

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

Inferential Statistics 2025/2026

Barbato Alberto, Naldoni Martina

2026-01-13

Contents

1	Introduction	1
2	Vocabulary	1
3	Perspective	2
4	Analysis	39
5	Results	39
6	Conclusions	39

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 Vocabulary

Before starting the analysis, the meaning of some terms must be clarified

- Flight : A flight a trip made by an aircraft that connects two airports. It is identified by a number. (Example : AirFrance 225 is the name of the regular service from New Deli to Paris)
- Accident : An Accident is an occurrence where a person is fatally or seriously injured, the aircraft sustains significant structural damage, or the aircraft goes missing.

- Incident : An incident is a dangerous situation where no one is seriously hurt and the plane isn't badly damaged.
- Crash : A crash is a type of accident where an aircraft strikes the ground, water, or an obstacle with enough force to cause severe damage or total destruction.
- Accident flight : when a serious accident happens on a flight, usually the companies save the flight number to refer to that flight, and change the flight number for new flights. (Example : Air France 447 was the name of the connection from Rio to Paris, after the 2009 accident the flight became Air France 445)
- Serious injury : Any injury requiring more than 48 hours of hospitalization or involving broken bones (other than fingers/toes)

3 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

```
data <- read.csv("AllCREEP_cleaned.csv")
# transform the variables that are categorical into factors
data$fonte <- as.factor(data$fonte)
data$PhaseOfFlight <- as.factor(data$PhaseOfFlight)
data$Time <- as.factor(data$Time)
data$Place <- as.factor(data$Place)
data$HasFire <- as.factor(data$HasFire)
data$Environment <- as.factor(data$Environment)
data$Energy_absorption <- as.factor(data$Energy_absorption)
```

```
str(data)
```

```
## 'data.frame':   47 obs. of  26 variables:
## $ Airline      : chr  "singapore airlines" "british airtours" "british midland" "china airlines"
## $ 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 : Factor w/ 4 levels "", "fr", "r", "w": 4 4 4 4 4 4 4 4 4 ...
## $ PhaseOfFlight : Factor w/ 2 levels "landing", "takeoff": 2 2 1 1 2 1 1 1 2 1 ...
## $ Time : Factor w/ 2 levels "day", "night": 2 1 2 1 2 1 1 2 1 1 ...
## $ Place : Factor w/ 2 levels "airport", "outside": 1 2 2 1 2 2 1 2 2 2 ...
## $ HasFire : Factor w/ 2 levels "fire", "no-fire": 1 1 1 1 1 1 1 2 1 1 ...
## $ Crushed_fuselage : int 1 1 1 1 1 1 1 1 1 1 ...
## $ Restraint_intact : int 0 0 0 1 0 0 0 0 0 0 ...
## $ Environment : Factor w/ 2 levels "clear", "dangerous": 2 2 2 1 2 2 2 2 1 2 ...
## $ Energy_absorption: Factor w/ 2 levels "gear", "nogear": 2 2 1 2 2 2 1 2 2 2 ...
```

```
summary(data)
```

```
##      Airline      NumVolo      X1.terzo.lievi      X1.terzo.gravi
## Length:47      Min.      : 6      Min.      : 0.00      Min.      : 0.000
## Class :character 1st Qu.: 227      1st Qu.: 0.00      1st Qu.: 0.000
## Mode  :character Median : 812      Median : 4.00      Median : 0.000
##                      Mean  :1591      Mean  : 12.87      Mean  : 2.553
##                      3rd Qu.:1603      3rd Qu.: 17.00      3rd Qu.: 4.000
##                      Max.   :9642      Max.   :141.00      Max.   :16.000
## X1.terzo.morti   X2.terzo.lievi   X2.terzo.gravi   X2.terzo.morti
## Min.      : 0.00      Min.      : 0.00      Min.      : 0.000      Min.      : 0.00
## 1st Qu.: 1.50      1st Qu.: 1.00      1st Qu.: 0.000      1st Qu.: 2.00
## Median : 11.00      Median : 7.00      Median : 2.000      Median : 12.00
## Mean  : 16.15      Mean  : 22.11      Mean  : 4.638      Mean  : 28.62
## 3rd Qu.: 19.00      3rd Qu.: 26.50      3rd Qu.: 8.000      3rd Qu.: 31.00
## Max.   :136.00      Max.   :174.00      Max.   :30.000      Max.   :214.00
## X3.terzo.lievi   X3.terzo.gravi   X3.terzo.morti   X1.meta.lievi
## Min.      : 0.0      Min.      : 0.000      Min.      : 0.00      Min.      : 0.00
## 1st Qu.: 3.0      1st Qu.: 0.000      1st Qu.: 1.50      1st Qu.: 1.50
## Median : 10.0      Median : 0.000      Median : 8.00      Median : 11.00
## Mean  : 20.4      Mean  : 3.872      Mean  : 19.83      Mean  : 29.23
## 3rd Qu.: 27.0      3rd Qu.: 4.500      3rd Qu.: 24.50      3rd Qu.: 40.50
## Max.   :142.0      Max.   :27.000      Max.   :113.00      Max.   :221.00
## X1.meta.gravi   X1.meta.morti   X2.meta.lievi   X2.meta.gravi
## Min.      : 0.00      Min.      : 0.00      Min.      : 0.00      Min.      : 0.000
## 1st Qu.: 0.00      1st Qu.: 5.00      1st Qu.: 6.50      1st Qu.: 0.000
## Median : 0.00      Median : 20.00      Median : 14.00      Median : 0.000
## Mean  : 4.34      Mean  : 28.57      Mean  : 28.98      Mean  : 5.234
## 3rd Qu.: 7.00      3rd Qu.: 35.00      3rd Qu.: 37.50      3rd Qu.: 7.500
## Max.   :34.00      Max.   :226.00      Max.   :184.00      Max.   :31.000
## X2.meta.morti   fonte   PhaseOfFlight   Time      Place      HasFire
## Min.      : 0.00      : 8      landing:30      day :36      airport:25      fire :33
## 1st Qu.: 3.50      fr: 4      takeoff:17      night:11      outside:22      no-fire:14
## Median : 15.00      r : 1
## Mean  : 36.53      w :34
## 3rd Qu.: 45.00
```

```
## Max.      :225.00
## Crushed_fuselage Restraint_intact Environment Energy_absorption
## Min.      :0.0000 Min.      :0.0000 clear      : 5 gear :29
## 1st Qu.:1.0000 1st Qu.:0.0000 dangerous:42 nogear:18
## Median :1.0000 Median :0.0000
## Mean    :0.9362 Mean    :0.1702
## 3rd Qu.:1.0000 3rd Qu.:0.0000
## Max.    :1.0000 Max.    :1.0000
```

```
#> data <- read.csv("Aerei_Final.csv")
## 'data.frame': 47 obs. of 26 variables:
## $ Airline : chr "Singapore Airlines" "British Airtours" "British Midland" "China Airlines"
## $ 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" ...
```

2

```
## [1] 2
## $ Place : chr "Airport " "Outside" "Outside" "Airport " ...
## $ HasFire : chr "Fire" "Fire" "Fire" "Fire" ...
## $ Crushed_fuselage : int 1 1 1 1 1 1 1 1 1 1 ...
## $ Restraint_intact : int 0 0 0 1 0 0 0 0 0 0 ...
## $ Environment : chr "dangerous" "dangerous" "dangerous" "clear" ...
## $ Energy_absorption: chr "nogear" "nogear" "gear" "nogear"
```

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
```

```
str(data)
```

```
## 'data.frame': 47 obs. of 31 variables:
## $ Airline : chr "singapore airlines" "british airtours" "british midland" "china airlines"
```

```

## $ 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           : Factor w/ 4 levels "", "fr", "r", "w": 4 4 4 4 4 4 4 4 4 4 ...
## $ PhaseOfFlight   : Factor w/ 2 levels "landing", "takeoff": 2 2 1 1 2 1 1 1 2 1 ...
## $ Time             : Factor w/ 2 levels "day", "night": 2 1 2 1 2 1 1 2 1 1 ...
## $ Place            : Factor w/ 2 levels "airport", "outside": 1 2 2 1 2 2 1 2 2 2 ...
## $ HasFire          : Factor w/ 2 levels "fire", "no-fire": 1 1 1 1 1 1 1 2 1 1 ...
## $ Crushed_fuselage : int  1 1 1 1 1 1 1 1 1 1 ...
## $ Restraint_intact : int  0 0 0 1 0 0 0 0 0 0 ...
## $ Environment      : Factor w/ 2 levels "clear", "dangerous": 2 2 2 1 2 2 2 2 1 2 ...
## $ Energy_absorption: Factor w/ 2 levels "gear", "nogear": 2 2 1 2 2 2 1 2 2 2 ...
## $ X1.third.total   : int  34 36 33 0 136 18 18 16 55 36 ...
## $ X2.third.total   : int  80 46 47 13 214 65 146 35 59 26 ...
## $ X3.third.total   : int  43 46 38 26 145 67 91 38 33 29 ...
## $ X1.half.total    : int  61 64 66 9 226 30 102 37 80 42 ...
## $ X2.half.total    : int  91 64 52 30 229 116 145 53 72 46 ...

```

`summary(data)`

```

##      Airline      NumVolo      X1.terzo.lievi      X1.terzo.gravi
## Length:47      Min.      : 6      Min.      : 0.00      Min.      : 0.000
## Class :character 1st Qu.: 227      1st Qu.: 0.00      1st Qu.: 0.000
## Mode  :character Median : 812      Median : 4.00      Median : 0.000
##                      Mean  :1591      Mean  : 12.87      Mean   : 2.553
##                      3rd Qu.:1603      3rd Qu.: 17.00      3rd Qu.: 4.000
##                      Max.   :9642      Max.   :141.00      Max.   :16.000
## X1.terzo.morti    X2.terzo.lievi    X2.terzo.gravi    X2.terzo.morti
## Min.      : 0.00      Min.      : 0.00      Min.      : 0.000      Min.      : 0.00
## 1st Qu.: 1.50      1st Qu.: 1.00      1st Qu.: 0.000      1st Qu.: 2.00
## Median : 11.00      Median : 7.00      Median : 2.000      Median : 12.00
## Mean  : 16.15      Mean  : 22.11      Mean  : 4.638      Mean  : 28.62
## 3rd Qu.: 19.00      3rd Qu.: 26.50      3rd Qu.: 8.000      3rd Qu.: 31.00
## Max.   :136.00      Max.   :174.00      Max.   :30.000      Max.   :214.00
## X3.terzo.lievi    X3.terzo.gravi    X3.terzo.morti    X1.meta.lievi
## Min.      : 0.0      Min.      : 0.000      Min.      : 0.00      Min.      : 0.00
## 1st Qu.: 3.0      1st Qu.: 0.000      1st Qu.: 1.50      1st Qu.: 1.50
## Median : 10.0      Median : 0.000      Median : 8.00      Median : 11.00
## Mean  : 20.4      Mean  : 3.872      Mean  : 19.83      Mean  : 29.23
## 3rd Qu.: 27.0      3rd Qu.: 4.500      3rd Qu.: 24.50      3rd Qu.: 40.50
## Max.   :142.0      Max.   :27.000      Max.   :113.00      Max.   :221.00
## X1.meta.gravi     X1.meta.morti     X2.meta.lievi     X2.meta.gravi

```

```
## Min. : 0.00 Min. : 0.00 Min. : 0.00 Min. : 0.000
## 1st Qu.: 0.00 1st Qu.: 5.00 1st Qu.: 6.50 1st Qu.: 0.000
## Median : 0.00 Median : 20.00 Median : 14.00 Median : 0.000
## Mean : 4.34 Mean : 28.57 Mean : 28.98 Mean : 5.234
## 3rd Qu.: 7.00 3rd Qu.: 35.00 3rd Qu.: 37.50 3rd Qu.: 7.500
## Max. :34.00 Max. :226.00 Max. :184.00 Max. :31.000
## X2.meta.morti fonte PhaseOfFlight Time Place HasFire
## Min. : 0.00 : 8 landing:30 day :36 airport:25 fire :33
## 1st Qu.: 3.50 fr: 4 takeoff:17 night:11 outside:22 no-fire:14
## Median : 15.00 r : 1
## Mean : 36.53 w :34
## 3rd Qu.: 45.00
## Max. :225.00
## Crushed_fuselage Restraint_intact Environment Energy_absorption
## Min. :0.0000 Min. :0.0000 clear : 5 gear :29
## 1st Qu.:1.0000 1st Qu.:0.0000 dangerous:42 nogear:18
## Median :1.0000 Median :0.0000
## Mean :0.9362 Mean :0.1702
## 3rd Qu.:1.0000 3rd Qu.:0.0000
## Max. :1.0000 Max. :1.0000
## X1.third.total X2.third.total X3.third.total X1.half.total
## Min. : 0.00 Min. : 4.00 Min. : 2.00 Min. : 5.00
## 1st Qu.: 14.50 1st Qu.: 22.00 1st Qu.: 21.50 1st Qu.: 27.50
## Median : 24.00 Median : 35.00 Median : 30.00 Median : 48.00
## Mean : 31.57 Mean : 55.36 Mean : 44.11 Mean : 62.15
## 3rd Qu.: 36.00 3rd Qu.: 73.50 3rd Qu.: 53.00 3rd Qu.: 68.00
## Max. :141.00 Max. :214.00 Max. :145.00 Max. :226.00
## X2.half.total
## Min. : 7.00
## 1st Qu.: 30.50
## Median : 46.00
## Mean : 70.74
## 3rd Qu.:106.50
## Max. :229.00
```

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)
```

```
##           Airline NumVolo X1.terzo.lievi X1.terzo.gravi X1.terzo.morti
## 1 singapore airlines      6           17           2          15
## 2 british airtours      28           36           0           0
## 3 british midland       92            0          11          22
## 4 china airlines      120            0           0           0
## 5 japan air lines     123            0           0          136
## 6 air china          129            4           0           14
## X2.terzo.lievi X2.terzo.gravi X2.terzo.morti X3.terzo.lievi X3.terzo.gravi
```

```

## 1      1      15      64      26      17
## 2     30      0      16     10      0
## 3      4     30     13      0     27
## 4      5      0      8      5      0
## 5      0      0     214     36      0
## 6      5      0     60     24      0
##  X3.terzo.morti X1.meta.lievi X1.meta.gravi X1.meta.morti X2.meta.lievi
## 1          0          17          3          41          26
## 2         36         56          0          8          20
## 3         11         32          0         34          39
## 4         21          3          0          6          7
## 5        109          0          0        226          4
## 6         43          7          0         23          26
##  X2.meta.gravi X2.meta.morti fonte PhaseOfFlight Time Place HasFire
## 1          31          34      w      takeoff night airport fire
## 2           0          44      w      takeoff  day outside fire
## 3           0          13      w      landing night outside fire
## 4           0          23      w      landing  day airport fire
## 5           0         225      w      takeoff night outside fire
## 6           0          90      w      landing  day outside fire
##  Crushed_fuselage Restraint_intact Environment Energy_absorption
## 1           1           0 dangerous      nogear
## 2           1           0 dangerous      nogear
## 3           1           0 dangerous      gear
## 4           1           1      clear      nogear
## 5           1           0 dangerous      nogear
## 6           1           0 dangerous      nogear
##  X1.third.total X2.third.total X3.third.total X1.half.total X2.half.total
## 1          34          80          43          61          91
## 2          36          46          46          64          64
## 3          33          47          38          66          52
## 4           0          13          26           9          30
## 5         136         214         145        226        229
## 6          18          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

```

```

## 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
)
str(mortality_data)

```

```

## 'data.frame':   141 obs. of  2 variables:
##  $ Section      : chr  "Front Third" "Front Third" "Front Third" "Front Third" ...
##  $ MortalityRate: num  0.441 0 0.667 NaN 1 ...

```

```

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

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

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)

#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
```

```
#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 = FALSE)
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
```

```
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
```

```
# 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)
```

```
##           Airline NumVolo X1.terzo.lievi X1.terzo.gravi X1.terzo.morti
## 1 singapore airlines      6           17           2          15
## 2 british airtours      28           36           0           0
## 3 british midland      92            0          11          22
## 4 china airlines     120            0           0           0
## 5 japan air lines    123            0           0          136
## 6 air china        129            4           0          14
## X2.terzo.lievi X2.terzo.gravi X2.terzo.morti X3.terzo.lievi X3.terzo.gravi
## 1             1           15           64           26           17
## 2             30            0           16           10            0
```

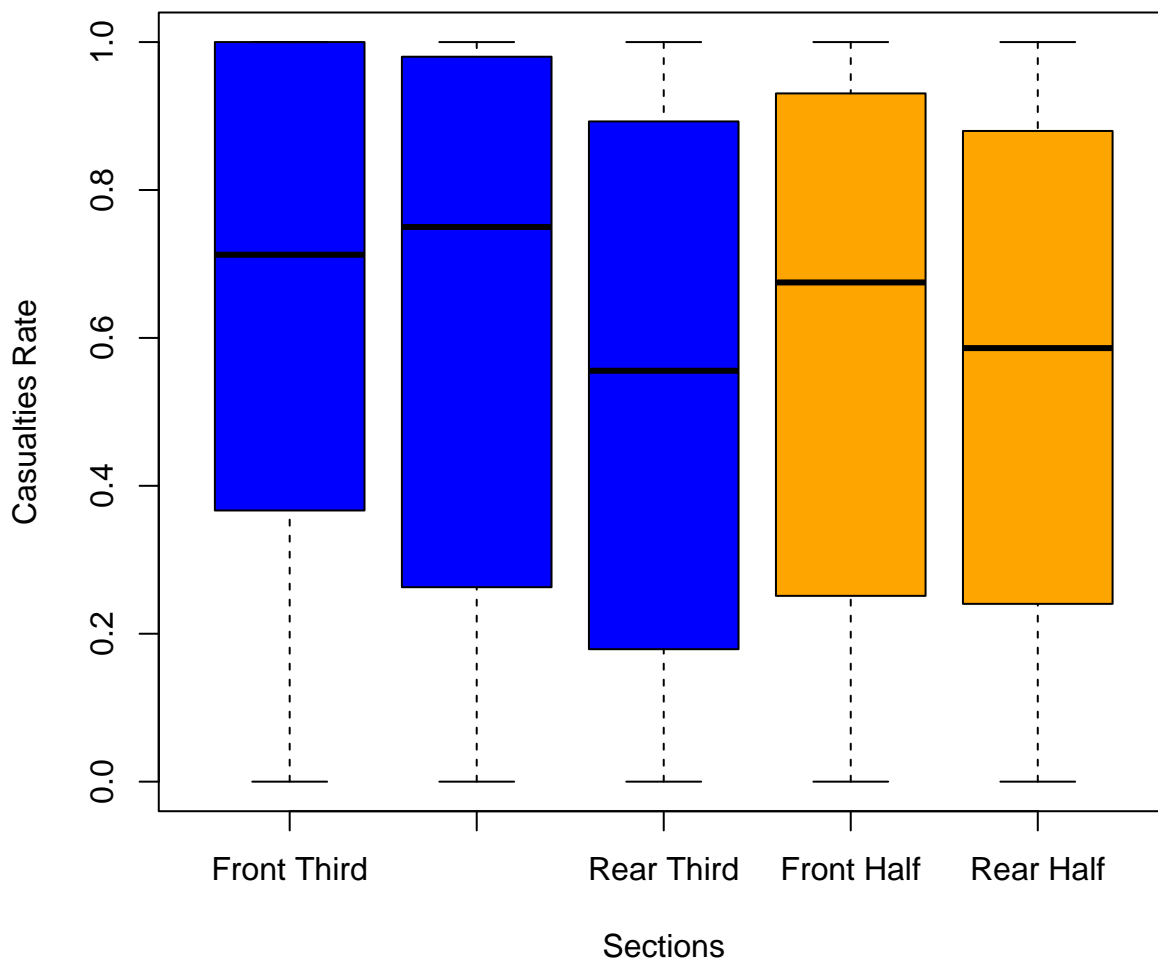
## 3	4	30	13	0	27		
## 4	5	0	8	5	0		
## 5	0	0	214	36	0		
## 6	5	0	60	24	0		
##	X3.terzo.morti	X1.meta.lievi	X1.meta.gravi	X1.meta.morti	X2.meta.lievi		
## 1	0	17	3	41	26		
## 2	36	56	0	8	20		
## 3	11	32	0	34	39		
## 4	21	3	0	6	7		
## 5	109	0	0	226	4		
## 6	43	7	0	23	26		
##	X2.meta.gravi	X2.meta.morti	fonte	PhaseOfFlight	Time	Place	HasFire
## 1	31	34	w	takeoff	night	airport	fire
## 2	0	44	w	takeoff	day	outside	fire
## 3	0	13	w	landing	night	outside	fire
## 4	0	23	w	landing	day	airport	fire
## 5	0	225	w	takeoff	night	outside	fire
## 6	0	90	w	landing	day	outside	fire
##	Crushed_fuselage	Restraint_intact	Environment	Energy_absorption			
## 1	1	0	dangerous	nogear			
## 2	1	0	dangerous	nogear			
## 3	1	0	dangerous	gear			
## 4	1	1	clear	nogear			
## 5	1	0	dangerous	nogear			
## 6	1	0	dangerous	nogear			
##	X1.third.total	X2.third.total	X3.third.total	X1.half.total	X2.half.total		
## 1	34	80	43	61	91		
## 2	36	46	46	64	64		
## 3	33	47	38	66	52		
## 4	0	13	26	9	30		
## 5	136	214	145	226	229		
## 6	18	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
```

```
# 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"))
```

Casualties Rates by Section (Including Serious Injuries)



```
# 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_new)
)
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

#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

#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.equal = F)
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
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
# 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
#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
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
# 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$HasFire + data$Environment + data$Energy_absorption + data$Crushed_fuselage + data$Restraint_intact)
summary(out)

##
## Call:
## lm(formula = data$X1.third.mortality.rate ~ data$PhaseOfFlight +
## data$Time + data$Place + data$HasFire + data$Environment +
## data$Energy_absorption + data$Crushed_fuselage + data$Restraint_intact)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.70350 -0.16771  0.04745  0.27778  0.59358
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.49130    0.31703   1.550  0.1297
## data$PhaseOfFlighttakeoff -0.20082    0.12363  -1.624  0.1128
## data$Timenight      0.05967    0.13300   0.449  0.6563
## data$Placeoutside    0.13799    0.13606   1.014  0.3171
```

```

## data$HasFireno-fire          -0.20853    0.13763   -1.515    0.1382
## data$Environmentdangerous    -0.05387    0.20839   -0.258    0.7975
## data$Energy_absorptionnogear  0.02816    0.14494    0.194    0.8470
## data$Crushed_fuselage        0.21220    0.26274    0.808    0.4245
## data$Restraint_intact        -0.32888    0.17213   -1.911    0.0638 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3645 on 37 degrees of freedom
## (1 observation deleted due to missingness)
## Multiple R-squared:  0.3156, Adjusted R-squared:  0.1676
## F-statistic: 2.133 on 8 and 37 DF,  p-value: 0.05704

##simplify the model susins stepwise regression

out_simple <- step(out)

## Start:  AIC=-84.86
## data$X1.third.mortality.rate ~ data$PhaseOfFlight + data$Time +
##      data$Place + data$HasFire + data$Environment + data$Energy_absorption +
##      data$Crushed_fuselage + data$Restraint_intact
##
##              Df Sum of Sq    RSS    AIC
## - data$Energy_absorption  1    0.00501  4.9218 -86.808
## - data$Environment        1    0.00888  4.9257 -86.772
## - data$Time                1    0.02675  4.9435 -86.606
## - data$Crushed_fuselage    1    0.08668  5.0035 -86.051
## - data$Place               1    0.13668  5.0535 -85.594
## <none>                    4.9168 -84.855
## - data$HasFire             1    0.30505  5.2218 -84.086
## - data$PhaseOfFlight       1    0.35065  5.2674 -83.686
## - data$Restraint_intact    1    0.48509  5.4019 -82.527
##
## Step:  AIC=-86.81
## data$X1.third.mortality.rate ~ data$PhaseOfFlight + data$Time +
##      data$Place + data$HasFire + data$Environment + data$Crushed_fuselage +
##      data$Restraint_intact
##
##              Df Sum of Sq    RSS    AIC
## - data$Environment        1    0.00855  4.9304 -88.729
## - data$Time                1    0.02934  4.9511 -88.535
## - data$Crushed_fuselage    1    0.08307  5.0049 -88.038
## <none>                    4.9218 -86.808
## - data$Place               1    0.22427  5.1461 -86.759
## - data$HasFire             1    0.30017  5.2220 -86.085
## - data$PhaseOfFlight       1    0.37687  5.2987 -85.414
## - data$Restraint_intact    1    0.49450  5.4163 -84.404
##
## Step:  AIC=-88.73
## data$X1.third.mortality.rate ~ data$PhaseOfFlight + data$Time +
##      data$Place + data$HasFire + data$Crushed_fuselage + data$Restraint_intact
##
##              Df Sum of Sq    RSS    AIC
## - data$Time                1    0.02464  4.9550 -90.499
## - data$Crushed_fuselage    1    0.07675  5.0071 -90.018

```

```

## - data$Place          1    0.21873 5.1491 -88.732
## <none>                  4.9304 -88.729
## - data$HasFire        1    0.31767 5.2480 -87.856
## - data$PhaseOfFlight  1    0.37350 5.3039 -87.369
## - data$Restraint_intact 1    0.49100 5.4214 -86.362
##
## Step: AIC=-90.5
## data$X1.third.mortality.rate ~ data$PhaseOfFlight + data$Place +
##   data$HasFire + data$Crushed_fuselage + data$Restraint_intact
##
##              Df Sum of Sq    RSS    AIC
## - data$Crushed_fuselage 1    0.09032 5.0453 -91.668
## <none>                  4.9550 -90.499
## - data$Place          1    0.22427 5.1793 -90.463
## - data$HasFire        1    0.34331 5.2983 -89.418
## - data$PhaseOfFlight  1    0.40733 5.3623 -88.865
## - data$Restraint_intact 1    0.47156 5.4266 -88.317
##
## Step: AIC=-91.67
## data$X1.third.mortality.rate ~ data$PhaseOfFlight + data$Place +
##   data$HasFire + data$Restraint_intact
##
##              Df Sum of Sq    RSS    AIC
## - data$Place          1    0.20996 5.2553 -91.793
## <none>                  5.0453 -91.668
## - data$PhaseOfFlight  1    0.41043 5.4557 -90.071
## - data$HasFire        1    0.54679 5.5921 -88.935
## - data$Restraint_intact 1    0.73385 5.7792 -87.422
##
## Step: AIC=-91.79
## data$X1.third.mortality.rate ~ data$PhaseOfFlight + data$HasFire +
##   data$Restraint_intact
##
##              Df Sum of Sq    RSS    AIC
## <none>                  5.2553 -91.793
## - data$HasFire        1    0.38822 5.6435 -90.514
## - data$PhaseOfFlight  1    0.44635 5.7016 -90.043
## - data$Restraint_intact 1    0.94626 6.2015 -86.177
summary(out_simple)

##
## Call:
## lm(formula = data$X1.third.mortality.rate ~ data$PhaseOfFlight +
##   data$HasFire + data$Restraint_intact)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.74242 -0.18669  0.06875  0.25758  0.66594
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.7424     0.0781   9.506 4.98e-12 ***
## data$PhaseOfFlighttakeoff -0.2059     0.1090  -1.889  0.06585 .
## data$HasFireeno-fire    -0.2025     0.1149  -1.761  0.08544 .

```

```
## data$Restraint_intact      -0.4028      0.1465  -2.750  0.00875 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3537 on 42 degrees of freedom
## (1 observation deleted due to missingness)
## Multiple R-squared:  0.2685, Adjusted R-squared:  0.2163
## F-statistic: 5.139 on 3 and 42 DF,  p-value: 0.004063

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$HasFire + data$Environment +
##      data$Energy_absorption + data$Crushed_fuselage + data$Restraint_intact)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.64500 -0.23262  0.06522  0.30156  0.56033
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.32793    0.32393   1.012  0.3178
## data$PhaseOfFlighttakeoff -0.02475    0.12161  -0.204  0.8398
## data$Timenight      0.09171    0.13546   0.677  0.5025
## data$Placeoutside    0.02624    0.13675   0.192  0.8488
## data$HasFireno-fire -0.11753    0.14040  -0.837  0.4078
## data$Environmentdangerous -0.03709    0.20188  -0.184  0.8552
## data$Energy_absorptionnogear 0.12627    0.13811   0.914  0.3663
## data$Crushed_fuselage  0.24011    0.26278   0.914  0.3666
## data$Restraint_intact -0.32094    0.17037  -1.884  0.0673 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3733 on 38 degrees of freedom
## Multiple R-squared:  0.2229, Adjusted R-squared:  0.05931
## F-statistic: 1.363 on 8 and 38 DF,  p-value: 0.244

out_2_simple <- step(out_2)

## Start:  AIC=-84.62
## data$X2.third.mortality.rate ~ data$PhaseOfFlight + data$Time +
##      data$Place + data$HasFire + data$Environment + data$Energy_absorption +
##      data$Crushed_fuselage + data$Restraint_intact
##
##              Df Sum of Sq    RSS    AIC
## - data$Environment      1  0.00470 5.2996 -86.578
## - data$Place             1  0.00513 5.3000 -86.575
## - data$PhaseOfFlight     1  0.00577 5.3006 -86.569
## - data$Time              1  0.06387 5.3587 -86.057
## - data$HasFire           1  0.09764 5.3925 -85.761
## - data$Crushed_fuselage  1  0.11633 5.4112 -85.599
## - data$Energy_absorption  1  0.11647 5.4113 -85.598
```

```

## <none> 5.2949 -84.620
## - data$Restraint_intact 1 0.49449 5.7894 -82.424
##
## Step: AIC=-86.58
## data$X2.third.mortality.rate ~ data$PhaseOfFlight + data$Time +
## data$Place + data$HasFire + data$Energy_absorption + data$Crushed_fuselage +
## data$Restraint_intact
##
## Df Sum of Sq RSS AIC
## - data$Place 1 0.00403 5.3036 -88.543
## - data$PhaseOfFlight 1 0.00637 5.3059 -88.522
## - data$Time 1 0.05930 5.3589 -88.055
## - data$HasFire 1 0.10700 5.4066 -87.639
## - data$Crushed_fuselage 1 0.11286 5.4124 -87.588
## - data$Energy_absorption 1 0.12287 5.4224 -87.501
## <none> 5.2996 -86.578
## - data$Restraint_intact 1 0.52252 5.8221 -84.159
##
## Step: AIC=-88.54
## data$X2.third.mortality.rate ~ data$PhaseOfFlight + data$Time +
## data$HasFire + data$Energy_absorption + data$Crushed_fuselage +
## data$Restraint_intact
##
## Df Sum of Sq RSS AIC
## - data$PhaseOfFlight 1 0.00883 5.3124 -90.465
## - data$Time 1 0.05980 5.3634 -90.016
## - data$HasFire 1 0.10298 5.4066 -89.639
## - data$Crushed_fuselage 1 0.11164 5.4152 -89.564
## - data$Energy_absorption 1 0.18847 5.4921 -88.902
## <none> 5.3036 -88.543
## - data$Restraint_intact 1 0.57974 5.8834 -85.667
##
## Step: AIC=-90.46
## data$X2.third.mortality.rate ~ data$Time + data$HasFire + data$Energy_absorption +
## data$Crushed_fuselage + data$Restraint_intact
##
## Df Sum of Sq RSS AIC
## - data$Time 1 0.06693 5.3794 -91.876
## - data$HasFire 1 0.09568 5.4081 -91.626
## - data$Crushed_fuselage 1 0.11097 5.4234 -91.493
## - data$Energy_absorption 1 0.18015 5.4926 -90.897
## <none> 5.3124 -90.465
## - data$Restraint_intact 1 0.58539 5.8978 -87.551
##
## Step: AIC=-91.88
## data$X2.third.mortality.rate ~ data$HasFire + data$Energy_absorption +
## data$Crushed_fuselage + data$Restraint_intact
##
## Df Sum of Sq RSS AIC
## - data$HasFire 1 0.11136 5.4907 -92.913
## - data$Crushed_fuselage 1 0.13348 5.5129 -92.724
## - data$Energy_absorption 1 0.18538 5.5647 -92.284
## <none> 5.3794 -91.876
## - data$Restraint_intact 1 0.56169 5.9411 -89.208

```

```
##
## Step: AIC=-92.91
## data$X2.third.mortality.rate ~ data$Energy_absorption + data$Crushed_fuselage +
## data$Restraint_intact
##
##           Df Sum of Sq    RSS    AIC
## - data$Energy_absorption  1  0.13638 5.6271 -93.760
## <none>                    5.4907 -92.913
## - data$Crushed_fuselage   1  0.28032 5.7710 -92.573
## - data$Restraint_intact   1  0.53410 6.0248 -90.550
##
## Step: AIC=-93.76
## data$X2.third.mortality.rate ~ data$Crushed_fuselage + data$Restraint_intact
##
##           Df Sum of Sq    RSS    AIC
## - data$Crushed_fuselage   1  0.22563 5.8527 -93.912
## <none>                    5.6271 -93.760
## - data$Restraint_intact   1  0.57197 6.1991 -91.210
##
## Step: AIC=-93.91
## data$X2.third.mortality.rate ~ data$Restraint_intact
##
##           Df Sum of Sq    RSS    AIC
## <none>                    5.8527 -93.912
## - data$Restraint_intact   1  0.96097 6.8137 -88.767
```

```
summary(out_2_simple)
```

```
##
## Call:
## lm(formula = data$X2.third.mortality.rate ~ data$Restraint_intact)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.56785 -0.25564  0.03215  0.31435  0.43215
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.56785    0.05775   9.833 8.78e-13 ***
## data$Restraint_intact -0.38048    0.13997  -2.718  0.00929 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3606 on 45 degrees of freedom
## Multiple R-squared:  0.141, Adjusted R-squared:  0.1219
## F-statistic: 7.389 on 1 and 45 DF, p-value: 0.009288
```

```
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$HasFire + data$Environment +
## data$Energy_absorption + data$Crushed_fuselage + data$Restraint_intact)
```

```
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.61335 -0.26873  0.02205  0.25488  0.63675
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.27928    0.32115   0.870   0.390
## data$PhaseOfFlighttakeoff -0.04307    0.12056  -0.357   0.723
## data$Timenight      -0.06231    0.13430  -0.464   0.645
## data$Placeoutside    0.02994    0.13557   0.221   0.826
## data$HasFireno-fire  -0.15616    0.13920  -1.122   0.269
## data$Environmentdangerous  0.10748    0.20015   0.537   0.594
## data$Energy_absorptionnogear 0.08976    0.13693   0.656   0.516
## data$Crushed_fuselage  0.10689    0.26053   0.410   0.684
## data$Restraint_intact  -0.15400    0.16891  -0.912   0.368
##
## Residual standard error: 0.3701 on 38 degrees of freedom
## Multiple R-squared:  0.119, Adjusted R-squared:  -0.06646
## F-statistic: 0.6417 on 8 and 38 DF,  p-value: 0.7378

out_3_simple <- step(out_3)

## Start:  AIC=-85.43
## data$X3.third.mortality.rate ~ data$PhaseOfFlight + data$Time +
##      data$Place + data$HasFire + data$Environment + data$Energy_absorption +
##      data$Crushed_fuselage + data$Restraint_intact
##
##              Df Sum of Sq    RSS    AIC
## - data$Place      1  0.006677  5.2111 -87.370
## - data$PhaseOfFlight  1  0.017477  5.2219 -87.272
## - data$Crushed_fuselage  1  0.023057  5.2275 -87.222
## - data$Time        1  0.029485  5.2339 -87.164
## - data$Environment  1  0.039495  5.2439 -87.075
## - data$Energy_absorption  1  0.058857  5.2633 -86.901
## - data$Restraint_intact  1  0.113848  5.3183 -86.413
## - data$HasFire      1  0.172367  5.3768 -85.899
## <none>                5.2044 -85.430
##
## Step:  AIC=-87.37
## data$X3.third.mortality.rate ~ data$PhaseOfFlight + data$Time +
##      data$HasFire + data$Environment + data$Energy_absorption +
##      data$Crushed_fuselage + data$Restraint_intact
##
##              Df Sum of Sq    RSS    AIC
## - data$Crushed_fuselage  1  0.022037  5.2331 -89.171
## - data$PhaseOfFlight      1  0.022982  5.2341 -89.163
## - data$Time                1  0.029871  5.2410 -89.101
## - data$Environment          1  0.044424  5.2555 -88.971
## - data$Energy_absorption     1  0.104431  5.3155 -88.437
## - data$Restraint_intact      1  0.128672  5.3398 -88.223
## - data$HasFire               1  0.165738  5.3768 -87.898
## <none>                5.2111 -87.370
##
## Step:  AIC=-89.17
```

```

## data$X3.third.mortality.rate ~ data$PhaseOfFlight + data$Time +
##   data$HasFire + data$Environment + data$Energy_absorption +
##   data$Restraint_intact
##
##
##           Df Sum of Sq   RSS   AIC
## - data$PhaseOfFlight      1  0.022625  5.2558 -90.969
## - data$Time                1  0.025610  5.2588 -90.942
## - data$Environment         1  0.050967  5.2841 -90.716
## - data$Energy_absorption    1  0.098403  5.3315 -90.296
## - data$Restraint_intact     1  0.176090  5.4092 -89.616
## <none>                      5.2331 -89.171
## - data$HasFire             1  0.245032  5.4782 -89.021
##
## Step: AIC=-90.97
## data$X3.third.mortality.rate ~ data$Time + data$HasFire + data$Environment +
##   data$Energy_absorption + data$Restraint_intact
##
##
##           Df Sum of Sq   RSS   AIC
## - data$Time                1  0.020026  5.2758 -92.790
## - data$Environment         1  0.048949  5.3047 -92.533
## - data$Energy_absorption    1  0.080509  5.3363 -92.254
## - data$Restraint_intact     1  0.181840  5.4376 -91.370
## - data$HasFire             1  0.225635  5.4814 -90.993
## <none>                      5.2558 -90.969
##
## Step: AIC=-92.79
## data$X3.third.mortality.rate ~ data$HasFire + data$Environment +
##   data$Energy_absorption + data$Restraint_intact
##
##
##           Df Sum of Sq   RSS   AIC
## - data$Environment         1  0.037371  5.3132 -94.458
## - data$Energy_absorption    1  0.078442  5.3542 -94.096
## - data$Restraint_intact     1  0.200180  5.4760 -93.040
## - data$HasFire             1  0.209267  5.4851 -92.962
## <none>                      5.2758 -92.790
##
## Step: AIC=-94.46
## data$X3.third.mortality.rate ~ data$HasFire + data$Energy_absorption +
##   data$Restraint_intact
##
##
##           Df Sum of Sq   RSS   AIC
## - data$Energy_absorption    1  0.07360  5.3868 -95.811
## - data$HasFire             1  0.19113  5.5043 -94.797
## <none>                      5.3132 -94.458
## - data$Restraint_intact     1  0.32988  5.6430 -93.627
##
## Step: AIC=-95.81
## data$X3.third.mortality.rate ~ data$HasFire + data$Restraint_intact
##
##
##           Df Sum of Sq   RSS   AIC
## - data$HasFire             1  0.14501  5.5318 -96.563
## <none>                      5.3868 -95.811
## - data$Restraint_intact     1  0.33883  5.7256 -94.944
##

```

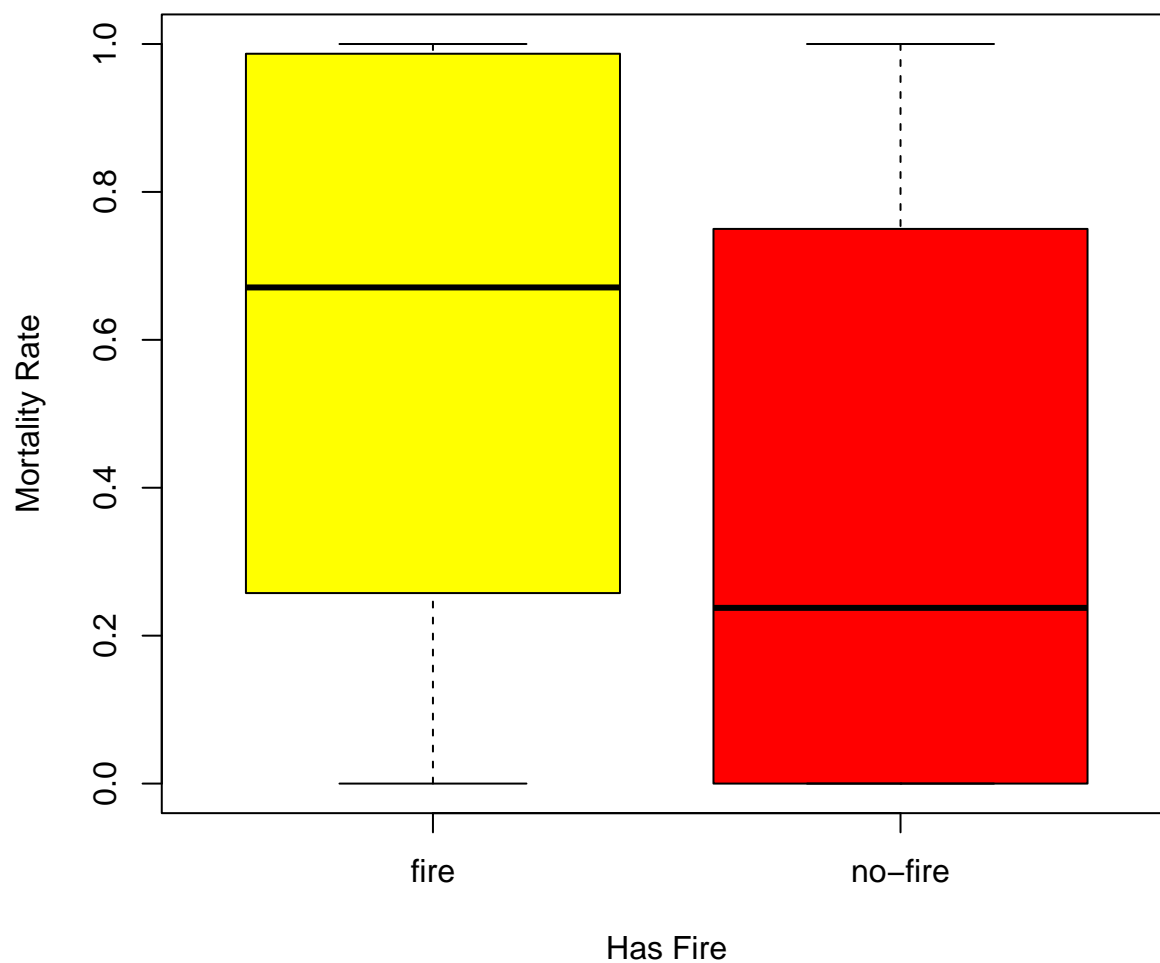
```
## Step: AIC=-96.56
## data$X3.third.mortality.rate ~ data$Restraint_intact
##
##              Df Sum of Sq    RSS    AIC
## <none>                5.5318 -96.563
## - data$Restraint_intact  1   0.37573 5.9075 -95.474

summary(out_3_simple)

##
## Call:
## lm(formula = data$X3.third.mortality.rate ~ data$Restraint_intact)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.46139 -0.25910 -0.03282  0.27811  0.70985
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.46139    0.05614   8.218 1.66e-10 ***
## data$Restraint_intact -0.23791    0.13608  -1.748   0.0872 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3506 on 45 degrees of freedom
## Multiple R-squared:  0.0636, Adjusted R-squared:  0.04279
## F-statistic: 3.056 on 1 and 45 DF,  p-value: 0.08723

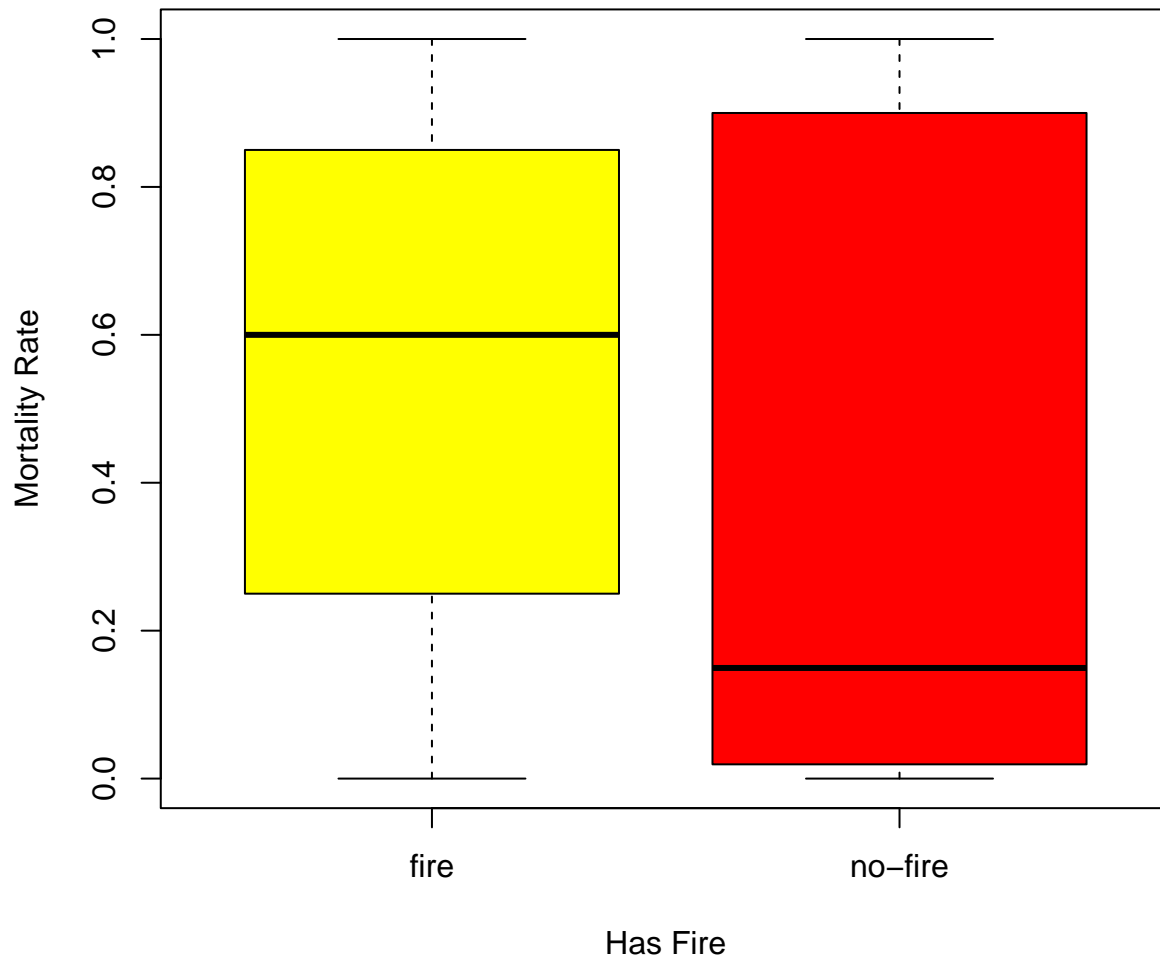
#plot the mortality rates based on HasFire
boxplot(data$X1.third.mortality.rate ~ data$HasFire,
        main = "Front Third Mortality Rate by Fire Presence",
        xlab = "Has Fire",
        ylab = "Mortality Rate",
        col = c("yellow", "red"))
```

Front Third Mortality Rate by Fire Presence



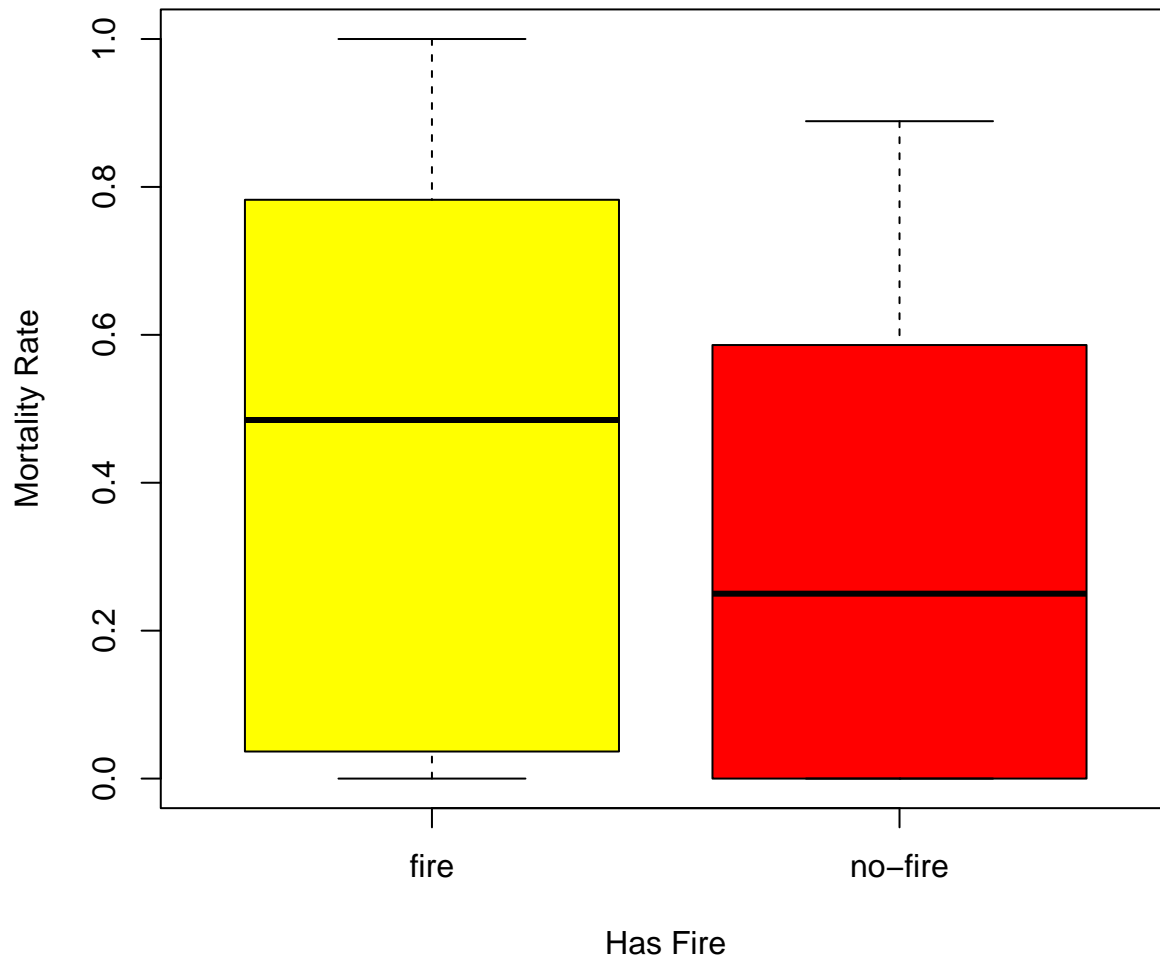
```
boxplot(data$X2.third.mortality.rate ~ data$HasFire,  
  main = "Middle Third Mortality Rate by Fire Presence",  
  xlab = "Has Fire",  
  ylab = "Mortality Rate",  
  col = c("yellow", "red"))
```

Middle Third Mortality Rate by Fire Presence



```
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"))
```

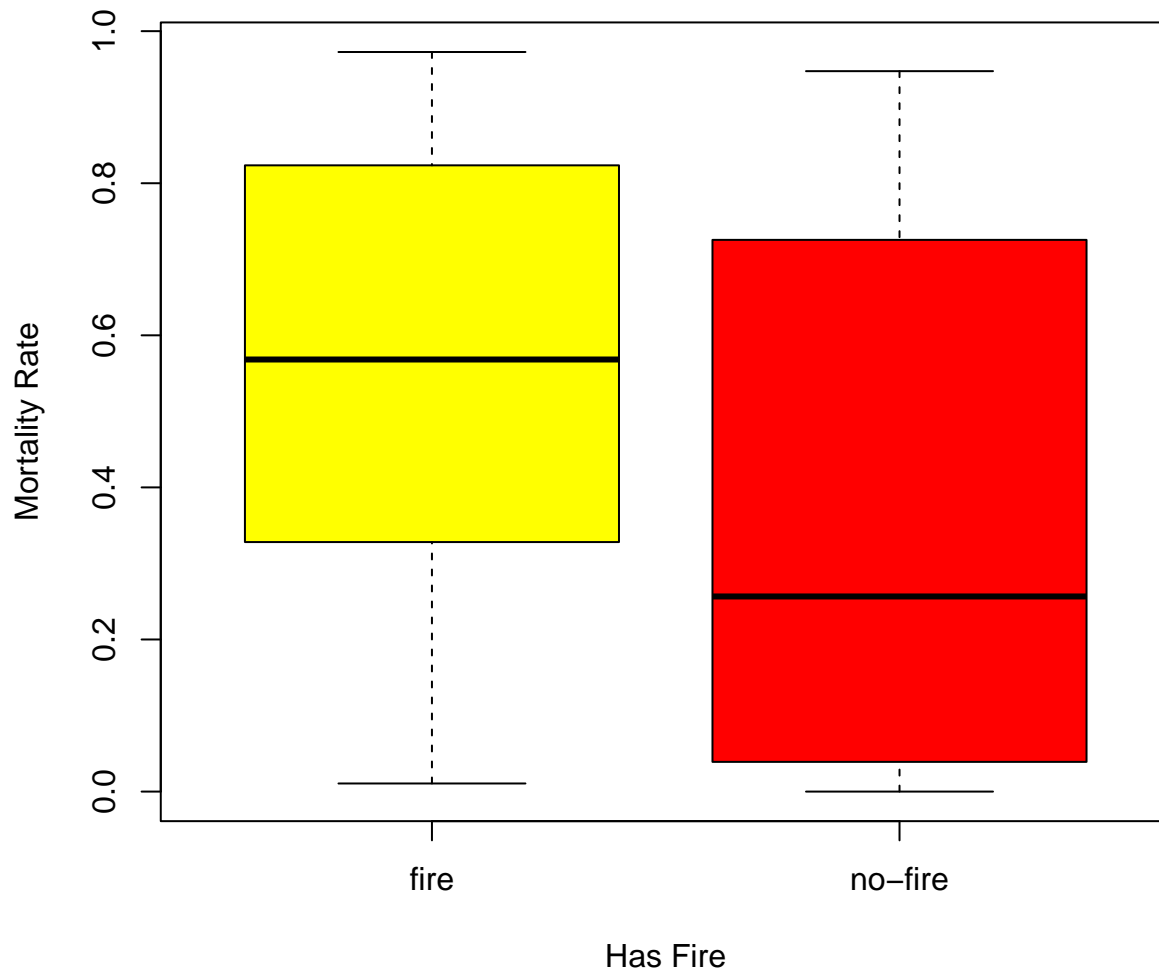
Rear Third Mortality Rate by Fire Presence



```
#and full plane mortality rate
data$Total.mortality.rate <- (data$X1.terzo.morti + data$X2.terzo.morti + data$X3.terzo.morti) / (data$X1.terzo.morti + data$X2.terzo.morti + data$X3.terzo.morti + 1)

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

Total Mortality Rate by Fire Presence



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

```
##
## Call:
## lm(formula = data$X1.half.mortality.rate ~ data$PhaseOfFlight +
##     data$Time + data$Place + data$HasFire + data$Environment +
##     data$Energy_absorption + data$Crushed_fuselage + data$Restraint_intact)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.64724 -0.27567  0.05323  0.24465  0.49986
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.39237    0.30871   1.271   0.211
## data$PhaseOfFlighttakeoff -0.10967    0.11590  -0.946   0.350
## data$Timenight    0.09219    0.12910   0.714   0.480
## data$Placeoutside  0.12520    0.13033   0.961   0.343
## data$HasFireno-fire -0.16634    0.13381  -1.243   0.221
```

```
## data$Environmentdangerous      -0.10597      0.19240     -0.551      0.585
## data$Energy_absorptionnogear    0.11735      0.13162      0.892      0.378
## data$Crushed_fuselage          0.25487      0.25044      1.018      0.315
## data$Restraint_intact           -0.30047      0.16237     -1.851      0.072 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3558 on 38 degrees of freedom
## Multiple R-squared:  0.2789, Adjusted R-squared:  0.1271
## F-statistic: 1.837 on 8 and 38 DF,  p-value: 0.1
```

```
#simplify the model using stepwise regression
out_half_1_simple <- step(out_half_1 , direction = "both", trace = 0)
summary(out_half_1_simple)
```

```
##
## Call:
## lm(formula = data$X1.half.mortality.rate ~ data$HasFire + data$Restraint_intact +
##      data$Place)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.7304 -0.2254  0.0637  0.2457  0.4820
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.55394    0.07845   7.061 1.05e-08 ***
## data$HasFireno-fire -0.21241    0.11879  -1.788  0.0808 .
## data$Restraint_intact -0.30672    0.13964  -2.196  0.0335 *
## data$Placeoutside    0.17643    0.11074   1.593  0.1185
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3486 on 43 degrees of freedom
## Multiple R-squared:  0.2164, Adjusted R-squared:  0.1617
## F-statistic: 3.958 on 3 and 43 DF,  p-value: 0.01408
```

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

```
##
## Call:
## lm(formula = data$X2.half.mortality.rate ~ data$PhaseOfFlight +
##      data$Time + data$Place + data$HasFire + data$Environment +
##      data$Energy_absorption + data$Crushed_fuselage + data$Restraint_intact)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.67138 -0.26653  0.05965  0.25924  0.55600
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.32146    0.31271   1.028  0.310
## data$PhaseOfFlighttakeoff -0.03733    0.11740  -0.318  0.752
## data$Timenight       -0.01337    0.13077  -0.102  0.919
```

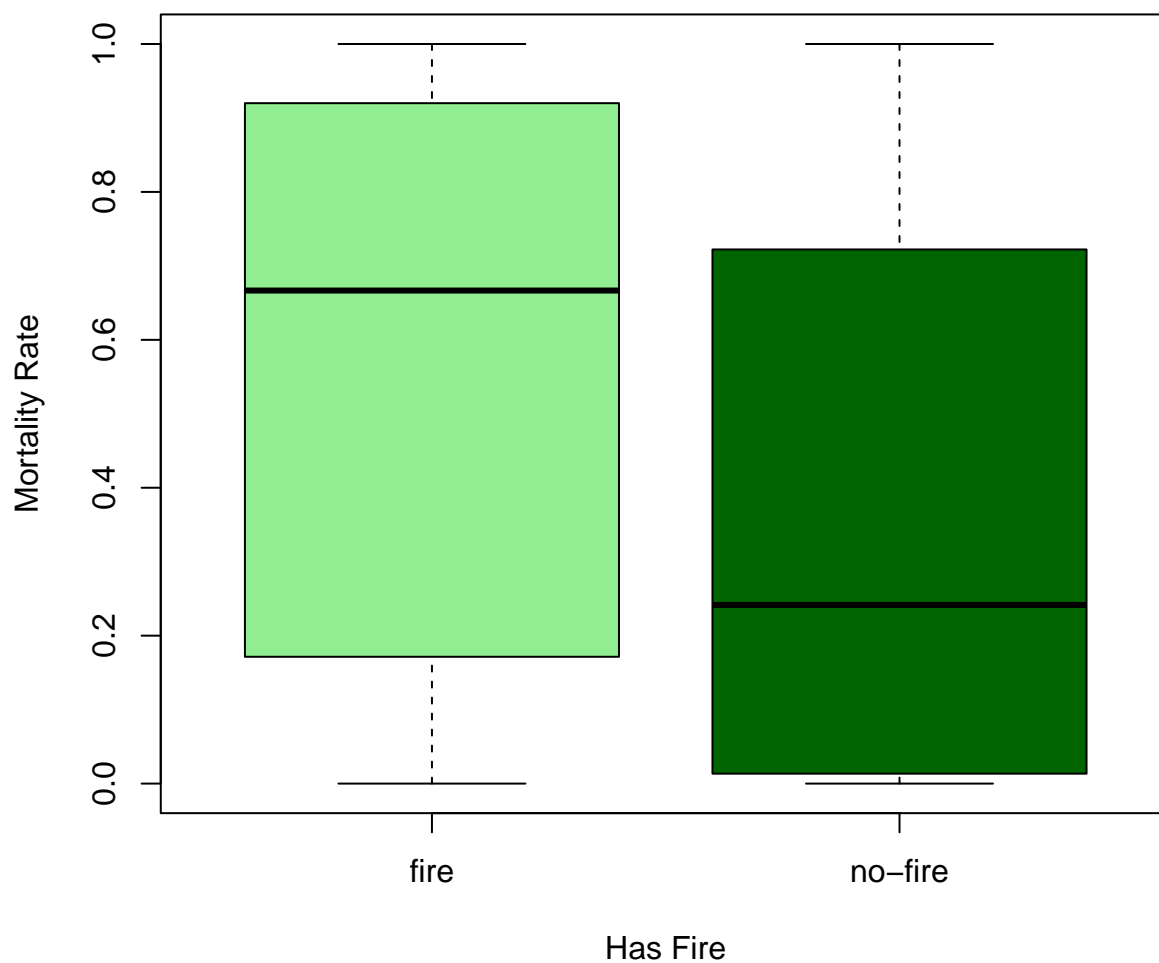
```
## data$Placeoutside          0.01529    0.13201    0.116    0.908
## data$HasFireno-fire       -0.16999    0.13554   -1.254    0.217
## data$Environmentdangerous  0.06064    0.19489    0.311    0.757
## data$Energy_absorptionnogear 0.11617    0.13333    0.871    0.389
## data$Crushed_fuselage     0.15782    0.25368    0.622    0.538
## data$Restraint_intact      -0.20860    0.16447   -1.268    0.212
##
## Residual standard error: 0.3604 on 38 degrees of freedom
## Multiple R-squared:  0.1653, Adjusted R-squared:  -0.01045
## F-statistic: 0.9405 on 8 and 38 DF,  p-value: 0.4954

#simplify the model using stepwise regression
out_half_2_simple <- step(out_half_2 , direction = "both", trace = 0)
summary(out_half_2_simple)

##
## Call:
## lm(formula = data$X2.half.mortality.rate ~ data$HasFire + data$Restraint_intact)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.56445 -0.29203  0.02176  0.30033  0.55661
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.5645     0.0628   8.987 1.62e-11 ***
## data$HasFireno-fire  -0.1545     0.1093  -1.414   0.1644
## data$Restraint_intact -0.2711     0.1330  -2.038   0.0476 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3417 on 44 degrees of freedom
## Multiple R-squared:  0.131, Adjusted R-squared:  0.09146
## F-statistic: 3.315 on 2 and 44 DF,  p-value: 0.04558

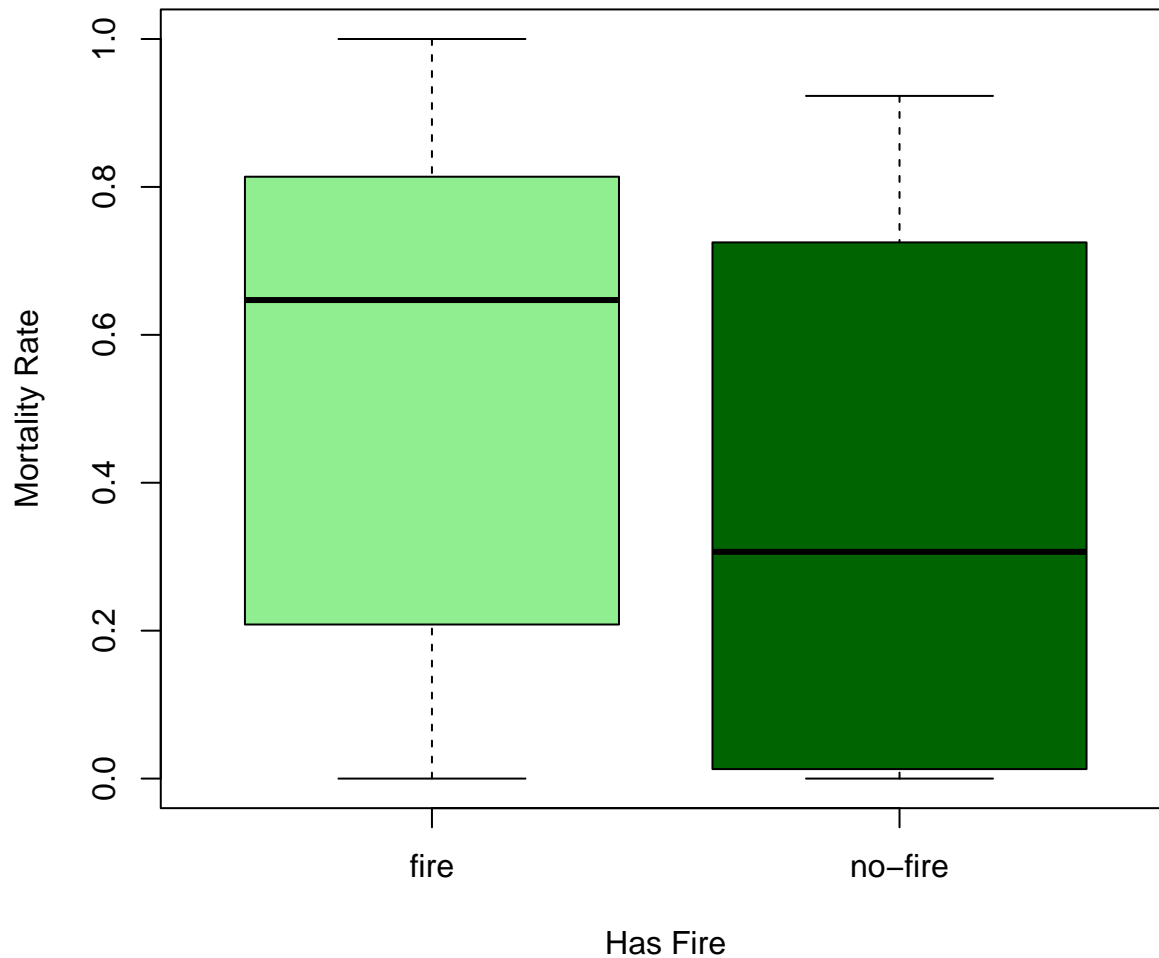
#boxplots for halves
boxplot(data$X1.half.mortality.rate ~ data$HasFire,
        main = "Front Half Mortality Rate by Fire Presence",
        xlab = "Has Fire",
        ylab = "Mortality Rate",
        col = c("lightgreen", "darkgreen"))
```

Front Half Mortality Rate by Fire Presence



```
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"))
```

Rear Half Mortality Rate by Fire Presence



```
#same lm analysis but considering casualties (deaths + serious injuries)
out_casualties <- lm(data$X1.casualties_rate_new ~ data$PhaseOfFlight + data$Time + data$Place + data$HasFire)
summary(out_casualties)
```

```
##
## Call:
## lm(formula = data$X1.casualties_rate_new ~ data$PhaseOfFlight +
##     data$Time + data$Place + data$HasFire + data$Environment +
##     data$Energy_absorption + data$Crushed_fuselage + data$Restraint_intact)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.64202 -0.23185  0.08631  0.22109  0.45624
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.424639   0.286966   1.480   0.1474
## data$PhaseOfFlighttakeoff -0.181421   0.111905  -1.621   0.1135
## data$Timenight      0.126730   0.120387   1.053   0.2993
## data$Placeoutside   0.082089   0.123158   0.667   0.5092
## data$HasFireno-fire -0.098266   0.124580  -0.789   0.4353
```

```
## data$Environmentdangerous      0.024271    0.188631    0.129    0.8983
## data$Energy_absorptionnogear    0.006028    0.131198    0.046    0.9636
## data$Crushed_fuselage          0.286416    0.237821    1.204    0.2361
## data$Restraint_intact           -0.296951    0.155808   -1.906    0.0645 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.33 on 37 degrees of freedom
## (1 observation deleted due to missingness)
## Multiple R-squared:  0.3317, Adjusted R-squared:  0.1872
## F-statistic: 2.296 on 8 and 37 DF,  p-value: 0.04157
```

```
#simplify the model susins stepwise regression
out_casualties_simple <- step(out_casualties, direction = "both", trace = 0)
summary(out_casualties_simple)
```

```
##
## Call:
## lm(formula = data$X1.casualties_rate_new ~ data$PhaseOfFlight +
##     data$Crushed_fuselage + data$Restraint_intact)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.69526 -0.21933  0.05932  0.21384  0.51919
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.41047    0.20915   1.963  0.0563 .
## data$PhaseOfFlighttakeoff -0.19215    0.09792  -1.962  0.0564 .
## data$Crushed_fuselage    0.37569    0.20652   1.819  0.0760 .
## data$Restraint_intact    -0.30535    0.14212  -2.148  0.0375 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3198 on 42 degrees of freedom
## (1 observation deleted due to missingness)
## Multiple R-squared:  0.2873, Adjusted R-squared:  0.2364
## F-statistic: 5.643 on 3 and 42 DF,  p-value: 0.002423
```

```
out_casualties_2 <- lm(data$X2.casualties_rate_new ~ data$PhaseOfFlight + data$Time + data$Place + data$
summary(out_casualties_2)
```

```
##
## Call:
## lm(formula = data$X2.casualties_rate_new ~ data$PhaseOfFlight +
##     data$Time + data$Place + data$HasFire + data$Environment +
##     data$Energy_absorption + data$Crushed_fuselage + data$Restraint_intact)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.69998 -0.17672  0.09558  0.22373  0.39474
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.43487    0.29447   1.477  0.1480
```

```
## data$PhaseOfFlighttakeoff    -0.01900    0.11055   -0.172    0.8645
## data$Timenight               0.17890    0.12314    1.453    0.1545
## data$Placeoutside            0.01615    0.12431    0.130    0.8973
## data$HasFireno-fire         -0.03324    0.12763   -0.260    0.7960
## data$Environmentdangerous    -0.07364    0.18353   -0.401    0.6905
## data$Energy_absorptionnogear  0.09803    0.12555    0.781    0.4398
## data$Crushed_fuselage       0.26302    0.23888    1.101    0.2778
## data$Restraint_intact       -0.38744    0.15488   -2.502    0.0168 *
```

```
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
```

```
## Residual standard error: 0.3393 on 38 degrees of freedom
```

```
## Multiple R-squared:  0.2929, Adjusted R-squared:  0.1441
```

```
## F-statistic: 1.968 on 8 and 38 DF,  p-value: 0.0777
```

```
out_casualties_2_simple <- step(out_casualties_2, direction = "both", trace = 0)
summary(out_casualties_2_simple)
```

```
##
```

```
## Call:
```

```
## lm(formula = data$X2.casualties_rate_new ~ data$Time + data$Restraint_intact)
```

```
##
```

```
## Residuals:
```

```
##      Min       1Q   Median       3Q      Max
## -0.64720 -0.16568  0.07195  0.27434  0.39970
```

```
##
```

```
## Coefficients:
```

```
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.64720    0.05813  11.134 2.19e-14 ***
## data$Timenight      0.19574    0.11202   1.747  0.08755 .
## data$Restraint_intact -0.43152    0.12621  -3.419  0.00136 **
```

```
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
```

```
## Residual standard error: 0.3251 on 44 degrees of freedom
```

```
## Multiple R-squared:  0.2484, Adjusted R-squared:  0.2143
```

```
## F-statistic: 7.272 on 2 and 44 DF,  p-value: 0.001868
```

```
out_casualties_3 <- lm(data$X3.casualties_rate_new ~ data$PhaseOfFlight + data$Time + data$Place + data$Energy_absorption + data$Crushed_fuselage + data$Restraint_intact)
summary(out_casualties_3)
```

```
##
```

```
## Call:
```

```
## lm(formula = data$X3.casualties_rate_new ~ data$PhaseOfFlight +
##      data$Time + data$Place + data$HasFire + data$Environment +
##      data$Energy_absorption + data$Crushed_fuselage + data$Restraint_intact)
```

```
##
```

```
## Residuals:
```

```
##      Min       1Q   Median       3Q      Max
## -0.74307 -0.21754  0.04781  0.28802  0.67549
```

```
##
```

```
## Coefficients:
```

```
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.42067    0.31490   1.336   0.190
## data$PhaseOfFlighttakeoff -0.05061    0.11822  -0.428   0.671
```

```
## data$Timenight          0.01145    0.13168    0.087    0.931
## data$Placeoutside       0.10401    0.13294    0.782    0.439
## data$HasFireno-fire    -0.20774    0.13649   -1.522    0.136
## data$Environmentdangerous 0.04541    0.19626    0.231    0.818
## data$Energy_absorptionnogear 0.04761    0.13426    0.355    0.725
## data$Crushed_fuselage   0.12538    0.25545    0.491    0.626
## data$Restraint_intact   -0.21634    0.16562   -1.306    0.199
##
## Residual standard error: 0.3629 on 38 degrees of freedom
## Multiple R-squared:  0.1886, Adjusted R-squared:  0.0178
## F-statistic: 1.104 on 8 and 38 DF,  p-value: 0.382

out_casualties_3_simple <- step(out_casualties_3, direction = "both", trace = 0)
summary(out_casualties_3_simple)

##
## Call:
## lm(formula = data$X3.casualties_rate_new ~ data$HasFire + data$Restraint_intact)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.63264 -0.29136  0.00915  0.32658  0.65800
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.63264    0.06346   9.970 7.33e-13 ***
## data$HasFireno-fire -0.17444    0.11044  -1.579   0.121
## data$Restraint_intact -0.29065    0.13439  -2.163   0.036 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3452 on 44 degrees of freedom
## Multiple R-squared:  0.1496, Adjusted R-squared:  0.1109
## F-statistic: 3.869 on 2 and 44 DF,  p-value: 0.02831
```

now for the halves

```
out_casualties_half_1 <- lm(data$X1.half.casualties_rate_new ~ data$PhaseOfFlight + data$Time + data$Placeoutside + data$Timenight + data$Energy_absorption + data$Crushed_fuselage + data$Restraint_intact)
summary(out_casualties_half_1)
```

```
##
## Call:
## lm(formula = data$X1.half.casualties_rate_new ~ data$PhaseOfFlight +
##      data$Time + data$Placeoutside + data$HasFire + data$Environment +
##      data$Energy_absorption + data$Crushed_fuselage + data$Restraint_intact)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.68911 -0.21398  0.07866  0.24077  0.43735
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.38679    0.28382   1.363   0.1810
## data$PhaseOfFlighttakeoff -0.11592    0.10655  -1.088   0.2835
## data$Timenight        0.12004    0.11869   1.011   0.3182
## data$Placeoutside      0.05593    0.11982   0.467   0.6433
```

```
## data$HasFireno-fire          -0.12375    0.12302   -1.006    0.3208
## data$Environmentdangerous    -0.05439    0.17689   -0.307    0.7601
## data$Energy_absorptionnogear  0.12646    0.12101    1.045    0.3026
## data$Crushed_fuselage        0.29808    0.23024    1.295    0.2033
## data$Restraint_intact        -0.27608    0.14927   -1.849    0.0722 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3271 on 38 degrees of freedom
## Multiple R-squared:  0.2912, Adjusted R-squared:  0.142
## F-statistic: 1.951 on 8 and 38 DF,  p-value: 0.08022

#simplify the model using stepwise regression
out_casualties_half_1_simple <- step(out_casualties_half_1 , direction = "both", trace = 0)
summary(out_casualties_half_1_simple)

##
## Call:
## lm(formula = data$X1.half.casualties_rate_new ~ data$Crushed_fuselage +
##     data$Restraint_intact)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.66325 -0.17536  0.05897  0.28004  0.33675
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.3073     0.2068   1.486   0.1444
## data$Crushed_fuselage  0.3559     0.2055   1.732   0.0903 .
## data$Restraint_intact -0.2725     0.1337  -2.038   0.0475 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3232 on 44 degrees of freedom
## Multiple R-squared:  0.1983, Adjusted R-squared:  0.1619
## F-statistic: 5.443 on 2 and 44 DF,  p-value: 0.007722

out_casualties_half_2 <- lm(data$X2.half.casualties_rate_new ~ data$PhaseOfFlight + data$Time + data$Place +
summary(out_casualties_half_2)

##
## Call:
## lm(formula = data$X2.half.casualties_rate_new ~ data$PhaseOfFlight +
##     data$Time + data$Place + data$HasFire + data$Environment +
##     data$Energy_absorption + data$Crushed_fuselage + data$Restraint_intact)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.73633 -0.25383  0.00972  0.25974  0.55273
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.452846     0.298985   1.515   0.138
## data$PhaseOfFlighttakeoff -0.043533     0.112245  -0.388   0.700
## data$Timenight        0.004543     0.125031   0.036   0.971
```

```
## data$Placeoutside      -0.054305    0.126219   -0.430    0.669
## data$HasFireno-fire    -0.169190    0.129591   -1.306    0.200
## data$Environmentdangerous -0.036055    0.186340   -0.193    0.848
## data$Energy_absorptionnogear 0.164263    0.127477    1.289    0.205
## data$Crushed_fuselage    0.209584    0.242548    0.864    0.393
## data$Restraint_intact    -0.210573    0.157251   -1.339    0.188
##
## Residual standard error: 0.3445 on 38 degrees of freedom
## Multiple R-squared:  0.187, Adjusted R-squared:  0.01582
## F-statistic: 1.092 on 8 and 38 DF,  p-value: 0.3895

#simplify the model using stepwise regression
out_casualties_half_2_simple <- step(out_casualties_half_2 , direction = "both", trace = 0)
summary(out_casualties_half_2_simple)

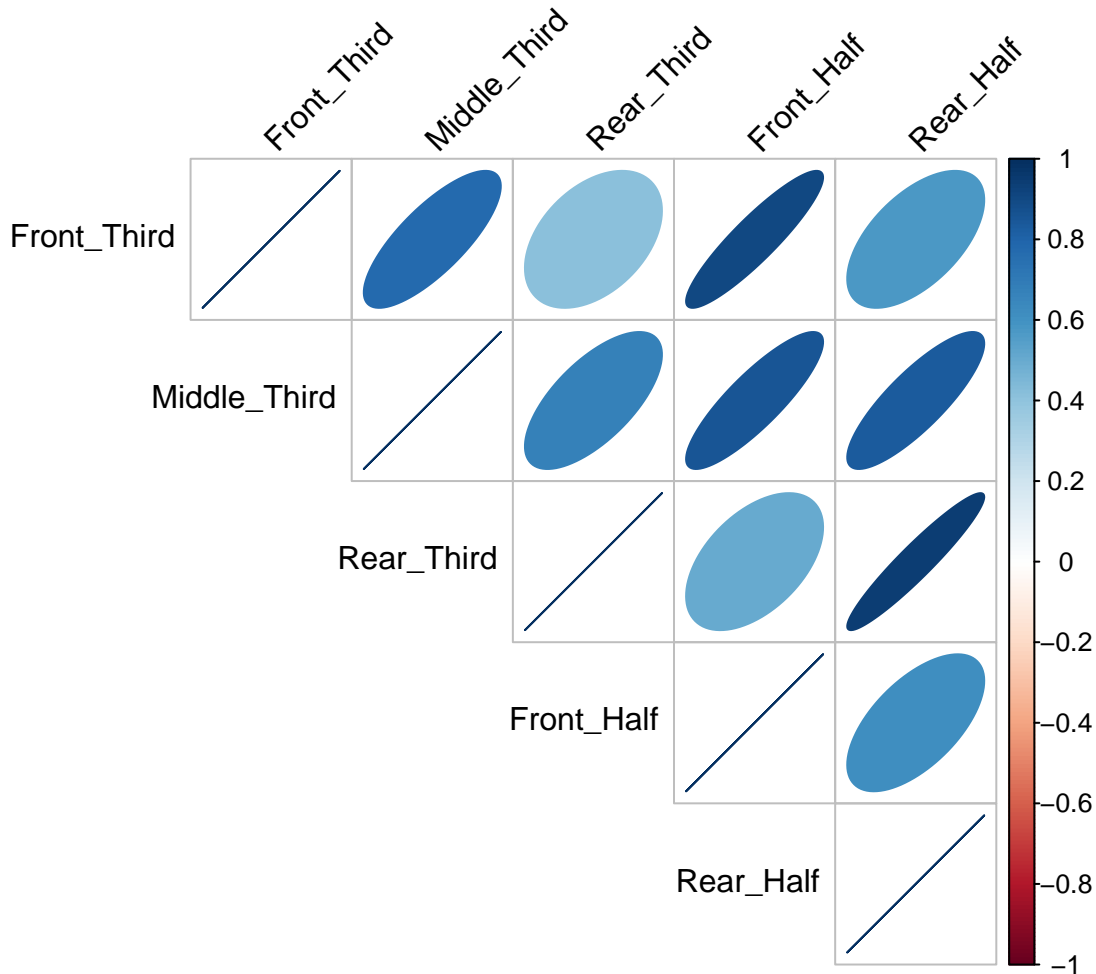
##
## Call:
## lm(formula = data$X2.half.casualties_rate_new ~ data$HasFire +
##     data$Restraint_intact)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.64816 -0.21341  0.03934  0.29064  0.50918
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.6482     0.0607  10.678 8.5e-14 ***
## data$HasFireno-fire  -0.1854     0.1056  -1.755  0.0863 .
## data$Restraint_intact -0.2323     0.1286  -1.807  0.0776 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3303 on 44 degrees of freedom
## Multiple R-squared:  0.135, Adjusted R-squared:  0.09572
## F-statistic: 3.435 on 2 and 44 DF,  p-value: 0.04111

# now we want to study the correlation between the different sections mortality rates
correlation_matrix <- data.frame(
  Front_Third = data$X1.third.mortality.rate,
  Middle_Third = data$X2.third.mortality.rate,
  Rear_Third = data$X3.third.mortality.rate,
  Front_Half = data$X1.half.mortality.rate,
  Rear_Half = data$X2.half.mortality.rate
)
correlation_results <- cor(correlation_matrix, use = "complete.obs")
correlation_results

##           Front_Third Middle_Third Rear_Third Front_Half Rear_Half
## Front_Third    1.0000000    0.7764119  0.4186839  0.9044428  0.5754482
## Middle_Third    0.7764119    1.0000000  0.6707028  0.8548248  0.8309324
## Rear_Third      0.4186839    0.6707028  1.0000000  0.5021073  0.9446734
## Front_Half      0.9044428    0.8548248  0.5021073  1.0000000  0.6126933
## Rear_Half       0.5754482    0.8309324  0.9446734  0.6126933  1.0000000
```

```
## cor plot
library(corrplot)

## corrplot 0.95 loaded
corrplot(correlation_results, method = "ellipse", type = "upper", tl.col = "black", tl.srt = 45)
```



```
## now study the correlation between the other variables and the total mortality rate
correlation_matrix_2 <- data.frame(
  Total_Mortality_Rate = data$Total.mortality.rate,
  PhaseOfFlight = as.numeric(as.factor(data$PhaseOfFlight)),
  Time = as.numeric(as.factor(data$Time)),
  Place = as.numeric(as.factor(data$Place)),
  HasFire = as.numeric(as.factor(data$HasFire)),
  Environment = as.numeric(as.factor(data$Environment)),
  Energy_absorption = as.numeric(as.factor(data$Energy_absorption)),
  Crushed_fuselage = as.numeric(as.factor(data$Crushed_fuselage)),
  Restraint_intact = as.numeric(as.factor(data$Restraint_intact))
)
correlation_results_2 <- cor(correlation_matrix_2, use = "complete.obs")
correlation_results_2
```

```
##           Total_Mortality_Rate PhaseOfFlight      Time      Place
## Total_Mortality_Rate           1.00000000  -0.08666607  0.08261213  0.13102220
```

```

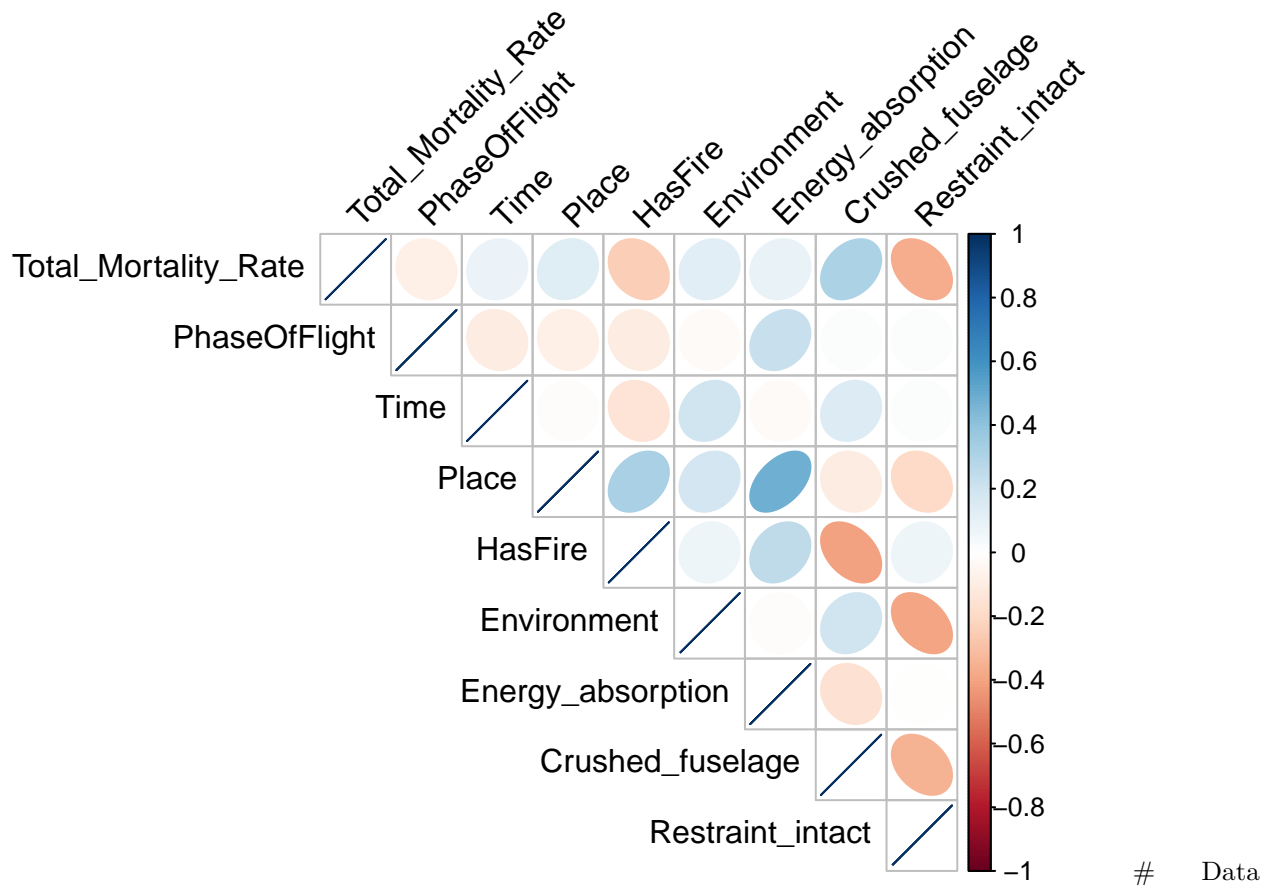
## PhaseOfFlight      -0.08666607    1.00000000 -0.10235879 -0.08496618
## Time               0.08261213   -0.10235879  1.00000000 -0.01499923
## Place              0.13102220   -0.08496618 -0.01499923  1.00000000
## HasFire            -0.24919279   -0.10300633 -0.14027577  0.32137544
## Environment         0.12500715   -0.02750095  0.19072405  0.18537431
## Energy_absorption   0.09710791    0.22675952 -0.02199466  0.48897279
## Crushed_fuselage    0.31917293    0.01541658  0.14433757 -0.10391775
## Restraint_intact    -0.36849986    0.01253452  0.01706972 -0.19794987
##
##               HasFire Environment Energy_absorption Crushed_fuselage
## Total_Mortality_Rate -0.24919279  0.12500715      0.097107907      0.31917293
## PhaseOfFlight        -0.10300633 -0.02750095      0.226759516      0.01541658
## Time                 -0.14027577  0.19072405     -0.021994655      0.14433757
## Place                0.32137544  0.18537431      0.488972793     -0.10391775
## HasFire              1.00000000  0.07384094      0.252502362     -0.40089186
## Environment          0.07384094  1.00000000     -0.012081340      0.19219999
## Energy_absorption     0.25250236 -0.01208134      1.000000000     -0.15238344
## Crushed_fuselage     -0.40089186  0.19219999     -0.152383440      1.00000000
## Restraint_intact      0.07638353 -0.39457942     -0.007433763     -0.34493223
##
##               Restraint_intact
## Total_Mortality_Rate    -0.368499856
## PhaseOfFlight           0.012534517
## Time                    0.017069719
## Place                   -0.197949865
## HasFire                 0.076383528
## Environment             -0.394579419
## Energy_absorption       -0.007433763
## Crushed_fuselage       -0.344932231
## Restraint_intact        1.000000000

```

```

corrplot(correlation_results_2, method = "ellipse", type = "upper", tl.col = "black", tl.srt = 45)

```

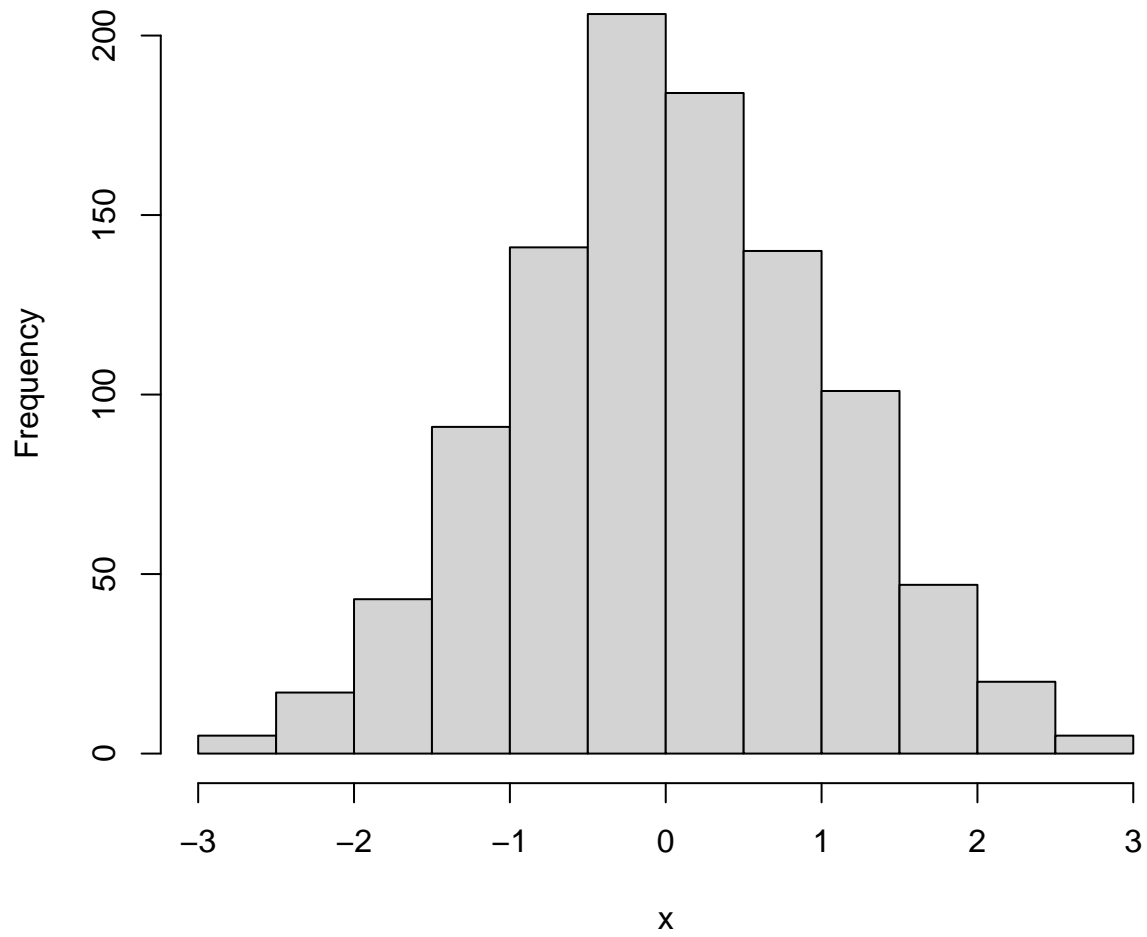


Description

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)
```

Histogram of x



4 Analysis

5 Results

6 Conclusions