

Obitos_confirmados_Covid19_GO.R

Robison_Nunes

2020-08-27

```
# ANÁLISE DESCRITIVA DOS CASOS CONFIRMADOS DE ÓBITOS NO ESTADO DE GOIÁS
```

```
# Fonte: SECRETARIA DE ESTADO DA SAÚDE - GOIÁS
```

```
# https://extranet.saude.go.gov.br/pentaho/api/repos/:coronavirus:paineis:painel.wcdf/generatedContent
```

```
# Data do estudo 24/08/2020
```

```
#Carregando os dados
```

```
dataset <- read.table(file = "C:/Users/robis/OneDrive/DSA/5-Business_analytics/Cap02/R/dados/obitos_confirmed.csv", sep=";", header = TRUE)
```

```
## Warning in scan(file = file, what = what, sep = sep, quote = quote, dec = dec, :  
## EOF within quoted string
```

```
## Warning in scan(file = file, what = what, sep = sep, quote = quote, dec = dec, :  
## número de itens não é múltiplo do número de colunas
```

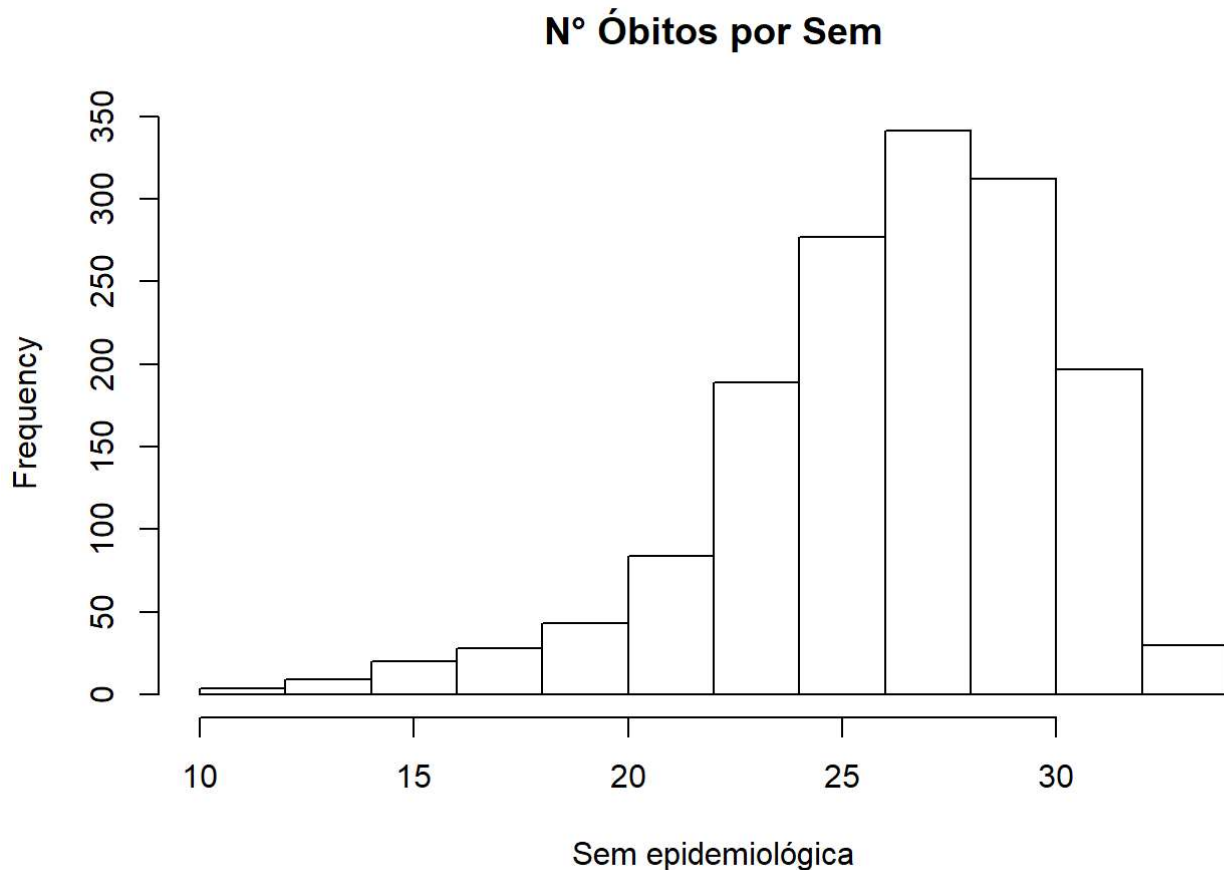
```
#Analisando o tipo das variáveis  
str(dataset)
```

```
## 'data.frame': 1534 obs. of 15 variables:  
## $ i..data_notificacao : int 20200725 20200728 20200729 20200624 20200528 20200803 20200620 202007  
21 20200801 20200807 ...  
## $ data_inicio_sintomas : int 20200715 20200720 20200703 20200618 20200521 20200720 20200613 202007  
13 20200710 20200730 ...  
## $ ano_epi : int 2020 2020 2020 2020 2020 2020 2020 2020 2020 2020 ...  
## $ semana_epi : int 29 30 27 25 21 30 24 29 28 31 ...  
## $ sexo : Factor w/ 2 levels "FEMININO","MASCULINO": 2 2 2 1 1 2 1 1 1 1 ...  
## $ raca_cor : Factor w/ 6 levels "AMARELA","BRANCA",...: 5 5 5 5 2 5 3 2 5 5 ...  
## $ faixa_etaria : Factor w/ 9 levels "< 10 anos", ">= 80 anos",...: 2 8 9 7 2 8 7 7 9 8 ...  
## $ diabetes : Factor w/ 3 levels "IGNORADO","NAO",...: 1 2 3 3 3 1 3 1 3 1 ...  
## $ doenca_cardiovascular: Factor w/ 3 levels "IGNORADO","NAO",...: 3 3 3 3 2 1 3 1 1 1 ...  
## $ doenca_respiratoria : Factor w/ 3 levels "IGNORADO","NAO",...: 1 2 2 2 3 1 2 1 1 1 ...  
## $ imunossupressao : Factor w/ 3 levels "IGNORADO","NAO",...: 1 2 2 2 2 1 2 1 1 1 ...  
## $ codigo_ibge : int 520870 520870 520140 520870 520870 520870 521880 520870 522045 521880  
...  
## $ municipio : Factor w/ 148 levels "ABADIA DE GOIAS",...: 55 55 11 55 55 55 118 55 135 11  
8 ...  
## $ regio_saude : Factor w/ 19 levels "", "CENTRAL", "CENTRO SUL",...: 2 2 3 2 2 2 17 2 3 17  
...  
## $ data_obito : int 20200803 20200807 20200729 20200623 20200606 20200802 20200706 202007  
21 20200803 20200812 ...
```

```
# Análise das variáveis numéricas  
summary(dataset$semana_epi)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    11.00  25.00   27.00   26.68  30.00   34.00
```

```
hist(dataset$semana_epi, main = "Nº Óbitos por Sem", xlab = "Sem epidemiológica")
```



```
#A função summary e o histogram mostram que 50% dos óbitos até a data estudada ficaram
#apartir da Sem 28 ate a 34 com um range de apenas 6 Sem, o range totaol do estudo
# é de 23
```

```
library(lubridate)
```

```
##
## Attaching package: 'lubridate'
```

```
## The following objects are masked from 'package:base':
##
##      date, intersect, setdiff, union
```

```
# Inclusao da coluna "sintomas_óbitos(dias) que o intervalo de dias entre o início dos sintomeas e o óbito.
dataset2 <- dataset
dataset2["sintomas_óbito(dias)"] <- (ymd(dataset2$data_inicio_sintomas) %--% ymd(dataset2$data_obito))/d
days(1))
```

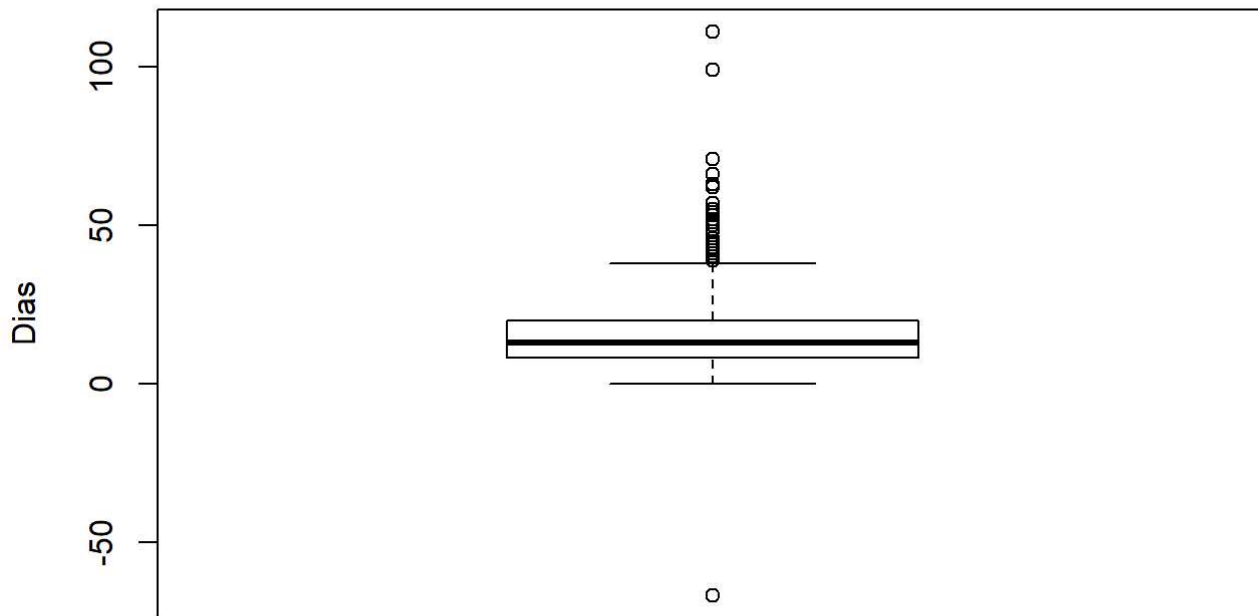
```
## Warning: 3 failed to parse.
```

```
summary(dataset2$sintomas_óbito(dias))
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.   NA's  
## -67.00    8.00   13.00   15.38   20.00   111.00     4
```

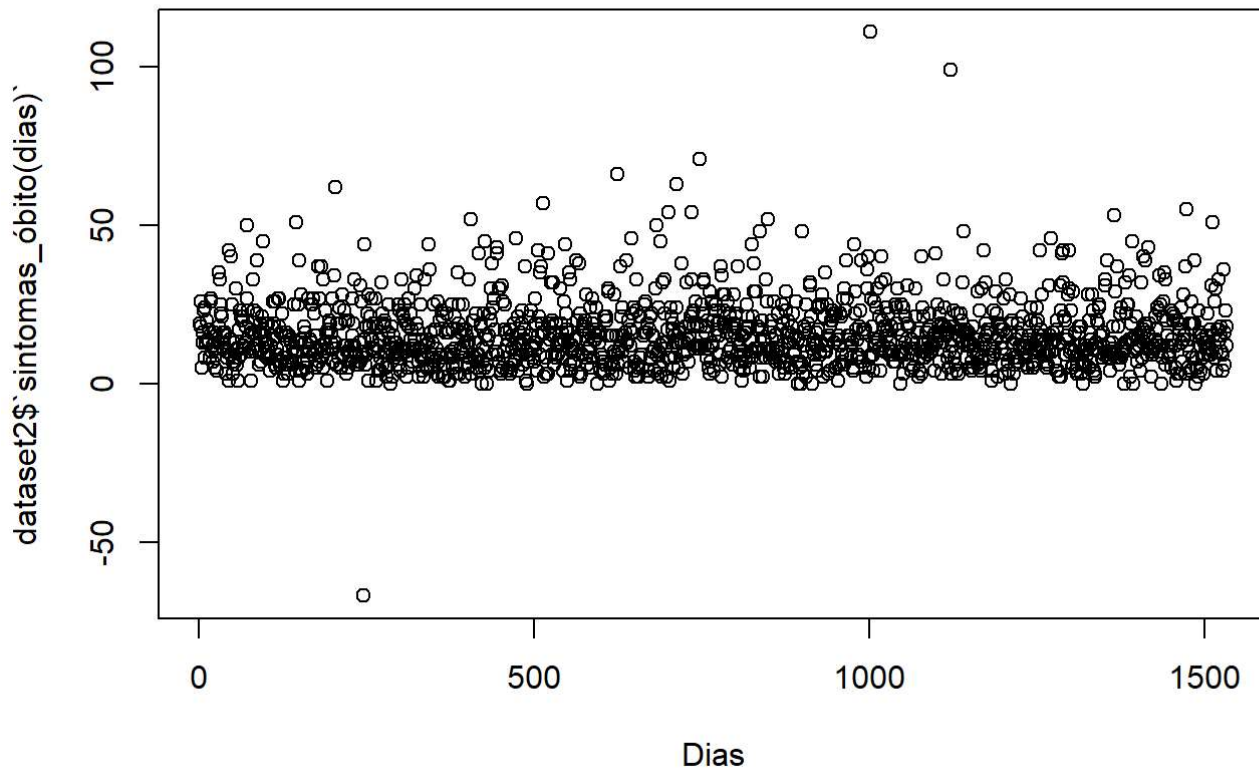
```
boxplot(dataset2$sintomas_óbito(dias), main = "Box Plot - Dias até o óbito", ylab = "Dias")
```

Box Plot - Dias até o óbito



```
plot(dataset2$sintomas_óbito(dias), main = "Scatter Plot - Dias até o óbito", xlab = "Dias")
```

Scatter Plot - Dias até o óbito



```
#Eliminação de outliers ou incorreções nos registros
outliers_correcoes <- c(245,1001,1121,236,442,1534,12)
dataset3 <- dataset2[-outliers_correcoes, ]
summary(dataset3$sintomas_óbito(dias)`)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.00   8.00   13.00   15.31   20.00   71.00
```

```
# Média de 15 dias do aparecimento dos sintomas até o óbito
mean(dataset3$sintomas_óbito(dias)`)
```

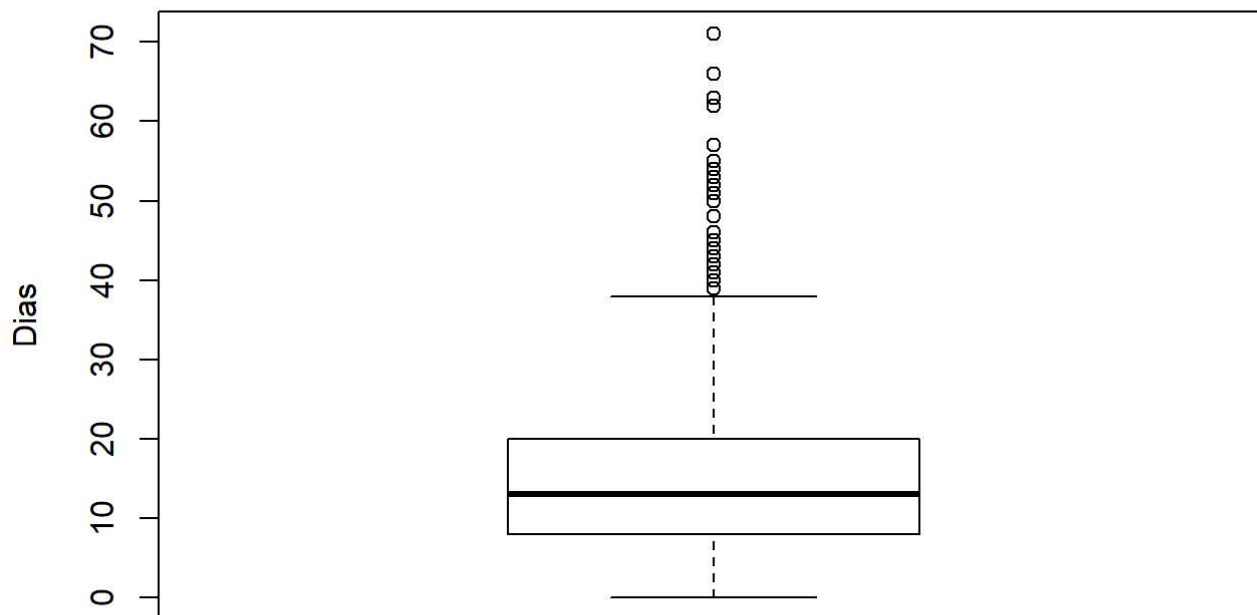
```
## [1] 15.315
```

```
# 75% dos casos estão no range de até 20 dias do aparecimento dos sintomas até o óbito
quantile(dataset3$sintomas_óbito(dias)`)
```

```
##      0%   25%   50%   75%  100%
##       0     8    13    20    71
```

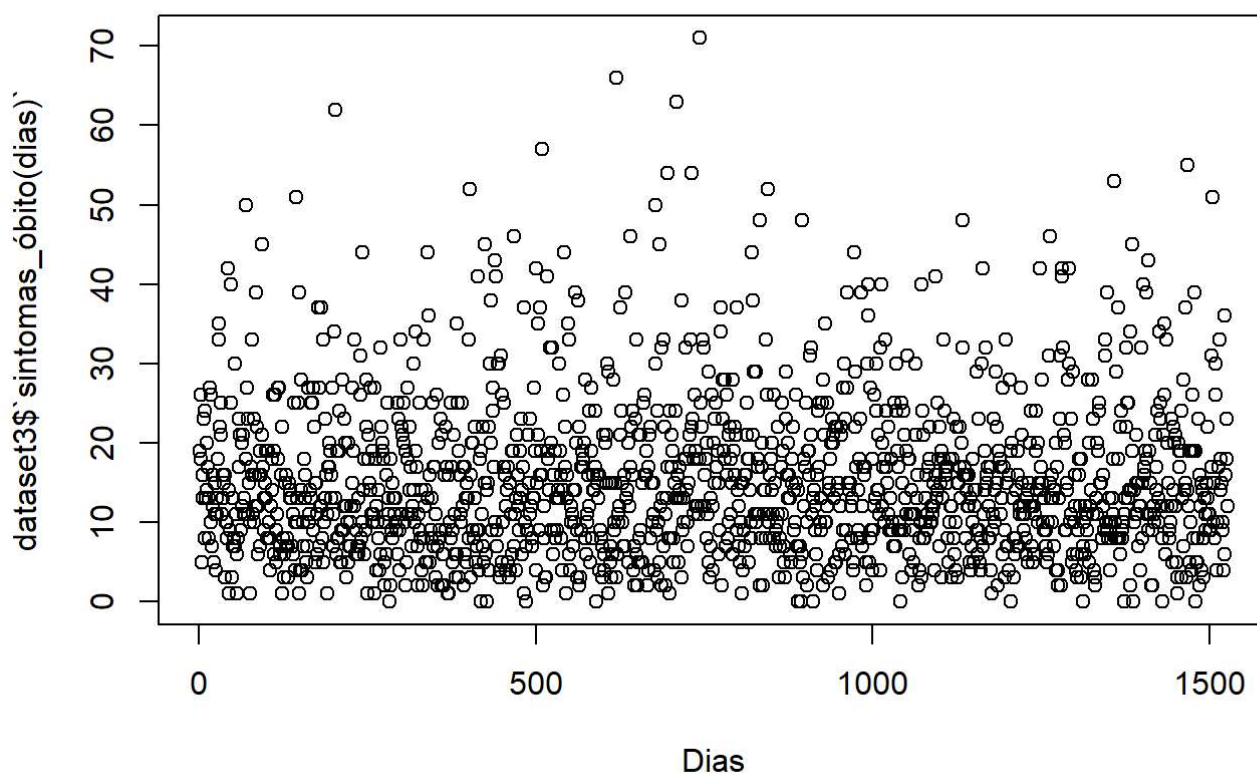
```
#Visualization
boxplot(dataset3$sintomas_óbito(dias)` , main = "Box Plot -Dias do aparecimento dos sintomas até o óbito", ylab = "Dias")
```

Box Plot -Dias do aparecimento dos sintomas até o óbito



```
plot(dataset3$sintomas_óbito(dias), main = "Scatter Plot -Dias do aparecimento dos sintomas até o óbito", xlab = "Dias")
```

Scatter Plot -Dias do aparecimento dos sintomas até o óbito



```
hist(dataset3$sintomas_óbito(dias)`, main = "Dias do aparecimento dos sintomas até o óbito", xlab = "Dias", ylab = "Nº de óbitos", axes=F)
axis(1, at = seq(0,111, by = 3), pos = 0)
axis(2, at = seq(0,1527, by = 10), pos = 0)
```

Análise das variáveis categóricas

#Proporção pro sexo em óbitos confirmados

```
prop_sex <- table(dataset3$sexo)
prop_sex1 <- prop.table(prop_sex)*100
round(prop_sex1, digits = 2)
```

```
##
##  FEMININO MASCULINO
##    36.87    63.13
```

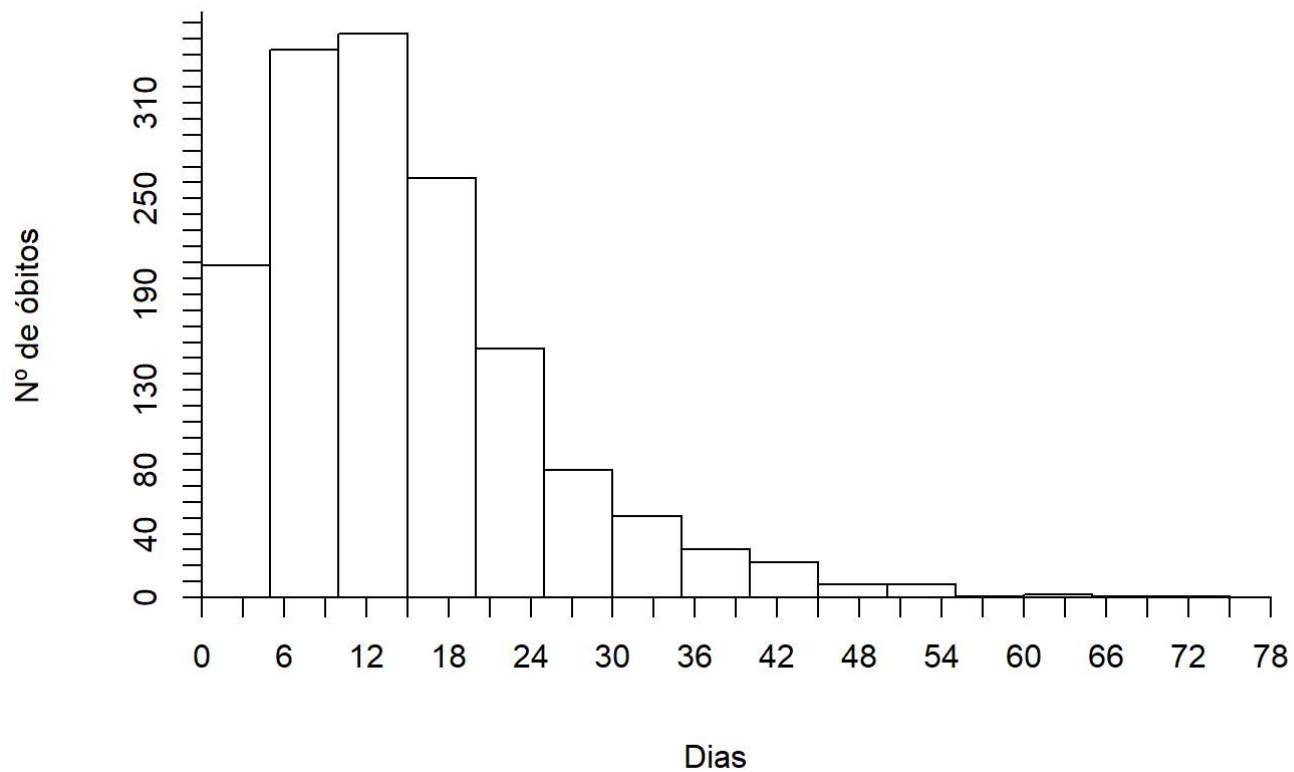
#Proporção pro sexo em óbitos confirmados

```
prop_fxetaria <- table(dataset3$faixa_etaria)
prop_fxetaria1 <- prop.table(prop_fxetaria)*100
round(prop_fxetaria1, digits = 2)
```

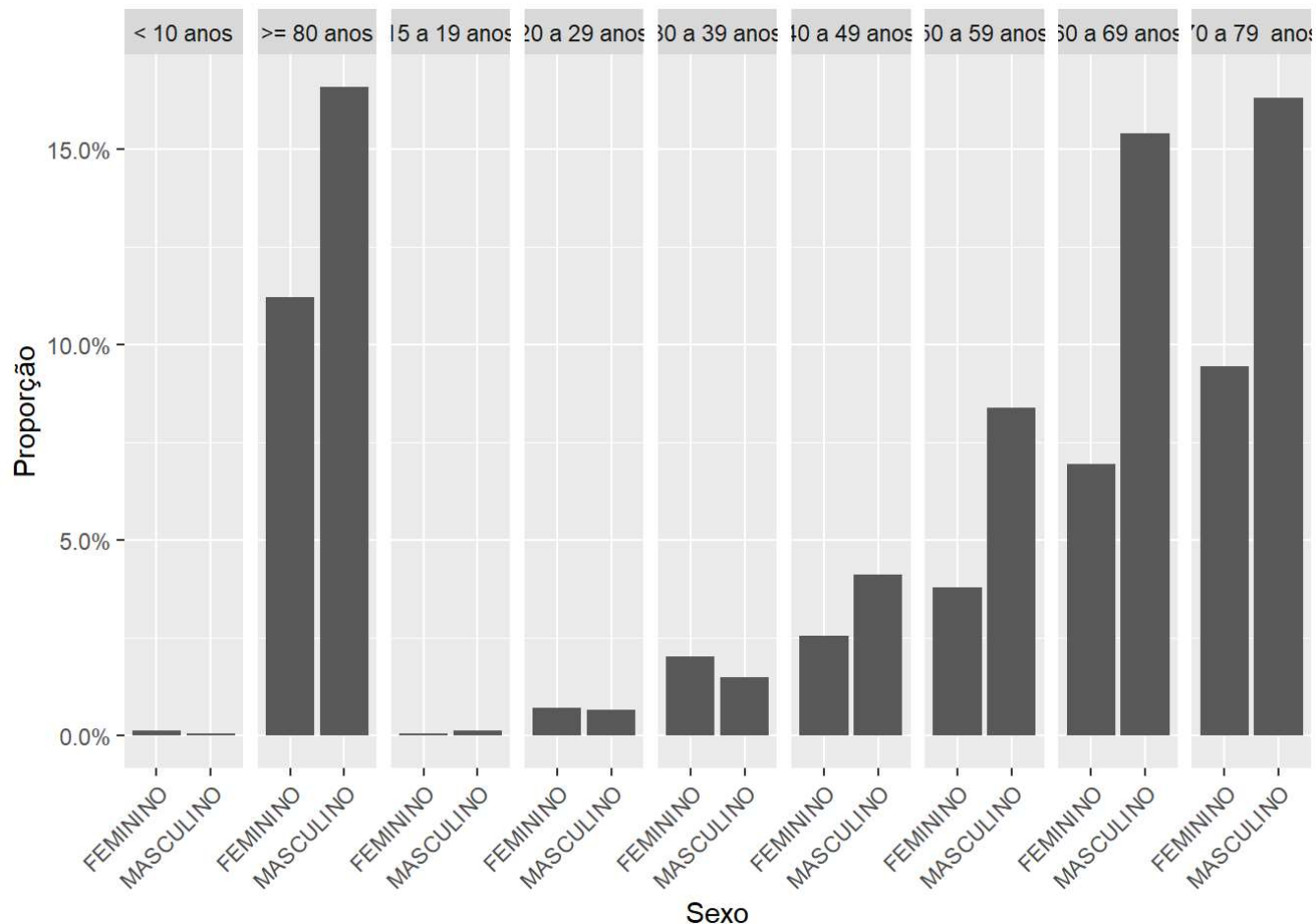
```
##
##    < 10 anos    >= 80 anos  15 a 19 anos  20 a 29 anos  30 a 39 anos
##         0.20         27.77         0.20         1.38         3.54
##  40 a 49 anos  50 a 59 anos  60 a 69 anos  70 a 79 anos
##         6.68         12.18         22.33         25.74
```

```
library(ggplot2)
```

Dias do aparecimento dos sintomas até o óbito



```
#Gráfico óbitos por sexo segmentado por faixa etária  
ggplot(dataset3, aes(x = sexo)) +  
  geom_bar(aes(y = (..count..)/sum(..count..))) +  
  xlab("Sexo") +  
  scale_y_continuous(labels = scales::percent, name = "Proporção") +  
  facet_grid(~ faixa_etaria) +  
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



#Óbitos no sexo feminino sao maiores na faixa etária de 20 a 40 anos

Média de dias dos sintomas até Óbito por apresentação de comorbidade

```
aggregate(dataset3$sintomas_óbito(dias)~ dataset3$diabetes, FUN = mean, data = dataset3)
```

```
## dataset3$diabetes dataset3`sintomas_óbito(dias)`
## 1          IGNORADO          14.93005
## 2             NAO          15.65198
## 3             SIM           15.51577
```

```
aggregate(dataset3`sintomas_óbito(dias)`~ dataset3$doenca_cardiovascular, FUN = mean, data = dataset3)
```

```
## dataset3$doenca_cardiovascular dataset3`sintomas_óbito(dias)`
## 1          IGNORADO          15.30645
## 2             NAO          15.78933
## 3             SIM          14.99060
```

```
aggregate(dataset3`sintomas_óbito(dias)`~ dataset3$doenca_respiratoria, FUN = mean, data = dataset3)
```

```
## dataset3$doenca_respiratoria dataset3`sintomas_óbito(dias)`
## 1          IGNORADO          15.20728
## 2             NAO          15.67804
## 3             SIM          14.10791
```

```
aggregate(dataset3`sintomas_óbito(dias)`~ dataset3$imunossupressao, FUN = mean, data = dataset3)
```



```
## dataset3$imunossupressao dataset3$sintomas_óbito(dias)"
## 1                IGNORADO                14.91565
## 2                NAO                15.74437
## 3                SIM                14.48649
```

total de casos por cidade

```
datasetcidades <- dataset3$municipio
datasetcidades <- as.data.frame(datasetcidades)
names(datasetcidades) <- "Cidade"
datasetcidades["Obitos_por_Covid-19"] <- 1
datasetcidades <- aggregate(datasetcidades$`Obitos_por_Covid-19` ~ datasetcidades$Cidade, FUN = sum, data = datasetcidades)
names(datasetcidades) <- c("Cidade", "Obitos_por_Covid-19")

library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
## filter, lag
```

```
## The following objects are masked from 'package:base':
##
## intersect, setdiff, setequal, union
```

```
datasetcidades1 <- datasetcidades
datasetcidades1["%"] <- datasetcidades1$`Obitos_por_Covid-19` / sum(datasetcidades1$`Obitos_por_Covid-19`)
datasetcidades1["%"] <- round(datasetcidades1$`%`*100, digits = 2)
datasetcidades1 <- datasetcidades1 %>% arrange(desc(datasetcidades1$`Obitos_por_Covid-19`))

print(datasetcidades1)
```

##	Cidade	Obitos_por_Covid-19	%
## 1	GOIANIA	444	29.08
## 2	APARECIDA DE GOIANIA	181	11.85
## 3	RIO VERDE	88	5.76
## 4	ANAPOLIS	78	5.11
## 5	AGUAS LINDAS DE GOIAS	40	2.62
## 6	SENADOR CANEDO	39	2.55
## 7	VALPARAISO DE GOIAS	39	2.55
## 8	LUZIANIA	35	2.29
## 9	TRINDADE	35	2.29
## 10	NOVO GAMA	28	1.83
## 11	FORMOSA	24	1.57
## 12	JATAI	22	1.44
## 13	MINEIROS	22	1.44
## 14	CATALAO	21	1.38
## 15	CIDADE OCIDENTAL	15	0.98
## 16	ITABERAI	14	0.92
## 17	CRISTALINA	13	0.85
## 18	IPORA	13	0.85
## 19	PLANALTINA	13	0.85
## 20	INHUMAS	12	0.79
## 21	ITUMBIARA	12	0.79
## 22	MORRINHOS	12	0.79
## 23	SANTA HELENA DE GOIAS	12	0.79
## 24	SANTO ANTONIO DO DESCOBERTO	12	0.79
## 25	GOIANESIA	11	0.72
## 26	GOIANIRA	10	0.65
## 27	ALEXANIA	9	0.59
## 28	PIRES DO RIO	9	0.59
## 29	CALDAS NOVAS	8	0.52
## 30	GOIATUBA	8	0.52
## 31	NIQUELANDIA	8	0.52
## 32	PADRE BERNARDO	8	0.52
## 33	SAO LUIS DE MONTES BELOS	8	0.52
## 34	ARAGARCAS	6	0.39
## 35	CERES	6	0.39
## 36	PORANGATU	6	0.39
## 37	BELA VISTA DE GOIAS	5	0.33
## 38	HIDROLANDIA	5	0.33
## 39	JARAGUA	5	0.33
## 40	MINACU	5	0.33
## 41	NEROPOLIS	5	0.33
## 42	PARANAIGUARA	5	0.33
## 43	ABADIANIA	4	0.26
## 44	ACREUNA	4	0.26
## 45	ITAPURANGA	4	0.26
## 46	PALMEIRAS DE GOIAS	4	0.26
## 47	QUIRINOPOLIS	4	0.26
## 48	SANTO ANTONIO DE GOIAS	4	0.26
## 49	ANHANGUERA	3	0.20
## 50	ANICUNS	3	0.20
## 51	ARAGUAPAZ	3	0.20
## 52	BONFINOPOLIS	3	0.20
## 53	IPAMERI	3	0.20
## 54	PIRANHAS	3	0.20
## 55	PROFESSOR JAMIL	3	0.20
## 56	RIALMA	3	0.20
## 57	SAO SIMAO	3	0.20
## 58	TURVANIA	3	0.20

## 59	TURVELANDIA	3	0.20
## 60	URUACU	3	0.20
## 61	VICENTINOPOLIS	3	0.20
## 62	ALOANDIA	2	0.13
## 63	AMERICANO DO BRASIL	2	0.13
## 64	ARAGOIANIA	2	0.13
## 65	BOM JESUS DE GOIAS	2	0.13
## 66	CACHOEIRA ALTA	2	0.13
## 67	CEZARINA	2	0.13
## 68	CHAPADAO DO CEU	2	0.13
## 69	COCALZINHO DE GOIAS	2	0.13
## 70	EDEIA	2	0.13
## 71	GOIANAPOLIS	2	0.13
## 72	GUARAITA	2	0.13
## 73	IACIARA	2	0.13
## 74	INACIOLANDIA	2	0.13
## 75	ITAGUARU	2	0.13
## 76	ITAPACI	2	0.13
## 77	ITAPIRAPUA	2	0.13
## 78	JANDAIA	2	0.13
## 79	JOVIANIA	2	0.13
## 80	MONTIVIDIU	2	0.13
## 81	MOZARLANDIA	2	0.13
## 82	NOVA CRIXAS	2	0.13
## 83	NOVA VENEZA	2	0.13
## 84	PARAUNA	2	0.13
## 85	PIRACANJUBA	2	0.13
## 86	SANTO ANTONIO DA BARRA	2	0.13
## 87	TEREZOPOLIS DE GOIAS	2	0.13
## 88	URUANA	2	0.13
## 89	VILA PROPICIO	2	0.13
## 90	ABADIA DE GOIAS	1	0.07
## 91	APARECIDA DO RIO DOCE	1	0.07
## 92	APORE	1	0.07
## 93	AVELINOPOLIS	1	0.07
## 94	BOM JARDIM DE GOIAS	1	0.07
## 95	BONOPOLIS	1	0.07
## 96	BRAZABRANTES	1	0.07
## 97	CABECEIRAS	1	0.07
## 98	CACU	1	0.07
## 99	CALDAZINHA	1	0.07
## 100	CAMPINORTE	1	0.07
## 101	CAMPO LIMPO DE GOIAS	1	0.07
## 102	CAMPOS BELOS	1	0.07
## 103	CATURAI	1	0.07
## 104	COLINAS DO SUL	1	0.07
## 105	CORUMBAIBA	1	0.07
## 106	CRISTIANOPOLIS	1	0.07
## 107	CRIXAS	1	0.07
## 108	CUMARI	1	0.07
## 109	DAMOLANDIA	1	0.07
## 110	DIVINOPOLIS DE GOIAS	1	0.07
## 111	FAZENDA NOVA	1	0.07
## 112	FIRMINOPOLIS	1	0.07
## 113	FLORES DE GOIAS	1	0.07
## 114	GOIANDIRA	1	0.07
## 115	GOIAS	1	0.07
## 116	HIDROLINA	1	0.07
## 117	INDIARA	1	0.07
## 118	ITAGUARI	1	0.07

## 119	JAUPACI	1	0.07
## 120	LEOPOLDO DE BULHOES	1	0.07
## 121	MAIRIPOTABA	1	0.07
## 122	MAURILANDIA	1	0.07
## 123	MOIPORA	1	0.07
## 124	MOSSAMEDES	1	0.07
## 125	NOVA AURORA	1	0.07
## 126	NOVA GLORIA	1	0.07
## 127	NOVA IGUACU DE GOIAS	1	0.07
## 128	ORIZONA	1	0.07
## 129	OUVIDOR	1	0.07
## 130	PALMINOPOLIS	1	0.07
## 131	PANAMA	1	0.07
## 132	PETROLINA DE GOIAS	1	0.07
## 133	PIRENOPOLIS	1	0.07
## 134	POSSE	1	0.07
## 135	RUBIATABA	1	0.07
## 136	SANCLERLANDIA	1	0.07
## 137	SANTA CRUZ DE GOIAS	1	0.07
## 138	SANTA ISABEL	1	0.07
## 139	SANTA RITA DO NOVO DESTINO	1	0.07
## 140	SANTA ROSA DE GOIAS	1	0.07
## 141	SAO LUIZ DO NORTE	1	0.07
## 142	SAO PATRICIO	1	0.07
## 143	SILVANIA	1	0.07
## 144	TAQUARAL DE GOIAS	1	0.07
## 145	TRES RANCHOS	1	0.07
## 146	VIANOPOLIS	1	0.07