Militares por Raça

Marcelo Ribeiro

08/12/2019

## -- Attaching packages --------------------------------------- tidyverse 1.2.1 --

## v ggplot2 3.2.1 v purrr 0.3.2  
## v tibble 2.1.3 v dplyr 0.8.3  
## v tidyr 1.0.0 v stringr 1.4.0  
## v readr 1.3.1 v forcats 0.4.0

## -- Conflicts ------------------------------------------ tidyverse\_conflicts() --  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag() masks stats::lag()

##   
## Attaching package: 'lubridate'

## The following object is masked from 'package:base':  
##   
## date

##   
## Attaching package: 'data.table'

## The following objects are masked from 'package:lubridate':  
##   
## hour, isoweek, mday, minute, month, quarter, second, wday,  
## week, yday, year

## The following objects are masked from 'package:dplyr':  
##   
## between, first, last

## The following object is masked from 'package:purrr':  
##   
## transpose

## Loading required package: carData

##   
## Attaching package: 'car'

## The following object is masked from 'package:dplyr':  
##   
## recode

## The following object is masked from 'package:purrr':  
##   
## some

## Loading required package: sandwich

## Loading required package: gplots

##   
## Attaching package: 'gplots'

## The following object is masked from 'package:stats':  
##   
## lowess

## R Markdown

Leitura dos dados

militares <- read.csv2("militares.csv")

Estatísticas descritivas

## NR\_REF POSTO\_GRADUACAO QUADRO RACA   
## Min. : 1 CB :15527 QPPM :33061 AMARELO : 481   
## 1st Qu.: 9498 3 SGT : 7069 QOPM : 2028 BRANCO :15698   
## Median :18996 SD 1 CL : 6668 QPE : 1119 INDIGENA : 159   
## Mean :18996 2 SGT : 3773 QOC : 757 NAO INF. : 871   
## 3rd Qu.:28494 2 TEN : 1207 QOS : 600 PARDO :17869   
## Max. :37991 1 TEN : 1017 QPEP : 360 PRETO : 2913   
## (Other) : 2730 (Other): 66

Tabelas de contigências e teste qui quadrado para verificar associaçoes entre as variaveis

local({  
 .Table <- with(militares, table(POSTO\_GRADUACAO))  
 cat("\ncounts:\n")  
 print(.Table)  
 cat("\npercentages:\n")  
 print(round(100 \* .Table/sum(.Table), 2))  
})

##   
## counts:  
## POSTO\_GRADUACAO  
## 1 SGT 1 TEN 2 SGT 2 TEN 3 SGT ALUNO ASP A OF CAD   
## 620 1017 3773 1207 7069 59 6 294   
## CAP CB CEL MAJ SD 1 CL SD 2 CL SUB TEN TEN CEL   
## 563 15527 45 386 6668 121 404 232   
##   
## percentages:  
## POSTO\_GRADUACAO  
## 1 SGT 1 TEN 2 SGT 2 TEN 3 SGT ALUNO ASP A OF CAD   
## 1.63 2.68 9.93 3.18 18.61 0.16 0.02 0.77   
## CAP CB CEL MAJ SD 1 CL SD 2 CL SUB TEN TEN CEL   
## 1.48 40.87 0.12 1.02 17.55 0.32 1.06 0.61

local({  
 .Table <- with(militares, table(QUADRO))  
 cat("\ncounts:\n")  
 print(.Table)  
 cat("\npercentages:\n")  
 print(round(100 \* .Table/sum(.Table), 2))  
})

##   
## counts:  
## QUADRO  
## QE QOC QOE QOPM QOS QPE QPEP QPPM QPR   
## 1 757 64 2028 600 1119 360 33061 1   
##   
## percentages:  
## QUADRO  
## QE QOC QOE QOPM QOS QPE QPEP QPPM QPR   
## 0.00 1.99 0.17 5.34 1.58 2.95 0.95 87.02 0.00

local({  
 .Table <- with(militares, table(RACA))  
 cat("\ncounts:\n")  
 print(.Table)  
 cat("\npercentages:\n")  
 print(round(100 \* .Table/sum(.Table), 2))  
})

##   
## counts:  
## RACA  
## AMARELO BRANCO INDIGENA NAO INF. PARDO PRETO   
## 481 15698 159 871 17869 2913   
##   
## percentages:  
## RACA  
## AMARELO BRANCO INDIGENA NAO INF. PARDO PRETO   
## 1.27 41.32 0.42 2.29 47.03 7.67

local({  
 .Table <- xtabs(~POSTO\_GRADUACAO + RACA, data = militares)  
 cat("\nFrequency table:\n")  
 print(.Table)  
 cat("\nTotal percentages:\n")  
 print(totPercents(.Table))  
 .Test <- chisq.test(.Table, correct = FALSE)  
 print(.Test)  
})

##   
## Frequency table:  
## RACA  
## POSTO\_GRADUACAO AMARELO BRANCO INDIGENA NAO INF. PARDO   
## 1 SGT 12 235 1 26 300  
## 1 TEN 13 543 5 20 377  
## 2 SGT 53 1510 14 137 1837  
## 2 TEN 17 599 5 24 506  
## 3 SGT 64 2706 32 341 3260  
## ALUNO 1 29 0 0 27  
## ASP A OF 0 3 0 0 2  
## CAD 2 156 1 0 125  
## CAP 7 322 0 10 195  
## CB 173 6367 79 222 7528  
## CEL 10 13 0 7 15  
## MAJ 10 211 0 14 132  
## SD 1 CL 93 2699 20 33 3230  
## SD 2 CL 1 46 1 0 65  
## SUB TEN 11 143 1 26 186  
## TEN CEL 14 116 0 11 84  
## RACA  
## POSTO\_GRADUACAO PRETO   
## 1 SGT 46  
## 1 TEN 59  
## 2 SGT 222  
## 2 TEN 56  
## 3 SGT 666  
## ALUNO 2  
## ASP A OF 1  
## CAD 10  
## CAP 29  
## CB 1158  
## CEL 0  
## MAJ 19  
## SD 1 CL 593  
## SD 2 CL 8  
## SUB TEN 37  
## TEN CEL 7  
##   
## Total percentages:  
## AMARELO BRANCO INDIGENA NAO INF. PARDO PRETO   
## 1 SGT 0.0 0.6 0.0 0.1 0.8 0.1  
## 1 TEN 0.0 1.4 0.0 0.1 1.0 0.2  
## 2 SGT 0.1 4.0 0.0 0.4 4.8 0.6  
## 2 TEN 0.0 1.6 0.0 0.1 1.3 0.1  
## 3 SGT 0.2 7.1 0.1 0.9 8.6 1.8  
## ALUNO 0.0 0.1 0.0 0.0 0.1 0.0  
## ASP A OF 0.0 0.0 0.0 0.0 0.0 0.0  
## CAD 0.0 0.4 0.0 0.0 0.3 0.0  
## CAP 0.0 0.8 0.0 0.0 0.5 0.1  
## CB 0.5 16.8 0.2 0.6 19.8 3.0  
## CEL 0.0 0.0 0.0 0.0 0.0 0.0  
## MAJ 0.0 0.6 0.0 0.0 0.3 0.1  
## SD 1 CL 0.2 7.1 0.1 0.1 8.5 1.6  
## SD 2 CL 0.0 0.1 0.0 0.0 0.2 0.0  
## SUB TEN 0.0 0.4 0.0 0.1 0.5 0.1  
## TEN CEL 0.0 0.3 0.0 0.0 0.2 0.0  
## Total 1.3 41.3 0.4 2.3 47.0 7.7  
## Total  
## 1 SGT 1.6  
## 1 TEN 2.7  
## 2 SGT 9.9  
## 2 TEN 3.2  
## 3 SGT 18.6  
## ALUNO 0.2  
## ASP A OF 0.0  
## CAD 0.8  
## CAP 1.5  
## CB 40.9  
## CEL 0.1  
## MAJ 1.0  
## SD 1 CL 17.6  
## SD 2 CL 0.3  
## SUB TEN 1.1  
## TEN CEL 0.6  
## Total 100.0

## Warning in chisq.test(.Table, correct = FALSE): Chi-squared approximation  
## may be incorrect

##   
## Pearson's Chi-squared test  
##   
## data: .Table  
## X-squared = 1043.4, df = 75, p-value < 2.2e-16

local({  
 .Table <- xtabs(~QUADRO + RACA, data = militares)  
 cat("\nFrequency table:\n")  
 print(.Table)  
 cat("\nTotal percentages:\n")  
 print(totPercents(.Table))  
 .Test <- chisq.test(.Table, correct = FALSE)  
 print(.Test)  
})

##   
## Frequency table:  
## RACA  
## QUADRO AMARELO BRANCO INDIGENA NAO INF. PARDO PRETO   
## QE 0 1 0 0 0 0  
## QOC 18 314 2 27 340 56  
## QOE 0 21 0 3 33 7  
## QOPM 49 1049 6 43 778 103  
## QOS 4 420 2 13 157 4  
## QPE 10 337 5 30 632 105  
## QPEP 3 188 1 0 155 13  
## QPPM 397 13368 143 754 15774 2625  
## QPR 0 0 0 1 0 0  
##   
## Total percentages:  
## AMARELO BRANCO INDIGENA NAO INF. PARDO PRETO   
## QE 0.0 0.0 0.0 0.0 0.0 0.0  
## QOC 0.0 0.8 0.0 0.1 0.9 0.1  
## QOE 0.0 0.1 0.0 0.0 0.1 0.0  
## QOPM 0.1 2.8 0.0 0.1 2.0 0.3  
## QOS 0.0 1.1 0.0 0.0 0.4 0.0  
## QPE 0.0 0.9 0.0 0.1 1.7 0.3  
## QPEP 0.0 0.5 0.0 0.0 0.4 0.0  
## QPPM 1.0 35.2 0.4 2.0 41.5 6.9  
## QPR 0.0 0.0 0.0 0.0 0.0 0.0  
## Total 1.3 41.3 0.4 2.3 47.0 7.7  
## Total  
## QE 0.0  
## QOC 2.0  
## QOE 0.2  
## QOPM 5.3  
## QOS 1.6  
## QPE 2.9  
## QPEP 0.9  
## QPPM 87.0  
## QPR 0.0  
## Total 100.0

## Warning in chisq.test(.Table, correct = FALSE): Chi-squared approximation  
## may be incorrect

##   
## Pearson's Chi-squared test  
##   
## data: .Table  
## X-squared = 507.38, df = 40, p-value < 2.2e-16

#{r} .df <- data.frame(x = militares$RACA, s = militares$POSTO\_GRADUACAO) .df <- as.data.frame(with(.df, prop.table(table(x, s), margin = 2))) .plot <- ggplot(data = .df, aes(x = x, y = Freq)) + geom\_bar(width = 0.9, stat = "identity") + scale\_y\_continuous(expand = c(0.01, 0), labels = scales::percent\_format()) + facet\_wrap(~s) + xlab("RACA") + ylab("Percent") + theme\_grey(base\_size = 6, base\_family = "sans") + theme(panel.spacing = unit(0.3, "lines")) print(.plot) rm(.df, .plot)

#{r} require("ggplot2") .df <- data.frame(x = militares$QUADRO, z = militares$RACA, s = militares$POSTO\_GRADUACAO) .df <- as.data.frame(with(.df, prop.table(table(x, z, s), margin = 3))) .plot <- ggplot(data = .df, aes(x = x, y = Freq, fill = z)) + geom\_bar(width = 0.9, position = "fill", stat = "identity") + scale\_fill\_brewer(palette = "Blues") + scale\_y\_continuous(expand = c(0.01, 0), labels = scales::percent\_format()) + facet\_wrap(~s) + xlab("QUADRO") + ylab("Percent") + labs(fill = "RACA") + theme\_grey(base\_size = 6, base\_family = "sans") + theme(panel.spacing = unit(0.3, "lines"), legend.position = "right") print(.plot) rm(.df, .plot)