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

```
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
##
##
##
```

```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 ...
# $ 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. . . .

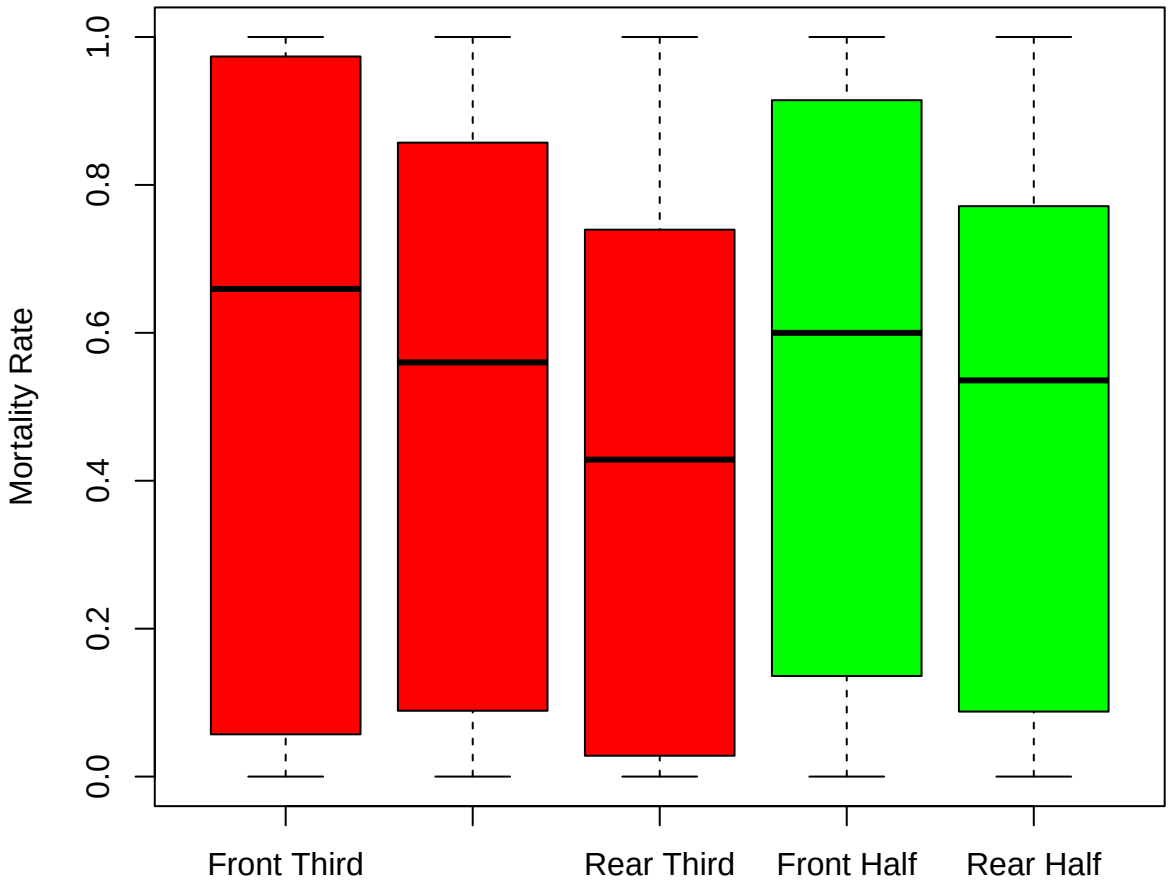
```
# add a colum of # of seat for each airplain section
```

```r
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



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"))
```




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

```
#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
```

```
#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
```

```
kruskal_result_half
```

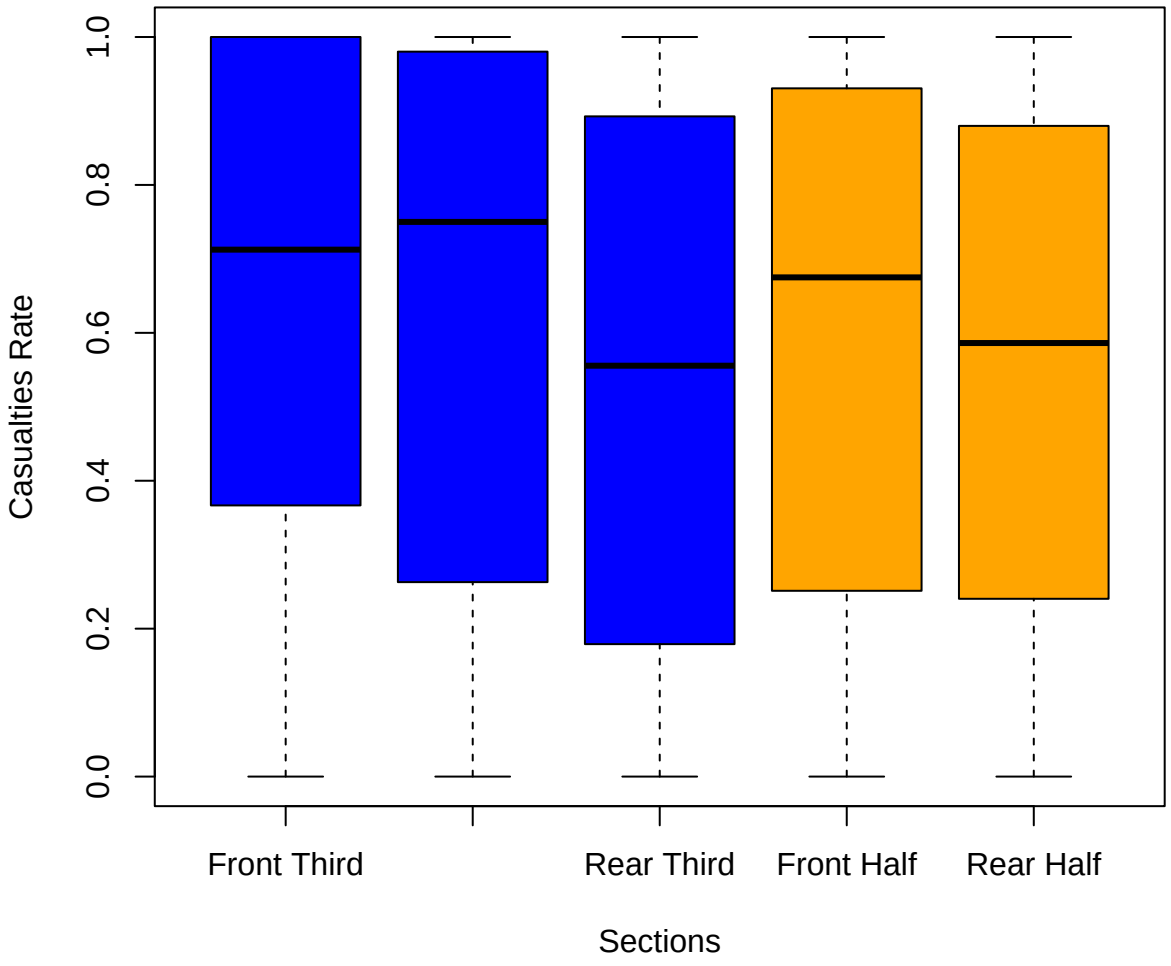
```
##
##  Kruskal-Wallis rank sum test
##
## data:  MortalityRate by Section
## Kruskal-Wallis chi-squared = 0.65567, df = 1, p-value = 0.4181
```

```
# 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_
)
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)),
  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
```

```
#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
```

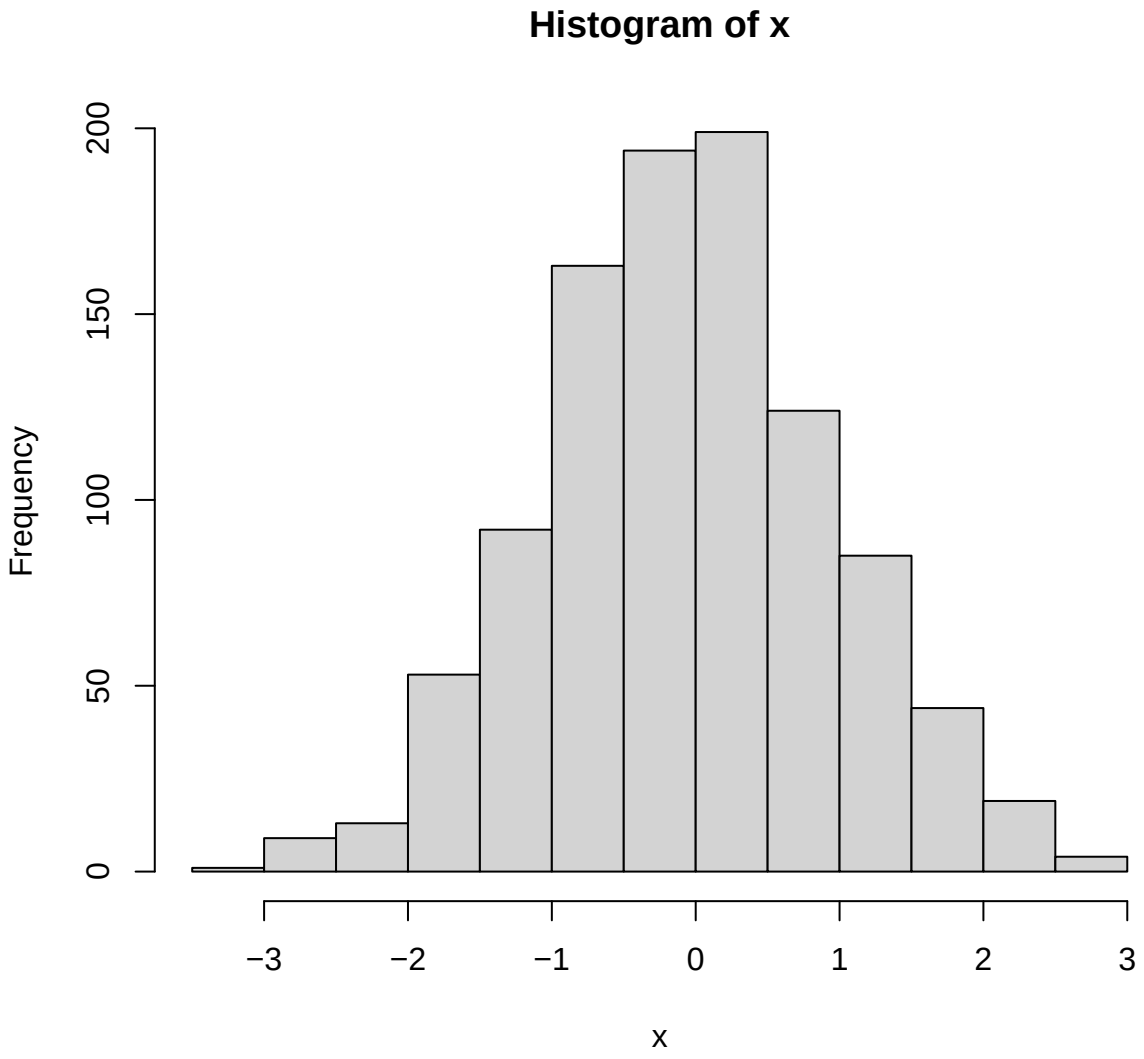
```
kruskal_casualties_result_half
```

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

2 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