

Análisis Exploratorio de los Datos GASTO SANITARIO

Alicia Perdices Guerra

17 de mayo, 2021

Contents

- 1.ANÁLISIS EXPLORATORIO POR PAISES.
 - 1.1 EN RELACIÓN CON EL GASTO SANITARIO
 - * 1.1.1 Análisis Descriptivo
 - * 1.1.2 Visualización y Distribución de la variable “Value”
 - Sobre la información de la FINANCIACIÓN
 - Sobre la información del GASTO SANITARIO:
 - * 1.1.3 Normalidad de la variable “Value”
 - Sobre la información de la FINANCIACIÓN
 - Sobre la información del GASTO SANITARIO:

1.ANÁLISIS EXPLORATORIO POR PAISES

Se procede en primer lugar a cargar todos los archivos para poder realizar el análisis.

```
gasto_fin<-read.csv("C:/temp/GastoSanitario_Financiacion_clean.csv",sep= ",")
gasto_fun<-read.csv("C:/temp/GastoSanitario_Funcion_clean.csv",sep= ",")
gasto_pro<-read.csv("C:/temp/GastoSanitario_Proveedor_clean.csv",sep= ",")
ingresos<-read.csv("C:/temp/IngresosSanitarios_Financiacion_clean.csv",sep= ",")
medicos_p<-read.csv("C:/temp/Medicos_Por_Paises_clean.csv",sep= ",")
medicos_e<-read.csv("C:/temp/Medicos_x_especialidad_clean.csv",sep= ",")
enfermeria<-read.csv("C:/temp/Personal_Enfermeria_Cuidados_clean.csv",sep= ",")
perosnal_ne<-read.csv("C:/temp/PersonalSanitario_No_enfermeria_clean.csv",sep= ",")
personal<-read.csv("C:/temp/PersonalSanitarioHospitalario_clean.csv",sep= ",")
camas_p<-read.csv("C:/temp/Camas_Propiedad_clean.csv",sep= ",")
recursos_tec<-read.csv("C:/temp/RecursosTecnicos_hospitalarios_clean.csv",sep= ",")
tecnologia<-read.csv("C:/temp/TecnologiaMedica_clean.csv",sep= ",")
camas_t<-read.csv("C:/temp/TiposCamasHospitalarias_clean.csv",sep= ",")
deteccion_cu<-read.csv("C:/temp/Deteccion_Cancer_Mama_Cervix_clean.csv",sep= ",")
deteccion_ecd<-read.csv("C:/temp/Deteccion_enfCardio_Diabetes_clean.csv",sep= ",")
vacunacion<-read.csv("C:/temp/Vacunacion_+65_clean.csv",sep= ",")
estado<-read.csv("C:/temp/EstadoDeSalud_Sexo_clean.csv",sep= ",")
cuidados<-read.csv("C:/temp/CuidadosDomiciliarios_clean.csv",sep= ",")
mortalidad_tp<-read.csv("C:/temp/Mortalidad_Tratable_Prevenible_clean.csv",sep= ",")
mortalidad_ei<-read.csv("C:/temp/Muertes_Enf_Infecciosas_clean.csv",sep= ",")
altas_es<-read.csv("C:/temp/Altas_Estancias_CamasHospitalarias_clean.csv",sep= ",")
Ocupacion_cama<-read.csv("C:/temp/OcupacionCamaHospitalaria_clean.csv",sep= ",")
dia_tras<-read.csv("C:/temp/Pacientes_Dialisis_Trasplantes_clean.csv",sep= ",")
altas_hdd<-read.csv("C:/temp/RatioAltas_HospitalDia_Diagnostico_clean.csv",sep= ",")
```

1.1.- EN RELACIÓN CON EL GASTO SANITARIO

- 1.1.1 Análisis Descriptivo

Se procede a realizar el análisis descriptivo:

```
summary(gasto_fin)
```

```
##          TIME          GEO          UNIT
## Min.      :2009    Austria          : 50    Million euro:2000
## 1st Qu.:2011    Belgium          : 50
## Median :2014    Bosnia and Herzegovina: 50
## Mean      :2014    Bulgaria          : 50
## 3rd Qu.:2016    Croatia          : 50
## Max.      :2018    Cyprus          : 50
##          (Other)          :1700
##
## ICHA11_HF
## All financing schemes :400
## Compulsory contributory health insurance schemes and compulsory medical saving accounts (CMSA):400
## Government schemes :400
## Government schemes and compulsory contributory health care financing schemes :400
## Social health insurance schemes :400
##
##
## Value      Value_imp
## Min.      :      0    Mode :logical
## 1st Qu.:    996    FALSE:1487
## Median :   8585    TRUE :513
## Mean      : 153275
## 3rd Qu.:  69961
## Max.      :1573543
##
```

```
summary(gasto_fun)
```

```
##          TIME          GEO          UNIT
## Min.      :2009    Austria          : 50    Million euro:2000
## 1st Qu.:2011    Belgium          : 50
## Median :2014    Bosnia and Herzegovina: 50
## Mean      :2014    Bulgaria          : 50
## 3rd Qu.:2016    Croatia          : 50
## Max.      :2018    Cyprus          : 50
##          (Other)          :1700
##
## ICHA11_HC      Value
## Curative care :400    Min.      :    70.5
## Curative care and rehabilitative care :400    1st Qu.:   1533.9
## Current health care expenditure (CHE) :400    Median :   9165.2
## Inpatient curative and rehabilitative care:400    Mean      : 123048.4
## Inpatient curative care :400    3rd Qu.:  41639.1
##          Max.      :1573542.9
##
## Value_imp
## Mode :logical
## FALSE:1405
## TRUE :595
##
##
##
```

```
summary(gasto_pro)
```

```
##          TIME          GEO          UNIT
```

```
## Min. :2009 Austria : 50 Million euro:2000
## 1st Qu.:2011 Belgium : 50
## Median :2014 Bosnia and Herzegovina: 50
## Mean :2014 Bulgaria : 50
## 3rd Qu.:2016 Croatia : 50
## Max. :2018 Cyprus : 50
## (Other) :1700
##
## ICHA11_HP
## All providers of health care :400
## General hospitals :400
## Hospitals :400
## Mental health hospitals :400
## Specialised hospitals (other than mental health hospitals):400
##
##
## Value Value_imp
## Min. : 0.0 Mode :logical
## 1st Qu.: 307.8 FALSE:1306
## Median : 2031.6 TRUE :694
## Mean : 99226.2
## 3rd Qu.: 19879.7
## Max. :1573542.9
##
```

Se filtra el dataframe para que la variable GEO aparezcan solo los países objeto de estudio. (Para cada archivo relacionado con el Gasto Sanitario, y unificamos la información en una variable, gasto). Además se selecciona la información relevante de las variables ICHA11_HF, ICHA11_HC y ICHA11_HP:

```
#En relación con la Financiación
#=====
gasto_fin_paises<-gasto_fin[((gasto_fin$GEO!="European Union - 27 countries (from 2020)")&
+(gasto_fin$GEO!="European Union - 28 countries (2013-2020)")&
+(gasto_fin$GEO!="European Union - 27 countries (2007-2013)")&
+(gasto_fin$GEO!="European Union - 15 countries (1995-2004)")&
+(gasto_fin$GEO!="Euro area - 12 countries (2001-2006)")&
+(gasto_fin$GEO!="Euro area - 19 countries (from 2015)")&
+(gasto_fin$GEO!="Euro area - 18 countries (2014)")&
+(gasto_fin$GEO!="Euro area - 12 countries (2001-2006)")),]

# Toda la financiación en conjunto
gasto_fin_paises<-
  gasto_fin_paises[(gasto_fin$ICHA11_HF=="All financing schemes"),]

#Se han creado registros con valores NA en el proceso de eliminación de
#factores GEO e ICHA11_HF . Se eliminan
idx=which(is.na(gasto_fin_paises))
gasto_fin_paises=gasto_fin_paises[-c(idx),]

#En relación con la Función.
#=====
gasto_fun_paises<-gasto_fun[(gasto_fun$GEO!="European Union - 27 countries (from 2020)")&
+(gasto_fun$GEO!="European Union - 28 countries (2013-2020)")&
+(gasto_fun$GEO!="European Union - 27 countries (2007-2013)")&
+(gasto_fun$GEO!="European Union - 15 countries (1995-2004)")&
+(gasto_fun$GEO!="Euro area - 19 countries (from 2015)")&
+(gasto_fun$GEO!="Euro area - 18 countries (2014)")&
```

```

      +(gasto_fun$GEO!="Euro area - 12 countries (2001-2006)" ),]
#Se selecciona todo lo relacionado con el gasto
#en Cuidados Curativos y de Rehabilitación en general y en pacientes ingresados.
gasto_fun_paises_cr_care<-
  gasto_fun_paises[(gasto_fun$ICHA11_HC=="Curative care and rehabilitative care"),]
gasto_fun_paises_ip_cr_care<-
  gasto_fun_paises[(gasto_fun$ICHA11_HC=="Inpatient curative and rehabilitative care"),]

#Se han creado registros con valores NA en el proceso de eliminación de
#factores GEO e ICHA11_HC . Se eliminan
idx=which(is.na(gasto_fun_paises_cr_care))
gasto_fun_paises_cr_care=gasto_fun_paises_cr_care[-c(idx),]

idx=which(is.na(gasto_fun_paises_ip_cr_care))
gasto_fun_paises_ip_cr_care=gasto_fun_paises_ip_cr_care[-c(idx),]

#En relación con el Proveedor.
#=====
gasto_pro_paises<-gasto_pro[(gasto_pro$GEO!="European Union - 27 countries (from 2020)"&
  +(gasto_pro$GEO!="European Union - 28 countries (2013-2020)"&
  +(gasto_pro$GEO!="European Union - 27 countries (2007-2013)"&
  +(gasto_pro$GEO!="Euro area - 19 countries (from 2015)"&
  +(gasto_pro$GEO!="European Union - 15 countries (1995-2004)"&
  +(gasto_pro$GEO!="Euro area - 18 countries (2014)"&
  +(gasto_pro$GEO!="Euro area - 12 countries (2001-2006)" ),]

#Se selecciona información sobre el gasto sanitario destinado a:
#Hospitales, Hospitales mentales, y Hospitales Especiales.
gasto_pro_paises_hospital<-
gasto_pro_paises[(gasto_pro$ICHA11_HP=="Hospitals"),]
gasto_pro_paises_mental_hospital<-
gasto_pro_paises[(gasto_pro$ICHA11_HP=="Mental health hospitals"),]
gasto_pro_paises_specialised_hospital<-
gasto_pro_paises[(gasto_pro$ICHA11_HP=="Specialised hospitals (other than mental health hospitals)" ),]

#Se han creado registros con valores NA en el proceso de eliminación de
#factores GEO e ICHA11_HP. Se eliminan
idx=which(is.na(gasto_pro_paises_hospital))
gasto_pro_paises_hospital=gasto_pro_paises_hospital[-c(idx),]

idx=which(is.na(gasto_pro_paises_mental_hospital))
gasto_pro_paises_mental_hospital=gasto_pro_paises_mental_hospital[-c(idx),]

idx=which(is.na(gasto_pro_paises_specialised_hospital))
gasto_pro_paises_specialised_hospital=gasto_pro_paises_specialised_hospital[-c(idx),]

```

Se crea un Dataframe con toda la información (Unidades: Millones de Euros):

```

year<-(gasto_fun_paises_cr_care$TIME) #Columna Year
country<-(gasto_fun_paises_cr_care$GEO) #Columna Países

#Dataframe con toda la información del Gasto Sanitario por Años y Países

```

```

gasto_df<-data.frame("TIME"=year,"Pais"=country,
                    "Gasto_Cuidado_Rehabilitacion"= gasto_fun_paises_cr_care$Value,
                    "Gasto_Cuidado_Rehabilitacion_Ingresados"=gasto_fun_paises_ip_cr_care$Value,
                    "Gasto_Hospitales"=gasto_pro_paises_hospital$Value,
                    "Gasto_Hospitales_Mentales"=gasto_pro_paises_mental_hospital$Value,
                    "Gasto_Hospitales_Especiales"=gasto_pro_paises_specialised_hospital$Value)

#Variable con la información de la financiación General por Años y Países
financiacion<-select(gasto_fin_paises,-UNIT)

#Ordenadmos el DataFrame por Año en sentido descendente.
gasto_df<-gasto_df[with(gasto_df, order(gasto_df$TIME )),]
head(gasto_df)

```

```

##      TIME                      Pais
## 1 2009                      Belgium
## 2 2009                      Bulgaria
## 3 2009                      Czechia
## 4 2009                      Denmark
## 5 2009 Germany (until 1990 former territory of the FRG)
## 6 2009                      Estonia
##      Gasto_Cuidado_Rehabilitacion Gasto_Cuidado_Rehabilitacion_Ingresados
## 1                      18019.63                      9760.23
## 2                      1608.72                      1047.44
## 3                      6545.97                      2982.27
## 4                      14164.09                      7162.92
## 5                      140782.00                     74687.00
## 6                      534.32                      252.35
##      Gasto_Hospitales Gasto_Hospitales_Mentales Gasto_Hospitales_Especiales
## 1          11661.88          1416.28          59.84
## 2          1115.41           21.01          249.81
## 3          4971.52          173.28          333.75
## 4          11092.27          981.36           0.00
## 5          79394.00         79394.00          8202.00
## 6           444.73           2.61           22.10

```

```
head(financiacion)
```

```

##      TIME                      GEO                      ICHA11_HF
## 36 2009                      Belgium All financing schemes
## 41 2009                      Bulgaria All financing schemes
## 46 2009                      Czechia All financing schemes
## 51 2009                      Denmark All financing schemes
## 56 2009 Germany (until 1990 former territory of the FRG) All financing schemes
## 61 2009                      Estonia All financing schemes
##      Value Value_imp
## 36 35879.39      FALSE
## 41  3003.51       TRUE
## 46 12202.11       TRUE
## 51 25167.02       TRUE
## 56 274841.00      FALSE
## 61   925.55      FALSE

```

Se reescalan los datos:

```
gasto_fin_paises["Value_norm_fin"]<- (rescale(gasto_fin_paises$Value, to=c(0,1)))
head(gasto_fin_paises)
```

```
##      TIME                                GEO      UNIT
## 36 2009                                Belgium Million euro
## 41 2009                                Bulgaria Million euro
## 46 2009                                Czechia Million euro
## 51 2009                                Denmark Million euro
## 56 2009 Germany (until 1990 former territory of the FRG) Million euro
## 61 2009                                Estonia Million euro
##      ICHA11_HF      Value Value_imp Value_norm_fin
## 36 All financing schemes 35879.39      FALSE 0.092870339
## 41 All financing schemes 3003.51       TRUE 0.007112987
## 46 All financing schemes 12202.11      TRUE 0.031107709
## 51 All financing schemes 25167.02      TRUE 0.064926920
## 56 All financing schemes 274841.00     FALSE 0.716206195
## 61 All financing schemes 925.55       FALSE 0.001692590
```

```
gasto_df["Gasto_Cuidado_Rehabilitacion_norm"]<-
  rescale(gasto_df$Gasto_Cuidado_Rehabilitacion, to=c(0,1))
gasto_df["Gasto_Cuidado_Rehabilitacion_Ingresados_norm"]<-
  rescale(gasto_df$Gasto_Cuidado_Rehabilitacion_Ingresados , to=c(0,1))
gasto_df["Gasto_Hospitales_norm"]<-
  rescale(gasto_df$Gasto_Hospitales , to=c(0,1))
gasto_df["Gasto_Hospitales_Mentales_norm"]<-
  rescale(gasto_df$Gasto_Hospitales_Mentales, to=c(0,1))
gasto_df["Gasto_Hospitales_Especiales_norm"]<-
  rescale(gasto_df$Gasto_Hospitales_Especiales , to=c(0,1))

head(gasto_df)
```

```
##      TIME                                Pais
## 1 2009                                Belgium
## 2 2009                                Bulgaria
## 3 2009                                Czechia
## 4 2009                                Denmark
## 5 2009 Germany (until 1990 former territory of the FRG)
## 6 2009                                Estonia
##      Gasto_Cuidado_Rehabilitacion Gasto_Cuidado_Rehabilitacion_Ingresados
## 1                                18019.63                                9760.23
## 2                                1608.72                                1047.44
## 3                                6545.97                                2982.27
## 4                                14164.09                                7162.92
## 5                                140782.00                                74687.00
## 6                                534.32                                252.35
##      Gasto_Hospitales Gasto_Hospitales_Mentales Gasto_Hospitales_Especiales
## 1                11661.88                1416.28                59.84
## 2                1115.41                 21.01                249.81
## 3                4971.52                 173.28                333.75
## 4                11092.27                 981.36                 0.00
## 5                79394.00                79394.00                8202.00
## 6                 444.73                  2.61                22.10
##      Gasto_Cuidado_Rehabilitacion_norm
## 1                                0.094442782
```

```
## 2          0.007592301
## 3          0.033721415
## 4          0.074038338
## 5          0.744130760
## 6          0.001906318
##  Gasto_Cuidado_Rehabilitacion_Ingresados_norm Gasto_Hospitales_norm
## 1          0.096343917          0.107092045
## 2          0.009635486          0.010041930
## 3          0.028890636          0.045526402
## 4          0.070495862          0.101850413
## 5          0.742485777          0.730372636
## 6          0.001722865          0.003870237
##  Gasto_Hospitales_Mentales_norm Gasto_Hospitales_Especiales_norm
## 1          1.783863e-02          0.005924166
## 2          2.646296e-04          0.024731215
## 3          2.182533e-03          0.033041283
## 4          1.236063e-02          0.000000000
## 5          1.000000e+00          0.811998812
## 6          3.287402e-05          0.002187902
```

• 1.1.2 Visualización y Distribución de la variable “Value”

Se visualiza la variable “Value” en función de TIME, y el gasto Sanitario en el Cuidado y Rehabilitación en general y en pacientes ingresados.

```
#Gasto Sanitario Cuidado y Rehabilitación en general y en pacientes ingresados.
#=====

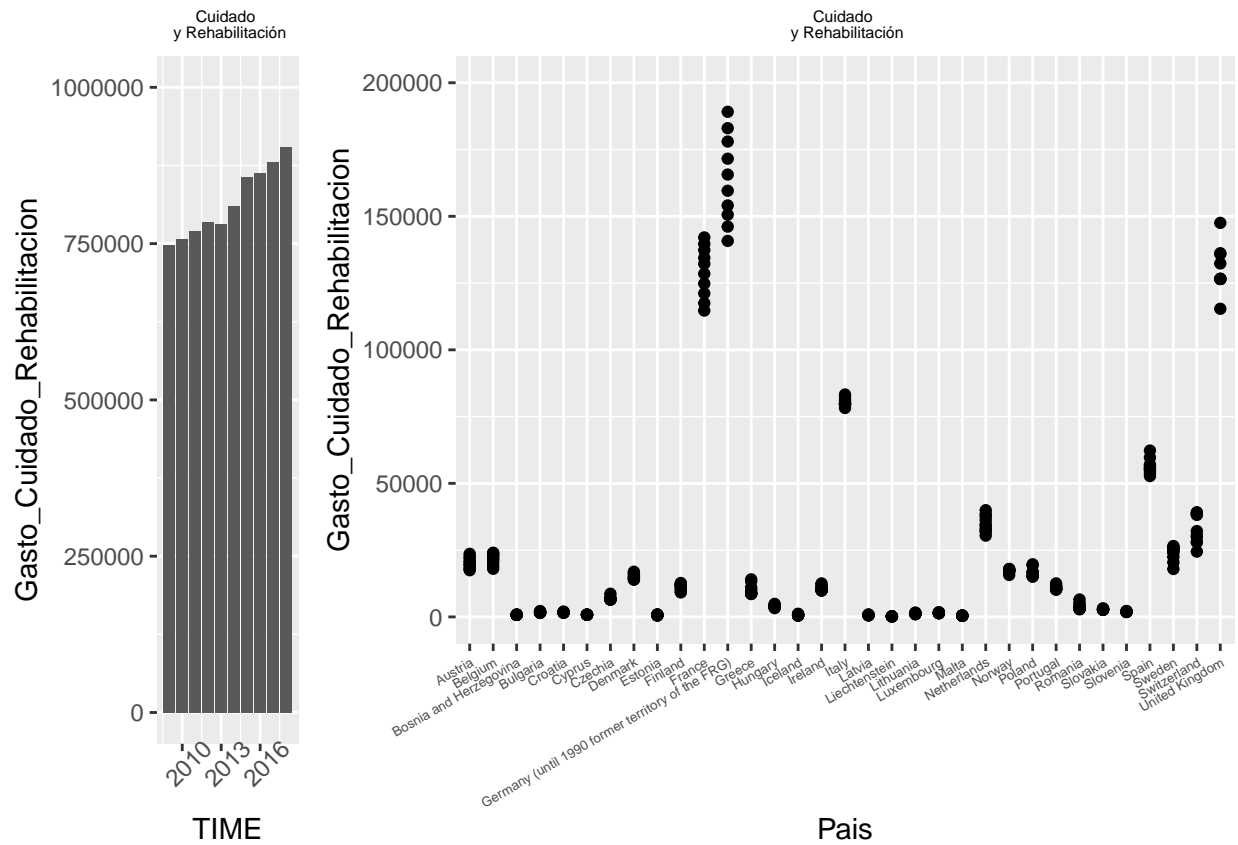
plot1=ggplot(data=gasto_df)+geom_col(aes(x=TIME,y=Gasto_Cuidado_Rehabilitacion))+
  theme(axis.text.x = element_text(angle = 45))+
  scale_y_continuous(limit=c(0,1000000))+
  ggtitle("Cuidado \n y Rehabilitación")+
  theme (plot.title = element_text(size=rel(0.5), hjust = 0.5))

plot2=ggplot(data=gasto_df)+geom_col(aes(x=TIME,y=Gasto_Cuidado_Rehabilitacion_Ingresados))+
  theme(axis.text.x = element_text(size= 5,angle = 30,vjust=1,hjust = 1))+
  scale_y_continuous(limit=c(0,1000000))+
  ggtitle("Cuidado \n y Rehabilitación \n Pacientes Ingresados")+
  theme (plot.title = element_text(size=rel(0.5),hjust = 0.5))

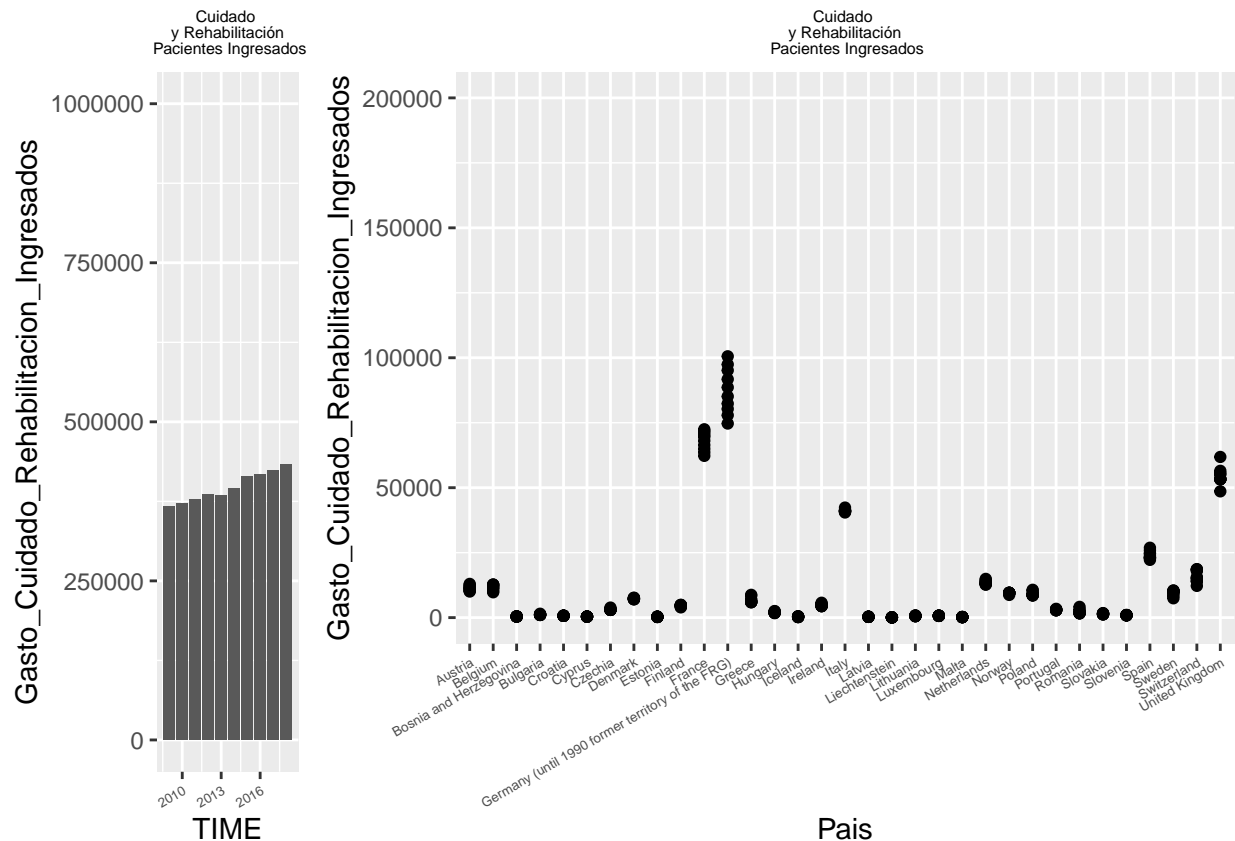
plot3=ggplot(data=gasto_df)+geom_point(aes(x=Pais,y=Gasto_Cuidado_Rehabilitacion))+
  theme(axis.text.x = element_text(size= 5,angle = 30,vjust=1,hjust = 1))+
  scale_y_continuous(limit=c(0,200000))+
  ggtitle("Cuidado \n y Rehabilitación")+
  theme (plot.title = element_text(size=rel(0.5),hjust=0.5))

plot4=ggplot(data=gasto_df)+geom_point(aes(x=Pais,y=Gasto_Cuidado_Rehabilitacion_Ingresados))+
  theme(axis.text.x = element_text(size= 5,angle = 30,vjust=1,hjust = 1))+
  scale_y_continuous(limit=c(0,200000))+
  ggtitle("Cuidado \n y Rehabilitación \n Pacientes Ingresados")+
  theme (plot.title = element_text(size=rel(0.5),hjust = 0.5))

grid.arrange(plot1,plot3,widths=c(1,3), ncol=2)
```



```
grid.arrange(plot2,plot4,widths=c(1,3), ncol=2)
```

Se grafica la información de la variable “Value” en función de TIME, y el gasto Sanitario en Hospitales, Hospitales Mentales y Hospitales Especiales.

```
#Gasto Sanitario en Hospitales, Hospitales Mentales y Hospitales Especiales.
#=====
plot1=ggplot(data=gasto_df)+geom_col(aes(x=TIME,y=Gasto_Hospitales))+
  theme(axis.text.x = element_text(angle = 45))+
  scale_y_continuous(limit=c(0,800000))+
  ggtitle("Hospitales")+
  theme (plot.title = element_text(size=rel(0.5), hjust = 0.5))

plot2=ggplot(data=gasto_df)+geom_col(aes(x=TIME,y=Gasto_Hospitales_Mentales))+
  theme(axis.text.x = element_text(size= 5,angle = 30,vjust=1,hjust = 1))+
  scale_y_continuous(limit=c(0,800000))+
  ggtitle("Hospitales \n y Mentales")+
  theme (plot.title = element_text(size=rel(0.5),hjust = 0.5))

plot3=ggplot(data=gasto_df)+geom_col(aes(x=TIME,y=Gasto_Hospitales_Especiales))+
  theme(axis.text.x = element_text(size= 5,angle = 30,vjust=1,hjust = 1))+
  scale_y_continuous(limit=c(0,800000))+
  ggtitle("Hospitales \n Especiales")+
  theme (plot.title = element_text(size=rel(0.5),hjust = 0.5))

plot4=ggplot(data=gasto_df)+geom_point(aes(x=Pais,y=Gasto_Hospitales))+
  theme(axis.text.x = element_text(size= 5,angle = 30,vjust=1,hjust = 1))+
  scale_y_continuous(limit=c(0,120000))+
```

```

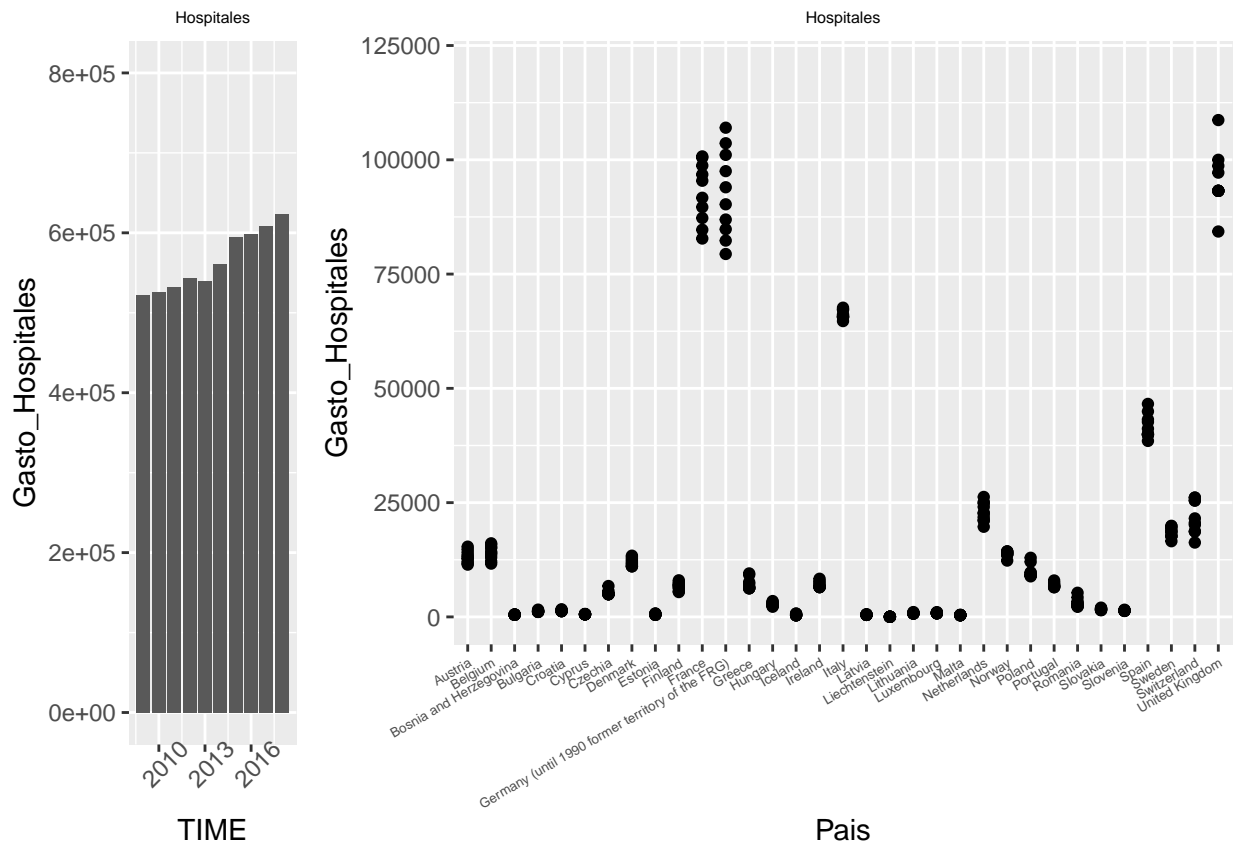
ggtitle("Hospitales")+
theme (plot.title = element_text(size=rel(0.5),hjust=0.5))

plot5=ggplot(data=gasto_df)+geom_point(aes(x=Pais,y=Gasto_Hospitales_Mentales))+
theme(axis.text.x = element_text(size= 5,angle = 30,vjust=1,hjust = 1))+
scale_y_continuous(limit=c(0,120000))+
ggtitle("Hospitales \n Mentales")+
theme (plot.title = element_text(size=rel(0.5),hjust = 0.5))

plot6=ggplot(data=gasto_df)+geom_point(aes(x=Pais,y=Gasto_Hospitales_Especiales))+
theme(axis.text.x = element_text(size= 5,angle = 30,vjust=1,hjust = 1))+
scale_y_continuous(limit=c(0,120000))+
ggtitle("Hospitales \n Especiales")+
theme (plot.title = element_text(size=rel(0.5),hjust = 0.5))

grid.arrange(plot1,plot4,widths=c(1,3), ncol=2)

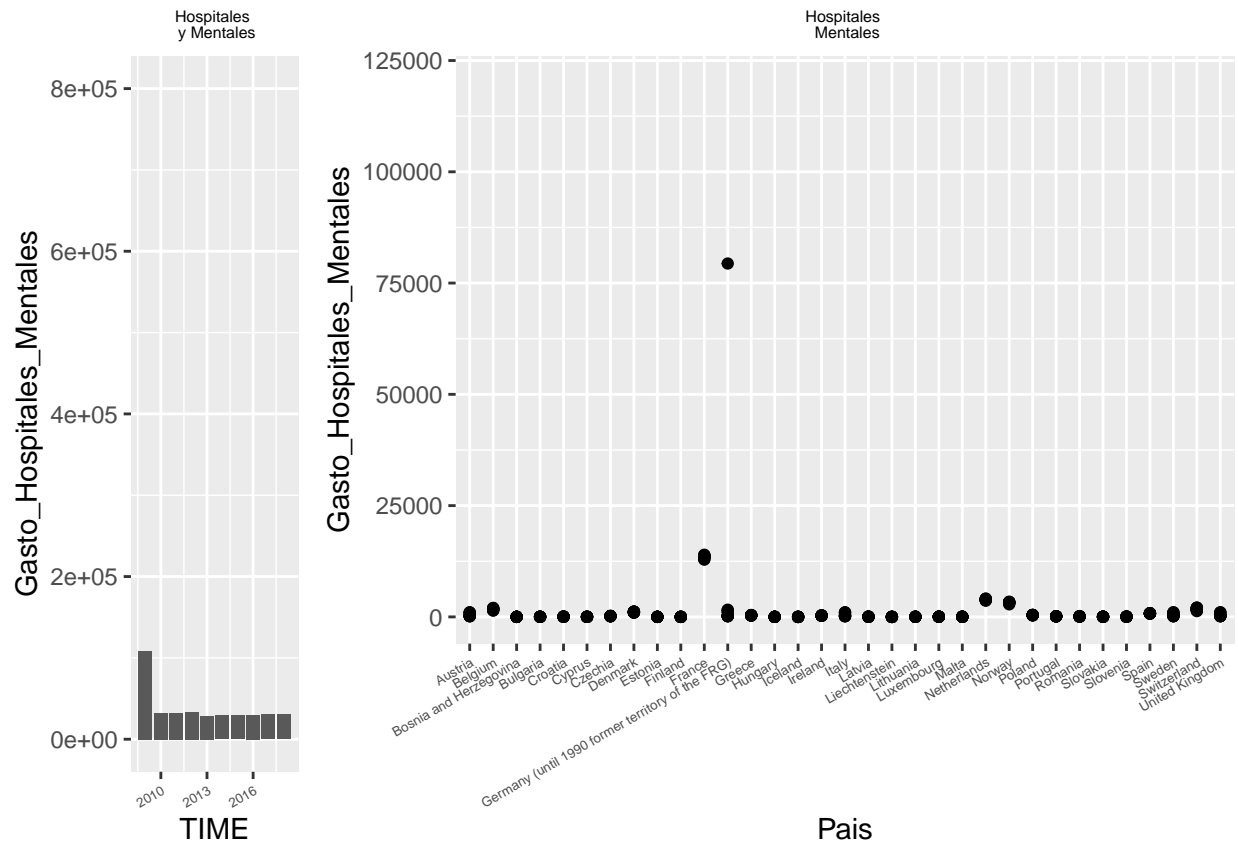
```



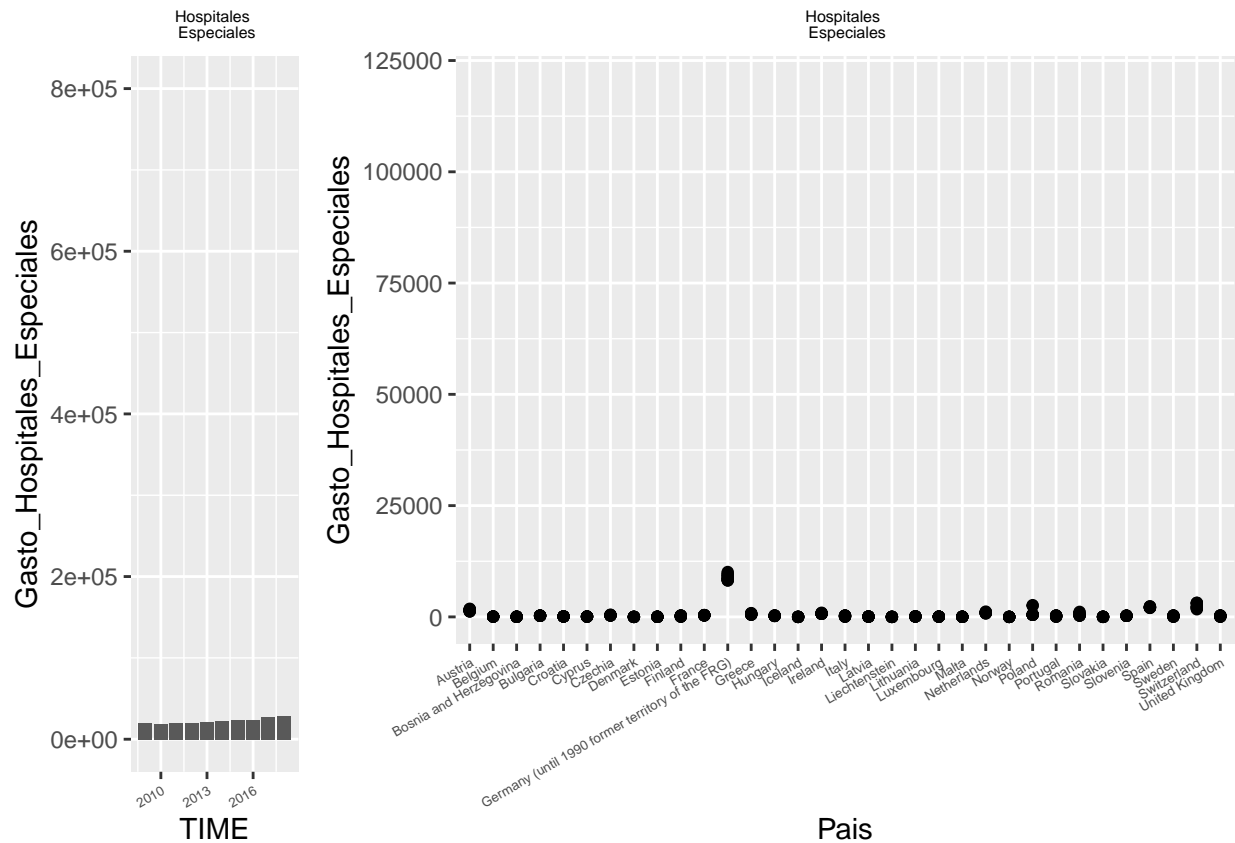
```

grid.arrange(plot2,plot5,widths=c(1,3), ncol=2)

```



```
grid.arrange(plot3,plot6,widths=c(1,3), ncol=2)
```

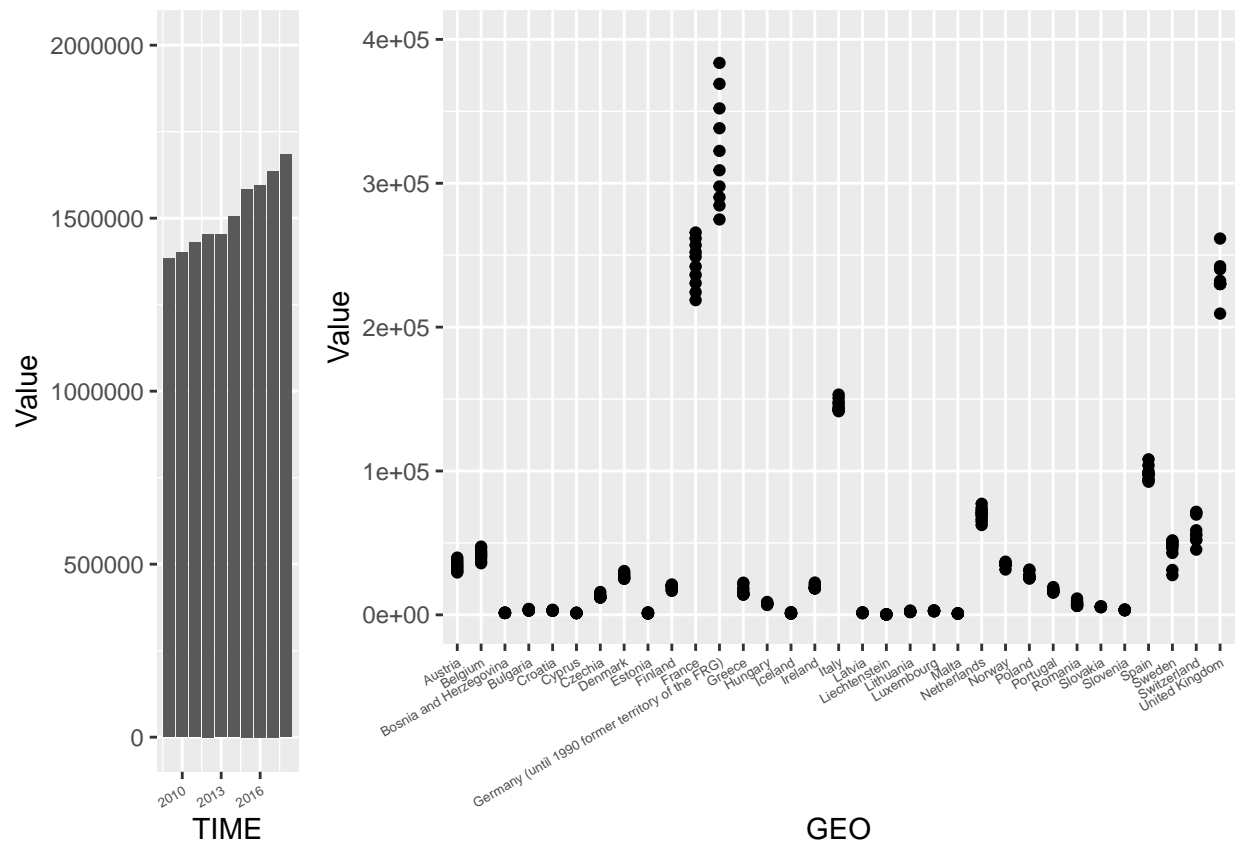


Se muestra la distribución de la variable “Value” en relación con la Financiación.

```
plot5=ggplot(data=gasto_fin_paises)+geom_col(aes(x=TIME,y=Value))+
  theme(axis.text.x = element_text(size= 5,angle = 30,vjust=1,hjust = 1))+
  scale_y_continuous(limit=c(0,2000000))

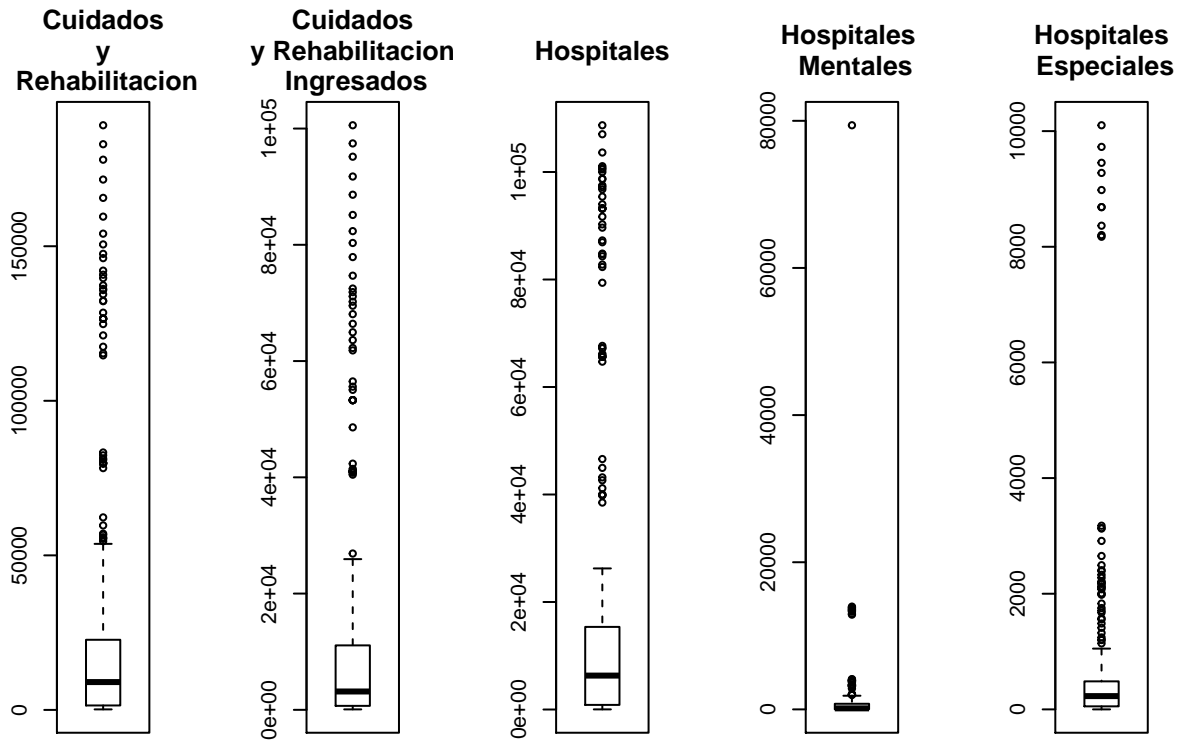
plot6=ggplot(data=gasto_fin_paises)+geom_point(aes(x=GEO,y=Value))+
  theme(axis.text.x = element_text(size= 5,angle = 30,vjust=1,hjust = 1))+
  scale_y_continuous(limit=c(0,400000))

grid.arrange(plot5,plot6,widths=c(1,3), ncol=2)
```



A continuación, se muestra la distribución de Value en cómputo global (años y países) para cada gasto aplicado:

```
par(mfrow=c(1,5))
boxplot(gasto_df$Gasto_Cuidado_Rehabilitacion,main="Cuidados\n y \n Rehabilitacion")
boxplot(gasto_df$Gasto_Cuidado_Rehabilitacion_Ingresados,main="Cuidados\n y Rehabilitacion \n Ingresados")
boxplot(gasto_df$Gasto_Hospitales,main="Hospitales")
boxplot(gasto_df$Gasto_Hospitales_Mentales,main="Hospitales \n Mentales")
boxplot(gasto_df$Gasto_Hospitales_Especiales,main="Hospitales\n Especiales")
```



Se obtienen los 5 países con más Financiación y gasto Sanitario especificando el organismo al que se destina por años.

```
y_2009<-filter(gasto_fin_paises, TIME==2009)#Se filtra por Año
y_2010<-filter(gasto_fin_paises, TIME==2010)
y_2011<-filter(gasto_fin_paises, TIME==2011)
y_2012<-filter(gasto_fin_paises, TIME==2012)
y_2013<-filter(gasto_fin_paises, TIME==2013)
y_2014<-filter(gasto_fin_paises, TIME==2014)
y_2015<-filter(gasto_fin_paises, TIME==2015)
y_2016<-filter(gasto_fin_paises, TIME==2016)
y_2017<-filter(gasto_fin_paises, TIME==2017)
y_2018<-filter(gasto_fin_paises, TIME==2018)

gasto_fin_5paises_2009<-y_2009[with(y_2009, order(-y_2009$Value )),]#Se ordena por Value
gasto_fin_5paises_2010<-y_2010[with(y_2010, order(-y_2010$Value )),]
gasto_fin_5paises_2011<-y_2011[with(y_2011, order(-y_2011$Value )),]
gasto_fin_5paises_2012<-y_2012[with(y_2012, order(-y_2012$Value )),]
gasto_fin_5paises_2013<-y_2013[with(y_2013, order(-y_2013$Value )),]
gasto_fin_5paises_2014<-y_2014[with(y_2014, order(-y_2014$Value )),]
gasto_fin_5paises_2015<-y_2015[with(y_2015, order(-y_2015$Value )),]
gasto_fin_5paises_2016<-y_2016[with(y_2016, order(-y_2016$Value )),]
gasto_fin_5paises_2017<-y_2017[with(y_2017, order(-y_2017$Value )),]
gasto_fin_5paises_2018<-y_2018[with(y_2018, order(-y_2018$Value )),]

#Se crea una tabla para cada año sobre la Financiación de los 5 Países un valor más alto.
```

```
kable(gasto_fin_5países_2009[0:5,c(2,5)], col.names = c("País","Financiación"), caption = "Países con m
```

Table 1: Países con mayor financiación en 2009

	País	Financiación
56	Germany (until 1990 former territory of the FRG)	274841.00
191	United Kingdom	229998.79
81	France	218751.88
91	Italy	142676.00
76	Spain	97384.01

```
kable(gasto_fin_5países_2010[0:5,c(2,5)], col.names = c("País","Financiación"), caption = "Países con m
```

Table 2: Países con mayor financiación en 2010

	País	Financiación
256	Germany (until 1990 former territory of the FRG)	284568.00
391	United Kingdom	229998.79
281	France	224272.75
291	Italy	142676.00
276	Spain	97815.78

```
kable(gasto_fin_5países_2011[0:5,c(2,5)], col.names = c("País","Financiación"), caption = "Países con m
```

Table 3: Países con mayor financiación en 2011

	País	Financiación
456	Germany (until 1990 former territory of the FRG)	290266.00
481	France	230575.03
591	United Kingdom	229998.79
491	Italy	142676.00
476	Spain	97532.09

```
kable(gasto_fin_5países_2012[0:5,c(2,5)], col.names = c("País","Financiación"), caption = "Países con m
```

Table 4: Países con mayor financiación en 2012

	País	Financiación
656	Germany (until 1990 former territory of the FRG)	297784.00
681	France	236311.46
791	United Kingdom	229998.79
691	Italy	142676.00
676	Spain	94417.66

```
kable(gasto_fin_5países_2013[0:5,c(2,5)], col.names = c("País","Financiación"), caption = "Países con m
```

Table 5: Países con mayor financiación en 2013

	País	Financiacion
856	Germany (until 1990 former territory of the FRG)	309020.0
881	France	242123.4
991	United Kingdom	209392.5
891	Italy	141526.0
876	Spain	92518.8

```
kable(gasto_fin_5países_2014[0:5,c(2,5)], col.names = c("País","Financiacion"), caption = "Países con m
```

Table 6: Países con mayor financiación en 2014

	País	Financiacion
1056	Germany (until 1990 former territory of the FRG)	322481.00
1081	France	248958.59
1191	United Kingdom	229998.79
1091	Italy	144317.00
1076	Spain	93824.25

```
kable(gasto_fin_5países_2015[0:5,c(2,5)], col.names = c("País","Financiacion"), caption = "Países con m
```

Table 7: Países con mayor financiación en 2015

	País	Financiacion
1256	Germany (until 1990 former territory of the FRG)	338267.00
1391	United Kingdom	261567.48
1281	France	252075.88
1291	Italy	146613.00
1276	Spain	98350.22

```
kable(gasto_fin_5países_2016[0:5,c(2,5)], col.names = c("País","Financiacion"), caption = "Países con m
```

Table 8: Países con mayor financiación en 2016

	País	Financiacion
1456	Germany (until 1990 former territory of the FRG)	352045.00
1481	France	256954.86
1591	United Kingdom	240259.87
1491	Italy	147963.00
1476	Spain	99715.25

```
kable(gasto_fin_5países_2017[0:5,c(2,5)], col.names = c("País","Financiacion"), caption = "Países con m
```

Table 9: Países con mayor financiación en 2017

	País	Financiacion
1656	Germany (until 1990 former territory of the FRG)	369091.0

	País	Financiacion
1681	France	261667.4
1791	United Kingdom	232178.1
1691	Italy	150697.0
1676	Spain	103899.9

```
kable(gasto_fin_5países_2018[0:5,c(2,5)], col.names = c("País","Financiacion"), caption = "Países con m
```

Table 10: Países con mayor financiación en 2018

	País	Financiacion
1856	Germany (until 1990 former territory of the FRG)	383636.0
1881	France	265763.7
1991	United Kingdom	242300.0
1891	Italy	153085.0
1876	Spain	108109.7

A continuación, se agrupa toda la financiación por países y se crea una tabla con toda la información:

```
a1<-group_by(gasto_fin_países,GEO)#Se agrupa por países
a2<-select(a1,TIME:GEO:Value)#Seleccionamos las variables GEO y Value

## Warning in x:y: numerical expression has 2 elements: only the first used
a3<-(summarize(a2,suma=sum(Value)))#Se muestra la información por cada país, con Value=suma de los valo

## `summarise()` ungrouping output (override with `.groups` argument)
a4<-data.frame(a3)#Se convierte la información en un dataframe
a5<-a4[with(a4,order(-a4$suma)),]#Se ordena el DataFrame por la variable Suma de forma descendente.

#Se crea una tabla con toda la información
kable(a5[0:5,c(1,2)], col.names = c("País","Financiacion_Millones de Euros"), caption = "Países con may
```

Table 11: Países con mayor financiación en 2009-2018

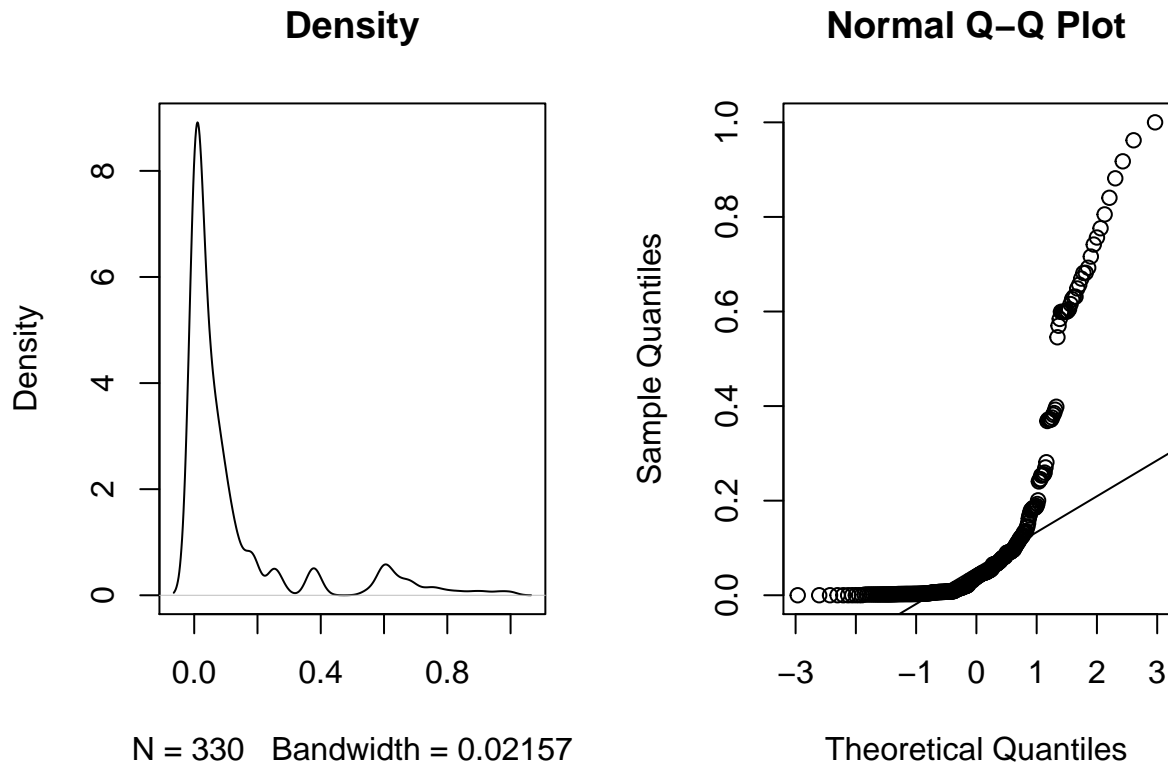
	País	Financiacion_Millones de Euros
12	Germany (until 1990 former territory of the FRG)	3221999.0
11	France	2437454.9
33	United Kingdom	2335692.0
17	Italy	1454905.0
30	Spain	983567.6

• 1.1.3 Normalidad de la variable “Value”

Se comprueba con métodos visuales si la variable tiene una distribución normal.

FINANCIACION

```
par(mfrow=c(1,2))
plot(density(gasto_fin_países$Value_norm_fin) ,main="Density")
qqnorm(gasto_fin_países$Value_norm_fin)
qqline(gasto_fin_países$Value_norm_fin)
```



Para estudiar si una muestra proviene de una población con distribución normal, se disponen de tres herramientas:

- Histograma o Densidad
- Gráficos cuantil cuantil (QQplot)
- Pruebas de hipótesis.

Si en la prueba de Densidad se observa sesgo hacia uno de los lados de la gráfica, sería indicio de que la muestra no proviene de una población normal. Si por otra parte, sí se observa simetría, **NO** se garantiza que la muestra provenga de una población normal. En estos casos sería necesario utilizar otras herramientas como **QQplot y pruebas de hipótesis**.

En la gráfica Densidad de la variable “Value” para La Financiación del Gasto Sanitario, se observa claramente sesgo hacia la izquierda, por lo que no se considera normalidad. Se puede confirmar observando la gráfica QQplot en la que la línea que grafica qqline sirve de referencia para interpretar el gráfico. Si se tuviese una muestra distribuida normalmente, se esperaría que los puntos del gráfico cuantil cuantil estuviesen perfectamente alineados con la línea de referencia, y observamos que para este caso, “Value” no se alinea. Confirmaría los resultados del gráfico de densidad.

Por otro lado, se realizan las pruebas de hipótesis:

- \$h_0\$: La muestra proviene de una población normal.
- \$h_1\$: La muestra NO proviene de una población normal.

Se aplica la prueba Shapiro-Wilk:

```
shapiro.test(gasto_fin_paises$Value_norm_fin)
```

```
##
## Shapiro-Wilk normality test
```

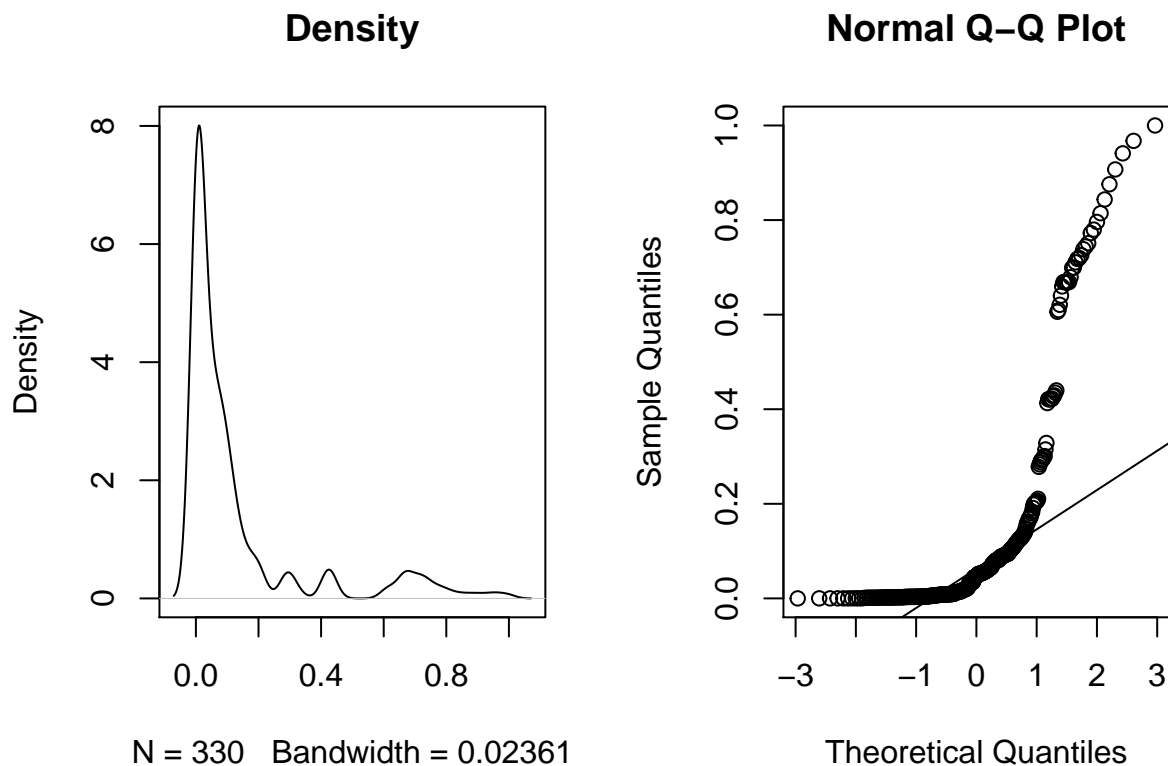
```
##
## data: gasto_fin_paises$Value_norm_fin
## W = 0.61273, p-value < 2.2e-16
```

Se observa un p-value muy pequeño, mucho más pequeño que cualquier nivel de significación ($\alpha=0.5$) por lo que se rechaza la hipótesis nula y asumimos **No normalidad** en la muestra.

GASTO EN RELACIÓN CON LA FUNCIÓN

Gasto en Cuidados y Rehabilitación

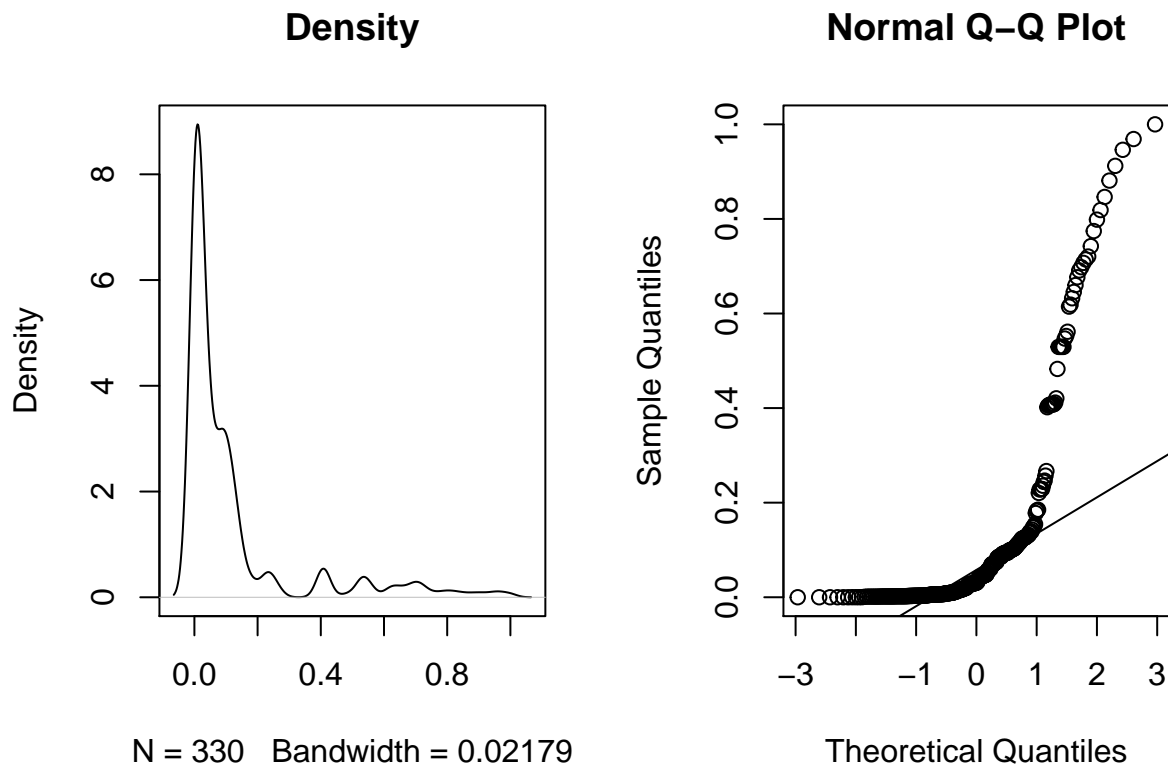
```
par(mfrow=c(1,2))
plot(density(gasto_df$Gasto_Cuidado_Rehabilitacion_norm) ,main="Density")
qqnorm(gasto_df$Gasto_Cuidado_Rehabilitacion_norm)
qqline(gasto_df$Gasto_Cuidado_Rehabilitacion_norm)
```



En la gráfica Densidad de la variable “Value” para El Gasto en Cuidados Y Rehabilitación, se observa claramente sesgo hacia la izquierda, por lo que no se considera normalidad. Se puede confirmar observando la gráfica QQplot en la que la línea que grafica qqline sirve de referencia para interpretar el gráfico, no se alinea con los puntos de los valores de la variable “Value”.

Gasto en Cuidados y Rehabilitación en Pacientes ingresados

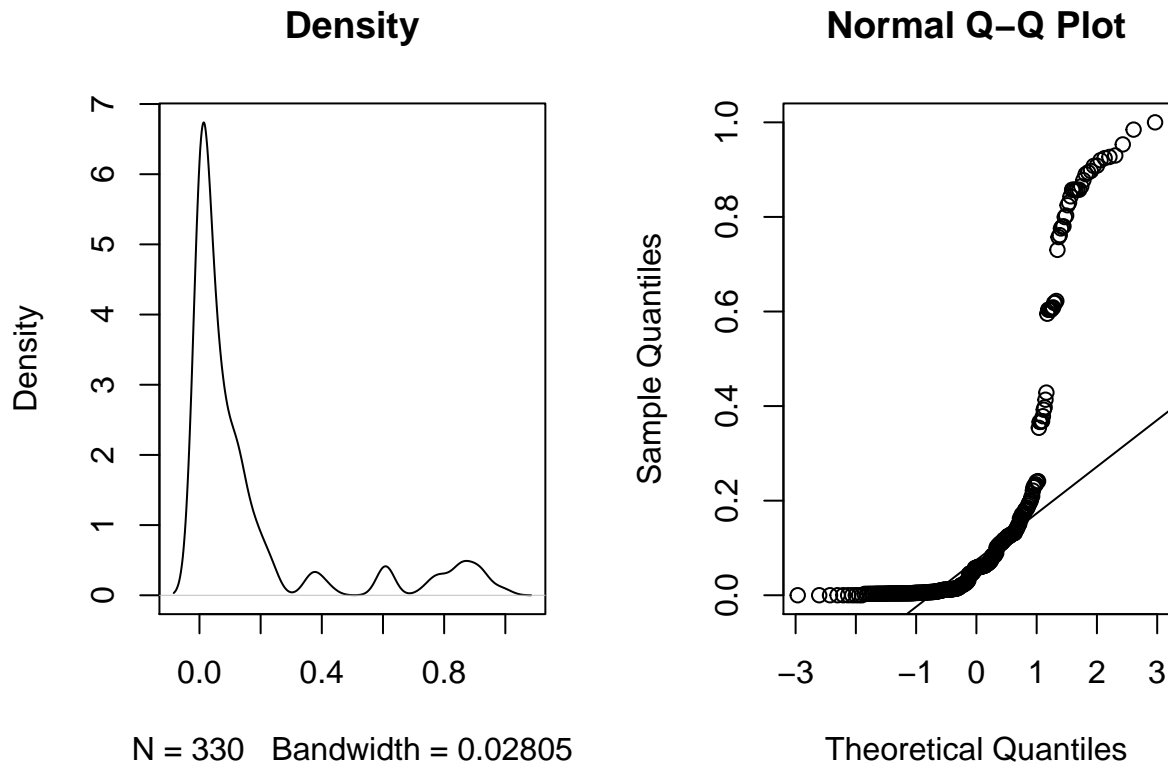
```
par(mfrow=c(1,2))
plot(density(gasto_df$Gasto_Cuidado_Rehabilitacion_Ingresados_norm) ,main="Density")
qqnorm(gasto_df$Gasto_Cuidado_Rehabilitacion_Ingresados_norm)
qqline(gasto_df$Gasto_Cuidado_Rehabilitacion_Ingresados_norm)
```



Al igual que en el caso anterior, en la gráfica Densidad de la variable “Value” para El Gasto en Cuidados Y Rehabilitación en pacientes Ingresados, se observa claramente sesgo hacia la izquierda, por lo que no se considera normalidad. Se puede confirmar observando la gráfica QQplot en la que la línea que grafica qqline sirve de referencia para interpretar el gráfico, no se alinea con los puntos de los valores de la variable “Value”.

Gasto en Hospitales

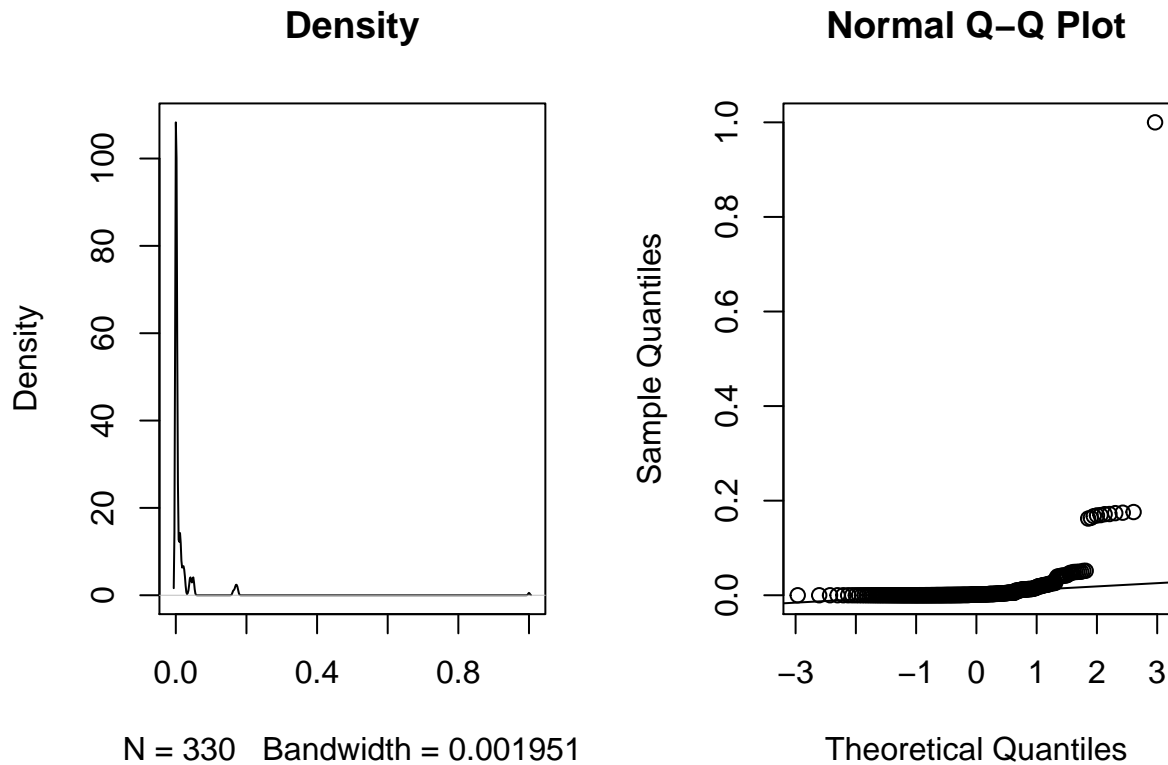
```
par(mfrow=c(1,2))
plot(density(gasto_df$Gasto_Hospitales_norm) ,main="Density")
qqnorm(gasto_df$Gasto_Hospitales_norm)
qqline(gasto_df$Gasto_Hospitales_norm)
```



Se vuelve a repetir la situación anterior, en la que la gráfica Densidad de la variable “Value” para los Hospitales Generales, se observa claramente sesgo hacia la izquierda, por lo que no se considera normalidad. Se puede confirmar observando la gráfica QQplot en la que la línea que grafica qqline sirve de referencia para interpretar el gráfico, no se alinea con los puntos de los valores de la variable “Value”.

Gasto en Hospitales Mentales

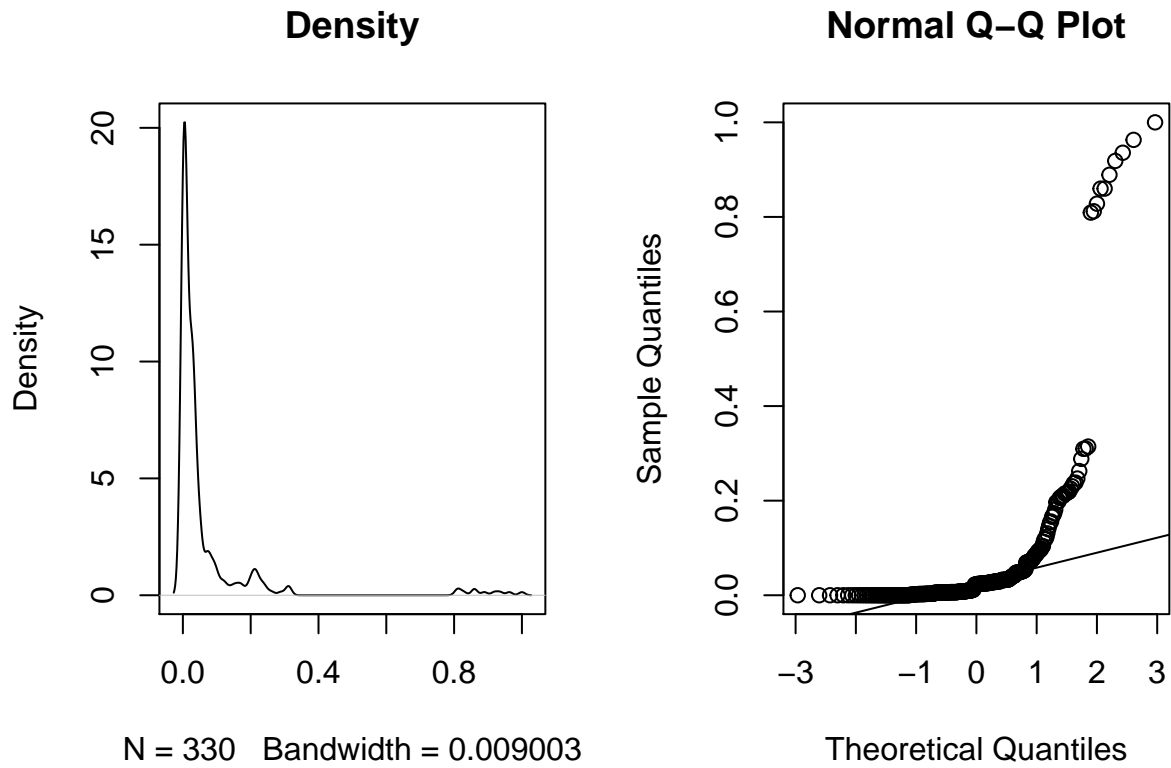
```
par(mfrow=c(1,2))
plot(density(gasto_df$Gasto_Hospitales_Mentales_norm) ,main="Density")
qqnorm(gasto_df$Gasto_Hospitales_Mentales_norm)
qqline(gasto_df$Gasto_Hospitales_Mentales_norm)
```



Se vuelve a repetir la situación anterior, en la que la gráfica Densidad de la variable “Value” para los Hospitales Mentales/Psiquiátricos, se observa claramente sesgo hacia la izquierda, por lo que no se considera normalidad. Se puede confirmar observando la gráfica QQplot en la que la línea que grafica qqline sirve de referencia para interpretar el gráfico, no se alinea con los puntos de los valores de la variable “Value”.

Gasto en Hospitales Especiales

```
par(mfrow=c(1,2))
plot(density(gasto_df$Gasto_Hospitales_Especiales_norm) ,main="Density")
qqnorm(gasto_df$Gasto_Hospitales_Especiales_norm)
qqline(gasto_df$Gasto_Hospitales_Especiales_norm)
```



Se vuelve a repetir la situación anterior, en la que la gráfica Densidad de la variable “Value” para los Hospitales Especiales, se observa claramente sesgo hacia la izquierda, por lo que no se considera normalidad. Se puede confirmar observando la gráfica QQplot en la que la línea que grafica qqline sirve de referencia para interpretar el gráfico, no se alinea con los puntos de los valores de la variable “Value”.