

CKME 136 Final Results

Data Description

#Read data into a dataframe

```
data <- read.csv(file="C:/Users/krist/Desktop/CKME136/03. Final Project/Data  
July 13/conposcovidloc (2).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:   658      2020-04-17:   750
## 1st Qu.: 9210      2020-04-13:   655      2020-04-15:   613
## Median :18420      2020-04-15:   650      2020-04-13:   608
## Mean    :18420      2020-04-14:   621      2020-05-29:   608
## 3rd Qu.:27630      2020-04-16:   603      2020-04-20:   604
## Max.    :36839      (Other)   :33650      2020-04-18:   603
##              NA's      :    2      (Other)   :33053
## Test_Reported_Date Specimen_Date Age_Group Client_Gender
## 2020-04-17:   689      2020-04-13:   746      50s      :5927 FEMALE      :19545
## 2020-04-18:   611      2020-04-15:   733      20s      :5730 MALE        :17014
## 2020-04-20:   607      2020-04-14:   690      40s      :5243 OTHER       :    9
## 2020-04-13:   604      2020-04-17:   687      30s      :5193 TRANSGENDER:    8
## 2020-04-15:   593      2020-04-16:   668      60s      :4150 UNKNOWN    :   263
## (Other)   :33327      (Other)   :33030      80s      :3520
## NA's      :   408      NA's      :   285      (Other):7076
## Case_AcquisitionInfo Outcome1 Outbreak_Related
## CC      :11405      Fatal      : 2722 Yes :14593
## No Epi-link : 7313      Not Resolved: 1454 NA's:22246
## No Info-Missing: 880      Resolved    :32663
## No Info-Unk :   851
## OB      :14582
## Travel    : 1808
##
##              Reporting_PHU
## Toronto Public Health      :13673
## Peel Public Health         : 6184
## York Region Public Health Services: 3130
## Ottawa Public Health       : 2158
## Windsor-Essex County Health Unit : 1805
## Durham Region Health Department : 1751
## (Other)                    : 8138
##              Reporting_PHU_Address Reporting_PHU_City
## 277 Victoria Street, 5th Floor:13673 Toronto      :13673
## 7120 Hurontario Street      : 6184 Mississauga: 6184
## 17250 Yonge Street          : 3130 Newmarket   : 3130
```

```

## 100 Constellation Drive      : 2158      Ottawa      : 2158
## 1005 Ouellette Avenue        : 1805      Windsor      : 1805
## 605 Rossland Road East       : 1751      Whitby       : 1751
## (Other)                      : 8138      (Other)      : 8138
## Reporting_PHU_Postal_Code
## M5B 1W2:13673
## L5W 1N4: 6184
## L3Y 6Z1: 3130
## K2G 6J8: 2158
## N9A 4J8: 1805
## L1N 0B2: 1751
## (Other): 8138
##
##                                     Reporting_PHU_Website
## www.toronto.ca/community-people/health-wellness-care/      :13673
## www.peelregion.ca/health/                                   : 6184
## www.york.ca/wps/portal/yorkhome/health/                     : 3130
## www.ottawapublichealth.ca                                   : 2158
## www.wechu.org                                                : 1805
## www.durham.ca/en/health-and-wellness/health-and-wellness.aspx: 1751
## (Other)                                                       : 8138
## 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     :-79.53
## 3rd Qu.:43.66      3rd Qu.: -79.38
## Max.     :49.77      Max.     :-74.74
##
str(data)

## 'data.frame': 36839 obs. of 17 variables:
## $ Row_ID : int 1 2 3 4 5 6 7 8 9 10 ...
## $ Accurate_Episode_Date : Factor w/ 158 levels "2020-01-01","2020-01-10",...: 31 26 30 26 27 30 33 33 33 35 ...
## $ Case_Reported_Date : Factor w/ 141 levels "2020-01-23","2020-01-24",...: 16 16 17 16 17 17 18 17 18 18 ...
## $ Test_Reported_Date : Factor w/ 140 levels "2020-01-27","2020-02-03",...: 17 15 16 18 17 16 18 17 16 17 ...
## $ Specimen_Date : Factor w/ 144 levels "2020-01-23","2020-01-24",...: 20 20 20 18 20 19 20 21 20 21 ...
## $ Age_Group : Factor w/ 10 levels "<20","20s","30s",...: 5
## 4 3 4 3 5 2 2 3 1 ...
## $ Client_Gender : Factor w/ 5 levels "FEMALE","MALE",...: 2 2 1
## 2 2 2 1 2 1 1 ...
## $ Case_AcquisitionInfo : Factor w/ 6 levels "CC","No Epi-link",...: 6
## 6 6 6 6 2 6 1 6 1 ...
## $ Outcome1 : Factor w/ 3 levels "Fatal","Not Resolved",...: 3
## 3 3 3 3 3 3 3 3 3 3 ...
## $ Outbreak_Related : Factor w/ 1 level "Yes": NA NA NA 1 NA NA NA

```

```

NA NA NA ...
## $ Reporting_PHU      : Factor w/ 34 levels "Algoma Public Health
Unit",...: 34 31 34 20 31 28 21 31 9 21 ...
## $ Reporting_PHU_Address : Factor w/ 34 levels "100 Constellation
Drive",...: 14 23 14 1 23 9 32 23 6 32 ...
## $ Reporting_PHU_City   : Factor w/ 34 levels
"Barrie","Belleville",...: 14 31 14 17 31 27 12 31 16 12 ...
## $ Reporting_PHU_Postal_Code: Factor w/ 34 levels "K2G 6J8","K6J 5T1",...:
11 16 11 1 16 30 13 16 14 13 ...
## $ Reporting_PHU_Website : Factor w/ 34 levels
"www.algomapublichealth.com",...: 34 31 34 19 31 22 20 31 6 20 ...
## $ Reporting_PHU_Latitude : num  44 43.7 44 45.3 43.7 ...
## $ Reporting_PHU_Longitude : num  -79.5 -79.4 -79.5 -75.8 -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             408
## Specimen_Date           Age_Group
##          285             0
## Client_Gender      Case_AcquisitionInfo
##           0             0
## Outcome1           Outbreak_Related
##           0             22246
## 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
##          36839             159
## Case_Reported_Date      Test_Reported_Date
##          141             141
## Specimen_Date           Age_Group
##          145             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] 21006

length(dataclean$Outbreak_Related)

## [1] 35385

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

##
##   Yes    No  <NA>
## 14379 21006     0

str(dataclean$Outbreak_Related)

```

```
## Factor w/ 2 levels "Yes","No": 2 2 2 1 2 2 2 2 2 2 ...

#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: 655	2020-04-17: 748
##	1st Qu.: 8872	2020-04-13: 653	2020-04-15: 611
##	Median :17759	2020-04-15: 647	2020-05-29: 608
##	Mean :18081	2020-04-14: 617	2020-04-13: 607
##	3rd Qu.:27424	2020-04-16: 602	2020-04-20: 601
##	Max. :36839	(Other) :32209	2020-04-18: 599
##		NA's : 2	(Other) :31611

##	Test_Reported_Date	Specimen_Date	Age_Group	Client_Gender
##	2020-04-17: 685	2020-04-13: 743	50s :5721	FEMALE :18903
##	2020-04-18: 606	2020-04-15: 728	20s :5396	MALE :16207
##	2020-04-20: 606	2020-04-14: 687	40s :5037	OTHER : 9
##	2020-04-13: 604	2020-04-17: 684	30s :4956	TRANSGENDER: 8
##	2020-04-15: 590	2020-04-16: 667	60s :4011	UNKNOWN : 258
##	(Other) :31903	(Other) :31607	80s :3457	
##	NA's : 391	NA's : 269	(Other):6807	

##	Case_AcquisitionInfo	Outcome1	Outbreak_Related
##	CC :10820	Fatal : 2722	Yes:14379
##	No Epi-link : 6957	Resolved:32663	No :21006
##	No Info-Missing: 661		
##	No Info-Unk : 815		
##	OB :14368		
##	Travel : 1764		

##	Reporting_PHU_City	Reporting_PHU_Latitude	Reporting_PHU_Longitude
##	Toronto :13183	Min. :42.31	Min. : -94.49
##	Mississauga: 5838	1st Qu.:43.65	1st Qu.: -79.71
##	Newmarket : 2997	Median :43.66	Median : -79.38
##	Ottawa : 2104	Mean :43.75	Mean : -79.52
##	Whitby : 1706	3rd Qu.:43.66	3rd Qu.: -79.38
##	Windsor : 1600	Max. :49.77	Max. : -74.74
##	(Other) : 7957		

```
str(dataclean)
```

```
## 'data.frame': 35385 obs. of 13 variables:
## $ Row_ID : int 1 2 3 4 5 6 7 8 9 10 ...
## $ Accurate_Episode_Date : Factor w/ 154 levels "2020-01-01","2020-01-10",...: 31 26 30 26 27 30 33 33 33 35 ...
## $ Case_Reported_Date : Factor w/ 141 levels "2020-01-23","2020-01-
```

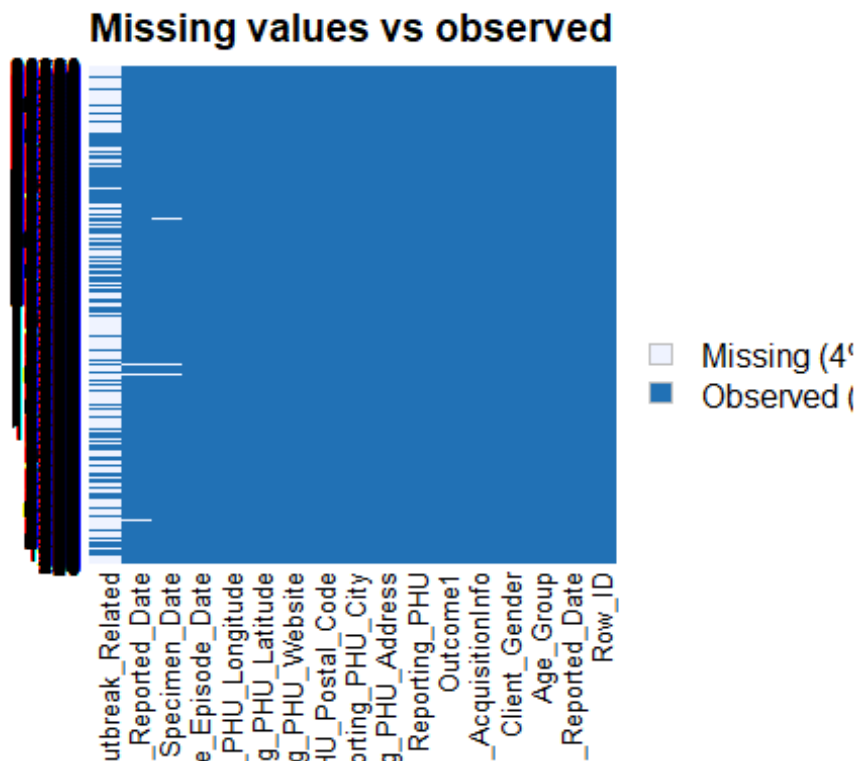
```

24",...: 16 16 17 16 17 17 18 17 18 18 ...
## $ Test_Reported_Date      : Factor w/ 140 levels "2020-01-27","2020-02-
03",...: 17 15 16 18 17 16 18 17 16 17 ...
## $ Specimen_Date          : Factor w/ 143 levels "2020-01-23","2020-01-
24",...: 20 20 20 18 20 19 20 21 20 21 ...
## $ Age_Group              : Factor w/ 10 levels "<20","20s","30s",...: 5 4
3 4 3 5 2 2 3 1 ...
## $ Client_Gender          : Factor w/ 5 levels "FEMALE","MALE",...: 2 2 1 2
2 2 1 2 1 1 ...
## $ Case_AcquisitionInfo    : Factor w/ 6 levels "CC","No Epi-link",...: 6 6
6 6 6 2 6 1 6 1 ...
## $ 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 1 2 2 2 2
2 2 ...
## $ Reporting_PHU_City      : Factor w/ 34 levels "Barrie","Belleville",...:
14 31 14 17 31 27 12 31 16 12 ...
## $ Reporting_PHU_Latitude  : num 44 43.7 44 45.3 43.7 ...
## $ Reporting_PHU_Longitude: num -79.5 -79.4 -79.5 -75.8 -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]
}

```

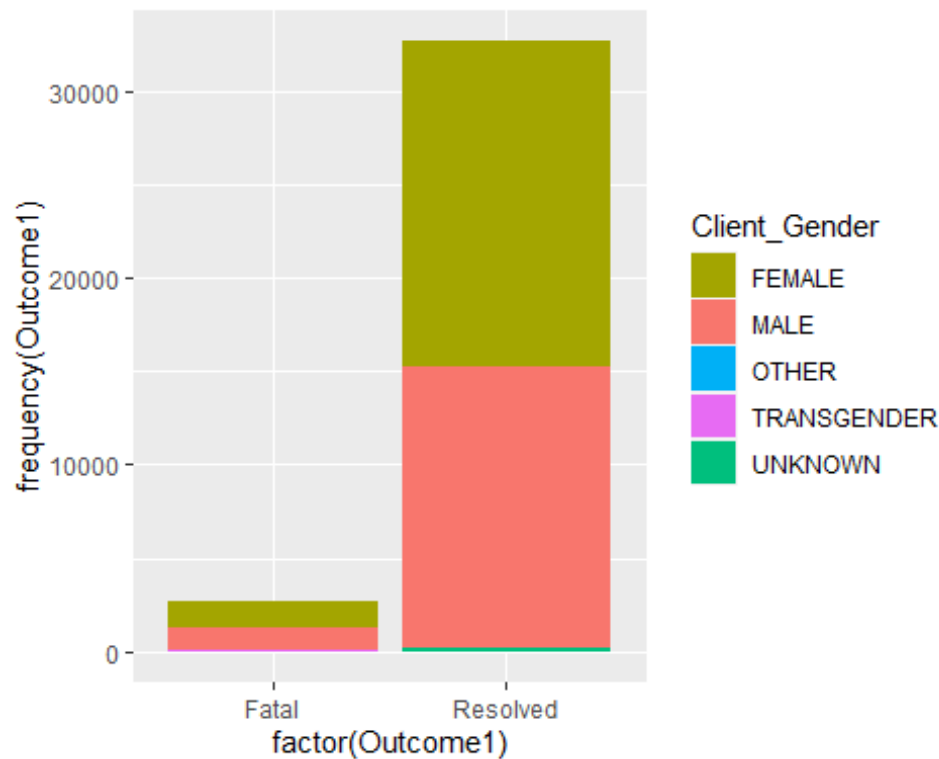
```

}

#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=Client_Gender)) + geom_bar(stat = 'identity') + scale_fill_manual(values
= mycols)

```



```

#Table of Client Genders in Outcome
Table.Gender <- table(dataclean$Outcome, dataclean$Client_Gender)
Table.Gender

##
##          FEMALE  MALE  OTHER  TRANSGENDER  UNKNOWN
##  Fatal      1442  1240     0           1      39
##  Resolved  17461 14967     9           7     219

prop.table(Table.Gender, 1) #as percentage

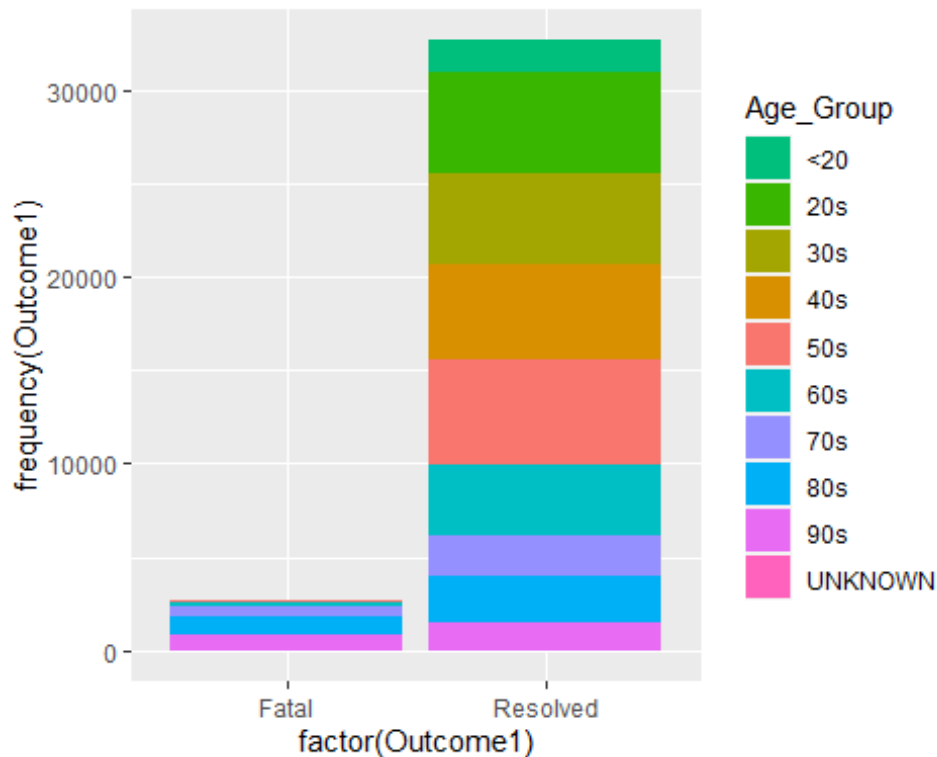
##
##          FEMALE          MALE          OTHER  TRANSGENDER
##  Fatal    0.5297575312 0.4555473916 0.0000000000 0.0003673769
##  Resolved 0.5345804121 0.4582249028 0.0002755411 0.0002143098
##

```

```
## UNKNOWN
## Fatal 0.0143277002
## Resolved 0.0067048342
```

#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)
```



#Table of Age Group in Outcome

```
Table.Age <- table(dataclean$Outcome, dataclean$Age_Group)
```

```
Table.Age
```

```
##
##      <20  20s  30s  40s  50s  60s  70s  80s  90s UNKNOWN
## Fatal      1    4    7   23   88  239  484  991  885      0
## Resolved 1723 5392 4949 5014 5633 3772 2169 2466 1537      8
```

```
prop.table(Table.Age, 1) #as percentage
```

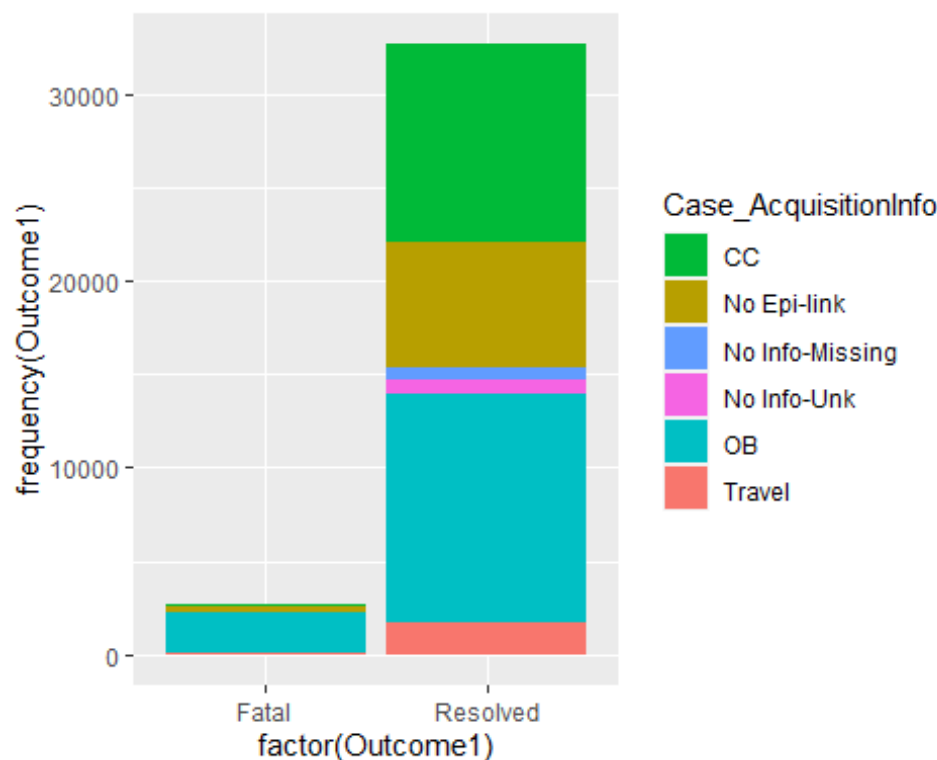
```
##
##      <20      20s      30s      40s
## Fatal 0.0003673769 0.0014695077 0.0025716385 0.0084496694
## Resolved 0.0527508190 0.1650797538 0.1515170070 0.1535070263
##
##      50s      60s      70s      80s
```



```
## Fatal      0.0323291697 0.0878030860 0.1778104335 0.3640705364
## Resolved 0.1724581331 0.1154823501 0.0664054129 0.0754982702
##
##          90s      UNKNOWN
## Fatal      0.3251285819 0.0000000000
## Resolved 0.0470563022 0.0002449255
```

#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=Case_AcquisitionInfo)) + geom_bar(stat = 'identity') +
scale_fill_manual(values = mycols)
```



#Table of Case Acquisition Info in Outcome

```
Table.CAI <- table(dataclean$Outcome, dataclean$Case_AcquisitionInfo)
Table.CAI
```

```
##
##          CC No Epi-link No Info-Missing No Info-Unk      OB Travel
## Fatal      178         252             23          54    2155     60
## Resolved 10642         6705             638         761   12213    1704
```

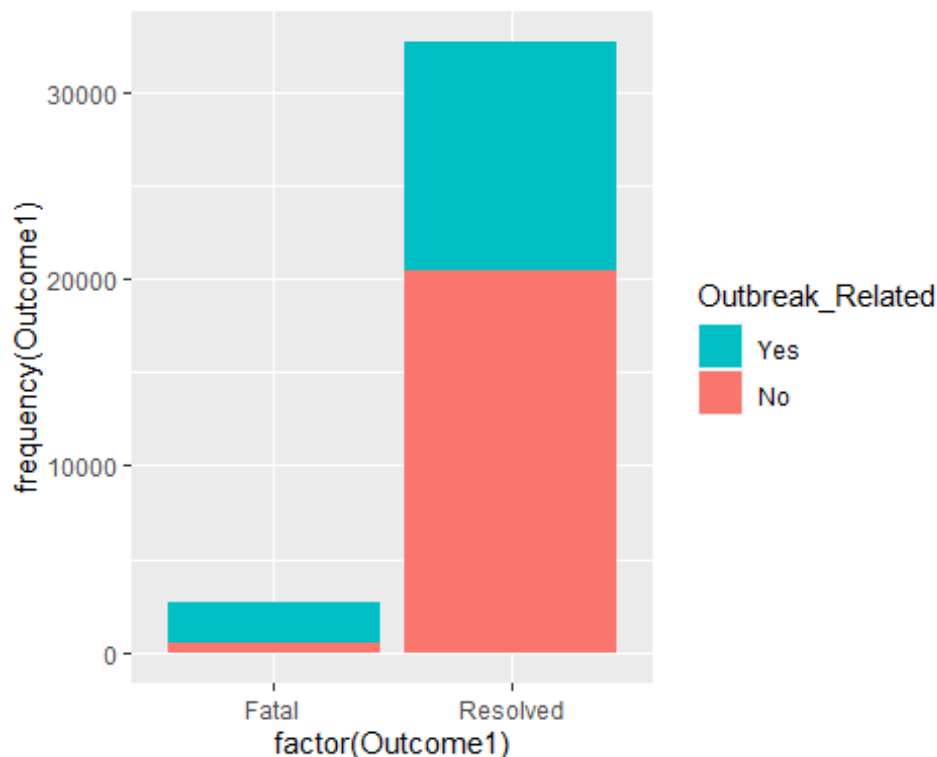
prop.table(Table.CAI, 1) #as percentage

```
##
##          CC No Epi-link No Info-Missing No Info-Unk      OB
## Fatal      0.065393093 0.092578986      0.008449669 0.019838354 0.791697281
```

```
## Resolved 0.325812081 0.205278143 0.019532805 0.023298534 0.373909316
##
## Travel
## Fatal 0.022042616
## Resolved 0.052169121
```

#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=Outbreak_Related)) + geom_bar(stat = 'identity') +
scale_fill_manual(values = mycols)
```



#Table of Outbreak Related in Outcome

```
Table.Outbreak <- table(dataclean$Outcome, dataclean$Outbreak_Related)
Table.Outbreak
```

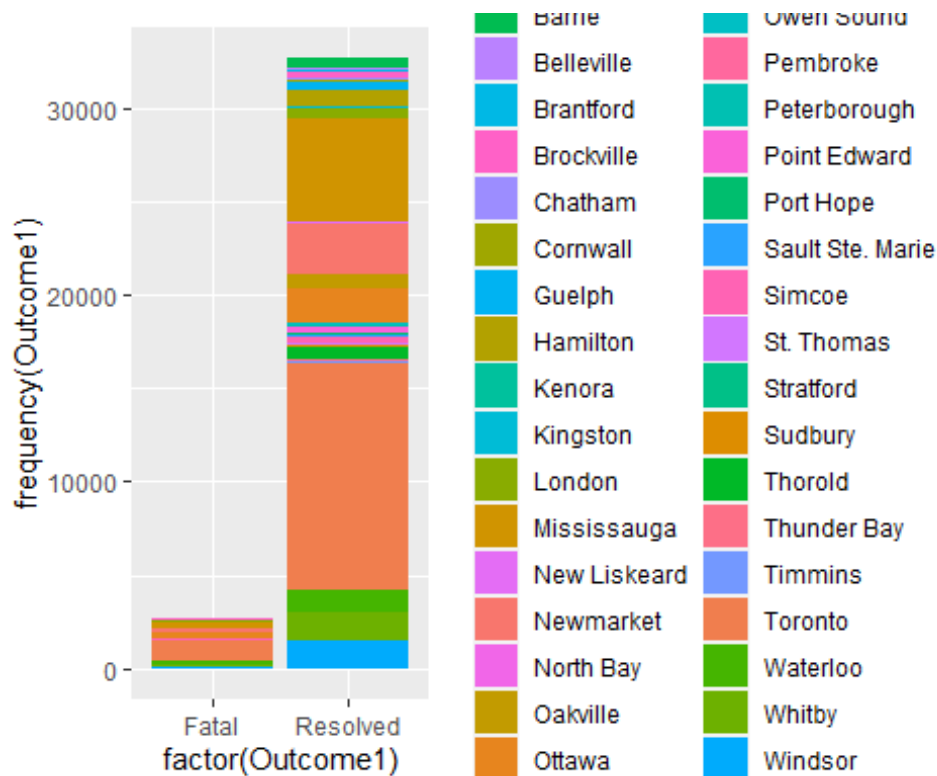
```
##
##           Yes    No
## Fatal    2155   567
## Resolved 12224 20439
```

prop.table(Table.Outbreak, 1) *#as percentage*

```
##
##           Yes      No
## Fatal    0.7916973 0.2083027
## Resolved 0.3742461 0.6257539
```

#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=Reporting_PHU_City)) + geom_bar(stat = 'identity') +
scale_fill_manual(values = mycols)
```



#Table of City in Outcome

```
Table.City <- table(dataclean$Outcome, dataclean$Reporting_PHU_City)
Table.City
```

```
##
##      Barrie Belleville Brantford Brockville Chatham Cornwall Guelph
## Fatal      34          5          5          52          1         12      36
## Resolved   572        39        127        300        157        151     446
##
##      Hamilton Kenora Kingston London Mississauga New Liskeard
## Fatal      44         0         0        57        308          0
## Resolved   792        40         96       569       5530         18
##
##      Newmarket North Bay Oakville Ottawa Owen Sound Pembroke
## Fatal      249         1         25       263          0          1
## Resolved   2748        33        737      1841        107         28
##
##      Peterborough Point Edward Port Hope Sault Ste. Marie Simcoe
## Fatal          2          25         20          0         37
## Resolved        93         260        181          25        394
```

```
##
##           St. Thomas Stratford Sudbury Thorold Thunder Bay Timmins
## Fatal           5           5           2           64           1           8
## Resolved        79          54          65          699          91          59
##
##           Toronto Waterloo Whitby Windsor
## Fatal          1092          118          181           69
## Resolved       12091         1185         1525         1531
```

```
prop.table(Table.City, 1) #as percentage
```

```
##
##           Barrie   Belleville   Brantford   Brockville
## Fatal    0.0124908156 0.0018368846 0.0018368846 0.0191036003
## Resolved 0.0175121697 0.0011940116 0.0038881915 0.0091847044
##
##           Chatham   Cornwall   Guelph   Hamilton
## Fatal    0.0003673769 0.0044085231 0.0132255694 0.0161645849
## Resolved 0.0048066620 0.0046229679 0.0136545939 0.0242476196
##
##           Kenora   Kingston   London   Mississauga
## Fatal    0.0000000000 0.0000000000 0.0209404849 0.1131520940
## Resolved 0.0012246273 0.0029391054 0.0174203227 0.1693047179
##
##           New Liskeard   Newmarket   North Bay   Oakville
## Fatal    0.0000000000 0.0914768553 0.0003673769 0.0091844232
## Resolved 0.0005510823 0.0841318924 0.0010103175 0.0225637572
##
##           Ottawa   Owen Sound   Pembroke   Peterborough
## Fatal    0.0966201323 0.0000000000 0.0003673769 0.0007347539
## Resolved 0.0563634694 0.0032758779 0.0008572391 0.0028472584
##
##           Point Edward   Port Hope   Sault Ste. Marie   Simcoe
## Fatal    0.0091844232 0.0073475386 0.0000000000 0.0135929464
## Resolved 0.0079600772 0.0055414383 0.0007653920 0.0120625785
##
##           St. Thomas   Stratford   Sudbury   Thorold
## Fatal    0.0018368846 0.0018368846 0.0007347539 0.0235121234
## Resolved 0.0024186388 0.0016532468 0.0019900193 0.0214003613
##
##           Thunder Bay   Timmins   Toronto   Waterloo
## Fatal    0.0003673769 0.0029390154 0.4011756062 0.0433504776
## Resolved 0.0027860270 0.0018063252 0.3701742032 0.0362795824
##
##           Whitby   Windsor
## Fatal    0.0664952241 0.0253490081
## Resolved 0.0466889141 0.0468726081
```

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

```

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

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

library(mapdata)

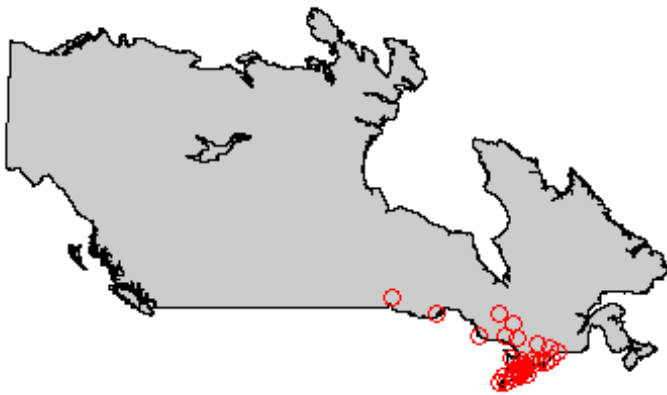
## Warning: package 'mapdata' was built under R version 3.5.3

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))
#convert points to projected lat/long
points(coord, pch=21, cex=1.2, col="red")

```



Step 2:

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,]

#check balance of training data
table(training$Outcome1)

##
##      Fatal Resolved
##      2450      29397

#very imbalanced data

#balance training data with function ROSE (Randomly Over Sampling Examples)
which creates a sample of synthetic data by enlarging the features space of
minority and majority class examples
#install.packages("ROSE")
library(ROSE)

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

## Loaded ROSE 0.0-3

training.balanced <- ROSE(Outcome1~., data = training, seed = 100)$data
table(training.balanced$Outcome1)

##
## Resolved      Fatal
##      15801      15659

#training data is balanced
```

Step 3: 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_AcquisitionInfo + Outbreak_Related + Reporting_PHU_City, family =
binomial(link = "logit"), data=training.balanced)
summary (model.LogReg)

##
## Call:
## glm(formula = Outcome1 ~ Client_Gender + Age_Group + Case_AcquisitionInfo
+

```

```

##      Outbreak_Related + Reporting_PHU_City, family = binomial(link =
"logit"),
##      data = training.balanced)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q        Max
## -2.5688   -0.3715   -0.0004    0.5915    3.2575
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      -17.35400  1028.15068  -0.017  0.98653
## Client_GenderMALE      0.71100    0.03550  20.027 < 2e-16
## Client_GenderOTHER     -10.91563   613.24134  -0.018  0.98580
## Client_GenderTRANSGENDER    1.98162    0.96008   2.064  0.03902
## Client_GenderUNKNOWN      0.36190    0.17397   2.080  0.03751
## Age_Group20s         -0.01072    0.50506  -0.021  0.98306
## Age_Group30s          0.74137    0.47648   1.556  0.11973
## Age_Group40s          1.92039    0.45870   4.187 2.83e-05
## Age_Group50s          3.13506    0.45217   6.933 4.11e-12
## Age_Group60s          4.52604    0.45098  10.036 < 2e-16
## Age_Group70s          5.79287    0.45122  12.838 < 2e-16
## Age_Group80s          6.23945    0.45123  13.828 < 2e-16
## Age_Group90s          6.64078    0.45214  14.687 < 2e-16
## Age_GroupUNKNOWN     -10.60679   727.69891  -0.015  0.98837
## Case_AcquisitionInfoNo Epi-link    0.58395    0.06280   9.299 < 2e-16
## Case_AcquisitionInfoNo Info-Missing 0.89289    0.15486   5.766 8.13e-09
## Case_AcquisitionInfoNo Info-Unk    1.36512    0.11305  12.075 < 2e-16
## Case_AcquisitionInfoOB    12.44801  1028.15057   0.012  0.99034
## Case_AcquisitionInfoTravel    0.01051    0.09401   0.112  0.91097
## Outbreak_RelatedNo    11.52260  1028.15057   0.011  0.99106
## Reporting_PHU_CityBelleville    0.26363    0.42370   0.622  0.53381
## Reporting_PHU_CityBrantford   -0.77801    0.37431  -2.079  0.03766
## Reporting_PHU_CityBrockville  -0.05779    0.18756  -0.308  0.75802
## Reporting_PHU_CityChatham    -0.45914    0.47258  -0.972  0.33127
## Reporting_PHU_CityCornwall    0.11797    0.28580   0.413  0.67978
## Reporting_PHU_CityGuelph     -0.02915    0.19426  -0.150  0.88073
## Reporting_PHU_CityHamilton   -0.37920    0.18059  -2.100  0.03574
## Reporting_PHU_CityKenora     -13.75803   307.43025  -0.045  0.96431
## Reporting_PHU_CityKingston   -13.86352   183.35900  -0.076  0.93973
## Reporting_PHU_CityLondon      0.59111    0.18402   3.212  0.00132
## Reporting_PHU_CityMississauga   0.08709    0.14407   0.604  0.54551
## Reporting_PHU_CityNew Liskeard -14.73273   355.64767  -0.041  0.96696
## Reporting_PHU_CityNewmarket    0.05666    0.14680   0.386  0.69951
## Reporting_PHU_CityNorth Bay   -0.45620    0.71758  -0.636  0.52494
## Reporting_PHU_CityOakville    -0.23570    0.19147  -1.231  0.21832
## Reporting_PHU_CityOttawa      0.42939    0.15021   2.859  0.00426
## Reporting_PHU_CityOwen Sound  -15.63187  177.53254  -0.088  0.92984
## Reporting_PHU_CityPembroke    -0.72437    0.57673  -1.256  0.20911
## Reporting_PHU_CityPeterborough -0.30463    0.41579  -0.733  0.46378
## Reporting_PHU_CityPoint Edward -0.17987    0.21077  -0.853  0.39344

```

## Reporting_PHU_CityPort Hope	-0.23734	0.22776	-1.042	0.29738
## Reporting_PHU_CitySault Ste. Marie	-14.09941	462.29383	-0.030	0.97567
## Reporting_PHU_CitySimcoe	-0.15064	0.19448	-0.775	0.43858
## Reporting_PHU_CitySt. Thomas	-0.08986	0.39124	-0.230	0.81834
## Reporting_PHU_CityStratford	-1.08957	0.38523	-2.828	0.00468
## Reporting_PHU_CitySudbury	-0.34699	0.43142	-0.804	0.42122
## Reporting_PHU_CityThorold	-0.16667	0.17336	-0.961	0.33635
## Reporting_PHU_CityThunder Bay	0.44743	0.45228	0.989	0.32252
## Reporting_PHU_CityTimmins	0.84083	0.32587	2.580	0.00987
## Reporting_PHU_CityToronto	-0.05328	0.13914	-0.383	0.70175
## Reporting_PHU_CityWaterloo	0.05896	0.15979	0.369	0.71212
## Reporting_PHU_CityWhitby	0.07348	0.15337	0.479	0.63187
## Reporting_PHU_CityWindsor	0.01918	0.16776	0.114	0.90898
##				
## (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: 43612  on 31459  degrees of freedom
## Residual deviance: 23412  on 31407  degrees of freedom
## AIC: 23518
##
## Number of Fisher Scoring iterations: 14

#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                      31459      43612
## Client_Gender           4      72.9      31455      43539 5.56e-15 ***
## Age_Group                9 19449.3      31446      24090 < 2.2e-16 ***
## Case_AcquisitionInfo     5      461.6      31441      23628 < 2.2e-16 ***
## Outbreak_Related         1         0.1      31440      23628  0.7753
## Reporting_PHU_City       33      216.1      31407      23412 < 2.2e-16 ***
## ---
## 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("Resolved", length(test.probs))
pred.log[test.probs>=0.5] <- "Fatal"

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

## Confusion Matrix and Statistics
##
##              Reference
## Prediction Fatal Resolved
##   Fatal      233      672
##   Resolved    39     2594
##
##              Accuracy : 0.799
##              95% CI : (0.7854, 0.8121)
##   No Information Rate : 0.9231
##   P-Value [Acc > NIR] : 1
##
##              Kappa : 0.3149
##
##  Mcnemar's Test P-Value : <2e-16
##
##              Sensitivity : 0.85662
##              Specificity : 0.79424
##              Pos Pred Value : 0.25746
##              Neg Pred Value : 0.98519
##              Prevalence : 0.07688
##              Detection Rate : 0.06586
##              Detection Prevalence : 0.25579
##              Balanced Accuracy : 0.82543
##
##              'Positive' Class : Fatal
##

#Recall = 0.8566
#Precision = 0.2575

#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

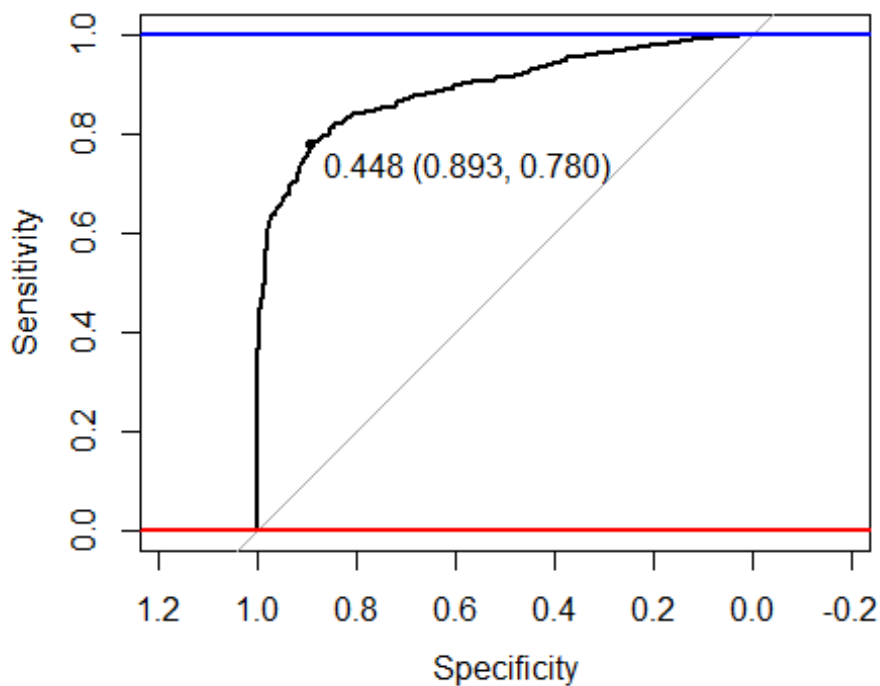
roc.curve <- roc(test$Outcome1, test.probs)

## Setting levels: control = Fatal, case = Resolved
## Setting direction: controls > cases

print(roc.curve)

##
## Call:
## roc.default(response = test$Outcome1, predictor = test.probs)
##
## Data: test.probs in 272 controls (test$Outcome1 Fatal) > 3266 cases
## (test$Outcome1 Resolved).
## Area under the curve: 0.8918

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 3: 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.balanced, training.balanced$Outcome1,
                             laplace = 1)
NBC.classifier

##
## Naive Bayes Classifier for Discrete Predictors
##
## Call:
## naiveBayes.default(x = training.balanced, y = training.balanced$Outcome1,
##      laplace = 1)
##
## A-priori probabilities:
## training.balanced$Outcome1
##   Resolved   Fatal
## 0.5022568 0.4977432
##
## Conditional probabilities:
##                                     Row_ID
## training.balanced$Outcome1      [,1]      [,2]
##               Resolved 18282.69 10873.92
##               Fatal   14836.12 10943.59
##
##                                     Accurate_Episode_Date
## training.balanced$Outcome1 2020-01-01 2020-01-10 2020-01-21
##               Resolved 1.253526e-04 6.267628e-05 1.880288e-04
##               Fatal   6.323911e-05 6.323911e-05 6.323911e-05
##                                     Accurate_Episode_Date
## training.balanced$Outcome1 2020-01-22 2020-01-24 2020-02-01
##               Resolved 6.267628e-05 6.267628e-05 6.267628e-05
##               Fatal   6.323911e-05 6.323911e-05 6.323911e-05
##                                     Accurate_Episode_Date
## training.balanced$Outcome1 2020-02-05 2020-02-07 2020-02-10
```

```

##          Resolved 1.253526e-04 6.267628e-05 1.880288e-04
##          Fatal   6.323911e-05 6.323911e-05 6.323911e-05
##          Accurate_Episode_Date
## training.balanced$Outcome1 2020-02-14 2020-02-15 2020-02-16
##          Resolved 6.267628e-05 6.267628e-05 6.267628e-05
##          Fatal   6.323911e-05 6.323911e-05 6.323911e-05
##          Accurate_Episode_Date
## training.balanced$Outcome1 2020-02-17 2020-02-19 2020-02-20
##          Resolved 6.267628e-05 1.253526e-04 1.253526e-04
##          Fatal   6.323911e-05 6.323911e-05 6.323911e-05
##          Accurate_Episode_Date
## training.balanced$Outcome1 2020-02-21 2020-02-22 2020-02-23
##          Resolved 6.267628e-05 1.880288e-04 6.267628e-05
##          Fatal   6.323911e-05 6.323911e-05 6.323911e-05
##          Accurate_Episode_Date
## training.balanced$Outcome1 2020-02-24 2020-02-25 2020-02-26
##          Resolved 1.880288e-04 1.253526e-04 6.267628e-05
##          Fatal   6.323911e-05 6.323911e-05 6.323911e-05
##          Accurate_Episode_Date
## training.balanced$Outcome1 2020-02-27 2020-02-28 2020-02-29
##          Resolved 1.253526e-04 3.133814e-04 1.253526e-04
##          Fatal   6.323911e-05 6.323911e-05 6.323911e-05
##          Accurate_Episode_Date
## training.balanced$Outcome1 2020-03-01 2020-03-02 2020-03-03
##          Resolved 3.760577e-04 5.014102e-04 5.640865e-04
##          Fatal   3.161955e-04 6.323911e-05 6.323911e-05
##          Accurate_Episode_Date
## training.balanced$Outcome1 2020-03-04 2020-03-05 2020-03-06
##          Resolved 5.014102e-04 3.760577e-04 1.065497e-03
##          Fatal   1.011826e-03 6.323911e-05 1.580978e-03
##          Accurate_Episode_Date
## training.balanced$Outcome1 2020-03-07 2020-03-08 2020-03-09
##          Resolved 4.387339e-04 1.316202e-03 2.507051e-03
##          Fatal   5.059129e-04 1.328021e-03 7.588693e-04
##          Accurate_Episode_Date
## training.balanced$Outcome1 2020-03-10 2020-03-11 2020-03-12
##          Resolved 2.883109e-03 3.321843e-03 3.635224e-03
##          Fatal   6.323911e-04 8.221084e-04 3.794346e-04
##          Accurate_Episode_Date
## training.balanced$Outcome1 2020-03-13 2020-03-14 2020-03-15
##          Resolved 3.948605e-03 4.011282e-03 6.142275e-03
##          Fatal   2.466325e-03 3.161955e-03 4.869411e-03
##          Accurate_Episode_Date
## training.balanced$Outcome1 2020-03-16 2020-03-17 2020-03-18
##          Resolved 7.959887e-03 9.401442e-03 8.586650e-03
##          Fatal   3.035477e-03 4.616455e-03 4.806172e-03
##          Accurate_Episode_Date
## training.balanced$Outcome1 2020-03-19 2020-03-20 2020-03-21
##          Resolved 7.771858e-03 9.150736e-03 7.897211e-03
##          Fatal   6.703345e-03 6.450389e-03 4.047303e-03

```

```

##                               Accurate_Episode_Date
## training.balanced$Outcome1  2020-03-22  2020-03-23  2020-03-24
##                               Resolved 5.828894e-03 9.213413e-03 6.831714e-03
##                               Fatal   7.335736e-03 8.410801e-03 5.565041e-03
##                               Accurate_Episode_Date
## training.balanced$Outcome1  2020-03-25  2020-03-26  2020-03-27
##                               Resolved 7.897211e-03 7.709182e-03 1.002820e-02
##                               Fatal   9.549105e-03 1.631569e-02 1.131980e-02
##                               Accurate_Episode_Date
## training.balanced$Outcome1  2020-03-28  2020-03-29  2020-03-30
##                               Resolved 8.962708e-03 7.207772e-03 1.040426e-02
##                               Fatal   1.113008e-02 1.675836e-02 1.720104e-02
##                               Accurate_Episode_Date
## training.balanced$Outcome1  2020-03-31  2020-04-01  2020-04-02
##                               Resolved 7.897211e-03 1.215920e-02 1.046694e-02
##                               Fatal   1.062417e-02 2.712958e-02 2.004680e-02
##                               Accurate_Episode_Date
## training.balanced$Outcome1  2020-04-03  2020-04-04  2020-04-05
##                               Resolved 1.153243e-02 1.153243e-02 9.464118e-03
##                               Fatal   2.346171e-02 1.606273e-02 1.916145e-02
##                               Accurate_Episode_Date
## training.balanced$Outcome1  2020-04-06  2020-04-07  2020-04-08
##                               Resolved 1.184582e-02 1.197117e-02 9.902852e-03
##                               Fatal   2.188073e-02 2.017328e-02 2.554860e-02
##                               Accurate_Episode_Date
## training.balanced$Outcome1  2020-04-09  2020-04-10  2020-04-11
##                               Resolved 1.278596e-02 1.566907e-02 1.491695e-02
##                               Fatal   1.947764e-02 2.510593e-02 3.073421e-02
##                               Accurate_Episode_Date
## training.balanced$Outcome1  2020-04-12  2020-04-13  2020-04-14
##                               Resolved 1.385146e-02 1.723598e-02 1.723598e-02
##                               Fatal   2.238664e-02 3.193575e-02 2.226017e-02
##                               Accurate_Episode_Date
## training.balanced$Outcome1  2020-04-15  2020-04-16  2020-04-17
##                               Resolved 1.717330e-02 1.692259e-02 1.911626e-02
##                               Fatal   3.199899e-02 2.820464e-02 3.180927e-02
##                               Accurate_Episode_Date
## training.balanced$Outcome1  2020-04-18  2020-04-19  2020-04-20
##                               Resolved 1.579442e-02 1.027891e-02 1.272328e-02
##                               Fatal   2.750901e-02 1.833934e-02 1.802315e-02
##                               Accurate_Episode_Date
## training.balanced$Outcome1  2020-04-21  2020-04-22  2020-04-23
##                               Resolved 1.234723e-02 1.240990e-02 1.184582e-02
##                               Fatal   1.663189e-02 2.226017e-02 2.156454e-02
##                               Accurate_Episode_Date
## training.balanced$Outcome1  2020-04-24  2020-04-25  2020-04-26
##                               Resolved 1.209652e-02 1.178314e-02 8.900031e-03
##                               Fatal   1.669512e-02 1.460823e-02 1.150952e-02
##                               Accurate_Episode_Date
## training.balanced$Outcome1  2020-04-27  2020-04-28  2020-04-29

```

```

##           Resolved 1.447822e-02 1.071764e-02 9.714823e-03
##           Fatal   1.277430e-02 1.113008e-02 1.802315e-02
##           Accurate_Episode_Date
## training.balanced$Outcome1 2020-04-30 2020-05-01 2020-05-02
##           Resolved 1.096835e-02 1.259793e-02 9.464118e-03
##           Fatal   7.968127e-03 7.904888e-03 8.347562e-03
##           Accurate_Episode_Date
## training.balanced$Outcome1 2020-05-03 2020-05-04 2020-05-05
##           Resolved 7.082419e-03 7.646506e-03 9.902852e-03
##           Fatal   3.161955e-03 7.082780e-03 9.802062e-03
##           Accurate_Episode_Date
## training.balanced$Outcome1 2020-05-06 2020-05-07 2020-05-08
##           Resolved 7.771858e-03 9.526794e-03 1.027891e-02
##           Fatal   1.056093e-02 8.474040e-03 1.056093e-02
##           Accurate_Episode_Date
## training.balanced$Outcome1 2020-05-09 2020-05-10 2020-05-11
##           Resolved 7.959887e-03 8.022563e-03 1.209652e-02
##           Fatal   9.359388e-03 3.984064e-03 8.284323e-03
##           Accurate_Episode_Date
## training.balanced$Outcome1 2020-05-12 2020-05-13 2020-05-14
##           Resolved 9.464118e-03 8.712003e-03 9.464118e-03
##           Fatal   7.272497e-03 6.576867e-03 5.628281e-03
##           Accurate_Episode_Date
## training.balanced$Outcome1 2020-05-15 2020-05-16 2020-05-17
##           Resolved 1.403949e-02 9.714823e-03 9.276089e-03
##           Fatal   7.841649e-03 5.248846e-03 9.485866e-03
##           Accurate_Episode_Date
## training.balanced$Outcome1 2020-05-18 2020-05-19 2020-05-20
##           Resolved 1.222187e-02 1.247258e-02 1.222187e-02
##           Fatal   4.932650e-03 6.640106e-03 4.110542e-03
##           Accurate_Episode_Date
## training.balanced$Outcome1 2020-05-21 2020-05-22 2020-05-23
##           Resolved 9.902852e-03 8.273269e-03 6.831714e-03
##           Fatal   5.817998e-03 2.592803e-03 3.098716e-03
##           Accurate_Episode_Date
## training.balanced$Outcome1 2020-05-24 2020-05-25 2020-05-26
##           Resolved 7.333124e-03 1.084300e-02 8.273269e-03
##           Fatal   2.213369e-03 2.339847e-03 3.414912e-03
##           Accurate_Episode_Date
## training.balanced$Outcome1 2020-05-27 2020-05-28 2020-05-29
##           Resolved 8.586650e-03 6.518333e-03 1.034159e-02
##           Fatal   4.173781e-03 2.276608e-03 3.351673e-03
##           Accurate_Episode_Date
## training.balanced$Outcome1 2020-05-30 2020-05-31 2020-06-01
##           Resolved 8.837355e-03 6.957067e-03 1.103102e-02
##           Fatal   1.454499e-03 1.138304e-03 1.707456e-03
##           Accurate_Episode_Date
## training.balanced$Outcome1 2020-06-02 2020-06-03 2020-06-04
##           Resolved 6.079599e-03 7.082419e-03 7.897211e-03
##           Fatal   1.201543e-03 1.391260e-03 1.138304e-03

```

```

##                               Accurate_Episode_Date
## training.balanced$Outcome1  2020-06-05  2020-06-06  2020-06-07
##                               Resolved 6.643685e-03 5.202131e-03 4.700721e-03
##                               Fatal   1.770695e-03 1.138304e-03 1.897173e-03
##                               Accurate_Episode_Date
## training.balanced$Outcome1  2020-06-08  2020-06-09  2020-06-10
##                               Resolved 6.267628e-03 4.136634e-03 6.455657e-03
##                               Fatal   5.059129e-04 1.517739e-03 1.075065e-03
##                               Accurate_Episode_Date
## training.balanced$Outcome1  2020-06-11  2020-06-12  2020-06-13
##                               Resolved 5.640865e-03 5.327484e-03 3.823253e-03
##                               Fatal   6.323911e-04 6.323911e-05 8.221084e-04
##                               Accurate_Episode_Date
## training.balanced$Outcome1  2020-06-14  2020-06-15  2020-06-16
##                               Resolved 3.572548e-03 7.270448e-03 4.826073e-03
##                               Fatal   4.426737e-04 4.426737e-04 1.517739e-03
##                               Accurate_Episode_Date
## training.balanced$Outcome1  2020-06-17  2020-06-18  2020-06-19
##                               Resolved 4.826073e-03 4.387339e-03 5.828894e-03
##                               Fatal   1.201543e-03 6.323911e-04 6.323911e-05
##                               Accurate_Episode_Date
## training.balanced$Outcome1  2020-06-20  2020-06-21  2020-06-22
##                               Resolved 5.202131e-03 3.635224e-03 5.452836e-03
##                               Fatal   5.059129e-04 1.897173e-04 7.588693e-04
##                               Accurate_Episode_Date
## training.balanced$Outcome1  2020-06-23  2020-06-24  2020-06-25
##                               Resolved 4.324663e-03 5.014102e-03 7.959887e-03
##                               Fatal   3.794346e-04 5.691520e-04 7.588693e-04
##                               Accurate_Episode_Date
## training.balanced$Outcome1  2020-06-26  2020-06-27  2020-06-28
##                               Resolved 3.823253e-03 3.823253e-03 3.760577e-04
##                               Fatal   1.075065e-03 1.897173e-04 6.323911e-05
##                               Accurate_Episode_Date
## training.balanced$Outcome1  2020-06-29  2020-06-30  2020-07-01
##                               Resolved 2.507051e-04 4.387339e-04 6.267628e-05
##                               Fatal   5.059129e-04 6.323911e-05 6.323911e-05
##                               Accurate_Episode_Date
## training.balanced$Outcome1  2020-07-02  2020-07-03  2020-07-04
##                               Resolved 3.760577e-04 1.253526e-04 1.253526e-04
##                               Fatal   1.897173e-04 6.323911e-05 6.323911e-05
##                               Accurate_Episode_Date
## training.balanced$Outcome1  2020-07-05  2020-07-06  2020-07-07
##                               Resolved 6.267628e-05 1.253526e-04 1.253526e-04
##                               Fatal   6.323911e-05 2.529564e-04 6.323911e-05
##                               Accurate_Episode_Date
## training.balanced$Outcome1  2020-07-08
##                               Resolved 1.880288e-04
##                               Fatal   6.323911e-05
##
##                               Case_Reported_Date

```



```

## training.balanced$Outcome1 2020-01-23 2020-01-24 2020-02-21
## Resolved 1.881822e-04 6.272739e-05 1.254548e-04
## Fatal 6.329114e-05 6.329114e-05 6.329114e-05
## Case_Reported_Date
## training.balanced$Outcome1 2020-02-25 2020-02-26 2020-02-27
## Resolved 6.272739e-05 6.272739e-05 6.272739e-05
## Fatal 6.329114e-05 6.329114e-05 6.329114e-05
## Case_Reported_Date
## training.balanced$Outcome1 2020-02-28 2020-02-29 2020-03-01
## Resolved 2.509095e-04 1.881822e-04 6.272739e-05
## Fatal 6.329114e-05 6.329114e-05 6.329114e-05
## Case_Reported_Date
## training.balanced$Outcome1 2020-03-03 2020-03-04 2020-03-05
## Resolved 1.254548e-04 6.272739e-05 1.881822e-04
## Fatal 6.329114e-05 6.329114e-05 6.329114e-05
## Case_Reported_Date
## training.balanced$Outcome1 2020-03-06 2020-03-07 2020-03-08
## Resolved 1.254548e-04 1.254548e-04 2.509095e-04
## Fatal 6.329114e-05 6.329114e-05 5.063291e-04
## Case_Reported_Date
## training.balanced$Outcome1 2020-03-09 2020-03-10 2020-03-11
## Resolved 1.254548e-04 2.509095e-04 3.763643e-04
## Fatal 6.329114e-05 6.329114e-05 6.329114e-05
## Case_Reported_Date
## training.balanced$Outcome1 2020-03-12 2020-03-13 2020-03-14
## Resolved 1.003638e-03 1.129093e-03 1.254548e-03
## Fatal 3.164557e-04 4.430380e-04 6.329114e-05
## Case_Reported_Date
## training.balanced$Outcome1 2020-03-15 2020-03-16 2020-03-17
## Resolved 8.781834e-04 2.070004e-03 2.258186e-03
## Fatal 6.329114e-05 5.696203e-04 3.797468e-04
## Case_Reported_Date
## training.balanced$Outcome1 2020-03-18 2020-03-19 2020-03-20
## Resolved 1.630912e-03 2.070004e-03 2.571823e-03
## Fatal 1.202532e-03 6.962025e-04 1.518987e-03
## Case_Reported_Date
## training.balanced$Outcome1 2020-03-21 2020-03-22 2020-03-23
## Resolved 2.509095e-03 2.760005e-03 3.826371e-03
## Fatal 1.898734e-03 1.772152e-03 3.164557e-03
## Case_Reported_Date
## training.balanced$Outcome1 2020-03-24 2020-03-25 2020-03-26
## Resolved 4.265462e-03 3.638188e-03 7.025467e-03
## Fatal 2.531646e-03 4.493671e-03 4.556962e-03
## Case_Reported_Date
## training.balanced$Outcome1 2020-03-27 2020-03-28 2020-03-29
## Resolved 8.970016e-03 5.582737e-03 6.335466e-03
## Fatal 7.721519e-03 6.962025e-03 8.860759e-03
## Case_Reported_Date
## training.balanced$Outcome1 2020-03-30 2020-03-31 2020-04-01
## Resolved 1.417639e-02 1.078911e-02 1.336093e-02

```

```

## Fatal 9.873418e-03 1.196203e-02 1.063291e-02
## Case_Reported_Date
## training.balanced$Outcome1 2020-04-02 2020-04-03 2020-04-04
## Resolved 1.141638e-02 1.242002e-02 9.032744e-03
## Fatal 2.082278e-02 2.727848e-02 1.537975e-02
## Case_Reported_Date
## training.balanced$Outcome1 2020-04-05 2020-04-06 2020-04-07
## Resolved 1.028729e-02 1.210639e-02 1.166729e-02
## Fatal 1.746835e-02 2.854430e-02 2.443038e-02
## Case_Reported_Date
## training.balanced$Outcome1 2020-04-08 2020-04-09 2020-04-10
## Resolved 1.066366e-02 1.380003e-02 1.066366e-02
## Fatal 2.303797e-02 1.734177e-02 1.791139e-02
## Case_Reported_Date
## training.balanced$Outcome1 2020-04-11 2020-04-12 2020-04-13
## Resolved 1.254548e-02 1.524275e-02 1.587003e-02
## Fatal 2.398734e-02 2.626582e-02 3.227848e-02
## Case_Reported_Date
## training.balanced$Outcome1 2020-04-14 2020-04-15 2020-04-16
## Resolved 1.492912e-02 1.781458e-02 1.486639e-02
## Fatal 2.848101e-02 3.379747e-02 3.151899e-02
## Case_Reported_Date
## training.balanced$Outcome1 2020-04-17 2020-04-18 2020-04-19
## Resolved 2.201731e-02 1.587003e-02 1.612094e-02
## Fatal 3.949367e-02 3.094937e-02 2.158228e-02
## Case_Reported_Date
## training.balanced$Outcome1 2020-04-20 2020-04-21 2020-04-22
## Resolved 1.725003e-02 1.373730e-02 1.449003e-02
## Fatal 2.025316e-02 2.183544e-02 2.367089e-02
## Case_Reported_Date
## training.balanced$Outcome1 2020-04-23 2020-04-24 2020-04-25
## Resolved 1.449003e-02 1.122820e-02 1.047547e-02
## Fatal 2.436709e-02 1.575949e-02 2.158228e-02
## Case_Reported_Date
## training.balanced$Outcome1 2020-04-26 2020-04-27 2020-04-28
## Resolved 1.016184e-02 1.380003e-02 1.035002e-02
## Fatal 1.373418e-02 1.645570e-02 1.607595e-02
## Case_Reported_Date
## training.balanced$Outcome1 2020-04-29 2020-04-30 2020-05-01
## Resolved 1.072638e-02 1.097729e-02 1.235730e-02
## Fatal 1.373418e-02 1.259494e-02 1.474684e-02
## Case_Reported_Date
## training.balanced$Outcome1 2020-05-02 2020-05-03 2020-05-04
## Resolved 9.283653e-03 9.910927e-03 9.534563e-03
## Fatal 8.544304e-03 9.556962e-03 8.101266e-03
## Case_Reported_Date
## training.balanced$Outcome1 2020-05-05 2020-05-06 2020-05-07
## Resolved 1.166729e-02 9.409108e-03 8.091833e-03
## Fatal 8.860759e-03 1.202532e-02 6.962025e-03
## Case_Reported_Date

```

```

## training.balanced$Outcome1  2020-05-08  2020-05-09  2020-05-10
##                               Resolved 8.970016e-03 8.781834e-03 8.029106e-03
##                               Fatal   7.974684e-03 9.430380e-03 1.075949e-02
##                               Case_Reported_Date
## training.balanced$Outcome1  2020-05-11  2020-05-12  2020-05-13
##                               Resolved 6.711830e-03 7.966378e-03 1.078911e-02
##                               Fatal   5.759494e-03 8.670886e-03 9.493671e-03
##                               Case_Reported_Date
## training.balanced$Outcome1  2020-05-14  2020-05-15  2020-05-16
##                               Resolved 8.342742e-03 9.283653e-03 1.110275e-02
##                               Fatal   1.018987e-02 1.037975e-02 6.835443e-03
##                               Case_Reported_Date
## training.balanced$Outcome1  2020-05-17  2020-05-18  2020-05-19
##                               Resolved 1.235730e-02 1.198093e-02 9.032744e-03
##                               Fatal   8.037975e-03 5.316456e-03 4.556962e-03
##                               Case_Reported_Date
## training.balanced$Outcome1  2020-05-20  2020-05-21  2020-05-22
##                               Resolved 1.041275e-02 1.292184e-02 1.210639e-02
##                               Fatal   8.797468e-03 4.620253e-03 8.291139e-03
##                               Case_Reported_Date
## training.balanced$Outcome1  2020-05-23  2020-05-24  2020-05-25
##                               Resolved 1.179275e-02 9.220926e-03 1.223184e-02
##                               Fatal   6.202532e-03 3.797468e-03 4.620253e-03
##                               Case_Reported_Date
## training.balanced$Outcome1  2020-05-26  2020-05-27  2020-05-28
##                               Resolved 7.840923e-03 9.785472e-03 1.016184e-02
##                               Fatal   2.784810e-03 1.455696e-03 3.037975e-03
##                               Case_Reported_Date
## training.balanced$Outcome1  2020-05-29  2020-05-30  2020-05-31
##                               Resolved 1.662276e-02 1.022456e-02 1.166729e-02
##                               Fatal   6.139241e-03 5.379747e-03 3.987342e-03
##                               Case_Reported_Date
## training.balanced$Outcome1  2020-06-01  2020-06-02  2020-06-03
##                               Resolved 1.066366e-02 1.041275e-02 1.072638e-02
##                               Fatal   1.329114e-03 6.329114e-05 2.721519e-03
##                               Case_Reported_Date
## training.balanced$Outcome1  2020-06-04  2020-06-05  2020-06-06
##                               Resolved 9.095471e-03 7.966378e-03 6.335466e-03
##                               Fatal   1.708861e-03 1.139241e-03 1.898734e-03
##                               Case_Reported_Date
## training.balanced$Outcome1  2020-06-07  2020-06-08  2020-06-09
##                               Resolved 7.088195e-03 5.582737e-03 6.335466e-03
##                               Fatal   2.531646e-04 1.835443e-03 2.025316e-03
##                               Case_Reported_Date
## training.balanced$Outcome1  2020-06-10  2020-06-11  2020-06-12
##                               Resolved 7.652741e-03 5.770920e-03 7.401832e-03
##                               Fatal   1.329114e-03 1.898734e-03 3.797468e-04
##                               Case_Reported_Date
## training.balanced$Outcome1  2020-06-13  2020-06-14  2020-06-15
##                               Resolved 4.955464e-03 4.390917e-03 5.269100e-03

```

```

## Fatal 6.329114e-04 6.329114e-05 5.063291e-04
## Case_Reported_Date
## training.balanced$Outcome1 2020-06-16 2020-06-17 2020-06-18
## Resolved 5.018191e-03 5.959102e-03 6.586376e-03
## Fatal 1.518987e-03 1.329114e-03 1.835443e-03
## Case_Reported_Date
## training.balanced$Outcome1 2020-06-19 2020-06-20 2020-06-21
## Resolved 4.830009e-03 4.390917e-03 5.143646e-03
## Fatal 7.594937e-04 1.708861e-03 3.164557e-04
## Case_Reported_Date
## training.balanced$Outcome1 2020-06-22 2020-06-23 2020-06-24
## Resolved 5.708192e-03 4.579099e-03 4.641827e-03
## Fatal 6.329114e-04 5.063291e-04 3.797468e-04
## Case_Reported_Date
## training.balanced$Outcome1 2020-06-25 2020-06-26 2020-06-27
## Resolved 4.704554e-03 5.833647e-03 5.959102e-03
## Fatal 9.493671e-04 5.696203e-04 6.329114e-05
## Case_Reported_Date
## training.balanced$Outcome1 2020-06-28 2020-06-29 2020-06-30
## Resolved 5.833647e-03 3.889098e-03 4.077280e-03
## Fatal 2.151899e-03 6.329114e-05 5.696203e-04
## Case_Reported_Date
## training.balanced$Outcome1 2020-07-01 2020-07-02 2020-07-03
## Resolved 2.132731e-03 1.881822e-03 1.066366e-03
## Fatal 6.329114e-05 6.329114e-04 6.329114e-05
## Case_Reported_Date
## training.balanced$Outcome1 2020-07-04 2020-07-05 2020-07-06
## Resolved 1.254548e-03 6.900013e-04 5.645465e-04
## Fatal 1.898734e-04 6.329114e-05 6.329114e-05
## Case_Reported_Date
## training.balanced$Outcome1 2020-07-07 2020-07-08 2020-07-09
## Resolved 1.881822e-04 3.136369e-04 3.763643e-04
## Fatal 6.329114e-05 6.329114e-05 6.329114e-05
## Case_Reported_Date
## training.balanced$Outcome1 2020-07-10 2020-07-11 2020-07-12
## Resolved 6.272739e-05 1.254548e-04 1.254548e-04
## Fatal 2.531646e-04 6.329114e-05 6.329114e-05
##
## Test_Reported_Date
## training.balanced$Outcome1 2020-01-27 2020-02-03 2020-02-24
## Resolved 1.881940e-04 6.273132e-05 1.254626e-04
## Fatal 6.329515e-05 6.329515e-05 6.329515e-05
## Test_Reported_Date
## training.balanced$Outcome1 2020-02-25 2020-02-27 2020-02-28
## Resolved 6.273132e-05 6.273132e-05 1.254626e-04
## Fatal 6.329515e-05 6.329515e-05 6.329515e-05
## Test_Reported_Date
## training.balanced$Outcome1 2020-02-29 2020-03-01 2020-03-02
## Resolved 6.273132e-05 3.136566e-04 6.273132e-05
## Fatal 6.329515e-05 6.329515e-05 6.329515e-05

```

```

##                               Test_Reported_Date
## training.balanced$Outcome1  2020-03-03  2020-03-04  2020-03-05
##                               Resolved 6.273132e-05 1.254626e-04 1.881940e-04
##                               Fatal   6.329515e-05 6.329515e-05 6.329515e-05
##                               Test_Reported_Date
## training.balanced$Outcome1  2020-03-07  2020-03-08  2020-03-09
##                               Resolved 1.254626e-04 1.254626e-04 2.509253e-04
##                               Fatal   6.329515e-05 6.329515e-05 6.329515e-05
##                               Test_Reported_Date
## training.balanced$Outcome1  2020-03-10  2020-03-11  2020-03-12
##                               Resolved 6.273132e-05 3.763879e-04 6.273132e-04
##                               Fatal   6.329515e-05 6.329515e-05 6.329515e-05
##                               Test_Reported_Date
## training.balanced$Outcome1  2020-03-13  2020-03-14  2020-03-15
##                               Resolved 1.003701e-03 1.317358e-03 6.900445e-04
##                               Fatal   6.962466e-04 6.329515e-05 6.329515e-05
##                               Test_Reported_Date
## training.balanced$Outcome1  2020-03-16  2020-03-17  2020-03-18
##                               Resolved 6.900445e-04 1.881940e-03 1.756477e-03
##                               Fatal   5.696563e-04 6.962466e-04 1.455788e-03
##                               Test_Reported_Date
## training.balanced$Outcome1  2020-03-19  2020-03-20  2020-03-21
##                               Resolved 1.003701e-03 2.634716e-03 1.191895e-03
##                               Fatal   6.329515e-05 6.329515e-04 1.835559e-03
##                               Test_Reported_Date
## training.balanced$Outcome1  2020-03-22  2020-03-23  2020-03-24
##                               Resolved 2.070134e-03 4.328461e-03 4.579386e-03
##                               Fatal   9.494272e-04 2.595101e-03 1.265903e-03
##                               Test_Reported_Date
## training.balanced$Outcome1  2020-03-25  2020-03-26  2020-03-27
##                               Resolved 3.889342e-03 6.649520e-03 9.911549e-03
##                               Fatal   5.696563e-03 3.544528e-03 6.709285e-03
##                               Test_Reported_Date
## training.balanced$Outcome1  2020-03-28  2020-03-29  2020-03-30
##                               Resolved 5.520356e-03 7.025908e-03 1.436547e-02
##                               Fatal   6.645990e-03 6.835876e-03 1.234255e-02
##                               Test_Reported_Date
## training.balanced$Outcome1  2020-03-31  2020-04-01  2020-04-02
##                               Resolved 1.041340e-02 1.191895e-02 1.235807e-02
##                               Fatal   1.170960e-02 1.088676e-02 2.392556e-02
##                               Test_Reported_Date
## training.balanced$Outcome1  2020-04-03  2020-04-04  2020-04-05
##                               Resolved 1.235807e-02 1.016247e-02 9.284236e-03
##                               Fatal   2.753339e-02 2.379897e-02 1.652003e-02
##                               Test_Reported_Date
## training.balanced$Outcome1  2020-04-06  2020-04-07  2020-04-08
##                               Resolved 1.104071e-02 1.323631e-02 1.160529e-02
##                               Fatal   2.411545e-02 2.614089e-02 1.715298e-02
##                               Test_Reported_Date
## training.balanced$Outcome1  2020-04-09  2020-04-10  2020-04-11

```

```

##          Resolved 1.367543e-02 1.009974e-02 1.141710e-02
##          Fatal   2.038104e-02 2.006456e-02 2.728021e-02
##          Test_Reported_Date
## training.balanced$Outcome1 2020-04-12 2020-04-13 2020-04-14
##          Resolved 1.543191e-02 1.681199e-02 1.706292e-02
##          Fatal   1.873536e-02 2.936895e-02 2.854611e-02
##          Test_Reported_Date
## training.balanced$Outcome1 2020-04-15 2020-04-16 2020-04-17
##          Resolved 1.687473e-02 1.511825e-02 1.863120e-02
##          Fatal   3.221723e-02 3.481233e-02 3.563517e-02
##          Test_Reported_Date
## training.balanced$Outcome1 2020-04-18 2020-04-19 2020-04-20
##          Resolved 1.587102e-02 1.530644e-02 1.674926e-02
##          Fatal   3.329325e-02 2.088740e-02 2.164694e-02
##          Test_Reported_Date
## training.balanced$Outcome1 2020-04-21 2020-04-22 2020-04-23
##          Resolved 1.430274e-02 1.392635e-02 1.518098e-02
##          Fatal   2.069751e-02 2.436863e-02 2.436863e-02
##          Test_Reported_Date
## training.balanced$Outcome1 2020-04-24 2020-04-25 2020-04-26
##          Resolved 1.116618e-02 1.317358e-02 1.009974e-02
##          Fatal   1.848218e-02 2.721691e-02 1.132983e-02
##          Test_Reported_Date
## training.balanced$Outcome1 2020-04-27 2020-04-28 2020-04-29
##          Resolved 1.336177e-02 9.284236e-03 9.911549e-03
##          Fatal   1.727957e-02 1.221596e-02 1.373505e-02
##          Test_Reported_Date
## training.balanced$Outcome1 2020-04-30 2020-05-01 2020-05-02
##          Resolved 1.122891e-02 1.304811e-02 9.284236e-03
##          Fatal   1.386164e-02 1.196278e-02 8.734730e-03
##          Test_Reported_Date
## training.balanced$Outcome1 2020-05-03 2020-05-04 2020-05-05
##          Resolved 1.047613e-02 1.009974e-02 1.141710e-02
##          Fatal   1.057029e-02 8.101779e-03 1.107665e-02
##          Test_Reported_Date
## training.balanced$Outcome1 2020-05-06 2020-05-07 2020-05-08
##          Resolved 8.845116e-03 7.214102e-03 9.346967e-03
##          Fatal   1.012722e-02 1.126654e-02 5.316792e-03
##          Test_Reported_Date
## training.balanced$Outcome1 2020-05-09 2020-05-10 2020-05-11
##          Resolved 8.155072e-03 8.719654e-03 6.774983e-03
##          Fatal   1.012722e-02 6.076334e-03 6.645990e-03
##          Test_Reported_Date
## training.balanced$Outcome1 2020-05-12 2020-05-13 2020-05-14
##          Resolved 8.845116e-03 1.091525e-02 7.966878e-03
##          Fatal   8.544845e-03 7.848598e-03 1.088676e-02
##          Test_Reported_Date
## training.balanced$Outcome1 2020-05-15 2020-05-16 2020-05-17
##          Resolved 9.409698e-03 1.053886e-02 1.285992e-02
##          Fatal   9.747452e-03 7.152351e-03 7.722008e-03

```

```

##                               Test_Reported_Date
## training.balanced$Outcome1  2020-05-18  2020-05-19  2020-05-20
##                               Resolved 1.229534e-02 8.845116e-03 1.041340e-02
##                               Fatal   5.316792e-03 6.962466e-03 7.785303e-03
##                               Test_Reported_Date
## training.balanced$Outcome1  2020-05-21  2020-05-22  2020-05-23
##                               Resolved 1.279719e-02 1.204441e-02 1.298538e-02
##                               Fatal   4.683841e-03 7.342237e-03 7.405532e-03
##                               Test_Reported_Date
## training.balanced$Outcome1  2020-05-24  2020-05-25  2020-05-26
##                               Resolved 9.284236e-03 1.580829e-02 6.837714e-03
##                               Fatal   4.873726e-03 6.266219e-03 1.012722e-03
##                               Test_Reported_Date
## training.balanced$Outcome1  2020-05-27  2020-05-28  2020-05-29
##                               Resolved 9.409698e-03 1.072706e-02 1.568283e-02
##                               Fatal   1.455788e-03 3.038167e-03 6.013039e-03
##                               Test_Reported_Date
## training.balanced$Outcome1  2020-05-30  2020-05-31  2020-06-01
##                               Resolved 1.191895e-02 1.016247e-02 9.221504e-03
##                               Fatal   5.823153e-03 3.164757e-03 2.784986e-03
##                               Test_Reported_Date
## training.balanced$Outcome1  2020-06-02  2020-06-03  2020-06-04
##                               Resolved 1.129164e-02 1.173076e-02 8.468728e-03
##                               Fatal   5.063612e-04 2.721691e-03 1.392493e-03
##                               Test_Reported_Date
## training.balanced$Outcome1  2020-06-05  2020-06-06  2020-06-07
##                               Resolved 8.343266e-03 6.335863e-03 6.963177e-03
##                               Fatal   1.139313e-03 1.898854e-03 6.329515e-05
##                               Test_Reported_Date
## training.balanced$Outcome1  2020-06-08  2020-06-09  2020-06-10
##                               Resolved 5.583088e-03 6.524057e-03 7.966878e-03
##                               Fatal   2.595101e-03 2.025445e-03 2.025445e-03
##                               Test_Reported_Date
## training.balanced$Outcome1  2020-06-11  2020-06-12  2020-06-13
##                               Resolved 5.708550e-03 6.963177e-03 5.143968e-03
##                               Fatal   1.202608e-03 3.797709e-04 6.329515e-04
##                               Test_Reported_Date
## training.balanced$Outcome1  2020-06-14  2020-06-15  2020-06-16
##                               Resolved 4.453924e-03 5.332162e-03 5.143968e-03
##                               Fatal   6.329515e-05 5.063612e-04 1.898854e-03
##                               Test_Reported_Date
## training.balanced$Outcome1  2020-06-17  2020-06-18  2020-06-19
##                               Resolved 5.959476e-03 6.900445e-03 4.704849e-03
##                               Fatal   1.329198e-03 1.455788e-03 7.595417e-04
##                               Test_Reported_Date
## training.balanced$Outcome1  2020-06-20  2020-06-21  2020-06-22
##                               Resolved 4.140267e-03 5.206700e-03 5.708550e-03
##                               Fatal   1.708969e-03 3.164757e-04 6.329515e-04
##                               Test_Reported_Date
## training.balanced$Outcome1  2020-06-23  2020-06-24  2020-06-25

```

```

##           Resolved 5.018506e-03 4.767580e-03 4.077536e-03
##           Fatal   5.063612e-04 6.962466e-04 6.329515e-04
##           Test_Reported_Date
## training.balanced$Outcome1 2020-06-26 2020-06-27 2020-06-28
##           Resolved 5.896744e-03 5.896744e-03 5.457625e-03
##           Fatal   5.696563e-04 6.329515e-05 2.152035e-03
##           Test_Reported_Date
## training.balanced$Outcome1 2020-06-29 2020-06-30 2020-07-01
##           Resolved 5.018506e-03 3.387491e-03 2.070134e-03
##           Fatal   6.329515e-05 5.696563e-04 6.329515e-05
##           Test_Reported_Date
## training.balanced$Outcome1 2020-07-02 2020-07-03 2020-07-04
##           Resolved 1.944671e-03 1.380089e-03 1.003701e-03
##           Fatal   6.329515e-04 6.329515e-05 1.898854e-04
##           Test_Reported_Date
## training.balanced$Outcome1 2020-07-05 2020-07-06 2020-07-07
##           Resolved 8.155072e-04 6.273132e-04 1.254626e-04
##           Fatal   6.329515e-05 6.329515e-05 6.329515e-05
##           Test_Reported_Date
## training.balanced$Outcome1 2020-07-08 2020-07-09 2020-07-10
##           Resolved 3.136566e-04 3.763879e-04 6.273132e-05
##           Fatal   6.329515e-05 6.329515e-05 2.531806e-04
##           Test_Reported_Date
## training.balanced$Outcome1 2020-07-11 2020-07-12
##           Resolved 1.254626e-04 1.254626e-04
##           Fatal   6.329515e-05 6.329515e-05
##
##           Specimen_Date
## training.balanced$Outcome1 2020-01-23 2020-01-24 2020-01-25
##           Resolved 1.881586e-04 6.271952e-05 6.271952e-05
##           Fatal   6.328313e-05 6.328313e-05 6.328313e-05
##           Specimen_Date
## training.balanced$Outcome1 2020-02-20 2020-02-22 2020-02-23
##           Resolved 6.271952e-05 2.508781e-04 1.254390e-04
##           Fatal   6.328313e-05 6.328313e-05 6.328313e-05
##           Specimen_Date
## training.balanced$Outcome1 2020-02-25 2020-02-26 2020-02-27
##           Resolved 6.271952e-05 6.271952e-05 1.254390e-04
##           Fatal   6.328313e-05 6.328313e-05 6.328313e-05
##           Specimen_Date
## training.balanced$Outcome1 2020-02-28 2020-02-29 2020-03-01
##           Resolved 6.271952e-05 3.135976e-04 6.271952e-05
##           Fatal   6.328313e-05 6.328313e-05 6.328313e-05
##           Specimen_Date
## training.balanced$Outcome1 2020-03-02 2020-03-03 2020-03-04
##           Resolved 1.254390e-04 1.881586e-04 1.881586e-04
##           Fatal   6.328313e-05 6.328313e-05 6.328313e-05
##           Specimen_Date
## training.balanced$Outcome1 2020-03-05 2020-03-06 2020-03-07
##           Resolved 1.254390e-04 1.881586e-04 1.254390e-04

```



```

## Fatal 6.328313e-05 6.328313e-05 6.328313e-05
## Specimen_Date
## training.balanced$Outcome1 2020-03-08 2020-03-09 2020-03-10
## Resolved 1.881586e-04 1.881586e-04 4.390366e-04
## Fatal 6.328313e-05 6.328313e-05 6.328313e-05
## Specimen_Date
## training.balanced$Outcome1 2020-03-11 2020-03-12 2020-03-13
## Resolved 4.390366e-04 2.446061e-03 1.693427e-03
## Fatal 3.164156e-04 6.328313e-05 1.139096e-03
## Specimen_Date
## training.balanced$Outcome1 2020-03-14 2020-03-15 2020-03-16
## Resolved 1.317110e-03 1.630707e-03 2.446061e-03
## Fatal 6.328313e-05 5.695482e-04 1.075813e-03
## Specimen_Date
## training.balanced$Outcome1 2020-03-17 2020-03-18 2020-03-19
## Resolved 3.575013e-03 4.014049e-03 4.578525e-03
## Fatal 7.593975e-04 1.455512e-03 2.088343e-03
## Specimen_Date
## training.balanced$Outcome1 2020-03-20 2020-03-21 2020-03-22
## Resolved 6.899147e-03 4.703964e-03 3.825891e-03
## Fatal 2.594608e-03 1.708644e-03 3.986837e-03
## Specimen_Date
## training.balanced$Outcome1 2020-03-23 2020-03-24 2020-03-25
## Resolved 6.146513e-03 6.397391e-03 7.777220e-03
## Fatal 5.189217e-03 3.796988e-03 4.050120e-03
## Specimen_Date
## training.balanced$Outcome1 2020-03-26 2020-03-27 2020-03-28
## Resolved 7.777220e-03 8.341696e-03 8.404415e-03
## Fatal 9.555752e-03 7.404126e-03 9.872168e-03
## Specimen_Date
## training.balanced$Outcome1 2020-03-29 2020-03-30 2020-03-31
## Resolved 6.460110e-03 1.292022e-02 9.596086e-03
## Fatal 1.037843e-02 1.322617e-02 1.841539e-02
## Specimen_Date
## training.balanced$Outcome1 2020-04-01 2020-04-02 2020-04-03
## Resolved 1.047416e-02 1.191671e-02 1.003512e-02
## Fatal 2.613593e-02 2.487027e-02 2.170611e-02
## Specimen_Date
## training.balanced$Outcome1 2020-04-04 2020-04-05 2020-04-06
## Resolved 9.094330e-03 8.467135e-03 1.342198e-02
## Fatal 1.980762e-02 1.822554e-02 2.740159e-02
## Specimen_Date
## training.balanced$Outcome1 2020-04-07 2020-04-08 2020-04-09
## Resolved 1.342198e-02 1.241846e-02 1.229303e-02
## Fatal 2.006075e-02 2.442729e-02 2.556638e-02
## Specimen_Date
## training.balanced$Outcome1 2020-04-10 2020-04-11 2020-04-12
## Resolved 1.542900e-02 1.574260e-02 1.348470e-02
## Fatal 2.461714e-02 3.024934e-02 1.980762e-02
## Specimen_Date

```

```

## training.balanced$Outcome1  2020-04-13  2020-04-14  2020-04-15
##                               Resolved 2.107376e-02 1.875314e-02 2.157551e-02
##                               Fatal   3.834958e-02 2.911024e-02 3.974180e-02
##                               Specimen_Date
## training.balanced$Outcome1  2020-04-16  2020-04-17  2020-04-18
##                               Resolved 1.605620e-02 2.044656e-02 1.335926e-02
##                               Fatal   3.556512e-02 2.461714e-02 2.524997e-02
##                               Specimen_Date
## training.balanced$Outcome1  2020-04-19  2020-04-20  2020-04-21
##                               Resolved 8.906172e-03 1.668339e-02 1.379829e-02
##                               Fatal   7.467409e-03 2.404759e-02 1.936464e-02
##                               Specimen_Date
## training.balanced$Outcome1  2020-04-22  2020-04-23  2020-04-24
##                               Resolved 1.248118e-02 1.241846e-02 1.473909e-02
##                               Fatal   2.461714e-02 1.993419e-02 1.866852e-02
##                               Specimen_Date
## training.balanced$Outcome1  2020-04-25  2020-04-26  2020-04-27
##                               Resolved 1.160311e-02 7.463623e-03 1.292022e-02
##                               Fatal   1.474497e-02 8.986204e-03 2.113656e-02
##                               Specimen_Date
## training.balanced$Outcome1  2020-04-28  2020-04-29  2020-04-30
##                               Resolved 1.160311e-02 1.304566e-02 1.335926e-02
##                               Fatal   1.544108e-02 1.594735e-02 6.771295e-03
##                               Specimen_Date
## training.balanced$Outcome1  2020-05-01  2020-05-02  2020-05-03
##                               Resolved 9.909684e-03 7.777220e-03 7.463623e-03
##                               Fatal   1.025187e-02 6.581445e-03 4.936084e-03
##                               Specimen_Date
## training.balanced$Outcome1  2020-05-04  2020-05-05  2020-05-06
##                               Resolved 1.072504e-02 1.160311e-02 9.031611e-03
##                               Fatal   1.082142e-02 1.012530e-02 9.555752e-03
##                               Specimen_Date
## training.balanced$Outcome1  2020-05-07  2020-05-08  2020-05-09
##                               Resolved 1.041144e-02 1.022328e-02 7.714501e-03
##                               Fatal   1.139096e-02 8.986204e-03 1.151753e-02
##                               Specimen_Date
## training.balanced$Outcome1  2020-05-10  2020-05-11  2020-05-12
##                               Resolved 5.456598e-03 1.304566e-02 9.157050e-03
##                               Fatal   5.125933e-03 1.278319e-02 1.164410e-02
##                               Specimen_Date
## training.balanced$Outcome1  2020-05-13  2020-05-14  2020-05-15
##                               Resolved 9.407928e-03 8.404415e-03 1.053688e-02
##                               Fatal   5.948614e-03 6.265030e-03 3.860271e-03
##                               Specimen_Date
## training.balanced$Outcome1  2020-05-16  2020-05-17  2020-05-18
##                               Resolved 1.053688e-02 9.157050e-03 7.150025e-03
##                               Fatal   5.062650e-03 5.948614e-03 5.885331e-03
##                               Specimen_Date
## training.balanced$Outcome1  2020-05-19  2020-05-20  2020-05-21
##                               Resolved 1.292022e-02 1.436277e-02 1.235575e-02

```

```

##          Fatal    5.822048e-03 8.669789e-03 5.125933e-03
##          Specimen_Date
## training.balanced$Outcome1 2020-05-22 2020-05-23 2020-05-24
##          Resolved 1.022328e-02 6.585549e-03 1.003512e-02
##          Fatal    4.366536e-03 5.505632e-03 3.417289e-03
##          Specimen_Date
## training.balanced$Outcome1 2020-05-25 2020-05-26 2020-05-27
##          Resolved 9.972403e-03 1.285750e-02 9.282489e-03
##          Fatal    2.531325e-03 2.911024e-03 4.366536e-03
##          Specimen_Date
## training.balanced$Outcome1 2020-05-28 2020-05-29 2020-05-30
##          Resolved 1.034872e-02 1.304566e-02 7.839940e-03
##          Fatal    3.796988e-03 3.354006e-03 2.974307e-03
##          Specimen_Date
## training.balanced$Outcome1 2020-05-31 2020-06-01 2020-06-02
##          Resolved 6.460110e-03 1.210487e-02 8.592574e-03
##          Fatal    8.859638e-04 1.582078e-03 1.202379e-03
##          Specimen_Date
## training.balanced$Outcome1 2020-06-03 2020-06-04 2020-06-05
##          Resolved 6.710988e-03 9.596086e-03 7.902659e-03
##          Fatal    1.771928e-03 1.455512e-03 1.202379e-03
##          Specimen_Date
## training.balanced$Outcome1 2020-06-06 2020-06-07 2020-06-08
##          Resolved 5.456598e-03 5.080281e-03 6.397391e-03
##          Fatal    7.593975e-04 1.898494e-03 8.226807e-04
##          Specimen_Date
## training.balanced$Outcome1 2020-06-09 2020-06-10 2020-06-11
##          Resolved 5.331159e-03 7.526342e-03 6.083793e-03
##          Fatal    1.139096e-03 1.265663e-03 3.796988e-04
##          Specimen_Date
## training.balanced$Outcome1 2020-06-12 2020-06-13 2020-06-14
##          Resolved 5.519318e-03 4.390366e-03 3.575013e-03
##          Fatal    8.226807e-04 1.708644e-03 4.429819e-04
##          Specimen_Date
## training.balanced$Outcome1 2020-06-15 2020-06-16 2020-06-17
##          Resolved 6.899147e-03 4.578525e-03 6.648269e-03
##          Fatal    4.429819e-04 1.898494e-03 3.164156e-04
##          Specimen_Date
## training.balanced$Outcome1 2020-06-18 2020-06-19 2020-06-20
##          Resolved 4.327647e-03 5.268440e-03 5.519318e-03
##          Fatal    6.328313e-04 6.328313e-04 5.062650e-04
##          Specimen_Date
## training.balanced$Outcome1 2020-06-21 2020-06-22 2020-06-23
##          Resolved 3.825891e-03 5.268440e-03 4.641244e-03
##          Fatal    6.328313e-04 1.075813e-03 6.328313e-04
##          Specimen_Date
## training.balanced$Outcome1 2020-06-24 2020-06-25 2020-06-26
##          Resolved 5.644757e-03 9.157050e-03 4.703964e-03
##          Fatal    5.695482e-04 1.075813e-03 6.328313e-05
##          Specimen_Date

```

```

## training.balanced$Outcome1 2020-06-27 2020-06-28 2020-06-29
##                               Resolved 3.700452e-03 2.007025e-03 2.383342e-03
##                               Fatal    7.593975e-04 5.695482e-04 5.062650e-04
##                               Specimen_Date
## training.balanced$Outcome1 2020-06-30 2020-07-01 2020-07-02
##                               Resolved 2.195183e-03 5.017561e-04 9.407928e-04
##                               Fatal    1.898494e-04 6.328313e-05 1.898494e-04
##                               Specimen_Date
## training.balanced$Outcome1 2020-07-03 2020-07-04 2020-07-05
##                               Resolved 3.763171e-04 8.153537e-04 1.881586e-04
##                               Fatal    6.328313e-05 6.328313e-05 6.328313e-05
##                               Specimen_Date
## training.balanced$Outcome1 2020-07-06 2020-07-07 2020-07-08
##                               Resolved 5.017561e-04 1.881586e-04 1.881586e-04
##                               Fatal    2.531325e-04 6.328313e-05 6.328313e-05
##                               Specimen_Date
## training.balanced$Outcome1 2020-07-09 2020-07-11
##                               Resolved 6.271952e-05 1.254390e-04
##                               Fatal    6.328313e-05 6.328313e-05
##
##                               Age_Group
## training.balanced$Outcome1 <20          20s          30s
##                               Resolved 5.274809e-02 1.606477e-01 1.491367e-01
##                               Fatal    3.829217e-04 1.276406e-03 2.680452e-03
##                               Age_Group
## training.balanced$Outcome1 40s          50s          60s
##                               Resolved 1.510973e-01 1.773449e-01 1.168174e-01
##                               Fatal    8.168996e-03 3.184632e-02 8.992278e-02
##                               Age_Group
## training.balanced$Outcome1 70s          80s          90s
##                               Resolved 6.533426e-02 7.804693e-02 4.851053e-02
##                               Fatal    1.841853e-01 3.573298e-01 3.241432e-01
##                               Age_Group
## training.balanced$Outcome1 UNKNOWN
##                               Resolved 3.162355e-04
##                               Fatal    6.382028e-05
##
##                               Client_Gender
## training.balanced$Outcome1 FEMALE        MALE        OTHER
##                               Resolved 5.346071e-01 4.593825e-01 3.796027e-04
##                               Fatal    5.211313e-01 4.645046e-01 6.384065e-05
##                               Client_Gender
## training.balanced$Outcome1 TRANSGENDER    UNKNOWN
##                               Resolved 1.898013e-04 5.440972e-03
##                               Fatal    6.384065e-04 1.366190e-02
##
##                               Case_AcquisitionInfo
## training.balanced$Outcome1 CC No Epi-link No Info-Missing
##                               Resolved 0.326437654 0.207629531 0.017840197
##                               Fatal    0.065943185 0.092818385 0.008426428

```

```

##                               Case_AcquisitionInfo
## training.balanced$Outcome1 No Info-Unk          OB          Travel
##                               Resolved 0.022711457 0.373378883 0.052002277
##                               Fatal   0.023428024 0.785636770 0.023747207
##
##                               Outcome1
## training.balanced$Outcome1      Resolved      Fatal
##                               Resolved 9.999367e-01 6.327912e-05
##                               Fatal   6.385288e-05 9.999361e-01
##
##                               Outbreak_Related
## training.balanced$Outcome1      Yes          No
##                               Resolved 0.3735999 0.6264001
##                               Fatal   0.7858374 0.2141626
##
##                               Reporting_PHU_City
## training.balanced$Outcome1      Barrie      Belleville      Brantford
##                               Resolved 1.837701e-02 1.263025e-03 3.978529e-03
##                               Fatal   1.204359e-02 1.784235e-03 1.401899e-03
##                               Reporting_PHU_City
## training.balanced$Outcome1      Brockville      Chatham      Cornwall
##                               Resolved 8.778023e-03 4.799495e-03 4.167982e-03
##                               Fatal   1.867074e-02 5.097814e-04 4.141974e-03
##                               Reporting_PHU_City
## training.balanced$Outcome1      Guelph      Hamilton      Kenora
##                               Resolved 1.439848e-02 2.330281e-02 1.136722e-03
##                               Fatal   1.306315e-02 1.605812e-02 6.372268e-05
##                               Reporting_PHU_City
## training.balanced$Outcome1      Kingston      London      Mississauga
##                               Resolved 2.841806e-03 1.597727e-02 1.739185e-01
##                               Fatal   6.372268e-05 2.223921e-02 1.152743e-01
##                               Reporting_PHU_City
## training.balanced$Outcome1      New Liskeard      Newmarket      North Bay
##                               Resolved 7.578150e-04 8.891696e-02 8.841175e-04
##                               Fatal   6.372268e-05 9.182438e-02 3.186134e-04
##                               Reporting_PHU_City
## training.balanced$Outcome1      Oakville      Ottawa      Owen Sound
##                               Resolved 2.216609e-02 5.386801e-02 2.652352e-03
##                               Fatal   1.076913e-02 9.583891e-02 6.372268e-05
##                               Reporting_PHU_City
## training.balanced$Outcome1      Pembroke      Peterborough      Point Edward
##                               Resolved 6.315125e-04 2.778655e-03 7.957057e-03
##                               Fatal   7.009495e-04 1.019563e-03 8.857452e-03
##                               Reporting_PHU_City
## training.balanced$Outcome1      Port Hope      Sault Ste. Marie      Simcoe
##                               Resolved 6.441427e-03      5.052100e-04 1.332491e-02
##                               Fatal   7.391831e-03      6.372268e-05 1.357293e-02
##                               Reporting_PHU_City
## training.balanced$Outcome1      St. Thomas      Stratford      Sudbury
##                               Resolved 2.336596e-03 2.083991e-03 2.210294e-03

```

```

##          Fatal    1.274454e-03 1.210731e-03 1.019563e-03
##          Reporting_PHU_City
## training.balanced$Outcome1    Thorold  Thunder Bay    Timmins
##          Resolved 2.109252e-02 3.410167e-03 2.462899e-03
##          Fatal    2.408717e-02 5.735041e-04 3.377302e-03
##          Reporting_PHU_City
## training.balanced$Outcome1    Toronto    Waterloo    Whitby
##          Resolved 3.688664e-01 3.542785e-02 4.414272e-02
##          Fatal    3.978207e-01 4.435098e-02 6.346779e-02
##          Reporting_PHU_City
## training.balanced$Outcome1    Windsor
##          Resolved 4.414272e-02
##          Fatal    2.701842e-02
##
##          Reporting_PHU_Latitude
## training.balanced$Outcome1    [,1]    [,2]
##          Resolved 43.74321 0.7433653
##          Fatal    43.82632 0.7272546
##
##          Reporting_PHU_Longitude
## training.balanced$Outcome1    [,1]    [,2]
##          Resolved -79.54768 1.588588
##          Fatal    -79.22124 1.592329

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

##
##
##      Cell Contents
## |-----|
## |                      N |
## |          N / Col Total |
## |-----|
##
##
## Total Observations in Table:  3538
##
##
##      predicted | actual
##      predicted |      Fatal |      Resolved | Row Total |
## -----|-----|-----|-----|
##      Resolved |          2 |          3266 |          3268 |
##              |         0.007 |          1.000 |              |
## -----|-----|-----|-----|
##      Fatal    |         270 |           0    |          270 |
##              |         0.993 |          0.000 |              |

```

```
## -----|-----|-----|-----|
## Column Total |      272 |      3266 |      3538 |
##              |      0.077 |      0.923 |              |
## -----|-----|-----|-----|
##
##
```

```
#recall = 0.9926
```

```
#precision = 1.000
```

```
pred.NBC <- ifelse(NBC.predict=="Fatal", 1, 0)
```

```
actual.NBC <- ifelse(test$Outcome1=="Fatal", 1, 0)
```

```
#ROC Curve
```

```
#install.packages("pROC")
```

```
library(pROC)
```

```
roc.curve2 <- roc(pred.NBC, actual.NBC)
```

```
## Setting levels: control = 0, case = 1
```

```
## Setting direction: controls < cases
```

```
print(roc.curve2)
```

```
##
```

```
## Call:
```

```
## roc.default(response = pred.NBC, predictor = actual.NBC)
```

```
##
```

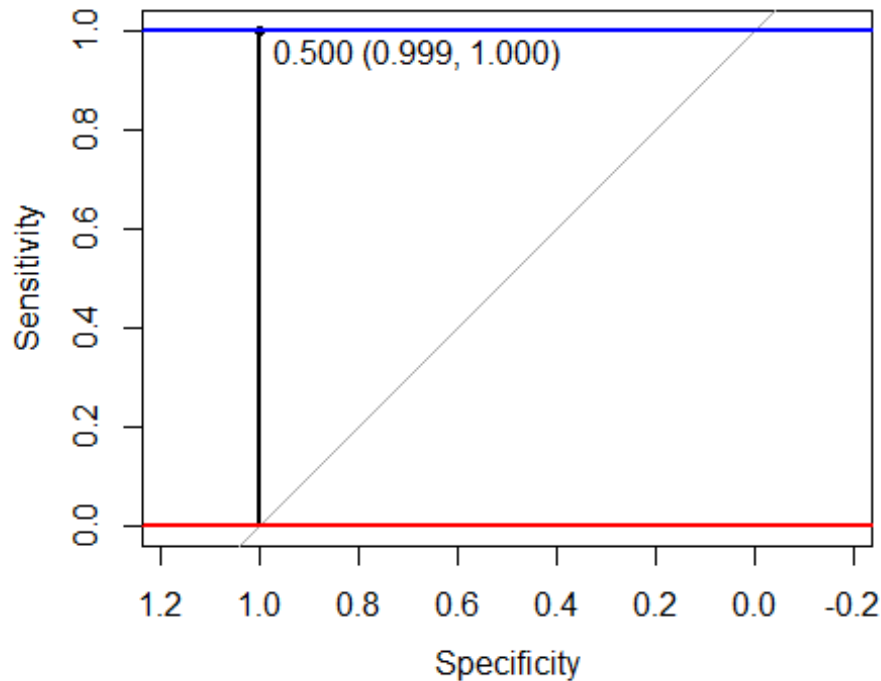
```
## Data: actual.NBC in 3268 controls (pred.NBC 0) < 270 cases (pred.NBC 1).
```

```
## Area under the curve: 0.9997
```

```
plot(roc.curve2, ylim=c(0,1), print.thres=TRUE)
```

```
abline(h=1,col='blue',lwd=2)
```

```
abline(h=0,col='red',lwd=2)
```



Step 3: 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 +
Case_AcquisitionInfo + Outbreak_Related + Reporting_PHU_City,
data=training.balanced, importance = TRUE)
model.randomForest

##
## Call:
## randomForest(formula = Outcome1 ~ Client_Gender + Age_Group +
```



```

Case_AcquisitionInfo + Outbreak_Related + Reporting_PHU_City,      data =
training.balanced, importance = TRUE)
##           Type of random forest: classification
##           Number of trees: 500
## No. of variables tried at each split: 2
##
##           OOB estimate of  error rate: 14.55%
## Confusion matrix:
##           Resolved Fatal class.error
## Resolved      12402  3399  0.21511297
## Fatal         1177 14482  0.07516444

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

##           5          16          37          43          53          62
## Resolved Resolved Resolved      Fatal Resolved Resolved
## Levels: Resolved Fatal

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

##
## predict.RF Fatal Resolved
##   Resolved      34      2578
##   Fatal       238      688

#recall = 0.8750
#precision = 0.2575

pred.RF <- ifelse(predict.RF=="Fatal", 1, 0)
actual.RF <- ifelse(test$Outcome1=="Fatal", 1, 0)

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

roc.curve3<- roc(pred.RF, actual.RF)

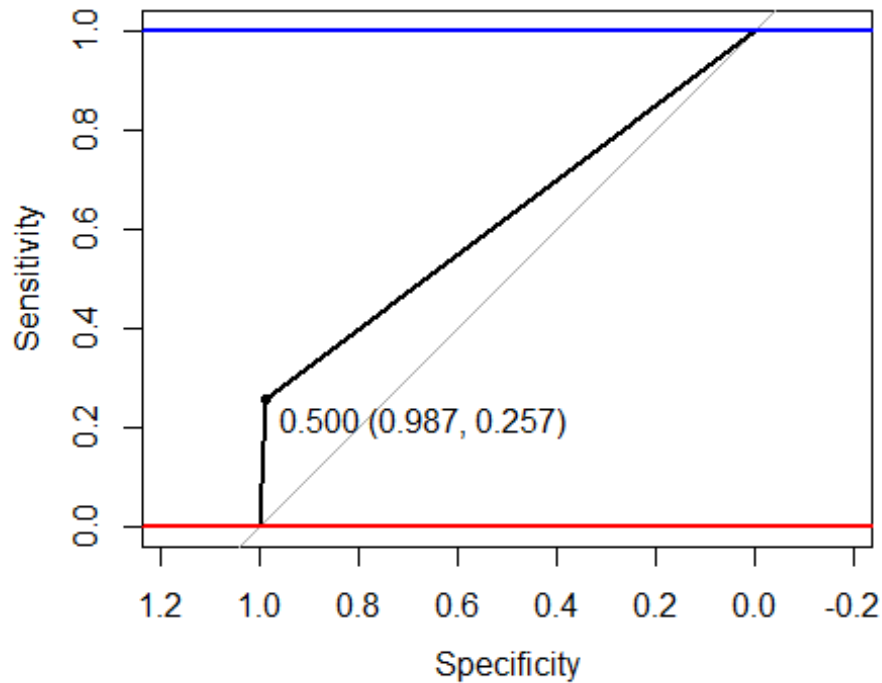
## Setting levels: control = 0, case = 1
## Setting direction: controls < cases

print(roc.curve3)

##
## Call:
## roc.default(response = pred.RF, predictor = actual.RF)
##
## Data: actual.RF in 2612 controls (pred.RF 0) < 926 cases (pred.RF 1).
## Area under the curve: 0.622

```

```
plot(roc.curve3, ylim=c(0,1), print.thres=TRUE)
abline(h=1,col='blue',lwd=2)
abline(h=0,col='red',lwd=2)
```



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

```
##           Resolved      Fatal
## Client_Gender    33.43182  33.43182
## Age_Group        142.81759 142.81759
## Case_AcquisitionInfo 13.80655 13.80655
## Outbreak_Related   9.91870  9.91870
## Reporting_PHU_City  48.36966 48.36966
```

varImpPlot(model.randomForest) *#plot*

model.randomForest

