195e-04
##                   Specimen_Date
## training$Outcome1   2020-03-02   2020-03-03   2020-03-04   2020-03-05
##          Fatal    3.957262e-04 3.957262e-04 3.957262e-04 3.957262e-04
##          Resolved 1.064396e-04 1.419195e-04 2.483591e-04 7.095973e-05
##                   Specimen_Date
## training$Outcome1   2020-03-06   2020-03-07   2020-03-08   2020-03-09
```

```
##          Fatal    3.957262e-04 3.957262e-04 3.957262e-04 3.957262e-04
##          Resolved 7.095973e-05 1.773993e-04 2.128792e-04 3.547987e-04
##                   Specimen_Date
## training$Outcome1   2020-03-10   2020-03-11   2020-03-12   2020-03-13
##          Fatal    3.957262e-04 7.914523e-04 7.914523e-04 7.914523e-04
##          Resolved 6.031577e-04 8.869966e-04 2.057832e-03 1.525634e-03
##                   Specimen_Date
## training$Outcome1   2020-03-14   2020-03-15   2020-03-16   2020-03-17
##          Fatal    3.957262e-04 1.187178e-03 1.582905e-03 1.187178e-03
##          Resolved 1.667554e-03 1.596594e-03 3.015789e-03 3.902785e-03
##                   Specimen_Date
## training$Outcome1   2020-03-18   2020-03-19   2020-03-20   2020-03-21
##          Fatal    1.582905e-03 1.978631e-03 3.561535e-03 2.374357e-03
##          Resolved 3.796346e-03 4.683342e-03 6.102537e-03 4.967181e-03
##                   Specimen_Date
## training$Outcome1   2020-03-22   2020-03-23   2020-03-24   2020-03-25
##          Fatal    4.352988e-03 6.727345e-03 3.165809e-03 5.540166e-03
##          Resolved 5.180060e-03 7.025013e-03 5.960617e-03 8.160369e-03
##                   Specimen_Date
## training$Outcome1   2020-03-26   2020-03-27   2020-03-28   2020-03-29
##          Fatal    7.914523e-03 6.331619e-03 1.028888e-02 7.914523e-03
##          Resolved 7.237892e-03 7.947490e-03 8.799007e-03 6.421856e-03
##                   Specimen_Date
## training$Outcome1   2020-03-30   2020-03-31   2020-04-01   2020-04-02
##          Fatal    1.305896e-02 1.622477e-02 2.690938e-02 2.295212e-02
##          Resolved 1.185027e-02 1.185027e-02 1.021820e-02 1.323399e-02
##                   Specimen_Date
## training$Outcome1   2020-04-03   2020-04-04   2020-04-05   2020-04-06
##          Fatal    1.978631e-02 1.978631e-02 1.939058e-02 2.532647e-02
##          Resolved 1.032464e-02 9.011886e-03 9.402164e-03 1.309207e-02
##                   Specimen_Date
## training$Outcome1   2020-04-07   2020-04-08   2020-04-09   2020-04-10
##          Fatal    2.097349e-02 2.532647e-02 2.651365e-02 2.690938e-02
##          Resolved 1.422743e-02 1.227603e-02 1.348235e-02 1.589498e-02
##                   Specimen_Date
## training$Outcome1   2020-04-11   2020-04-12   2020-04-13   2020-04-14
##          Fatal    2.453502e-02 1.939058e-02 3.680253e-02 2.809656e-02
##          Resolved 1.664006e-02 1.185027e-02 2.072024e-02 1.944297e-02
##                   Specimen_Date
## training$Outcome1   2020-04-15   2020-04-16   2020-04-17   2020-04-18
##          Fatal    3.561535e-02 3.561535e-02 2.374357e-02 2.334784e-02
##          Resolved 2.001064e-02 1.798829e-02 2.004612e-02 1.472414e-02
##                   Specimen_Date
## training$Outcome1   2020-04-19   2020-04-20   2020-04-21   2020-04-22
##          Fatal    8.705975e-03 2.255639e-02 1.859913e-02 2.532647e-02
##          Resolved 9.898882e-03 1.727869e-02 1.397907e-02 1.344687e-02
##                   Specimen_Date
## training$Outcome1   2020-04-23   2020-04-24   2020-04-25   2020-04-26
##          Fatal    1.780768e-02 1.662050e-02 1.582905e-02 9.893154e-03
##          Resolved 1.387263e-02 1.337591e-02 1.234699e-02 7.912010e-03
```

```
##                   Specimen_Date
## training$Outcome1   2020-04-27   2020-04-28   2020-04-29   2020-04-30
##          Fatal    2.334784e-02 1.582905e-02 1.464187e-02 8.310249e-03
##          Resolved 1.380167e-02 1.245343e-02 1.468866e-02 1.376619e-02
##                   Specimen_Date
## training$Outcome1   2020-05-01   2020-05-02   2020-05-03   2020-05-04
##          Fatal    8.705975e-03 6.331619e-03 4.352988e-03 1.028888e-02
##          Resolved 1.167288e-02 8.799007e-03 7.060493e-03 1.153096e-02
##                   Specimen_Date
## training$Outcome1   2020-05-05   2020-05-06   2020-05-07   2020-05-08
##          Fatal    9.893154e-03 1.028888e-02 1.147606e-02 8.705975e-03
##          Resolved 1.209863e-02 9.721483e-03 1.085684e-02 1.043108e-02
##                   Specimen_Date
## training$Outcome1   2020-05-09   2020-05-10   2020-05-11   2020-05-12
##          Fatal    1.108033e-02 5.144440e-03 1.147606e-02 9.497428e-03
##          Resolved 7.770090e-03 5.357460e-03 1.227603e-02 1.032464e-02
##                   Specimen_Date
## training$Outcome1   2020-05-13   2020-05-14   2020-05-15   2020-05-16
##          Fatal    7.123071e-03 5.935892e-03 4.748714e-03 5.935892e-03
##          Resolved 9.969842e-03 9.863403e-03 1.224055e-02 9.686003e-03
##                   Specimen_Date
## training$Outcome1   2020-05-17   2020-05-18   2020-05-19   2020-05-20
##          Fatal    7.123071e-03 4.352988e-03 5.540166e-03 9.893154e-03
##          Resolved 8.728047e-03 7.166933e-03 1.447578e-02 1.305659e-02
##                   Specimen_Date
## training$Outcome1   2020-05-21   2020-05-22   2020-05-23   2020-05-24
##          Fatal    4.352988e-03 4.748714e-03 6.727345e-03 3.957262e-03
##          Resolved 1.220507e-02 1.067944e-02 8.302288e-03 9.650523e-03
##                   Specimen_Date
## training$Outcome1   2020-05-25   2020-05-26   2020-05-27   2020-05-28
##          Fatal    3.957262e-03 2.770083e-03 3.957262e-03 3.165809e-03
##          Resolved 1.153096e-02 1.174384e-02 9.721483e-03 1.082136e-02
##                   Specimen_Date
## training$Outcome1   2020-05-29   2020-05-30   2020-05-31   2020-06-01
##          Fatal    2.374357e-03 3.957262e-03 1.187178e-03 1.978631e-03
##          Resolved 1.401455e-02 8.408728e-03 7.770090e-03 1.330495e-02
##                   Specimen_Date
## training$Outcome1   2020-06-02   2020-06-03   2020-06-04   2020-06-05
##          Fatal    1.187178e-03 1.978631e-03 2.374357e-03 7.914523e-04
##          Resolved 8.799007e-03 7.486252e-03 9.082845e-03 7.947490e-03
##                   Specimen_Date
## training$Outcome1   2020-06-06   2020-06-07   2020-06-08   2020-06-09
##          Fatal    1.187178e-03 1.978631e-03 1.187178e-03 1.978631e-03
##          Resolved 6.208976e-03 4.470463e-03 7.308852e-03 6.279936e-03
##                   Specimen_Date
## training$Outcome1   2020-06-10   2020-06-11   2020-06-12   2020-06-13
##          Fatal    1.187178e-03 7.914523e-04 1.187178e-03 1.187178e-03
##          Resolved 7.166933e-03 7.095973e-03 5.570339e-03 4.612382e-03
##                   Specimen_Date
## training$Outcome1   2020-06-14   2020-06-15   2020-06-16   2020-06-17
```

```
##          Fatal    7.914523e-04 7.914523e-04 2.770083e-03 7.914523e-04
##          Resolved 3.796346e-03 6.599255e-03 5.676778e-03 6.315416e-03
##                   Specimen_Date
## training$Outcome1   2020-06-18   2020-06-19   2020-06-20   2020-06-21
##          Fatal    1.978631e-03 7.914523e-04 7.914523e-04 7.914523e-04
##          Resolved 5.038141e-03 5.321980e-03 5.180060e-03 1.596594e-03
##                   Specimen_Date
## training$Outcome1   2020-06-22   2020-06-23   2020-06-24   2020-06-25
##          Fatal    7.914523e-04 7.914523e-04 1.187178e-03 7.914523e-04
##          Resolved 2.412631e-03 1.525634e-03 1.667554e-03 1.241795e-03
##                   Specimen_Date
## training$Outcome1   2020-06-26   2020-06-27   2020-06-28   2020-06-29
##          Fatal    3.957262e-04 3.957262e-04 7.914523e-04 3.957262e-04
##          Resolved 7.450772e-04 5.676778e-04 3.193188e-04 5.321980e-04
##                   Specimen_Date
## training$Outcome1   2020-06-30   2020-07-01   2020-07-02   2020-07-03
##          Fatal    7.914523e-04 7.914523e-04 3.957262e-04 3.957262e-04
##          Resolved 4.612382e-04 1.419195e-04 1.773993e-04 7.095973e-05
##
##                   Age_Group
## training$Outcome1          <20          20s          30s          40s
##          Fatal    0.0008227067 0.0016454134 0.0028794735 0.0090497738
##          Resolved 0.0511769280 0.1613769704 0.1485827384 0.1548031385
##                   Age_Group
## training$Outcome1          50s          60s          70s          80s
##          Fatal    0.0316742081 0.0888523241 0.1789387084 0.3615795969
##          Resolved 0.1741005160 0.1166678448 0.0670460168 0.0783204920
##                   Age_Group
## training$Outcome1          90s      UNKNOWN
##          Fatal    0.3241464418 0.0004113534
##          Resolved 0.0476072666 0.0003180886
##
##                   Client_Gender
## training$Outcome1       FEMALE         MALE        OTHER  TRANSGENDER
##          Fatal    0.5272052762 0.4554822754 0.0004122012 0.0008244023
##          Resolved 0.5390434444 0.4535685249 0.0002827954 0.0002474460
##                   Client_Gender
## training$Outcome1      UNKNOWN
##          Fatal    0.0160758450
##          Resolved 0.0068577892
##
##                   Case_AcquisitionInfo
## training$Outcome1          CC No Epi-link No Info-Missing No Info-Unk
##          Fatal    0.062216728 0.087350639     0.008240626 0.028842192
##          Resolved 0.321138211 0.191869919     0.023365147 0.035418876
##                   Case_AcquisitionInfo
## training$Outcome1          OB      Travel
##          Fatal    0.792336218 0.021013597
##          Resolved 0.375114882 0.053092966
##
```

```
##                   Outcome1
## training$Outcome1        Fatal     Resolved
##          Fatal    9.995873e-01 4.127115e-04
##          Resolved 3.535318e-05 9.999646e-01
##
##                   Outbreak_Related
## training$Outcome1       Yes        No
##          Fatal    0.7936442 0.2063558
##          Resolved 0.3754154 0.6245846
##
##                   Reporting_PHU_City
## training$Outcome1       Barrie   Belleville    Brantford   Brockville
##          Fatal    0.0122199593 0.0024439919 0.0024439919 0.0199592668
##          Resolved 0.0180450597 0.0013065894 0.0039550816 0.0096051981
##                   Reporting_PHU_City
## training$Outcome1      Chatham     Cornwall       Guelph     Hamilton
##          Fatal    0.0008146640 0.0044806517 0.0134419552 0.0158859470
##          Resolved 0.0047672858 0.0048732255 0.0132777739 0.0240129953
##                   Reporting_PHU_City
## training$Outcome1       Kenora     Kingston       London  Mississauga
##          Fatal    0.0004073320 0.0004073320 0.0203665988 0.1140529532
##          Resolved 0.0012359630 0.0021541069 0.0179391200 0.1676672081
##                   Reporting_PHU_City
## training$Outcome1 New Liskeard    Newmarket    North Bay     Oakville
##          Fatal    0.0004073320 0.0908350305 0.0008146640 0.0093686354
##          Resolved 0.0006356381 0.0838336041 0.0009534572 0.0228123455
##                   Reporting_PHU_City
## training$Outcome1       Ottawa   Owen Sound     Pembroke Peterborough
##          Fatal    0.0965376782 0.0004073320 0.0008146640 0.0012219959
##          Resolved 0.0575958754 0.0033194435 0.0009181439 0.0030369376
##                   Reporting_PHU_City
## training$Outcome1 Point Edward    Port Hope Sault Ste. Marie       Simcoe
##          Fatal    0.0101832994 0.0073319756     0.0004073320 0.0126272912
##          Resolved 0.0083339219 0.0056148033     0.0008475175 0.0120771241
##                   Reporting_PHU_City
## training$Outcome1   St. Thomas    Stratford      Sudbury      Thorold
##          Fatal    0.0024439919 0.0024439919 0.0012219959 0.0244399185
##          Resolved 0.0025072392 0.0016244085 0.0020481672 0.0217176354
##                   Reporting_PHU_City
## training$Outcome1  Thunder Bay      Timmins      Toronto     Waterloo
##          Fatal    0.0008146640 0.0028513238 0.3885947047 0.0419551935
##          Resolved 0.0029309980 0.0016597217 0.3735080161 0.0364079384
##                   Reporting_PHU_City
## training$Outcome1       Whitby      Windsor
##          Fatal    0.0712830957 0.0260692464
##          Resolved 0.0476022318 0.0411752242
##
##                   Reporting_PHU_Latitude
## training$Outcome1     [,1]      [,2]
##          Fatal    43.82831 0.6808654
```

```
##          Resolved 43.74926 0.7165834
##
##                  Reporting_PHU_Longitude
## training$Outcome1      [,1]     [,2]
##          Fatal    -79.20247 1.519114
##          Resolved -79.51879 1.565332
```

```
#Confusion matrix
NBC.predict <- predict(NBC.classifier, test)
CrossTable(NBC.predict, test$Outcome1, prop.chisq = FALSE, prop.t = FALSE, pr
op.r = FALSE,
dnn = c('predicted', 'actual'))
```

```
##
##
##    Cell Contents
## |-------------------------|
## |                       N |
## |           N / Col Total |
## |-------------------------|
##
##
## Total Observations in Table:  3410
##
##
##              | actual
##    predicted |     Fatal |  Resolved | Row Total |
## -------------|-----------|-----------|-----------|
##        Fatal |       264 |         0 |       264 |
##              |     0.985 |     0.000 |           |
## -------------|-----------|-----------|-----------|
##     Resolved |         4 |      3142 |      3146 |
##              |     0.015 |     1.000 |           |
## -------------|-----------|-----------|-----------|
## Column Total |       268 |      3142 |      3410 |
##              |     0.079 |     0.921 |           |
## -------------|-----------|-----------|-----------|
##
##
```

Step 4: Modeling - (3) Random Forest

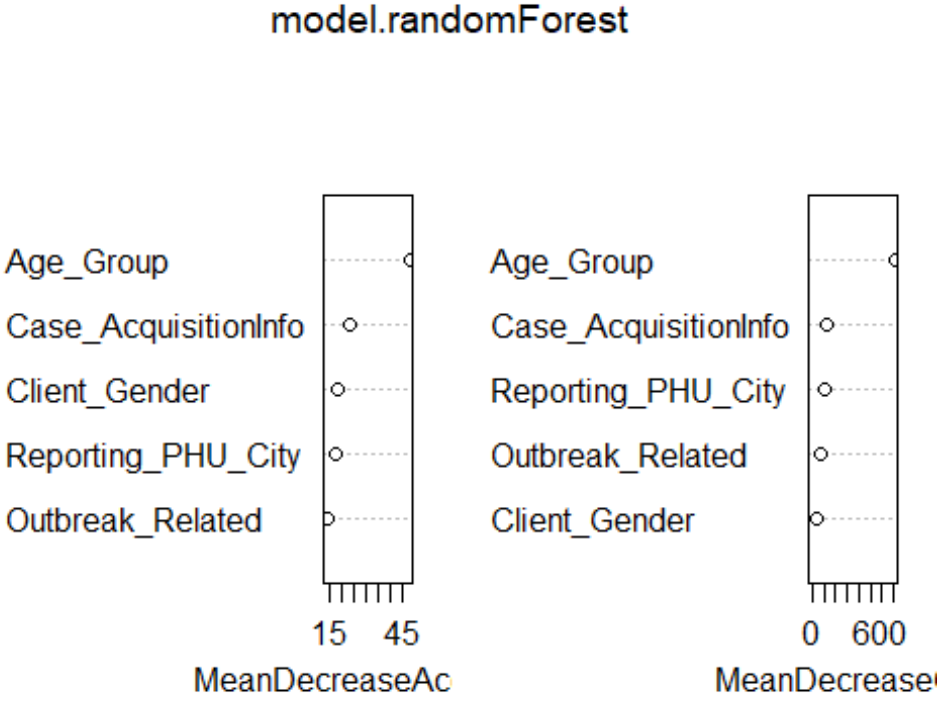
```
#install.packages("randomForest")
library(randomForest)
```

```
## Warning: package 'randomForest' was built under R version 3.5.3
```

```
## randomForest 4.6-14
```

```
## Type rfNews() to see new features/changes/bug fixes.
```

```
##
## Attaching package: 'randomForest'

## The following object is masked from 'package:ggplot2':
##
##     margin

#Building Random Forest model on training set
model.randomForest <- randomForest(Outcome1 ~ Client_Gender + Age_Group + Cas
e_AcquisitionInfo + Outbreak_Related + Reporting_PHU_City, data=training, imp
ortance = TRUE)
model.randomForest

##
## Call:
##  randomForest(formula = Outcome1 ~ Client_Gender + Age_Group +      Case_A
cquisitionInfo + Outbreak_Related + Reporting_PHU_City,      data = training,
importance = TRUE)
##                Type of random forest: classification
##                      Number of trees: 500
## No. of variables tried at each split: 2
##
##         OOB estimate of  error rate: 7.89%
## Confusion matrix:
##          Fatal Resolved  class.error
## Fatal        8     2413 0.9966955803
## Resolved    11    28273 0.0003889125

#Predicting on test set
predict.RF <- predict(model.randomForest, test, type = "class")
head(predict.RF)

##       16       31       66       99      102      110
## Resolved Resolved Resolved Resolved Resolved Resolved
## Levels: Fatal Resolved

#Confusion Matrix
table(predict.RF, test$Outcome1)

##
## predict.RF Fatal Resolved
##   Fatal        0        0
##   Resolved   268     3142

#importance that model has assigned to each variable
varImpPlot(model.randomForest)
```


model.randomForest
Age_Group
Case_AcquisitionInfo
Client_Gender
Reporting_PHU_City
Outbreak_Related
15
45
MeanDecreaseAc
Age_Group
Case_AcquisitionInfo
Reporting_PHU_City
Outbreak_Related
Client_Gender
0
600
MeanDecrease