Plane Crash Analysis: There exist such a thing as a safest seat?

Inferential Statistics 2025/2026

Barbato Alberto, Naldoni Martina

2026-01-12

Contents

1 Introduction

The Aviation industry is one of the safest in the world. The systems that it has in place to learn lessons from accidents and make sure they never happen again is well-established and highly regarded. This has made travelling by plane the safest way to travel. https://flyfright.com/plane-crash-statistics/#:~:text=Based%20on%20statistics%20from%202015,unharmed%2C%20injured%2C%20or%20killed

But perhaps due to these high safety standards, when an aircraft accident happens it makes headlines all over the world. Anxious passengers fear it's going to happen to them as well and they can't help but ask themselves if there is something they could do to have a safer flight.

One of the most frequent question that gets asked is : is there a part of the plane that is "safer" than other parts? Can the seating location make a difference in an accident? There are many articles where experts in the field give their opinions, and most answer that "yes, there are areas that give a higher chance of survival in case of an aircraft accident". But there are also studies that suggest that there is no safest seats on an airplane.

So, who is right? We are going to try to answer this question using statistics.

2 Perspective

We are studying if the seating in an airplane has an effect on the survivability of an aircraft accident. To study this, we need to look at all aircraft accidents, then rule out the accidents where every passenger survived, and every accident in which every passenger died. This is an extremely narrow data set of accidents. Of these accidents, we gathered the seatings arrangements and survivor seating maps for 47 crashes.

If 47 crashes look like a lot, we should consider the greater perspective of air travel safety in general.

https://flyfright.com/plane-crash-statistics/#tve-jump-18c020d9166 reported these figures studying the US General aviation data between 2015 and 2020 :

- $\frac{1}{260256}$: chance of boarding any flight and it being an accident flight
- $\frac{1}{6,864,250}$: chance of being on a plane involved in an accident that results in at least 1 fatality (possible case study of this study)
- $\frac{1}{816,545,929}$ chance of you specifically, dying in a plane crash

```r
data <- read.csv("Aerei_Final.csv")

str(data)
```

```
## 'data.frame':    47 obs. of  21 variables:
##  $ NumVolo       : int  6 28 92 120 123 129 140 148 191 204 ...
##  $ X1.terzo.lievi: int  17 36 0 0 0 4 0 0 0 3 ...
##  $ X1.terzo.gravi: int  2 0 11 0 0 0 0 0 0 0 ...
##  $ X1.terzo.morti: int  15 0 22 0 136 14 18 16 55 33 ...
##  $ X2.terzo.lievi: int  1 30 4 5 0 5 7 1 0 25 ...
##  $ X2.terzo.gravi: int  15 0 30 0 0 0 0 0 8 0 ...
##  $ X2.terzo.morti: int  64 16 13 8 214 60 139 34 51 1 ...
##  $ X3.terzo.lievi: int  26 10 0 5 36 24 0 8 10 29 ...
##  $ X3.terzo.gravi: int  17 0 27 0 0 0 0 0 7 0 ...
##  $ X3.terzo.morti: int  0 36 11 21 109 43 91 30 16 0 ...
##  $ X1.meta.lievi : int  17 56 32 3 0 7 7 1 0 7 ...
##  $ X1.meta.gravi : int  3 0 0 0 0 0 0 0 1 0 ...
##  $ X1.meta.morti : int  41 8 34 6 226 23 95 36 79 35 ...
##  $ X2.meta.lievi : int  26 20 39 7 4 26 0 7 10 46 ...
##  $ X2.meta.gravi : int  31 0 0 0 0 0 0 0 14 0 ...
##  $ X2.meta.morti : int  34 44 13 23 225 90 145 46 48 0 ...
##  $ fonte         : chr  "W" "W" "W" "W" ...
##  $ PhaseOfFlight : chr  "Takeoff" "Takeoff" "Landing" "Landing" ...
##  $ Time          : chr  "Night" "Day" "Night" "Day" ...
##  $ Place         : chr  "Airport " "Outside" "Outside" "Airport " ...
##  $ HasFire       : chr  "Fire" "Fire" "Fire" "Fire" ...
```

```r
summary(data)
```

```
##     NumVolo     X1.terzo.lievi   X1.terzo.gravi   X1.terzo.morti
##  Min.   :   6   Min.   :  0.00   Min.   : 0.000   Min.   :  0.00
##  1st Qu.: 227   1st Qu.:  0.00   1st Qu.: 0.000   1st Qu.:  1.50
##  Median : 812   Median :  4.00   Median : 0.000   Median : 11.00
##  Mean   :1591   Mean   : 12.87   Mean   : 2.553   Mean   : 16.15
##  3rd Qu.:1603   3rd Qu.: 17.00   3rd Qu.: 4.000   3rd Qu.: 19.00
##  Max.   :9642   Max.   :141.00   Max.   :16.000   Max.   :136.00
##  X2.terzo.lievi   X2.terzo.gravi   X2.terzo.morti   X3.terzo.lievi
##  Min.   :  0.00   Min.   : 0.000   Min.   :  0.00   Min.   :  0.0
##  1st Qu.:  1.00   1st Qu.: 0.000   1st Qu.:  2.00   1st Qu.:  3.0
##  Median :  7.00   Median : 2.000   Median : 12.00   Median : 10.0
##  Mean   : 22.11   Mean   : 4.638   Mean   : 28.62   Mean   : 20.4
```

```
##  3rd Qu.: 26.50   3rd Qu.: 8.000   3rd Qu.: 31.00   3rd Qu.: 27.0
##  Max.   :174.00   Max.   :30.000   Max.   :214.00   Max.   :142.0
##  X3.terzo.gravi   X3.terzo.morti   X1.meta.lievi    X1.meta.gravi
##  Min.   : 0.000   Min.   :  0.00   Min.   :  0.00   Min.   : 0.00
##  1st Qu.: 0.000   1st Qu.:  1.50   1st Qu.:  1.50   1st Qu.: 0.00
##  Median : 0.000   Median :  8.00   Median : 11.00   Median : 0.00
##  Mean   : 3.872   Mean   : 19.83   Mean   : 29.23   Mean   : 4.34
##  3rd Qu.: 4.500   3rd Qu.: 24.50   3rd Qu.: 40.50   3rd Qu.: 7.00
##  Max.   :27.000   Max.   :113.00   Max.   :221.00   Max.   :34.00
##  X1.meta.morti    X2.meta.lievi    X2.meta.gravi    X2.meta.morti
##  Min.   :  0.00   Min.   :  0.00   Min.   : 0.000   Min.   :  0.00
##  1st Qu.:  5.00   1st Qu.:  6.50   1st Qu.: 0.000   1st Qu.:  3.50
##  Median : 20.00   Median : 14.00   Median : 0.000   Median : 15.00
##  Mean   : 28.57   Mean   : 28.98   Mean   : 5.234   Mean   : 36.53
##  3rd Qu.: 35.00   3rd Qu.: 37.50   3rd Qu.: 7.500   3rd Qu.: 45.00
##  Max.   :226.00   Max.   :184.00   Max.   :31.000   Max.   :225.00
##     fonte           PhaseOfFlight          Time              Place
##  Length:47          Length:47          Length:47          Length:47
##  Class :character   Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character
##
##
##
##    HasFire
##  Length:47
##  Class :character
##  Mode  :character
##
##
##
```

```
#> data <- read.csv("Aerei_Final.csv")
#>
#> str(data)
#'data.frame':   47 obs. of  21 variables:
# $ NumVolo         : int  6 28 92 120 123 129 140 148 191 204 ...
# $ X1.terzo.lievi  : int  17 36 0 0 0 4 0 0 0 3 ...
# $ X1.terzo.gravi  : int  2 0 11 0 0 0 0 0 0 0 ...
# $ X1.terzo.morti  : int  15 0 22 0 136 14 18 16 55 33 ...
# $ X2.terzo.lievi  : int  1 30 4 5 0 5 7 1 0 25 ...
# $ X2.terzo.gravi  : int  15 0 30 0 0 0 0 0 8 0 ...
# $ X2.terzo.morti  : int  64 16 13 8 214 60 139 34 51 1 ...
# $ X3.terzo.lievi  : int  26 10 0 5 36 24 0 8 10 29 ...
# $ X3.terzo.gravi  : int  17 0 27 0 0 0 0 0 7 0 ...
# $ X2.terzo.morti.1: int  0 36 11 21 109 43 91 30 16 0 ...
# $ X1.meta.lievi   : int  17 56 32 3 0 7 7 1 0 7 ...
# $ X1.meta.gravi   : int  3 0 0 0 0 0 0 0 1 0 ...
# $ X1.meta.morti   : int  41 8 34 6 226 23 95 36 79 35 ...
# $ X2.meta.lievi   : int  26 20 39 7 4 26 0 7 10 46 ...
# $ X2.meta.gravi   : int  31 0 0 0 0 0 0 0 14 0 ...
# $ X2.meta.morti   : int  34 44 13 23 225 90 145 46 48 0 ...
# $ fonte           : chr  "W" "W" "W" "W" ...
# $ PhaseOfFlight   : chr  "Takeoff" "Takeoff" "Landing" "Landing" ...
# $ Time            : chr  "Night" "Day" "Night" "Day" ...
```

```
# $ Place           : chr  "Airport " "Outside" "Outside" "Airport " ...
# $ HasFire         : chr  "Fire" "Fire" "Fire" "Fire" ...
#> summary(data)
#    NumVolo     X1.terzo.lievi   X1.terzo.gravi   X1.terzo.morti
# Min.   :   6   Min.   :  0.00   Min.   : 0.000   Min.   :  0.00
# 1st Qu.: 227   1st Qu.:  0.00   1st Qu.: 0.000   1st Qu.:  1.50
# Median : 812   Median :  4.00   Median : 0.000   Median : 11.00
# Mean   :1591   Mean   : 12.87   Mean   : 2.553   Mean   : 16.15
# 3rd Qu.:1603   3rd Qu.: 17.00   3rd Qu.: 4.000   3rd Qu.: 19.00
# Max.   :9642   Max.   :141.00   Max.   :16.000   Max.   :136.00
# X2.terzo.lievi   X2.terzo.gravi   X2.terzo.morti   X3.terzo.lievi
# Min.   :  0.00   Min.   : 0.000   Min.   :  0.00   Min.   :  0.0
# 1st Qu.:  1.00   1st Qu.: 0.000   1st Qu.:  2.00   1st Qu.:  3.0
# Median :  7.00   Median : 2.000   Median : 12.00   Median : 10.0
# Mean   : 22.11   Mean   : 4.638   Mean   : 28.62   Mean   : 20.4
# 3rd Qu.: 26.50   3rd Qu.: 8.000   3rd Qu.: 31.00   3rd Qu.: 27.0
# Max.   :174.00   Max.   :30.000   Max.   :214.00   Max.   :142.0
# X3.terzo.gravi   X2.terzo.morti.1 X1.meta.lievi    X1.meta.gravi
# Min.   : 0.000   Min.   :  0.00   Min.   :  0.00   Min.   : 0.00
# 1st Qu.: 0.000   1st Qu.:  1.50   1st Qu.:  1.50   1st Qu.: 0.00
# Median : 0.000   Median :  8.00   Median : 11.00   Median : 0.00
# Mean   : 3.872   Mean   : 19.83   Mean   : 29.23   Mean   : 4.34
# 3rd Qu.: 4.500   3rd Qu.: 24.50   3rd Qu.: 40.50   3rd Qu.: 7.00
# Max.   :27.000   Max.   :113.00   Max.   :221.00   Max.   :34.00
# X1.meta.morti    X2.meta.lievi    X2.meta.gravi    X2.meta.morti
# Min.   :  0.00   Min.   :  0.00   Min.   : 0.000   Min.   :  0.00
# 1st Qu.:  5.00   1st Qu.:  6.50   1st Qu.: 0.000   1st Qu.:  3.50
# Median : 20.00   Median : 14.00   Median : 0.000   Median : 15.00
# Mean   : 28.57   Mean   : 28.98   Mean   : 5.234   Mean   : 36.53
# 3rd Qu.: 35.00   3rd Qu.: 37.50   3rd Qu.: 7.500   3rd Qu.: 45.00
# Max.   :226.00   Max.   :184.00   Max.   :31.000   Max.   :225.00
#    fonte            PhaseOfFlight          Time               Place
# Length:47          Length:47          Length:47          Length:47
# Class :character   Class :character   Class :character   Class :character
# Mode  :character   Mode  :character   Mode  :character   Mode  :character
```

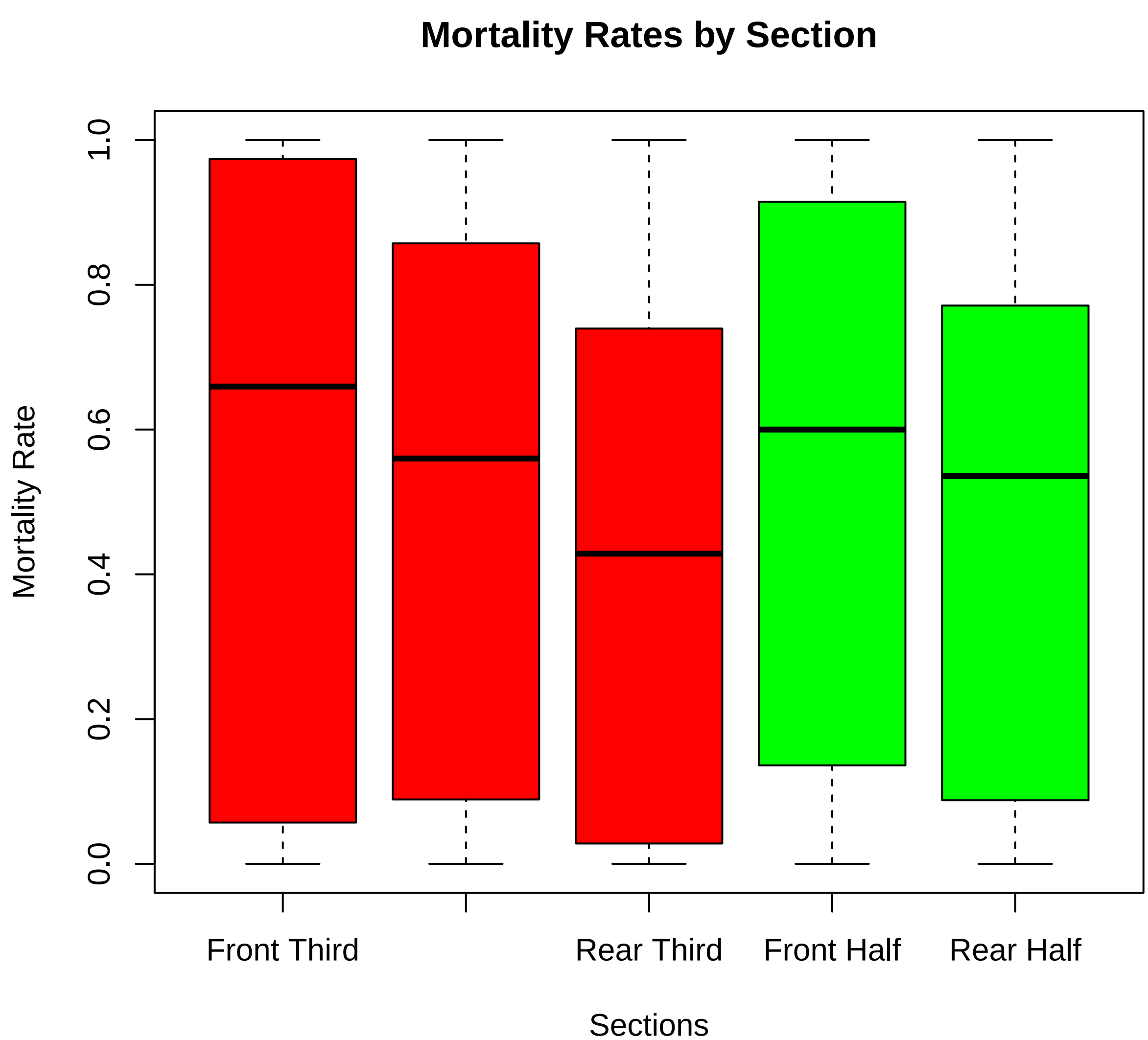
Some random text. . . .

```
# add a colum of # of seat for each airplain section

data$X1.third.total <- data$X1.terzo.lievi + data$X1.terzo.gravi + data$X1.terzo.morti
data$X2.third.total <- data$X2.terzo.lievi + data$X2.terzo.gravi + data$X2.terzo.morti
data$X3.third.total <- data$X3.terzo.lievi + data$X3.terzo.gravi + data$X3.terzo.morti
data$X1.half.total <- data$X1.meta.lievi + data$X1.meta.gravi + data$X1.meta.morti
data$X2.half.total <- data$X2.meta.lievi + data$X2.meta.gravi + data$X2.meta.morti

# now make a colum of mortality rate for each section

data$X1.third.mortality.rate <- data$X1.terzo.morti / data$X1.third.total
data$X2.third.mortality.rate <- data$X2.terzo.morti / data$X2.third.total
data$X3.third.mortality.rate <- data$X3.terzo.morti / data$X3.third.total
data$X1.half.mortality.rate <- data$X1.meta.morti / data$X1.half.total
data$X2.half.mortality.rate <- data$X2.meta.morti / data$X2.half.total
```

```r
head(data)
```

```
##   NumVolo X1.terzo.lievi X1.terzo.gravi X1.terzo.morti X2.terzo.lievi
## 1       6             17              2             15              1
## 2      28             36              0              0             30
## 3      92              0             11             22              4
## 4     120              0              0              0              5
## 5     123              0              0            136              0
## 6     129              4              0             14              5
##   X2.terzo.gravi X2.terzo.morti X3.terzo.lievi X3.terzo.gravi X3.terzo.morti
## 1             15             64             26             17              0
## 2              0             16             10              0             36
## 3             30             13              0             27             11
## 4              0              8              5              0             21
## 5              0            214             36              0            109
## 6              0             60             24              0             43
##   X1.meta.lievi X1.meta.gravi X1.meta.morti X2.meta.lievi X2.meta.gravi
## 1            17             3            41            26            31
## 2            56             0             8            20             0
## 3            32             0            34            39             0
## 4             3             0             6             7             0
## 5             0             0           226             4             0
## 6             7             0            23            26             0
##   X2.meta.morti fonte PhaseOfFlight  Time    Place HasFire X1.third.total
## 1            34     W       Takeoff Night Airport     Fire             34
## 2            44     W       Takeoff   Day  Outside    Fire             36
## 3            13     W       Landing Night  Outside    Fire             33
## 4            23     W       Landing   Day Airport     Fire              0
## 5           225     W       Takeoff Night  Outside    Fire            136
## 6            90     W       Landing   Day  Outside    Fire             18
##   X2.third.total X3.third.total X1.half.total X2.half.total
## 1             80             43            61            91
## 2             46             46            64            64
## 3             47             38            66            52
## 4             13             26             9            30
## 5            214            145           226           229
## 6             65             67            30           116
##   X1.third.mortality.rate X2.third.mortality.rate X3.third.mortality.rate
## 1               0.4411765               0.8000000               0.0000000
## 2               0.0000000               0.3478261               0.7826087
## 3               0.6666667               0.2765957               0.2894737
## 4                     NaN               0.6153846               0.8076923
## 5               1.0000000               1.0000000               0.7517241
## 6               0.7777778               0.9230769               0.6417910
##   X1.half.mortality.rate X2.half.mortality.rate
## 1              0.6721311              0.3736264
## 2              0.1250000              0.6875000
## 3              0.5151515              0.2500000
## 4              0.6666667              0.7666667
## 5              1.0000000              0.9825328
## 6              0.7666667              0.7758621
```

```r
## Plot mortality rates as scatter plots for each section
## they must be separated on the x axis by group

boxplot(data$X1.third.mortality.rate,
  data$X2.third.mortality.rate,
  data$X3.third.mortality.rate,
  data$X1.half.mortality.rate,
  data$X2.half.mortality.rate,
  names = c("Front Third", "Middle Third", "Rear Third", "Front Half", "Rear Half"),
  main = "Mortality Rates by Section",
  ylab = "Mortality Rate",
  xlab = "Sections",
  col = c("red", "red", "red", "green", "green"))
```




```r
# Perform ANOVA to test if there are significant differences in mortality rates between sections (only
mortality_data <- data.frame(
```

```r
  Section = rep(c("Front Third", "Middle Third", "Rear Third"), each = nrow(data)),
  MortalityRate = c(data$X1.third.mortality.rate, data$X2.third.mortality.rate, data$X3.third.mortality
)
mortality_data$Section <- as.factor(mortality_data$Section)

anova_result <- aov(MortalityRate ~ Section, data = mortality_data)
summary(anova_result)
```

```
##              Df Sum Sq Mean Sq F value Pr(>F)
## Section       2  0.364  0.1818   1.251  0.289
## Residuals   137 19.906  0.1453
## 1 observation deleted due to missingness
```

```r
library(multcomp)
```

```
## Loading required package: mvtnorm

## Loading required package: survival

## Loading required package: TH.data

## Loading required package: MASS

##
## Attaching package: 'TH.data'

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

```r
test_result <- glht(anova_result, linfct = mcp(Section = "Tukey"))
summary(test_result)
```

```
##
##   Simultaneous Tests for General Linear Hypotheses
##
## Multiple Comparisons of Means: Tukey Contrasts
##
##
## Fit: aov(formula = MortalityRate ~ Section, data = mortality_data)
##
## Linear Hypotheses:
##                                  Estimate Std. Error t value Pr(>|t|)
## Middle Third - Front Third == 0 -0.04032    0.07906  -0.510    0.867
## Rear Third - Front Third == 0   -0.12252    0.07906  -1.550    0.271
## Rear Third - Middle Third == 0  -0.08220    0.07863  -1.045    0.550
## (Adjusted p values reported -- single-step method)
```

```r
#half sections
mortality_data_half <- data.frame(
  Section = rep(c("Front Half", "Rear Half"), each = nrow(data)),
  MortalityRate = c(data$X1.half.mortality.rate, data$X2.half.mortality.rate)
)
anova_result_half <- aov(MortalityRate ~ Section, data = mortality_data_half)
summary(anova_result_half)
```

```
##             Df Sum Sq Mean Sq F value Pr(>F)
## Section      1  0.056 0.05591   0.409  0.524
## Residuals   92 12.581 0.13675
```

```r
#try using non-parametric test if ANOVA assumptions are not met
kruskal_result <- kruskal.test(MortalityRate ~ Section, data = mortality_data)
kruskal_result_half <- kruskal.test(MortalityRate ~ Section, data = mortality_data_half)
kruskal_result
```

```
##
##  Kruskal-Wallis rank sum test
##
## data:  MortalityRate by Section
## Kruskal-Wallis chi-squared = 2.9755, df = 2, p-value = 0.2259
```

```r
kruskal_result_half
```

```
##
##  Kruskal-Wallis rank sum test
##
## data:  MortalityRate by Section
## Kruskal-Wallis chi-squared = 0.65567, df = 1, p-value = 0.4181
```

```r
#oneway test

oneway_result <- oneway.test(MortalityRate ~ Section, data = mortality_data, var.equal = FALSE)
oneway_result_half <- oneway.test(MortalityRate ~ Section, data = mortality_data_half, var.equal = FALSI
oneway_result
```

```
##
##  One-way analysis of means (not assuming equal variances)
##
## data:  MortalityRate and Section
## F = 1.2857, num df = 2.000, denom df = 91.026, p-value = 0.2814
```

```r
oneway_result_half
```

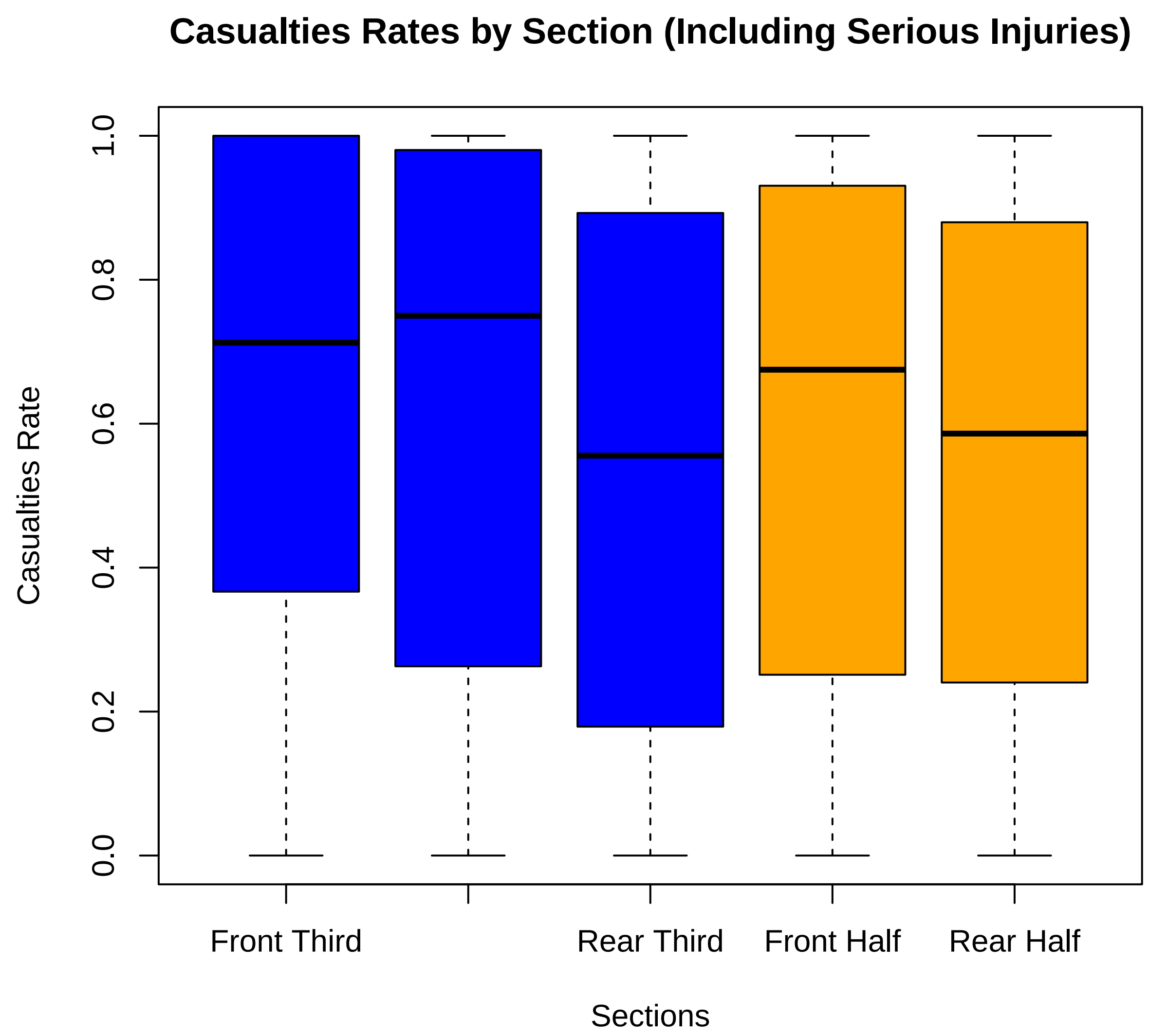
```
##
##  One-way analysis of means (not assuming equal variances)
##
## data:  MortalityRate and Section
## F = 0.40883, num df = 1.000, denom df = 91.667, p-value = 0.5242
```

```r
# now we do the same thing but considering the "gravi" as casualties too

data$X1.casualties_rate_new <- (data$X1.terzo.morti + data$X1.terzo.gravi) / data$X1.third.total
data$X2.casualties_rate_new <- (data$X2.terzo.morti + data$X2.terzo.gravi) / data$X2.third.total
data$X3.casualties_rate_new <- (data$X3.terzo.morti + data$X3.terzo.gravi) / data$X3.third.total
data$X1.half.casualties_rate_new <- (data$X1.meta.morti + data$X1.meta.gravi) / data$X1.half.total
data$X2.half.casualties_rate_new <- (data$X2.meta.morti + data$X2.meta.gravi) / data$X2.half.total

head(data)
```

```
##   NumVolo X1.terzo.lievi X1.terzo.gravi X1.terzo.morti X2.terzo.lievi
## 1       6             17              2             15              1
## 2      28             36              0              0             30
## 3      92              0             11             22              4
## 4     120              0              0              0              5
## 5     123              0              0            136              0
## 6     129              4              0             14              5
##   X2.terzo.gravi X2.terzo.morti X3.terzo.lievi X3.terzo.gravi X3.terzo.morti
## 1             15             64             26             17              0
## 2              0             16             10              0             36
## 3             30             13              0             27             11
## 4              0              8              5              0             21
## 5              0            214             36              0            109
## 6              0             60             24              0             43
##   X1.meta.lievi X1.meta.gravi X1.meta.morti X2.meta.lievi X2.meta.gravi
## 1            17             3            41            26            31
## 2            56             0             8            20             0
## 3            32             0            34            39             0
## 4             3             0             6             7             0
## 5             0             0           226             4             0
## 6             7             0            23            26             0
##   X2.meta.morti fonte PhaseOfFlight  Time    Place HasFire X1.third.total
## 1            34     W       Takeoff Night Airport     Fire             34
## 2            44     W       Takeoff   Day  Outside    Fire             36
## 3            13     W       Landing Night  Outside    Fire             33
## 4            23     W       Landing   Day Airport     Fire              0
## 5           225     W       Takeoff Night  Outside    Fire            136
## 6            90     W       Landing   Day  Outside    Fire             18
##   X2.third.total X3.third.total X1.half.total X2.half.total
## 1             80             43            61            91
## 2             46             46            64            64
## 3             47             38            66            52
## 4             13             26             9            30
## 5            214            145           226           229
## 6             65             67            30           116
##   X1.third.mortality.rate X2.third.mortality.rate X3.third.mortality.rate
## 1               0.4411765               0.8000000               0.0000000
## 2               0.0000000               0.3478261               0.7826087
## 3               0.6666667               0.2765957               0.2894737
## 4                     NaN               0.6153846               0.8076923
## 5               1.0000000               1.0000000               0.7517241
## 6               0.7777778               0.9230769               0.6417910
##   X1.half.mortality.rate X2.half.mortality.rate X1.casualties_rate_new
```

```
## 1                0.6721311               0.3736264               0.5000000
## 2                0.1250000               0.6875000               0.0000000
## 3                0.5151515               0.2500000               1.0000000
## 4                0.6666667               0.7666667                     NaN
## 5                1.0000000               0.9825328               1.0000000
## 6                0.7666667               0.7758621               0.7777778
##   X2.casualties_rate_new X3.casualties_rate_new X1.half.casualties_rate_new
## 1                0.9875000               0.3953488                    0.7213115
## 2                0.3478261               0.7826087                    0.1250000
## 3                0.9148936               1.0000000                    0.5151515
## 4                0.6153846               0.8076923                    0.6666667
## 5                1.0000000               0.7517241                    1.0000000
## 6                0.9230769               0.6417910                    0.7666667
##   X2.half.casualties_rate_new
## 1                     0.7142857
## 2                     0.6875000
## 3                     0.2500000
## 4                     0.7666667
## 5                     0.9825328
## 6                     0.7758621
```

```r
# Plot new casualties rates as scatter plots for each section
boxplot(data$X1.casualties_rate_new,
  data$X2.casualties_rate_new,
  data$X3.casualties_rate_new,
  data$X1.half.casualties_rate_new,
  data$X2.half.casualties_rate_new,
  names = c("Front Third", "Middle Third", "Rear Third", "Front Half", "Rear Half"),
  main = "Casualties Rates by Section (Including Serious Injuries)",
  ylab = "Casualties Rate",
  xlab = "Sections",
  col = c("blue", "blue", "blue", "orange", "orange"))
```




```r
# Perform ANOVA to test if there are significant differences in casualties rates between sections (only
casualties_data <- data.frame(
  Section = rep(c("Front Third", "Middle Third", "Rear Third"), each = nrow(data)),
  CasualtiesRate = c(data$X1.casualties_rate_new, data$X2.casualties_rate_new, data$X3.casualties_rate_r
)
anova_casualties_result <- aov(CasualtiesRate ~ Section, data = casualties_data)
summary(anova_casualties_result)
```

```
##              Df Sum Sq Mean Sq F value Pr(>F)
## Section       2   0.33  0.1648   1.228  0.296
## Residuals   137  18.38  0.1342
## 1 observation deleted due to missingness
```

```r
#half sections
casualties_data_half <- data.frame(
  Section = rep(c("Front Half", "Rear Half"), each = nrow(data)),
```

```r
  CasualtiesRate = c(data$X1.half.casualties_rate_new, data$X2.half.casualties_rate_new)
)
anova_casualties_result_half <- aov(CasualtiesRate ~ Section, data = casualties_data_half)
summary(anova_casualties_result_half)
```

```
##             Df Sum Sq Mean Sq F value Pr(>F)
## Section      1  0.039 0.03904   0.318  0.574
## Residuals   92 11.283 0.12264
```

```r
#try using non-parametric test if ANOVA assumptions are not met
kruskal_casualties_result <- kruskal.test(CasualtiesRate ~ Section, data = casualties_data)
kruskal_casualties_result_half <- kruskal.test(CasualtiesRate ~ Section, data = casualties_data_half)
kruskal_casualties_result
```

```
##
##  Kruskal-Wallis rank sum test
##
## data:  CasualtiesRate by Section
## Kruskal-Wallis chi-squared = 2.6064, df = 2, p-value = 0.2717
```

```r
kruskal_casualties_result_half
```

```
##
##  Kruskal-Wallis rank sum test
##
## data:  CasualtiesRate by Section
## Kruskal-Wallis chi-squared = 0.4796, df = 1, p-value = 0.4886
```

```r
#oneway test
oneway_casualties_result <- oneway.test(CasualtiesRate ~ Section, data = casualties_data, var.equal = F
oneway_casualties_result_half <- oneway.test(CasualtiesRate ~ Section, data = casualties_data_half, var
oneway_casualties_result
```

```
##
##  One-way analysis of means (not assuming equal variances)
##
## data:  CasualtiesRate and Section
## F = 1.2205, num df = 2.000, denom df = 91.313, p-value = 0.2999
```

```r
oneway_casualties_result_half
```

```
##
##  One-way analysis of means (not assuming equal variances)
##
## data:  CasualtiesRate and Section
## F = 0.31829, num df = 1.000, denom df = 91.975, p-value = 0.574
```

```r
# t.test only the fisrts third against the rears third
t_test_result <- t.test(data$X1.third.mortality.rate, data$X3.third.mortality.rate, var.equal = FALSE)
t_test_result
```

```
##
##  Welch Two Sample t-test
##
## data:  data$X1.third.mortality.rate and data$X3.third.mortality.rate
## t = 1.5556, df = 89.487, p-value = 0.1233
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.03396324  0.27900140
## sample estimates:
## mean of x mean of y
## 0.5434137 0.4208946
```

```r
#use non-parametric test if t-test assumptions are not met
wilcox_test_result <- wilcox.test(data$X1.third.mortality.rate, data$X3.third.mortality.rate)
```

```
## Warning in wilcox.test.default(data$X1.third.mortality.rate,
## data$X3.third.mortality.rate): cannot compute exact p-value with ties
```

```r
wilcox_test_result
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  data$X1.third.mortality.rate and data$X3.third.mortality.rate
## W = 1291, p-value = 0.1052
## alternative hypothesis: true location shift is not equal to 0
```

```r
# now we want to test if there is a difference in mortality rate based on the other variables
# PhaseOfFlight, Time, Place, HasFire
# using the total mortality rate, only deaths, do this for evwey third

out <- lm(data$X1.third.mortality.rate ~ data$PhaseOfFlight + data$Time + data$Place + data$HasFire)
summary(out)
```

```
##
## Call:
## lm(formula = data$X1.third.mortality.rate ~ data$PhaseOfFlight +
##     data$Time + data$Place + data$HasFire)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.62619 -0.27399  0.07094  0.26658  0.66495
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 0.55633    0.12745   4.365 9.43e-05 ***
## data$PhaseOfFlightLanding   0.31337    0.34121   0.918   0.3642
## data$PhaseOfFlightTakeoff  -0.17701    0.12568  -1.408   0.1671
## data$TimeDay               -0.10427    0.26774  -0.389   0.6991
## data$TimeNight              0.01703    0.14142   0.120   0.9048
## data$PlaceAirport           0.06205    0.17287   0.359   0.7216
## data$PlaceOutside           0.24686    0.14262   1.731   0.0916 .
```

```
## data$HasFireNo-Fire        -0.29114    0.13331  -2.184   0.0352 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3851 on 38 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.2155, Adjusted R-squared:  0.07096
## F-statistic: 1.491 on 7 and 38 DF,  p-value: 0.1999
```

```r
out_2 <- lm(data$X2.third.mortality.rate ~ data$PhaseOfFlight + data$Time + data$Place + data$HasFire)
summary(out_2)
```

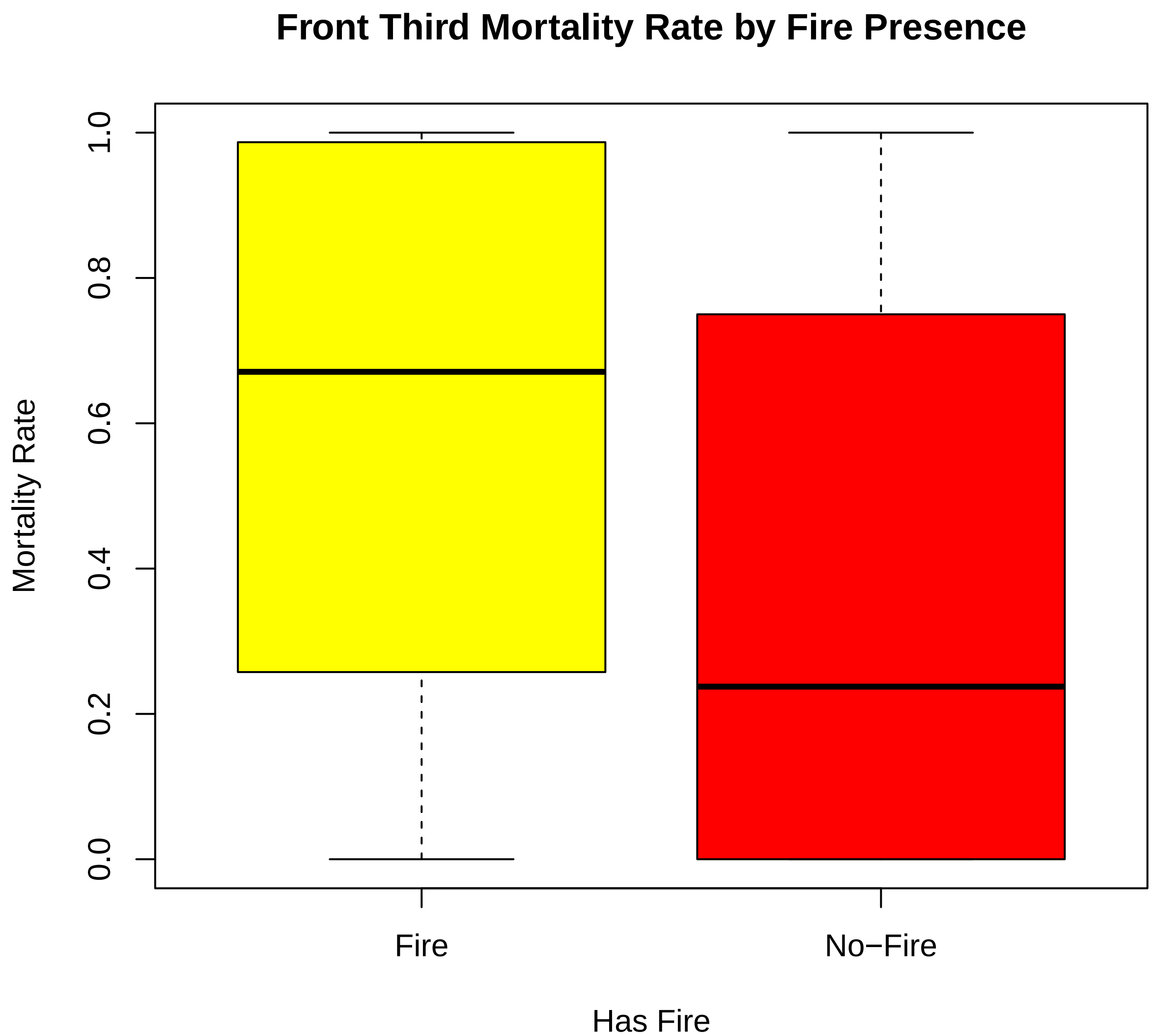
```
##
## Call:
## lm(formula = data$X2.third.mortality.rate ~ data$PhaseOfFlight +
##     data$Time + data$Place + data$HasFire)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.63420 -0.37099  0.01981  0.31813  0.59990
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 0.38769    0.12962   2.991   0.0048 **
## data$PhaseOfFlightLanding   0.08039    0.34876   0.230   0.8189
## data$PhaseOfFlightTakeoff   0.01241    0.12735   0.097   0.9229
## data$TimeDay               -0.03338    0.27323  -0.122   0.9034
## data$TimeNight              0.06944    0.14302   0.486   0.6300
## data$PlaceAirport           0.20788    0.17042   1.220   0.2298
## data$PlaceOutside           0.23411    0.14574   1.606   0.1163
## data$HasFireNo-Fire        -0.20291    0.13559  -1.497   0.1426
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3937 on 39 degrees of freedom
## Multiple R-squared:  0.1128, Adjusted R-squared:  -0.04639
## F-statistic: 0.7086 on 7 and 39 DF,  p-value: 0.6648
```

```r
out_3 <- lm(data$X3.third.mortality.rate ~ data$PhaseOfFlight + data$Time + data$Place + data$HasFire)
summary(out_3)
```

```
##
## Call:
## lm(formula = data$X3.third.mortality.rate ~ data$PhaseOfFlight +
##     data$Time + data$Place + data$HasFire)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.56827 -0.23612 -0.04599  0.25262  0.58917
##
## Coefficients:
##                             Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 0.413075   0.120342   3.433  0.00143 **
```

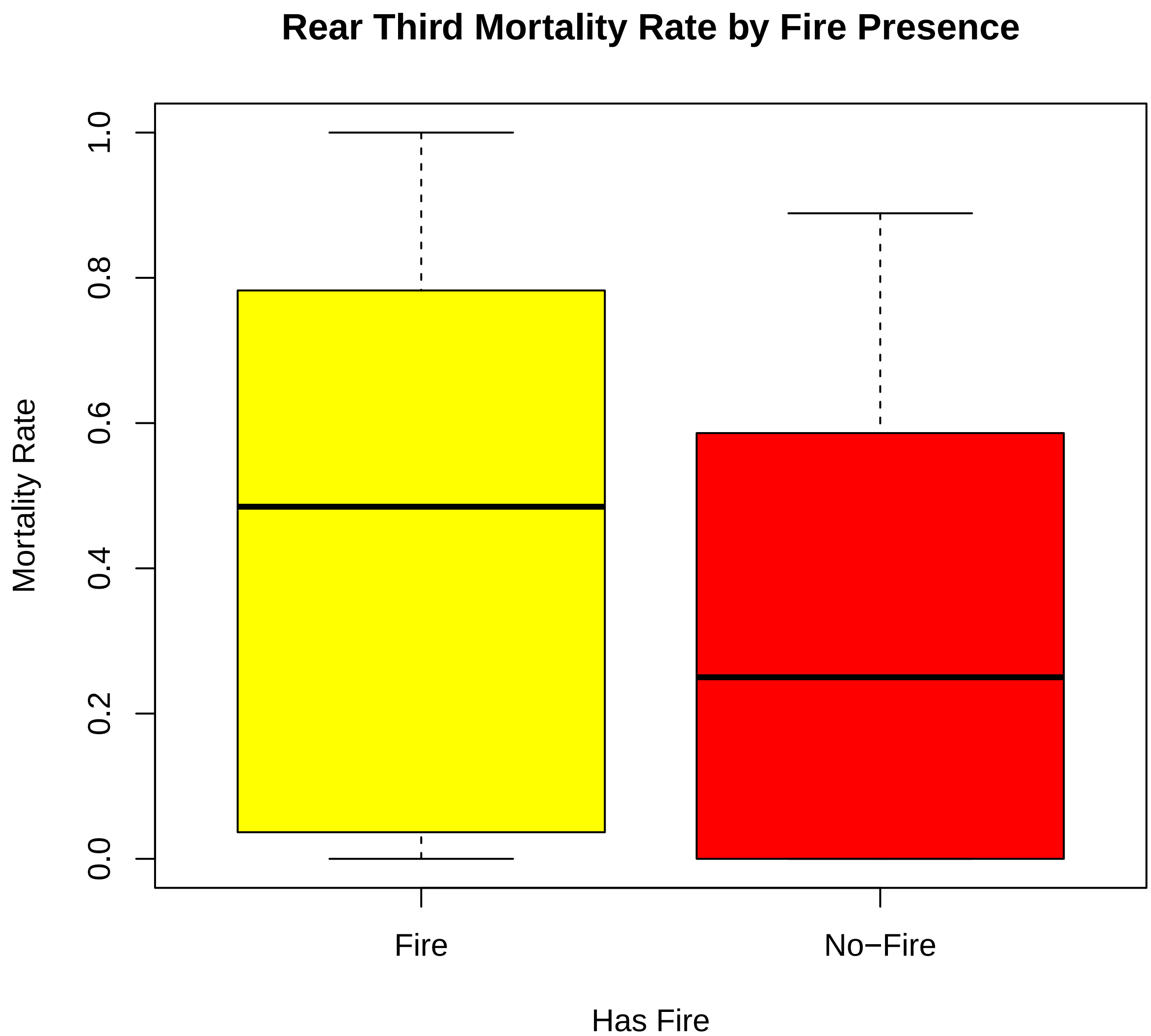
```
## data$PhaseOfFlightLanding  -0.057355   0.323798  -0.177  0.86032
## data$PhaseOfFlightTakeoff  -0.002242   0.118232  -0.019  0.98497
## data$TimeDay               -0.288042   0.253673  -1.135  0.26310
## data$TimeNight             -0.071894   0.132786  -0.541  0.59129
## data$PlaceAirport           0.137487   0.158220   0.869  0.39019
## data$PlaceOutside           0.155196   0.135307   1.147  0.25837
## data$HasFireNo-Fire        -0.186705   0.125886  -1.483  0.14608
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3655 on 39 degrees of freedom
## Multiple R-squared:  0.118,  Adjusted R-squared:  -0.04035
## F-statistic: 0.7452 on 7 and 39 DF,  p-value: 0.6355
```

```r
#plot the mortality rates based on HasFire
boxplot(data$X1.third.mortality.rate ~ data$HasFire,
        main = "Front Third Mortality Rate by Fire Presence",
        xlab = "Has Fire",
        ylab = "Mortality Rate",
        col = c("yellow", "red"))
```

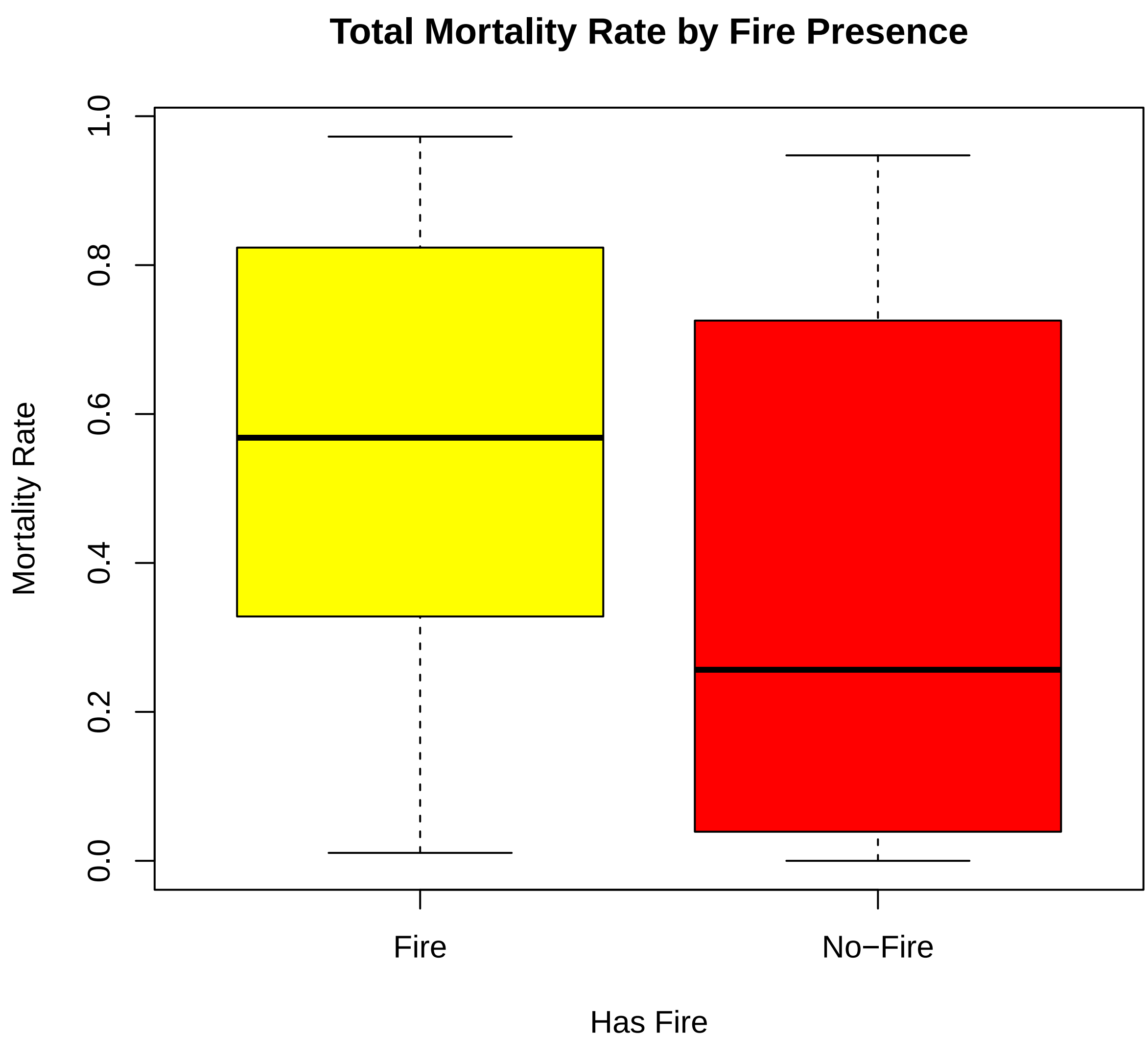



```
boxplot(data$X2.third.mortality.rate ~ data$HasFire,
        main = "Middle Third Mortality Rate by Fire Presence",
        xlab = "Has Fire",
        ylab = "Mortality Rate",
        col = c("yellow", "red"))
```


```r
boxplot(data$X3.third.mortality.rate ~ data$HasFire,
        main = "Rear Third Mortality Rate by Fire Presence",
        xlab = "Has Fire",
        ylab = "Mortality Rate",
        col = c("yellow", "red"))
```




```
#and full plane mortality rate
data$Total.mortality.rate <- (data$X1.terzo.morti + data$X2.terzo.morti + data$X3.terzo.morti) / (data$X

boxplot(data$Total.mortality.rate ~ data$HasFire,
        main = "Total Mortality Rate by Fire Presence",
        xlab = "Has Fire",
        ylab = "Mortality Rate",
        col = c("yellow", "red"))
```




```
# now for the halves
out_half_1 <- lm(data$X1.half.mortality.rate ~ data$PhaseOfFlight + data$Time + data$Place + data$HasFi
summary(out_half_1)
```

```
##
## Call:
## lm(formula = data$X1.half.mortality.rate ~ data$PhaseOfFlight +
##     data$Time + data$Place + data$HasFire)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.6724 -0.3041  0.0530  0.2636  0.5802
##
## Coefficients:
##                              Estimate Std. Error t value Pr(>|t|)
## (Intercept)                   0.47111    0.12149   3.878 0.000394 ***
```

```
## data$PhaseOfFlightLanding  -0.08327    0.32688  -0.255 0.800265
## data$PhaseOfFlightTakeoff  -0.06912    0.11936  -0.579 0.565829
## data$TimeDay               -0.26135    0.25609  -1.021 0.313758
## data$TimeNight              0.05772    0.13405   0.431 0.669130
## data$PlaceAirport           0.14255    0.15973   0.892 0.377613
## data$PlaceOutside           0.27041    0.13659   1.980 0.054827 .
## data$HasFireNo-Fire        -0.25262    0.12708  -1.988 0.053881 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.369 on 39 degrees of freedom
## Multiple R-squared:  0.2038, Adjusted R-squared:  0.06089
## F-statistic: 1.426 on 7 and 39 DF,  p-value: 0.2228
```

```
out_half_2 <- lm(data$X2.half.mortality.rate ~ data$PhaseOfFlight + data$Time + data$Place + data$HasFi
summary(out_half_2)
```

```
##
## Call:
## lm(formula = data$X2.half.mortality.rate ~ data$PhaseOfFlight +
##     data$Time + data$Place + data$HasFire)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.60915 -0.23925  0.02419  0.28219  0.58485
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 0.41515    0.12107   3.429  0.00144 **
## data$PhaseOfFlightLanding   0.13238    0.32575   0.406  0.68669
## data$PhaseOfFlightTakeoff   0.01443    0.11894   0.121  0.90409
## data$TimeDay               -0.17559    0.25520  -0.688  0.49549
## data$TimeNight             -0.03249    0.13359  -0.243  0.80910
## data$PlaceAirport           0.17936    0.15917   1.127  0.26669
## data$PlaceOutside           0.19400    0.13612   1.425  0.16205
## data$HasFireNo-Fire        -0.21447    0.12664  -1.693  0.09834 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3677 on 39 degrees of freedom
## Multiple R-squared:  0.1079, Adjusted R-squared:  -0.05217
## F-statistic: 0.6742 on 7 and 39 DF,  p-value: 0.6926
```

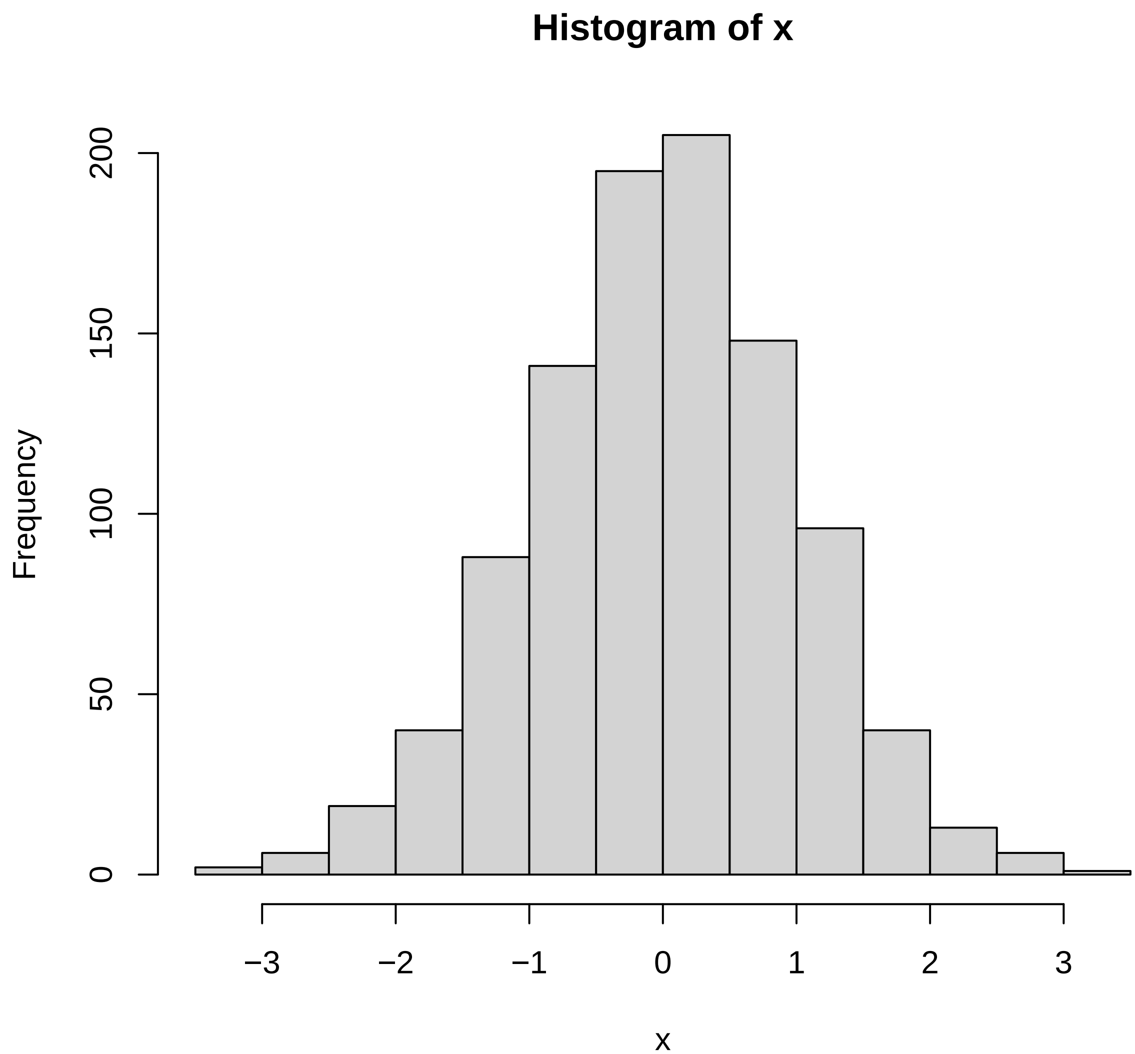
```
#boxplots for halves
boxplot(data$X1.half.mortality.rate ~ data$HasFire,
        main = "Front Half Mortality Rate by Fire Presence",
        xlab = "Has Fire",
        ylab = "Mortality Rate",
        col = c("lightgreen", "darkgreen"))
```


```r
boxplot(data$X2.half.mortality.rate ~ data$HasFire,
        main = "Rear Half Mortality Rate by Fire Presence",
        xlab = "Has Fire",
        ylab = "Mortality Rate",
        col = c("lightgreen", "darkgreen"))
```

Data Description

This is an example of r-chuch. You can instrt the R code, Knit-it and your report is compiled in pdf.

```
x <- rnorm(1000)
hist(x)
```



3 Analysis

4 Results

5 Conclusions