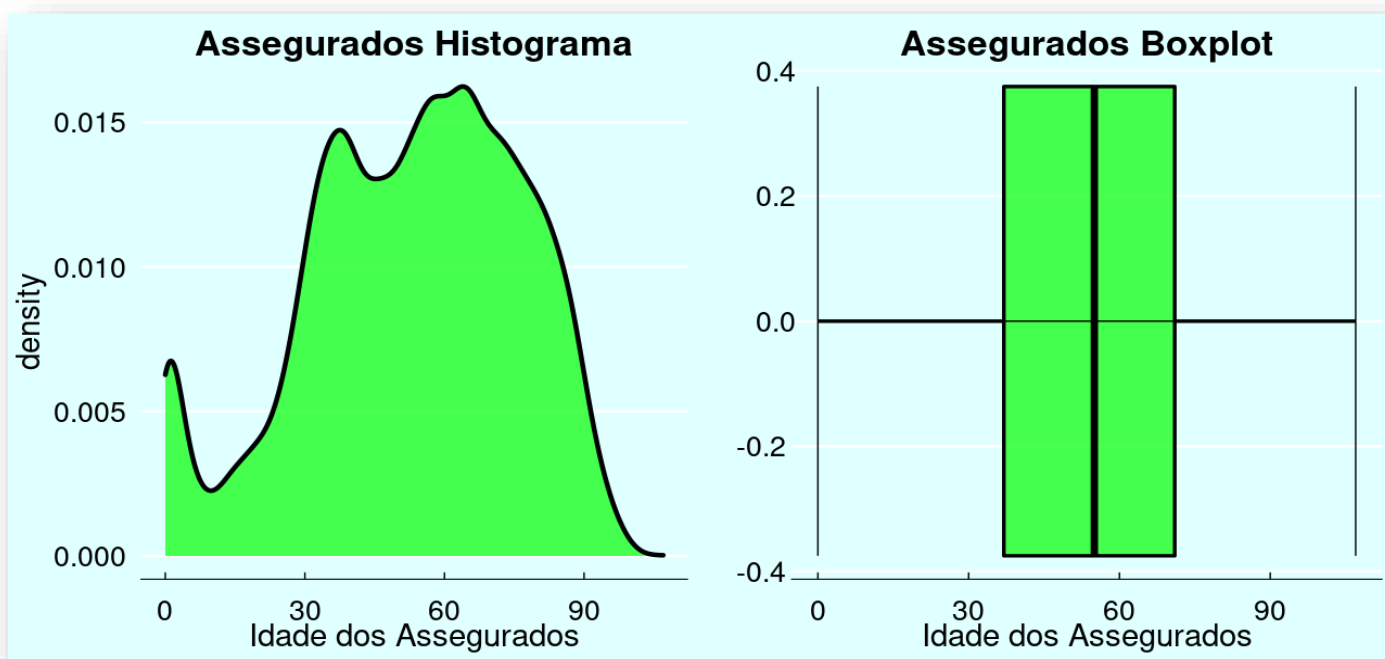


Atividade I – Projeto integrador transdisciplinar em ciência de dados II



```
> summary(idade)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  0.00  37.00   55.00   53.03  71.00  107.00
```

Observações:

O dataset tinha algumas colunas que a princípio achei irrelevante levando em consideração o objetivo da análise inicialmente por isso escolhi as variáveis “idade dos Assegurados” e “Valor Liberado”. Depois de algumas análises utilizando o software Rstudio percebi que existia uma distribuição interessante dos dados na variável “idade dos Assegurados”, assim, a mesma foi levada em consideração para a análise.

Consegui perceber a distribuição das faixa de idades dos pacientes que mais procuram alguma forma de tratamento (Interpretação levando em conta que são pacientes).

Como sugerido plotei um histograma que no qual revelou uma alta de pacientes supondo dos primeiros meses de vida e logo depois vai caindo e começa a subir muito a partir dos 25 anos. Aos 35 anos houve a tendência de queda, assim, encontramos o 1º quartil da análise chegando em 45 e subindo novamente. Chegando em 53 temos a media e 55

a mediana e o ponto mais crítico da análise chegamos a 60. Isso mostra que pessoas com faixa etária de 60 procuram mais tratamentos de saúde. Por fim, a tendência de queda ocorre chegando ao 4º quartil.

Plotei um boxplot dos mesmos dados para melhor identificar a média e demais informações pertinentes.

Não houve encontro de dados discrepantes.

Pensei em colocar um histograma com um tema mais simples, porém achei o theme_economist muito interessante para essa análise.

Segue o Código:

```
# Instalação dos pacotes
install.packages("hrbrthemes")

# Execução dos pacotes
library(hrbrthemes)
library(tidyverse)
library(ggribes)
library(ggthemes)
library(cowplot)
library(viridis)
library(GGally)
library(dplyr)
library(ggplot2)

# Importação dos dados
dados <- read.csv("pacientes.csv", sep = ";")

# Visualização dos dados
View(dados)

# Verificando os tipos de dados
str(dados)

# Verificando os tipos de dados para numéricos
idade = as.numeric(dados$Idade.do.Segurado )
valor = as.numeric(dados$Valor.Total.Liberado)

#Verificando os tipos de dados após transformação
str(idade)
str(valor)

valor

any(is.na(valor)) # verifica se tem qualquer valor nulo

is.na(valor) # mostra todos os valores nulos
```

```

valor2 = valor[is.na(valor)] <- mean(valor, na.rm = TRUE) # Substitui os valores nulos pela
média

any(is.na(valor)) # verifica se tem qualquer valor nulo

View(valor) # Visualizar dataframe

# Verificando o tipo de dados
str(valor)

mean(idade) # média
median(idade) # mediana
range(idade) # limites inferiores e superiores

# Primeiro método de visualização estatística dos dados
Q1 <- quantile(idade, probs = 0.25) # 1º quartil ( = percentil 25%)
Q2 <- quantile(idade, probs = 0.50) # 2º quartil ( = percentil 50% = mediana)
Q3 <- quantile(idade, probs = 0.75) # 3º quartil ( = percentil 75%)

Q1
Q2
Q3

# Informações estatísticas das variáveis selecionadas
summary(idade)
summary(valor)

# Edição da visualização
tema = theme(plot.background=element_rect(fill="#E0FFFF"),
             plot.title = element_text(size=25, hjust=.5, vjust=1),
             axis.title.x = element_text(size=22),
             axis.title.y = element_text(size=22, vjust=2),
             axis.text.x = element_text(size=20),
             axis.text.y = element_text(size=20, hjust=2))

# Ajustes finais de visualização

options(repr.plot.width=14, repr.plot.height=12)

hist1<-ggplot(idade = idade, mapping = aes(x = idade)) +
  geom_density(fill = "green", color = "black", size = 1.6, alpha = .7) +
  theme_economist() +
  xlab("Idade dos Assegurados") +
  ggtitle("Assegurados Histograma") +
  tema

# Criação do boxplot
box1 <- ggplot(idade = idade, mapping = aes(x = idade)) +
  geom_boxplot(fill = "green", color = "black", size = 1.2, alpha = .7) +
  stat_boxplot(geom="errorbar") +

```

```
theme_economist() +  
xlab("Idade dos Assegurados") +  
ggtitle("Assegurados Boxplot") +  
tema  
  
# Graficos  
plot_grid(hist1, box1, nrow=1, ncol=2)
```