

Prueba Especial Programada 2

Ricardo Ramírez y Rodrigo Vidal

18/07/2019

```
1 source("source/asignar_hospital.R") #Script para asignar hospital
2 nombre_alumno = "Ricardo Ramirez - Rodrigo Vidal"
3 mihospital=asignar_hospital(nombre_alumno)
4 cat(paste("\t - El establecimiento a cargo de",crayon::bold((nombre_alumno)),
5 "\n es el:", crayon::bold((as.character(mihospital$Hospital))),
6 "y \n su código es:",crayon::bold((as.character(mihospital$C.digo)))))

## - El establecimiento a cargo de Ricardo Ramirez - Rodrigo Vidal
## es el: Hospital Clínico de Magallanes Dr. Lautaro Navarro Avaria y
## su código es: 126100

1 library(knitr)
2 library(kableExtra)
3 datos = read.csv("source/BD/hospitales.csv", sep = ",",header = T, encoding = "latin1")
4 kable(head(datos,20))>%
5   kable_styling(full_width = F) %>%
6   column_spec(1, bold = T, border_right = T) %>%
7   column_spec(2, width = "30em")
```

C.digo	Hospital
101100	Hospital Dr. Juan No Crevanni (Arica)
102100	Hospital Dr. Ernesto Torres Galdames (Iquique)
103100	Hospital Dr. Leonardo Guzmán (Antofagasta)
103101	Hospital Dr. Carlos Cisternas (Calama)
104100	Hospital San José del Carmen (Copiapó)
105100	Hospital San Juan de Dios (La Serena)
105101	Hospital San Pablo (Coquimbo)
105102	Hospital Dr. Antonio Tirado Lanas (Ovalle)
106100	Hospital Carlos Van Buren (Valparaíso)
106102	Hospital Dr. Eduardo Pereira Ramírez (Valparaíso)
106103	Hospital Claudio Vicuña (San Antonio)
107100	Hospital Dr. Gustavo Fricke (Viña del Mar)
107101	Hospital San Martín (Quillota)
107102	Hospital de Quilpu
108100	Hospital de San Camilo (San Felipe)
108101	Hospital San Juan de Dios (Los Andes)
109100	Complejo Hospitalario San José (Santiago, Independencia)
109101	Hospital Clínico de Niños Dr. Roberto del Río (Santiago, Independencia)
109102	Instituto Psiquiátrico Dr. José Horwitz Barak (Santiago, Recoleta)
109103	Instituto Nacional del Cáncer Dr. Caupolicán Pardo Correa (Santiago, Recoleta)

Actividad 1

```
1 source("source/asignar_hospital.R") #Script para asignar hospital
2 nombre_alumno = "Ricardo Ramírez - Rodrigo Vidal"
3 mihospital=asignar_hospital(nombre_alumno)
4 cat(paste("\t - El establecimiento a cargo de",crayon::bold((nombre_alumno)),
5 "\n es el:", crayon::bold((as.character(mihospital$Hospital))),
6 "y \n su código es:",crayon::bold((as.character(mihospital$C.digo)))))

## - El establecimiento a cargo de Ricardo Ramírez - Rodrigo Vidal
## es el: Hospital San Jos (Melipilla) y
## su código es: 110150

1 library(knitr)
2 library(kableExtra)
3 # Lista de los hospitales
4 hospitales_general = read.csv("source/BD/hospitales.csv", sep = ",",header = T, encoding = "latin1")
5 #kable(head(hospitales_general, 100))%>%
6 #kable_styling(full_width = T)

1 library(knitr)
2 library(kableExtra)
3 # Datos totales de los hospitales
4 datos_general = read.csv("datos/BD_principal.csv", sep = ",",header = T, encoding = "latin1")
5 #kable(head(datos_general, 20))%>%
6 #kable_styling(full_width = T)
7 #kable_styling(latex_options = "striped", full_width = T) # Aplicación de estilos a la tabla

1 library("dplyr",warn.conflicts = F)
2 filtro_codigo=filter(datos_general, Hospital == mihospital$C.digo)
3 filtro_codigo=filtro_codigo[order(filtro_codigo$DIAS_ESTAD, decreasing = TRUE),]
4 #kable(filtro_codigo)
5 #kable(head(filtro_codigo, 20))%>%
6 #kable_styling(latex_options = "striped")# Aplicación de estilos a la tabla
```

Tabla de frecuencia de los días de Estada de los pacientes en general en el hospital San José de Melipilla

```
1 library(agricolae)
2 tabla_freq_DE=table.freq(hist(filtro_codigo$DIAS_ESTAD,breaks = "Sturges",plot=FALSE))# Tabla de frecu
3 kable(tabla_freq_DE)%>%
4 kable_styling(latex_options = "striped") # Aplicación de estilos a la tabla
```

Lower	Upper	Main	Frequency	Percentage	CF	CPF
0	20	10	6693	96.5	6693	96.5
20	40	30	192	2.8	6885	99.3
40	60	50	31	0.4	6916	99.7
60	80	70	11	0.2	6927	99.9
80	100	90	6	0.1	6933	100.0
100	120	110	0	0.0	6933	100.0
120	140	130	0	0.0	6933	100.0
140	160	150	0	0.0	6933	100.0
160	180	170	0	0.0	6933	100.0
180	200	190	0	0.0	6933	100.0
200	220	210	0	0.0	6933	100.0
220	240	230	0	0.0	6933	100.0
240	260	250	0	0.0	6933	100.0
260	280	270	1	0.0	6934	100.0

Actividad 1 - Estadística inferencial: comparación de medidas de localización

Histograma de servicios vs días de estada

```

1 #ggplot(data=filtro_codigo, aes(x=factor(filtro_codigo$SERC_EGR),y=filtro_codigo$DIAS_ESTAD)) +
2 #geom_bar(stat="identity", position="stack") +
3 #ylab("Días Estada") + xlab("Servicios")

```

tabla de días estada según servicio

```

1 tabla_dias_servicio = aggregate(filtro_codigo$DIAS_ESTAD, by=list(SERVICIO=filtro_codigo$SERC_EGR), FUN
2
3 print(tabla_dias_servicio)

```

```

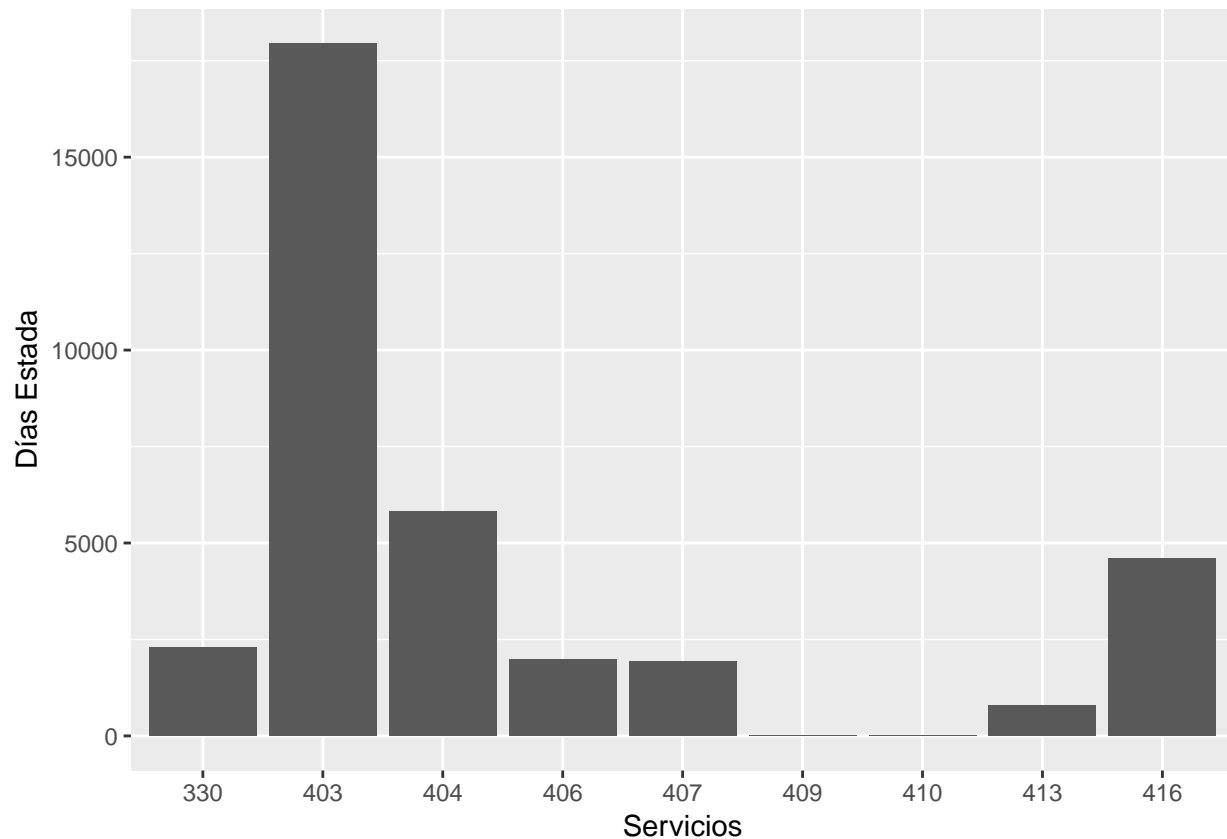
##  SERVICIO      x
## 1      330 2303
## 2      403 17953
## 3      404  5817
## 4      406  1987
## 5      407  1935
## 6      409    12
## 7      410     5
## 8      413   799
## 9      416 4590

```

```

1 library("ggplot2")
2 ggplot(data=tabla_dias_servicio, aes(x=factor(tabla_dias_servicio$SERVICIO),y=tabla_dias_servicio$x)) +
3 geom_bar(stat="identity", position="stack") +
4 ylab("Días Estada") + xlab("Servicios")

```



Pruebas de normalidad

Aplicando un nivel de significancia de 0.05, se plante la hipótesis de que la distribución es normal H_0 .

```

1 library("nortest")
2 tabla_dias_servicio_norm = tabla_dias_servicio
3 mean = mean(tabla_dias_servicio_norm$x)
4 sd = sd(tabla_dias_servicio_norm$x)
5 tabla_dias_servicio_norm$x=rnorm(9, mean = mean, sd = sd)
6
7 cat("Prueba Shapiro-Wilk ", shapiro.test(tabla_dias_servicio_norm$x)$p.value, "\n")

## Prueba Shapiro-Wilk  0.811784

1 cat("Prueba Lillie (Kolmogorov-Smirnov) ", lillie.test(tabla_dias_servicio_norm$x)$p.value, "\n")

## Prueba Lillie (Kolmogorov-Smirnov)  0.6812832

1 cat("Prueba Anderson-Darling ", ad.test(tabla_dias_servicio_norm$x)$p.value, "\n")

## Prueba Anderson-Darling  0.6276877

```

```

1 cat("Prueba Cramer-von Mises ", cvm.test(tabla_dias_servicio_norm$x)$p.value, "\n")

## Prueba Cramer-von Mises  0.5942455

1 cat("Prueba Pearson chi-square ", pearson.test(tabla_dias_servicio_norm$x)$p.value, "\n")

## Prueba Pearson chi-square  0.4594258

1 cat("Prueba Shapiro-Francia ", sf.test(tabla_dias_servicio_norm$x)$p.value, "\n")

## Prueba Shapiro-Francia  0.7255707

```

Según resultados, el p value es inferior al 0.1 por lo que se rechaza la hipótesis H_0 , lo que indica que la distribución NO es normal.

Prueba de Homocedasticidad

```

fligner.test(x = list(tabla_dias_servicio_norm$x, tabla_dias_servicio_norm$SERVICIO))

##
## Fligner-Killeen test of homogeneity of variances
##
## data:  list(tabla_dias_servicio_norm$x, tabla_dias_servicio_norm$SERVICIO)
## Fligner-Killeen:med chi-squared = 8.0735, df = 1, p-value =
## 0.004492

```

Test de Kruskal Walls para prueba de contraste

```

kruskal.test(tabla_dias_servicio_norm$x ~ as.factor(tabla_dias_servicio_norm$SERVICIO), data=tabla_dias_s

##
## Kruskal-Wallis rank sum test
##
## data:  tabla_dias_servicio_norm$x by as.factor(tabla_dias_servicio_norm$SERVICIO)
## Kruskal-Wallis chi-squared = 8, df = 8, p-value = 0.4335

```

Conclusión: Según pruebas aplicadas, los valores del pvalue obtenidos, están por sobre el margen de significancia del 0.05, por lo que se rechaza la hipótesis que indica que la distribución de los datos es normal.

Comparación con los demás hospitales

```

1 #tabla_dias_servicio_general = aggregate(datos_general$DIAS_ESTAD, by=list(SERVICIO=datos_general$SERC_
2 tabla_dias_servicio_general = aggregate(datos_general$DIAS_ESTAD, by=list(SERVICIO=datos_general$SERC_E
3
4 filtro_servicios_otros_hospitales=filter(datos_general, datos_general$Hospital != mihospital$C.digo & (
5 tabla_servicios_otros_hospitales_agr = aggregate(filtro_servicios_otros_hospitales$DIAS_ESTAD, by=list(
6
7 filtro_servicios_mi_hospital=filter(datos_general,datos_general$Hospital == mihospital$C.digo & (datos_
8 tabla_servicios_mi_hospitale_agr = aggregate(filtro_servicios_mi_hospital$DIAS_ESTAD, by=list(SERVICIO=
9
10
11 tabla_dias_servicio_general_desc = tabla_dias_servicio_general[order(tabla_dias_servicio_general$x, dec
12
13 #kable(tabla_dias_servicio_general_desc)
14 kable(head(tabla_dias_servicio_general_desc, 20))%>%
15 kable_styling(latex_options = "striped")# Aplicación de estilos a la tabla

```

	SERVICIO	x
53	403	1656268
51	401	655107
54	404	594514
66	416	500622
52	402	360826
50	330	221115
56	406	194548
67	418	169246
57	407	159171
59	409	130596
26	212	114217
65	415	108729
69	420	95004
55	405	91793
3	110	91596
25	211	91215
9	120	75644
63	413	75438
60	410	74860
58	408	72746

```

1 print(tabla_servicios_otros_hospitales_agr)

```

```

##      SERVICIO      x
## 1      403 1638315
## 2      404  588697
## 3      416  496032

```

```

1 print(tabla_servicios_mi_hospitale_agr)

```

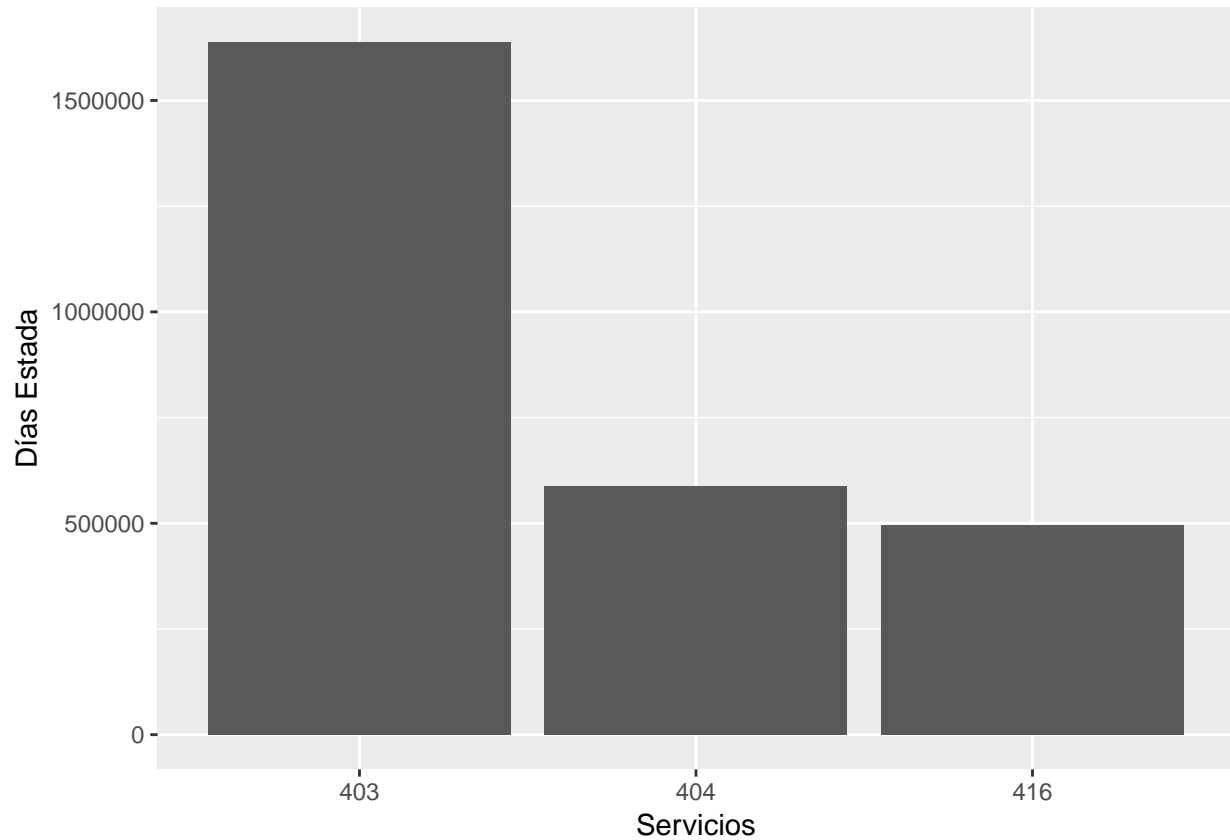
```

##      SERVICIO      x
## 1      403 17953
## 2      404  5817
## 3      416  4590

```

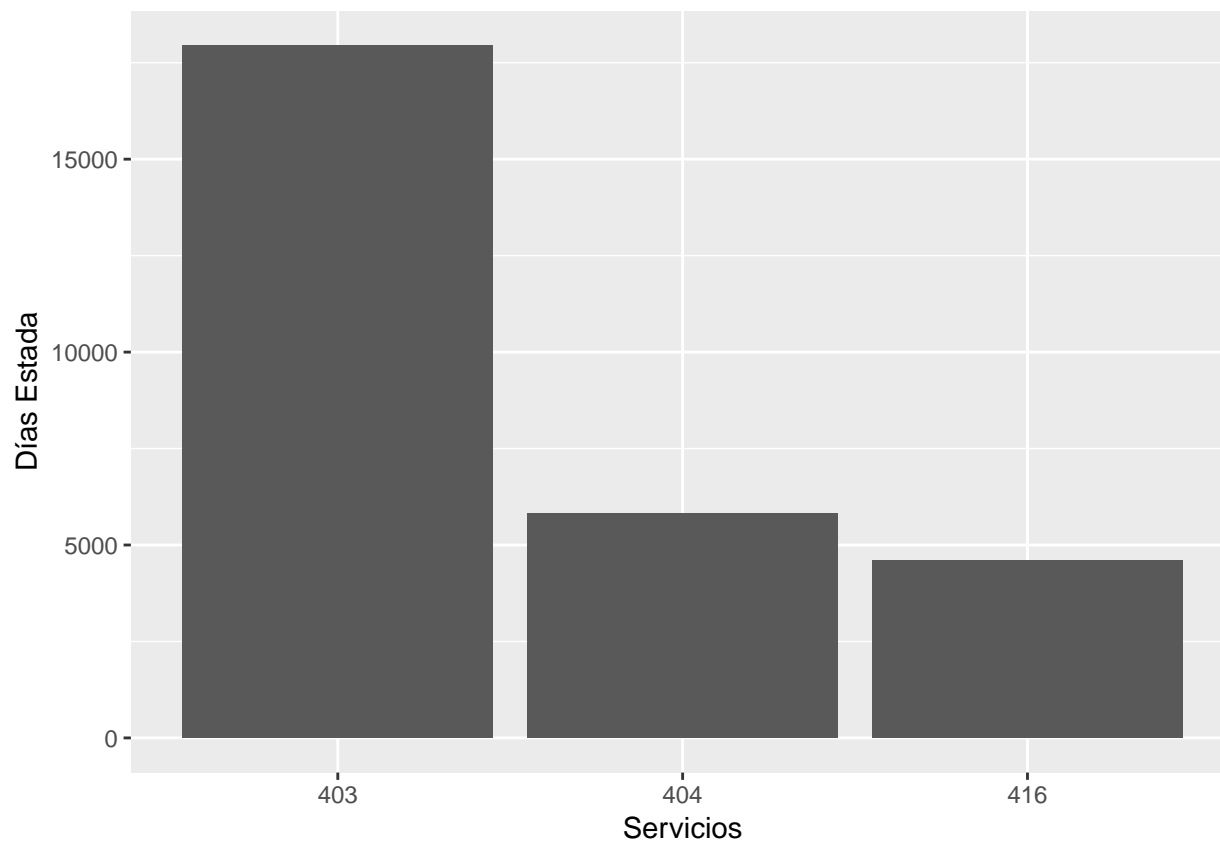
Histograma general servicios 403, 404, 416

```
ggplot(data=tabla_servicios_otros_hospitales Agr, aes(x=factor(tabla_servicios_otros_hospitales Agr$SERV), y=tabla_servicios_otros_hospitales Agr$Días Estada)) +  
geom_bar(stat="identity", position="stack") +  
ylab("Días Estada") + xlab("Servicios")
```



Histograma mi hospital servicios 403, 404, 416

```
ggplot(data=tabla_servicios_mi_hospitale Agr, aes(x=factor(tabla_servicios_mi_hospitale Agr$SERVICIO), y=tabla_servicios_mi_hospitale Agr$Días Estada)) +  
geom_bar(stat="identity", position="stack") +  
ylab("Días Estada") + xlab("Servicios")
```



Actividad 2

Lista de servicios prestado solo por el hospital asignado

```

1 library("dplyr",warn.conflicts = F)
2 principal = read.csv("datos/BD_principal.csv", sep = ",",header = T, encoding = "latin1")
3 filtro1=filter(principal, Hospital == mihospital$C.digo)
4 kable(head(filtro1[,c(1,11,15,17)], 20))%>%
5   kable_styling(full_width = F) %>%
6   kable_styling(latex_options = "striped") # Aplicación de estilos a la tabla

```


Hospital	FECHA_EGR	DIAS_ESTAD	INTERV_Q
110150	10/02/2017	4	1
110150	10/02/2017	12	2
110150	04/02/2017	4	2
110150	04/02/2017	1	2
110150	04/02/2017	3	2
110150	24/05/2017	2	1
110150	25/05/2017	6	2
110150	24/05/2017	4	1
110150	24/05/2017	2	1
110150	25/05/2017	2	2
110150	05/02/2017	3	2
110150	05/02/2017	2	2
110150	06/02/2017	10	2
110150	06/02/2017	3	2
110150	06/02/2017	7	2
110150	06/02/2017	17	2
110150	06/02/2017	9	2
110150	06/02/2017	6	2
110150	10/02/2017	7	2
110150	06/02/2017	4	2

INTERV_Q: Intervención Quirúrgica, los valores aceptados son: 1 = Sí 2 = No ## Rango de fecha

```
1 rango_fecha = range(as.Date(filtro1$FECHA_EGR, format = "%d/%m/%Y"))
2 print(rango_fecha)
```

```
## [1] "2017-01-01" "2017-12-31"
```

Total de intervenciones

```
1 filtro_intervencion = filter(filtro1, INTERV_Q == 1)
2 filtro_intervencion2 = filter(filtro1, INTERV_Q == 2)
3 cat("De un total de ", count(filtro1)$n, "atenciones, se logró llevar a cabo un total ", count(filtro_intervencion2)$n, "intervenciones quirúrgicas.")
```

```
## De un total de 6934 atenciones, se logró llevar a cabo un total 3352 de intervenciones
## quirúrgicas, mientras que 3582 no acabaron en una intervención.
```

tabla de intervenciones por mes durante el año 2017

```
1 library(kableExtra)
2 tabla_por_mes = aggregate(filtro_intervencion$INTERV_Q, by=list(MESES= months(as.Date(filtro_intervencion$FECHA_EGR, format = "%d/%m/%Y"))), FUN=function(x) sum(x))
3 tabla_por_mes2 = aggregate(filtro_intervencion2$INTERV_Q, by=list(MESES= months(as.Date(filtro_intervencion2$FECHA_EGR, format = "%d/%m/%Y"))), FUN=function(x) sum(x))
4
5 kable(head(tabla_por_mes, 12))%>%
6   kable_styling(full_width = F) %>%
7   kable_styling(latex_options = "striped") # Aplicación de estilos a la tabla
```

MESES	x
abril	304
agosto	264
diciembre	261
enero	292
febrero	256
julio	228
junio	282
marzo	296
mayo	303
noviembre	273
octubre	287
septiembre	306

```

1 kable(head(tabla_por_mes2, 12))%>%
2   kable_styling(full_width = F) %>%
3   kable_styling(latex_options = "striped") # Aplicación de estilos a la tabla

```

MESES	x
abril	478
agosto	712
diciembre	636
enero	544
febrero	518
julio	656
junio	628
marzo	566
mayo	550
noviembre	578
octubre	630
septiembre	668

```

1 library("ggplot2")
2 ggplot(data=tabla_por_mes, aes(x=factor(tabla_por_mes$MESES),y=tabla_por_mes$x)) +
3 geom_bar(stat="identity", position="stack") +
4 ylab("Total de intervenciones") + xlab("Meses")

```

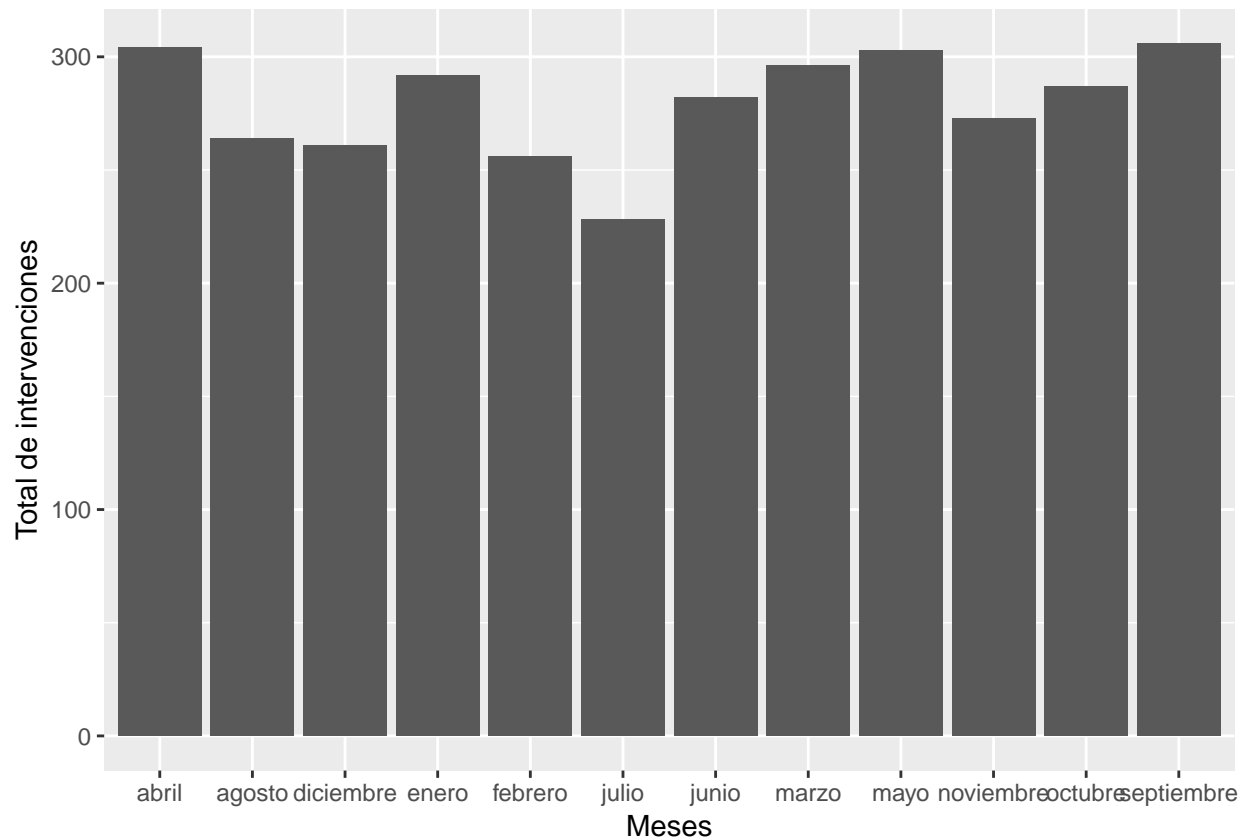


Tabla de contingencia

```

1 suma_total = sum(tabla_por_mes$x)
2 suma_total2 = count(filtro_intervencion2)$n
3
4 suma_meses_vacaciones = sum(tabla_por_mes$x[4]+tabla_por_mes$x[5]+tabla_por_mes$x[8])
5 suma_meses_vacaciones2 = sum(tabla_por_mes2$x[4]+tabla_por_mes2$x[5]+tabla_por_mes2$x[8])
6
7 suma_resto_del_anio = sum(tabla_por_mes$x)-suma_meses_vacaciones
8 suma_resto_del_anio2 = count(filtro_intervencion2)$n-suma_meses_vacaciones2
9
10 cat("De un total de ",count(filtro1)$n,"prestaciones realizadas por el hospital, solo" ,suma_total," fu

```

```

## De un total de 6934 prestaciones realizadas por el hospital, solo 3352 fueron intervenciones y
## 3582 no llegaron a una. Durante los meses de vacaciones(Enero y Febrero) se llevaron a
## cabo 844 intervesiones mientras que el 1628 prestaciones en la misma fecha no
## llegaron a una, mientras que el resto del años se lograron 2508 intervencionesm, mientras que
## 1954 prestaciones en ese mismo tiempo no llegaron a una

```

```

1 #Datos
2 filas=c("Intervension","no Intervension")
3 vacaciones = c(suma_meses_vacaciones,suma_meses_vacaciones2)
4 resto= c(suma_resto_del_anio,suma_resto_del_anio2)

```

5
6
7
8
9
10
11
12

1
2

1
2

1
2

1

```
#Creación de tabla
tabla=data.matrix(data.frame(vacaciones,resto))
rownames(tabla)=filas

#Prueba de Chi cuadrado
resultado=chisq.test(tabla)
print(resultado)

##
## Pearson's Chi-squared test with Yates' continuity correction
##
## data:  tabla
## X-squared = 309.26, df = 1, p-value < 2.2e-16

#Frecuencia observada
resultado$observed

##              vacaciones resto
## Intervension           844 2508
## no Intervension        1628 1954

#Frecuencia esperada
resultado$expected

##              vacaciones    resto
## Intervension      1195.002 2156.998
## no Intervension   1276.998 2305.002

#Residuos
resultado$residuals

##              vacaciones    resto
## Intervension      -10.15372  7.557615
## no Intervension    9.82233 -7.310952

cat("De lo anterior podemos concluir que se rechaza H0(los médicos no se coordinan y la mayor parte se va")

## De lo anterior podemos concluir que se rechaza H0(los médicos no se coordinan y la mayor parte se va
```

Actividad 3 - Análisis de regresión lineal

```
library(agricolae)

tabla_fallecidos = filter(datos_general, datos_general$COND_EGR != 2)

tabla_fallecidos_suma = aggregate(tabla_fallecidos$SEXO, by=list(DIAG1=tabla_fallecidos$DIAG1), FUN=sum)

tabla_fallecidos_suma_desc = tabla_fallecidos_suma[order(tabla_fallecidos_suma$x, decreasing = TRUE),]
```

```
kable(head(tabla_fallecidos_suma_desc, 5))%>%
  kable_styling(full_width = F) %>%
  kable_styling(latex_options = "striped") # Aplicación de estilos a la tabla
```

	DIAG1	x
4558	O800	67598
4572	O829	47768
4561	O809	44858
3223	K802	33349
4568	O820	33178

La mayor tasa de muertes lo tiene el diagnóstico PARTO UNICO ESPONTANEO, PRESENTACION CEFALICA DE VERTICE

Tabla muertes por Sexo

```
tabla_fallecidos_sexo = filter(tabla_fallecidos, DIAG1 == 'O800')
tabla_fallecidos_suma_sexo = aggregate(tabla_fallecidos_sexo$SEXO, by=list(SEXO=tabla_fallecidos_sexo$SEXO),
kable(head(tabla_fallecidos_suma_sexo, 5))%>%
  kable_styling(full_width = F) %>%
  kable_styling(latex_options = "striped") # Aplicación de estilos a la tabla
```

SEXO	x
2	67598

Tabla muertes por Edad

```
tabla_fallecidos_edad = filter(tabla_fallecidos, DIAG1 == 'O800')
tabla_fallecidos_suma_edad = aggregate(tabla_fallecidos_edad$EDAD, by=list(EDAD=tabla_fallecidos_edad$EDAD),
kable(head(tabla_fallecidos_suma_edad, 5))%>%
  kable_styling(full_width = F) %>%
  kable_styling(latex_options = "striped") # Aplicación de estilos a la tabla
```

EDAD	x
10	10
12	36
13	273
14	1092
15	3390

Tabla muertes por Previsión de salud

```
tabla_fallecidos_prev = filter(tabla_fallecidos, DIAG1 == '0800')
tabla_fallecidos_suma_prev = aggregate(tabla_fallecidos_prev$PREVI, by=list(PREVI=tabla_fallecidos_prev$PREVI), FUN=sum)
kable(head(tabla_fallecidos_suma_prev, 5))%>%
  kable_styling(full_width = F) %>%
  kable_styling(latex_options = "striped") # Aplicación de estilos a la tabla
```

PREVI	x
1	28201
2	9174
3	1461
5	1360
6	150

Tabla muertes por Servicio Clínico

```
tabla_fallecidos_serv = filter(tabla_fallecidos, DIAG1 == '0800')
tabla_fallecidos_suma_serv = aggregate(tabla_fallecidos_serv$SERC_EGR, by=list(SERC_EGR=tabla_fallecidos_serv$SERC_EGR), FUN=sum)
kable(head(tabla_fallecidos_suma_serv, 5))%>%
  kable_styling(full_width = F) %>%
  kable_styling(latex_options = "striped") # Aplicación de estilos a la tabla
```

SERC_EGR	x
119	952
150	150
160	66560
161	200928
330	671220

Tabla muertes por Intervención Quirúrgica

```
tabla_fallecidos_inter = filter(tabla_fallecidos, DIAG1 == '0800')
tabla_fallecidos_suma_inter = aggregate(tabla_fallecidos_inter$INTERV_Q, by=list(INTERV_Q=tabla_fallecidos_inter$INTERV_Q), FUN=sum)
kable(head(tabla_fallecidos_suma_inter, 5))%>%
  kable_styling(full_width = F) %>%
  kable_styling(latex_options = "striped") # Aplicación de estilos a la tabla
```

INTERV_Q	x
1	12776
2	42046