

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

# Análisis de encuestas de hogares con R

## Módulo 5: Análisis gráfico en R

CEPAL - Unidad de Estadísticas Sociales

- 1 Gráficas de variables continuas.
- 2 Diagrama de barras para variables categoricas
- 3 Creando mapas

# Lectura de la base

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

```
data(BigCity, package = "TeachingSampling")  
encuesta <- readRDS("../Data/encuesta.rds")
```

# Definir diseño de la muestra con srvyr

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

Definiendo el diseño muestral, esto se hace de forma análoga a la anterior.

```
library(srvyr)
diseno <- encuesta %>%
  as_survey_design(
    strata = Stratum,
    ids = PSU,
    weights = wk,
    nest = T
  )
```

# Definir nuevas variables

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

Creando nuevas variables, para ello se hace uso de la función `mutate`.

```
disenio <- disenio %>% mutate(  
  pobreza = ifelse(Poverty != "NotPoor", 1, 0),  
  desempleo = ifelse(Employment == "Unemployed", 1, 0),  
  edad_18 = case_when(  
    Age < 18 ~ "< 18 años",  
    TRUE ~ ">= 18 años"  
  )  
)
```

# Sub-grupos

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

Dividiendo la muestra en sub grupos.

```
sub_Urbano <- diseno %>% filter(Zone == "Urban")
sub_Rural <- diseno %>% filter(Zone == "Rural")
sub_Mujer <- diseno %>% filter(Sex == "Female")
sub_Hombre <- diseno %>% filter(Sex == "Male")
```

# Creando tema para las gráficas

Para tener un estilo estandar las gráficas se define el siguiente tema.

```
theme_cepal <- function(...) {  
  theme_light(10) +  
    theme(  
      axis.text.x = element_blank(),  
      axis.ticks.x = element_blank(),  
      axis.text.y = element_blank(),  
      axis.ticks.y = element_blank(),  
      legend.position = "bottom",  
      legend.justification = "left",  
      legend.direction = "horizontal",  
      plot.title = element_text(size = 20, hjust = 0.  
      ...  
    )  
}
```

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

## Gráficas de variables continuas.



# Histogramas

Análisis de encuestas de hogares con R

Gráficas de variables continuas.

Diagrama de barras para variables categoricas

Creando mapas

Las gráficas son realizadas principalmente con la librería `ggplot2` y nos apoyamos en la librería `patchwork` para organizar la visual de las gráficas.

```
library(ggplot2)
library(patchwork)
plot1_Ponde <- ggplot(
  data = encuesta,                # Fuente de datos.
  aes(x = Income, weight = wk)   # Parámetros gráficos
) +
  geom_histogram(                # Parámetro geométrico
    aes(y = ..density..) +      # Parámetros del gráfico
    ylab("") +                  # Nombre para el eje
    ggtitle("Ponderado") +      # Título.
    theme_cpal()                # Aplicando tema
```

# Histogramas

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

De forma análoga se define el gráfico siguiente, note que en este caso se omitió el parámetro `weight`.

```
plot1_SinPonde <-  
  ggplot(encuesta, aes(x = Income)) +  
  geom_histogram(aes(y = ..density..)) +  
  ylab("") +  
  ggtitle("Sin ponderar") +  
  theme_cepal()
```

# Histogramas

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

Para el siguiente dato se tomó la base poblacional (BigCity).

```
plot1_censo <- ggplot(BigCity, aes(x = Income)) +  
  geom_histogram(aes(y = ..density..)) +  
  ylab("") +  
  ggtitle("Poblacional") +  
  theme_cepal() +  
  xlim(0, 2500)  
# Organizando la salida gráfica  
plot1_censo | plot1_Ponde | plot1_SinPonde
```

# Histograma

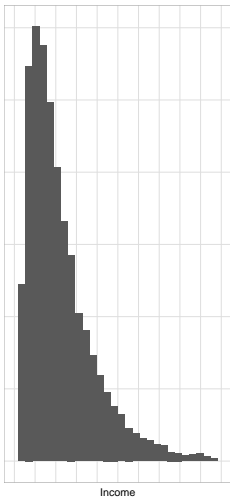
Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

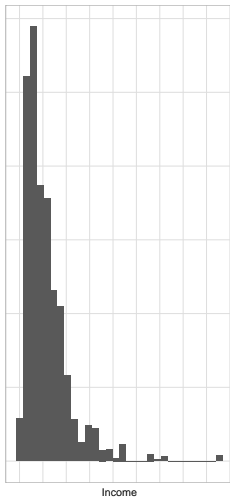
Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

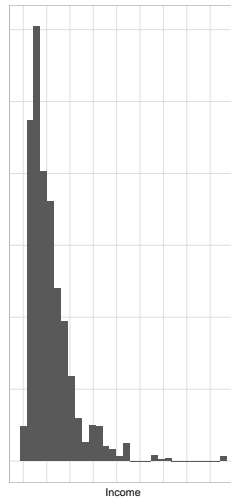
Poblacional



Ponderado



Sin ponderar



# Histogramas

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

Ahora, repetimos la secuencia de gráficos para la variable *Expenditure*

```
plot2_Ponde <- ggplot(  
  data = encuesta,  
  aes(x = Expenditure, weight = wk)  
) +  
  geom_histogram(aes(y = ..density..)) +  
  ylab("") +  
  ggtitle("Ponderado") +  
  theme_cepal()
```

# Histogramas

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

```
plot2_SinPonde <- ggplot(data = encuesta,  
  aes(x = Expenditure)) +  
  geom_histogram(aes(y = ..density..)) +  
  ylab("") +  
  ggtitle("Sin ponderar") +  
  theme_cepal()
```

# Histogramas

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

```
plot2_censo <- ggplot(BigCity, aes(x = Expenditure)) +  
  geom_histogram(aes(y = ..density..)) +  
  ylab("") +  
  ggtitle("Poblacional") +  
  theme_cepal() +  
  xlim(0, 1500)
```

```
plot2_censo | plot2_Ponde | plot2_SinPonde
```

# Histogramas

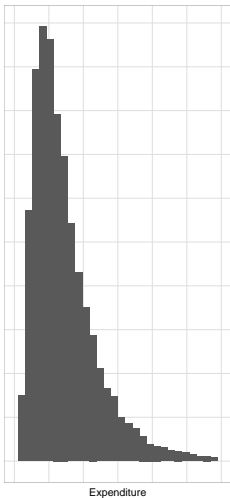
Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

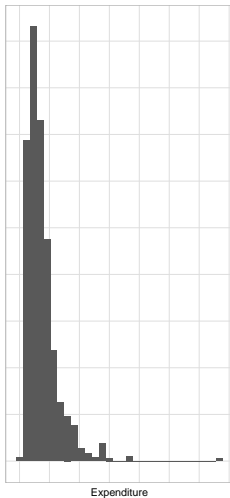
Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

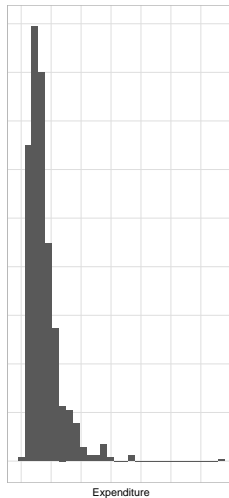
Poblacional



Ponderado



Sin ponderar





# Histogramas por sub-grupos

Análisis de encuestas de hogares con R

Gráficas de variables continuas.

Diagrama de barras para variables categoricas

Creando mapas

Quando el interés es realizar comparaciones entre dos o más agrupaciones, es posible hacer uso del parámetro `fill`, el cual “rellena” las barras del histograma con diferentes colores según sea el grupo.

```
plot3_Ponde <- ggplot(
  encuesta,
  aes(x = Income, weight = wk)
) +
  geom_histogram(
    aes(y = ..density.., fill = Zone),
    alpha = 0.5,
    position = "identity" # Para que las barras no estén apiladas.
  ) +
  ylab("") +
  ggtitle("Ponderado") +
  theme_cepal()
```

# Histogramas por sub-grupos

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

La sintaxis es homologa a la anterior, sin embargo, se retiro el parámetro weight.

```
plot3_SinPonde <- ggplot(encuesta, aes(x = Income)) +  
  geom_histogram(aes(y = ..density.., fill = Zone),  
    alpha = 0.5, position = "identity"  
  ) +  
  ggtitle("Sin ponderar") +  
  theme_cepal() +  
  ylab("")
```

# Histogramas por sub-grupos

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

```
plot3_censo <- ggplot(BigCity, aes(x = Income)) +  
  geom_histogram(aes(y = ..density.., fill = Zone),  
    alpha = 0.5, position = "identity"  
  ) +  
  ggtitle("Poblacional") +  
  theme_cepal() +  
  xlim(0, 1500) +  
  ylab("")  
plot3_censo | plot3_Ponde | plot3_SinPonde
```

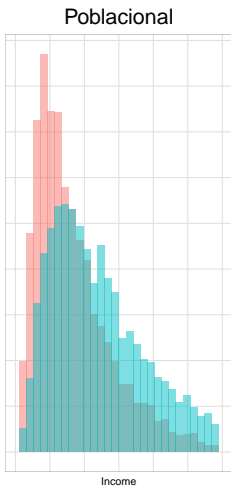
# Histogramas por sub-grupos

Análisis de  
encuestas de  
hogares con R

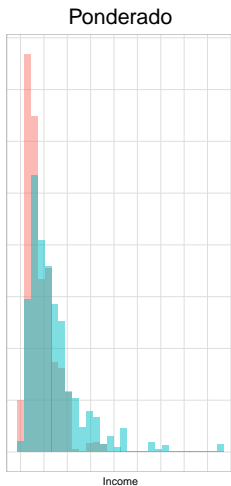
Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

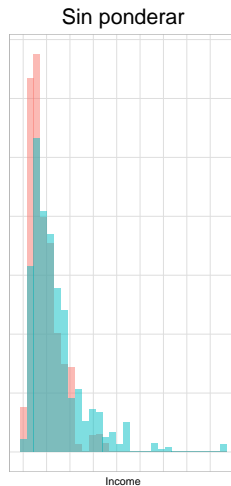
Creando  
mapas



Zone Rural Urban



Zone Rural Urban



Zone Rural Urban

# Histogramas por sub-grupos

Análisis de encuestas de hogares con R

Gráficas de variables continuas.

Diagrama de barras para variables categoricas

Creando mapas

Ahora, repetimos la secuencia de gráficos para la variable *Expenditure*

```
plot4_Ponde <- ggplot(
  encuesta,
  aes(x = Expenditure, weight = wk)
) +
  geom_histogram(aes(y = ..density.., fill = Zone),
    alpha = 0.5, position = "identity"
  ) +
  ylab("") +
  ggtitle("Ponderado") +
  theme_cepal()
```

# Histogramas por sub-grupos

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

```
plot4_SinPonde <- ggplot(  
  encuesta,  
  aes(x = Expenditure)  
) +  
  geom_histogram(aes(y = ..density.., fill = Zone),  
    alpha = 0.5, position = "identity"  
) +  
  ggtitle("Sin ponderar") +  
  theme_cepal() +  
  ylab("")
```

# Histogramas por sub-grupos

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

```
plot4_censo <- ggplot(BigCity, aes(x = Expenditure))  
  geom_histogram(aes(y = ..density.., fill = Zone),  
    alpha = 0.5, position = "identity"  
  ) +  
  ggtitle("Poblacional") +  
  theme_cepal() +  
  xlim(0, 1500) +  
  ylab("")  
plot4_censo | plot4_Ponde | plot4_SinPonde
```

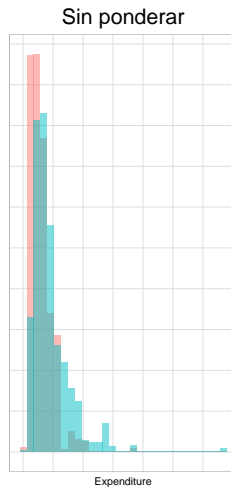
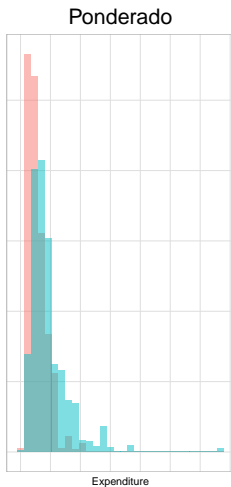
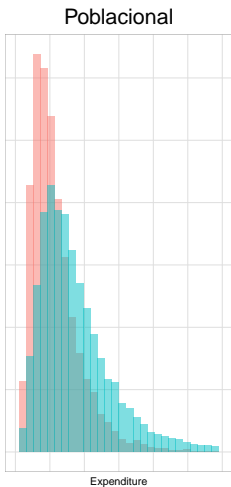
# Histogramas por sub-grupos

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas





# Histogramas por sub-grupos

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

Ahora, repetimos la secuencia de gráficos para la variable *Income*, pero haremos el relleno por la variable *sexo*.

```
plot5_Ponde <- ggplot(  
  encuesta,  
  aes(x = Income, weight = wk)  
) +  
  geom_histogram(aes(y = ..density.., fill = Sex),  
    alpha = 0.5, position = "identity"  
) +  
  ylab("") +  
  ggtitle("Ponderado") +  
  theme_cepal()
```

# Histogramas por sub-grupos

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

```
plot5_SinPonde <- ggplot(encuesta, aes(x = Income)) +  
  geom_histogram(aes(y = ..density.., fill = Sex),  
    alpha = 0.5, position = "identity"  
  ) +  
  ggtitle("Sin ponderar") +  
  theme_cepal() +  
  ylab("")
```

# Histogramas por sub-grupos

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

```
plot5_censo <- ggplot(BigCity, aes(x = Income)) +  
  geom_histogram(aes(y = ..density.., fill = Sex),  
    alpha = 0.5, position = "identity"  
  ) +  
  ggtitle("Poblacional") +  
  theme_cepal() +  
  xlim(0, 1500) +  
  ylab("")  
plot5_censo | plot5_Ponde | plot5_SinPonde
```

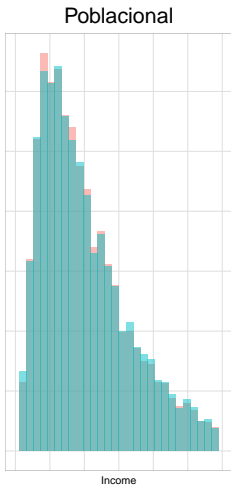
# Histogramas por sub-grupos

Análisis de  
encuestas de  
hogares con R

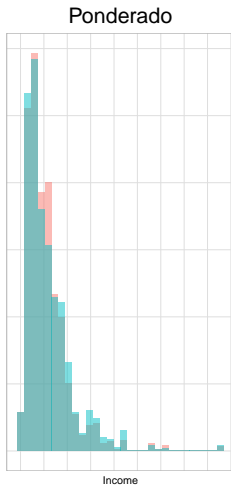
Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

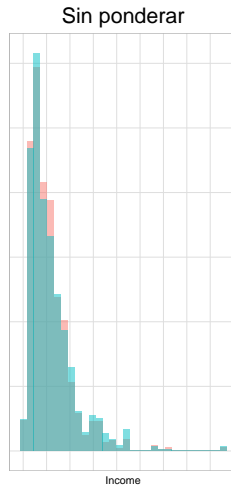
Creando  
mapas



Sex Female Male



Sex Female Male



Sex Female Male

# Histogramas por sub-grupos

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

Ahora, repetimos la secuencia de gráficos para la variable *Expenditure* y el relleno por la variable *sexo*.

```
plot6_Ponde <- ggplot(  
  encuesta,  
  aes(x = Expenditure, weight = wk)  
) +  
  geom_histogram(aes(y = ..density.., fill = Sex),  
    alpha = 0.5, position = "identity"  
) +  
  ylab("") +  
  ggtitle("Ponderado") +  
  theme_cepal()
```

# Histogramas por sub-grupos

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

```
plot6_SinPonde <- ggplot(encuesta, aes(x = Expenditure)) +  
  geom_histogram(aes(y = ..density.., fill = Sex),  
    alpha = 0.5, position = "identity"  
  ) +  
  ggtitle("Sin ponderar") +  
  theme_cepal() +  
  ylab("")
```

```
plot6_censo <- ggplot(BigCity, aes(x = Expenditure)) +  
  geom_histogram(aes(y = ..density.., fill = Sex),  
    alpha = 0.5, position = "identity"  
  ) +  
  ggtitle("Poblacional") +  
  theme_cepal() +  
  xlim(0, 1500) +  
  ylab("")  
plot6_censo | plot6_Ponde | plot6_SinPonde
```

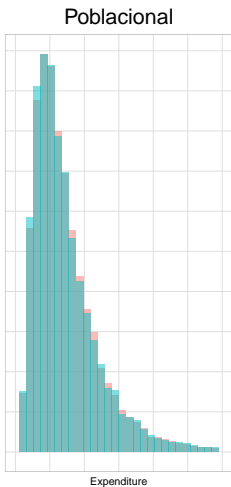
# Histogramas por sub-grupos

Análisis de  
encuestas de  
hogares con R

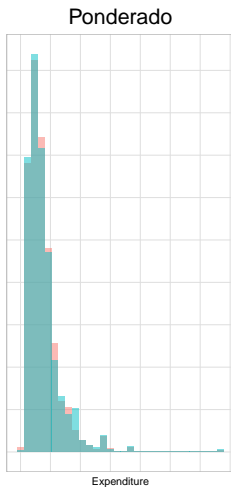
Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

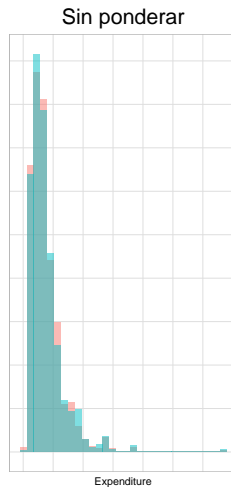
Creando  
mapas



Sex Female Male



Sex Female Male



Sex Female Male

# Agregando densidad

Análisis de  
encuestas de  
hogares con R

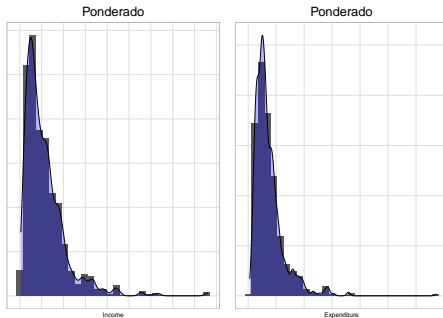
Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

Dadas las cualidades de la librería ggplot2, podemos agregar nuevas capas a la gráfica. Por ejemplo, la densidad con la función `geom_density` e incorporamos el parámetro `alpha` que regula la transparencia del relleno.

```
plot1_Ponde + geom_density(fill = "blue", alpha = 0.3) |  
plot2_Ponde + geom_density(fill = "blue", alpha = 0.3)
```





# Agregando densidad

Análisis de encuestas de hogares con R

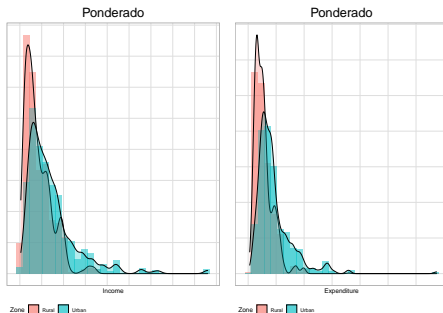
Al hacer `aes(fill = Zone)` permite que la densidad sea agregada para cada una de las agrupaciones.

```
plot3_Ponde + geom_density(aes(fill = Zone), alpha = 0.3) |  
plot4_Ponde + geom_density(aes(fill = Zone), alpha = 0.3)
```

Gráficas de variables continuas.

Diagrama de barras para variables categóricas

Creando mapas



# Agregando densidad

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

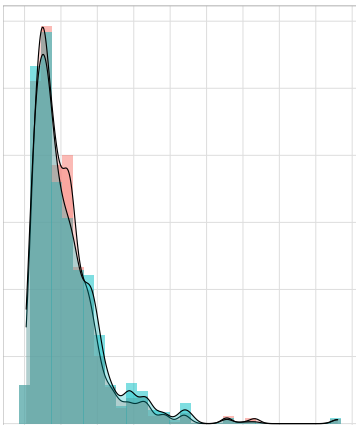
Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

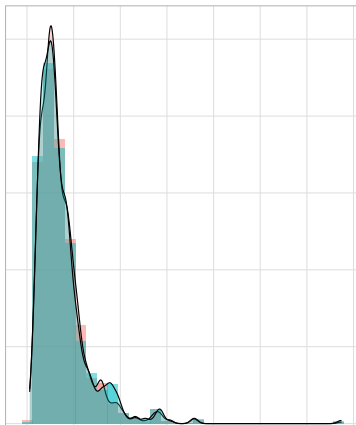
En está oportunidad se agrega la densidad por sexo

```
plot5_Ponde + geom_density(aes(fill = Sex), alpha = 0.3) |  
plot6_Ponde + geom_density(aes(fill = Sex), alpha = 0.3)
```

Ponderado



Ponderado



# Boxplot

Otro gráfico que podemos hacer son los diagramas de caja, para esto deben emplear la función `geom_boxplot`.

```
plot7_Ponde <- ggplot(
  encuesta,
  aes(x = Income, weight = wk)
) +
  geom_boxplot() +
  ggtitle("Ponderado") +
  coord_flip() +
  theme_cepal()

plot8_Ponde <- ggplot(
  encuesta,
  aes(x = Expenditure, weight = wk)
) +
  geom_boxplot() +
  ggtitle("Ponderado") +
  coord_flip() +
  theme_cepal()
```

```
plot7_Ponde | plot8_Ponde
```

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

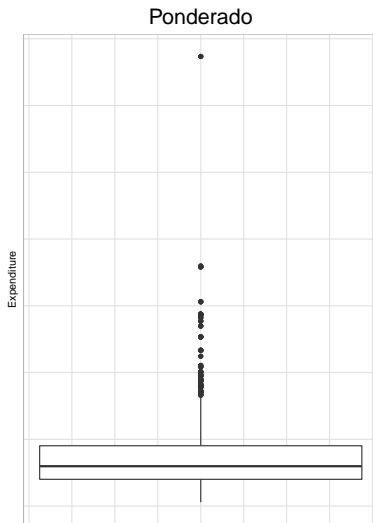
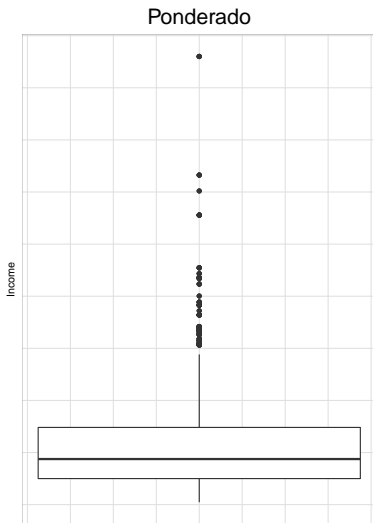
# Boxplot

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas



# Boxplot

Estos diagramas también permiten la comparación entre dos o más niveles de agrupamiento.

```
plot9_Ponde <- ggplot(
  encuesta,
  aes(x = Income, weight = wk)
) +
  geom_boxplot(aes(fill = Zone)) +
  ggtitle("Ponderado") +
  coord_flip() +
  theme_cepal()

plot10_Ponde <- ggplot(
  encuesta,
  aes(x = Expenditure, weight = wk)
) +
  geom_boxplot(aes(fill = Zone)) +
  ggtitle("Ponderado") +
  coord_flip() +
  theme_cepal()

plot9_Ponde | plot10_Ponde
```

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

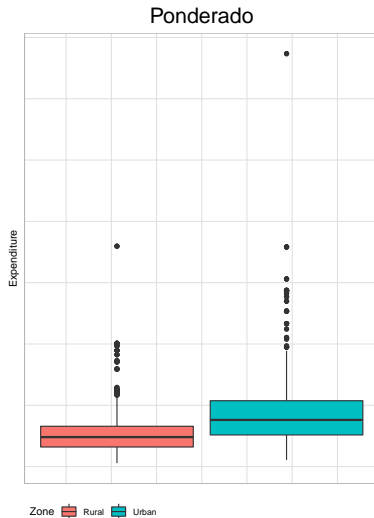
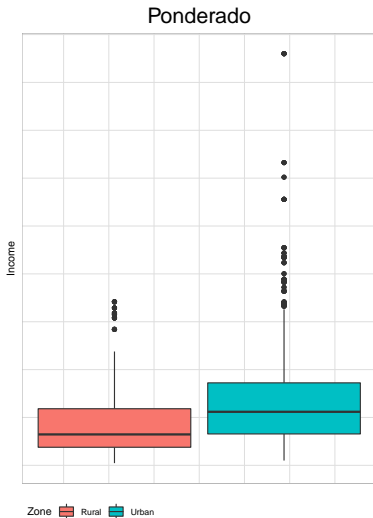
# Boxplot

Análisis de encuestas de hogares con R

Gráficas de variables continuas.

Diagrama de barras para variables categóricas

Creando mapas



# Boxplot

Análisis de encuestas de hogares con R

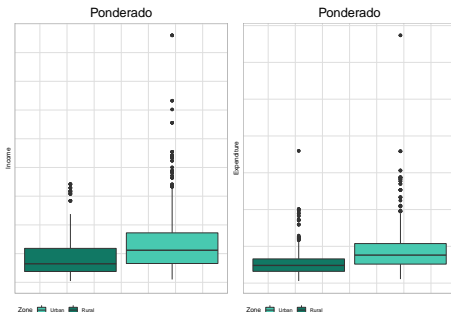
Gráficas de variables continuas.

Diagrama de barras para variables categóricas

Creando mapas

Ahora, si desean personalizar los colores del relleno debe hacer uso de la función `scale_fill_manual`.

```
colorZona <- c(Urban = "#48C9B0", Rural = "#117864")  
plot9_Ponde + scale_fill_manual(values = colorZona) |  
plot10_Ponde + scale_fill_manual(values = colorZona)
```



# Boxplot

Comparando los ingresos y gastos por sexo.

```
plot11_Ponde <- ggplot(  
  encuesta,  
  aes(x = Income, weight = wk)  
) +  
  geom_boxplot(aes(fill = Sex)) +  
  ggtitle("Ponderado") +  
  coord_flip() +  
  theme_cepal()  
  
plot12_Ponde <- ggplot(  
  encuesta,  
  aes(x = Expenditure, weight = wk)  
) +  
  geom_boxplot(aes(fill = Sex)) +  
  ggtitle("Ponderado") +  
  coord_flip() +  
  theme_cepal()  
  
plot11_Ponde | plot12_Ponde
```

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas



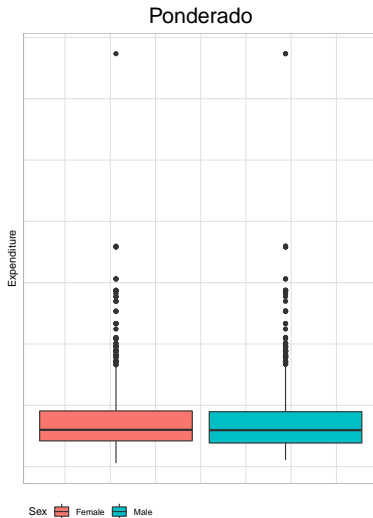
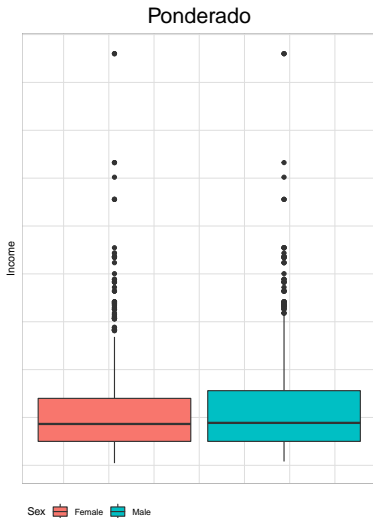
# Boxplot

Análisis de encuestas de hogares con R

Gráficas de variables continuas.

Diagrama de barras para variables categóricas

Creando mapas



# Boxplot

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

Definiendo el color del relleno para hombres y mujeres.

```
colorSex <- c(Male = "#5DADE2", Female = "#2874A6")  
plot11_Ponde + scale_fill_manual(values = colorSex) |  
  plot12_Ponde + scale_fill_manual(values = colorSex)
```

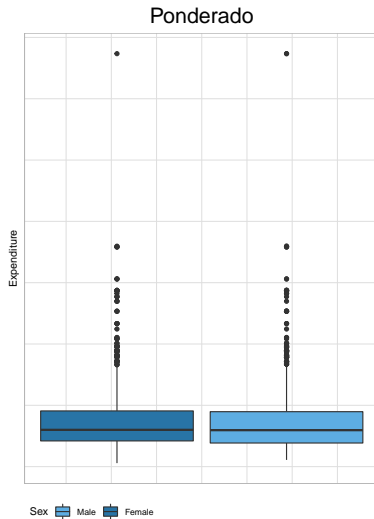
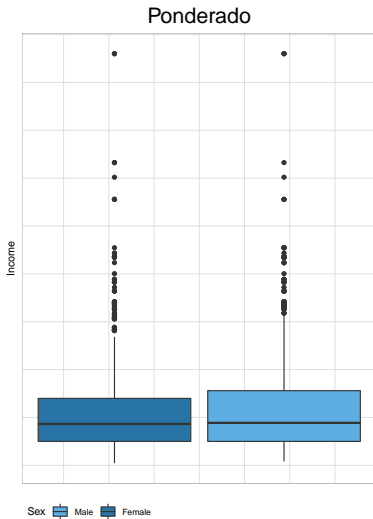
# Boxplot

Análisis de encuestas de hogares con R

Gráficas de variables continuas.

Diagrama de barras para variables categóricas

Creando mapas



# Boxplot

Realizando la comparación para más de dos categorías.

```
plot13_Ponde <- ggplot(  
  encuesta,  
  aes(x = Income, weight = wk)  
) +  
  geom_boxplot(aes(fill = Region)) +  
  ggtitle("Ponderado") +  
  coord_flip() +  
  theme_cepal()
```

```
plot14_Ponde <- ggplot(  
  encuesta,  
  aes(x = Expenditure, weight = wk)  
) +  
  geom_boxplot(aes(fill = Region)) +  
  ggtitle("Ponderado") +  
  coord_flip() +  
  theme_cepal()
```

```
plot13_Ponde | plot14_Ponde
```

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

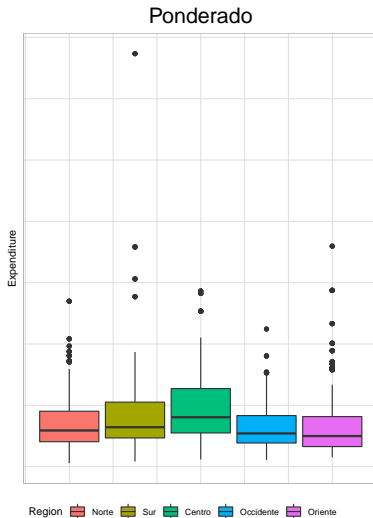
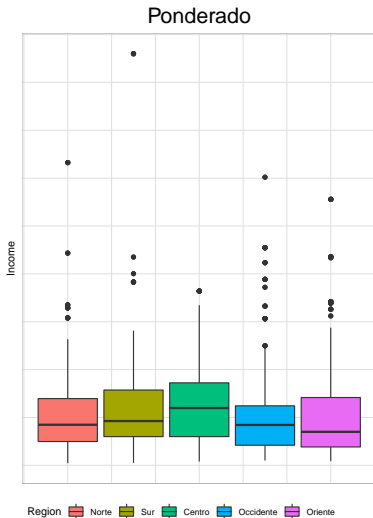
# Boxplot

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas



# Boxplot

Análisis de encuestas de hogares con R

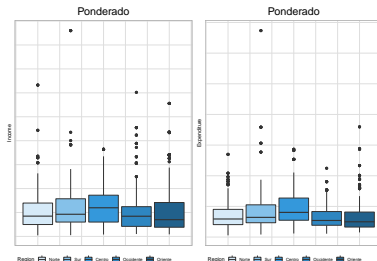
Gráficas de variables continuas.

Diagrama de barras para variables categoricas

Creando mapas

Personalizando los coles cuando hay más de dos categorías.

```
colorRegion <- c(
  Norte = "#D6EAF8", Sur = "#85C1E9",
  Centro = "#3498DB", Occidente = "#2E86C1", Oriente = "#21618C"
)
plot13_Ponde + scale_fill_manual(values = colorRegion) |
plot14_Ponde + scale_fill_manual(values = colorRegion)
```



# Boxplot

Análisis de encuestas de hogares con R

Gráficas de variables continuas.

Diagrama de barras para variables categoricas

Creando mapas

La función `geom_boxplot` permite realizar comparaciones con más de dos variables al tiempo. A continuación se compara los ingresos por sexo en las diferentes zonas.

```
plot15_Ponde <-  
  ggplot(  
    encuesta,  
    aes(x = Income, y = Zone, weight = wk)  
  ) +  
  geom_boxplot(aes(fill = Sex)) +  
  ggtitle("Ponderado") +  
  scale_fill_manual(values = colorSex) +  
  coord_flip()
```

# Boxplot

Análisis de encuestas de hogares con R

Gráficas de variables continuas.

Diagrama de barras para variables categóricas

Creando mapas

De forma análoga podemos realizar la comparación de los gastos por sexo en las diferentes zonas.

```
plot16_Ponde <-  
  ggplot(  
    encuesta,  
    aes(x = Expenditure, y = Zone, weight = wk)  
  ) +  
    geom_boxplot(aes(fill = Sex)) +  
    ggtitle("Ponderado") +  
    scale_fill_manual(values = colorSex) +  
    coord_flip()  
plot15_Ponde / plot16_Ponde
```



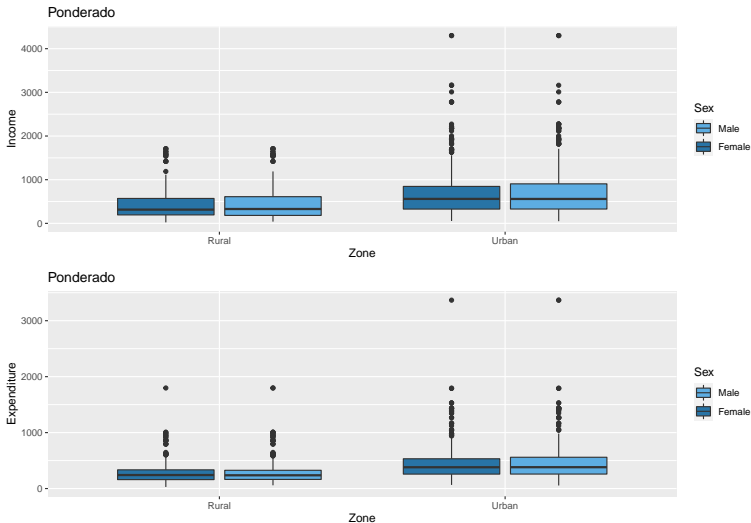
# Boxplot

Análisis de encuestas de hogares con R

Gráficas de variables continuas.

Diagrama de barras para variables categoricas

Creando mapas



# Boxplot

Análisis de encuestas de hogares con R

Gráficas de variables continuas.

Diagrama de barras para variables categoricas

Creando mapas

Podemos extender las comparaciones variables que tienen más de dos categorías.

```
plot17_Ponde <-  
  ggplot(  
    encuesta,  
    aes(x = Income, y = Region, weight = wk)  
  ) +  
  geom_boxplot(aes(fill = Sex)) +  
  ggtitle("Ponderado") +  
  scale_fill_manual(values = colorSex) +  
  coord_flip()
```

# Boxplot

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

```
plot18_Ponde <-  
  ggplot(  
    encuesta,  
    aes(  
      x = Expenditure,  
      y = Region, weight = wk  
    )  
  ) +  
  geom_boxplot(aes(fill = Sex)) +  
  ggtitle("Ponderado") +  
  scale_fill_manual(values = colorSex) +  
  coord_flip()  
  
plot17_Ponde / plot18_Ponde
```

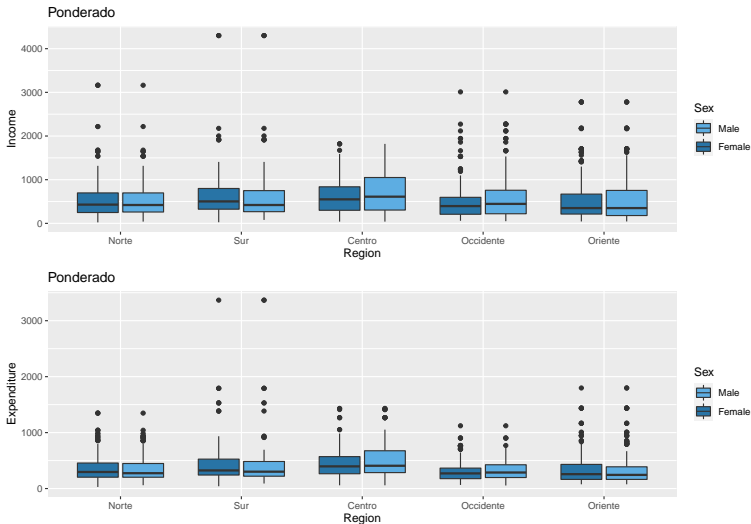
# Boxplot

Análisis de encuestas de hogares con R

Gráficas de variables continuas.

Diagrama de barras para variables categoricas

Creando mapas



# Scaterplot

Un gráfico de interés son los de puntos o Scaterplot, que nos ayudan a ver correlación entre las variables, para realizar este tipo de gráfico usamos la función `geom_point`

```
plot19_Ponde <-  
  ggplot(  
    encuesta,  
    aes(  
      y = Income,  
      x = Expenditure,  
      weight = wk  
    )  
  ) +  
  geom_point() +  
  theme_cepal()  
plot19_Ponde
```

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

# Scatterplot

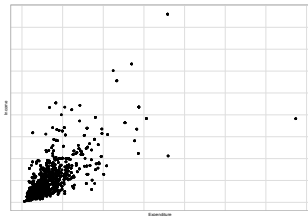
Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

Note, que este caso el parámetro `weight` no esta aportando información visual al gráfico.



# Scaterplot

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

El parámetro `weight` lo podemos usar controlar el tamaño de los puntos de esa forma tener un mejor panorama del comportamiento de la muestra.

```
plot20_Ponde <-  
  ggplot(  
    encuesta,  
    aes(y = Income, x = Expenditure)  
  ) +  
  geom_point(aes(size = wk), alpha = 0.3) +  
  theme_cepal()  
plot20_Ponde
```

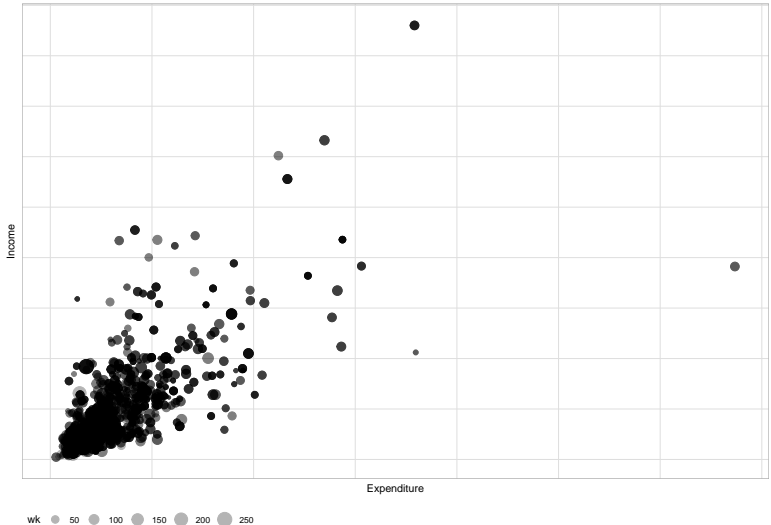
# Scaterplot

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas





# Scaterplot

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

Otra forma de usar la variable `wk`, es asignar la intensidad del color según el valor de la variable.

```
plot21_Ponde <-  
  ggplot(  
    encuesta,  
    aes(y = Income, x = Expenditure)  
  ) +  
  geom_point(aes(col = wk), alpha = 0.3) +  
  theme_cepal()  
plot21_Ponde
```

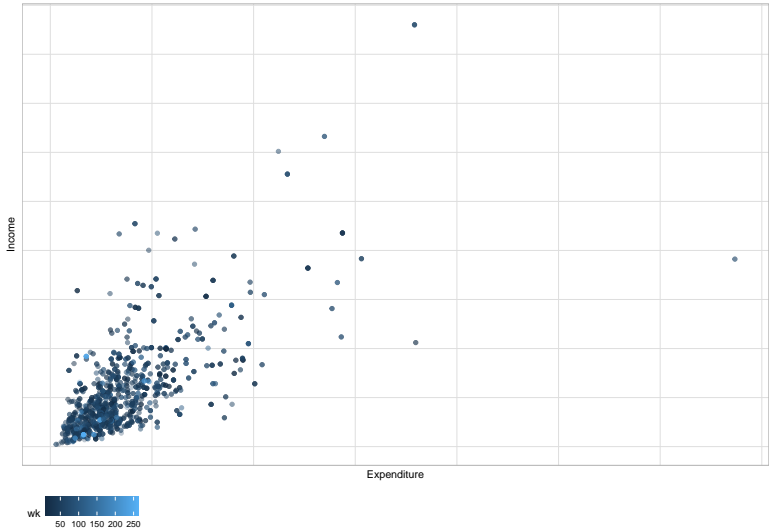
# Scaterplot

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas



# Scaterplot

Análisis de encuestas de hogares con R

Gráficas de variables continuas.

Diagrama de barras para variables categoricas

Creando mapas

Podemos extender las bondades de los gráfico de ggplot2 para obtener mayor información de las muestra. Por ejemplo, agrupar los datos por Zona, para lograr esto se introduce el parámetro shape.

```
plot22_Ponde <-  
  ggplot(  
    encuesta,  
    aes(  
      y = Income, x = Expenditure,  
      shape = Zone) # Formas por zona  
  ) + geom_point(aes(  
    size = wk, color = Zone  
  ), alpha = 0.3) +  
  labs(size = "Peso") +  
  scale_color_manual(values = colorZona) +  
  theme cepal()
```

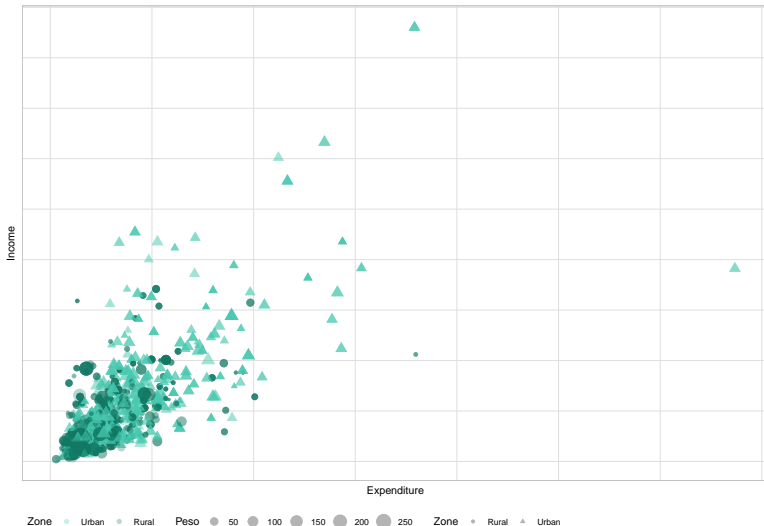
# Scaterplot

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas



# Scaterplot

De forma similar podemos obtener el resultado por sexo.

```
plot23_Ponde <-  
  ggplot(  
    encuesta,  
    aes(  
      y = Income,  
      x = Expenditure,  
      shape = Sex  
    )  
  ) +  
  geom_point(aes(  
    size = wk,  
    color = Sex  
  ),  
  alpha = 0.3  
  ) +  
  labs(size = "Peso") +  
  scale_color_manual(values = colorSex) +  
  theme_cepal()  
plot23_Ponde
```

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

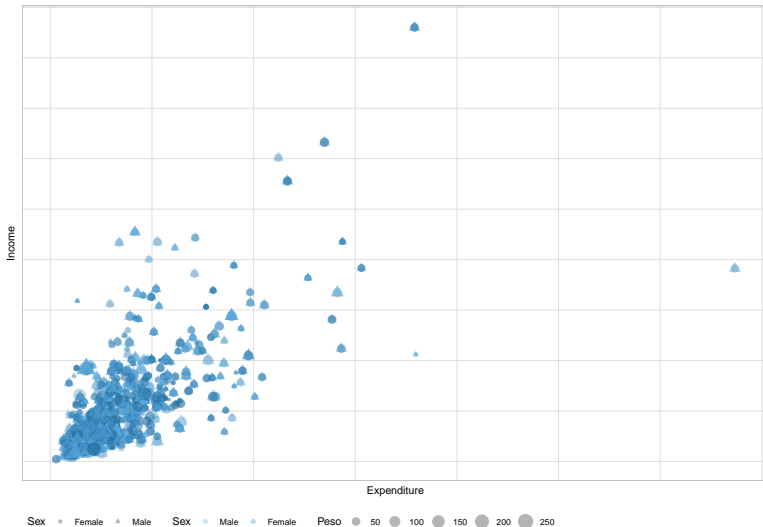
# Scaterplot

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas



# Scaterplot

Un resultado equivalente se obtiene por región.

```
plot24_Ponde <-  
  ggplot(  
    encuesta,  
    aes(  
      y = Income,  
      x = Expenditure,  
      shape = Region  
    )  
  ) +  
  geom_point(aes(  
    size = wk,  
    color = Region  
  ),  
  alpha = 0.3  
  ) +  
  labs(size = "Peso") +  
  scale_color_manual(values = colorRegion) +  
  theme_cepal()  
plot24_Ponde
```

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

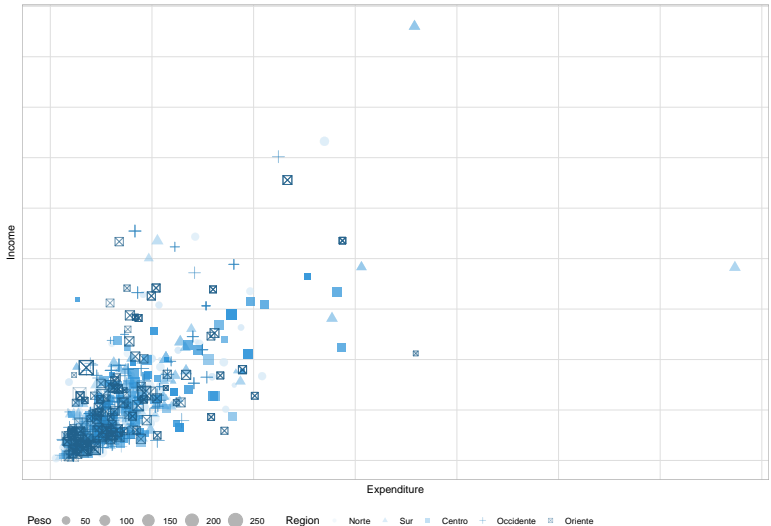
# Scaterplot

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas





Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

## Diagrama de barras para variables categoricas

# Diagrama de barras

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

Para realizar estos gráfico un primer paso es realizar las estimaciones puntuales.

```
(tamano_zona <- disenio %>%  
  group_by(Zone) %>%  
  summarise(  
    Nd = survey_total(vartype = c("se", "ci"))  
  ))
```

Zone	Nd	Nd_se	Nd_low	Nd_upp
Rural	72102	3062	66039	78165
Urban	78164	2847	72526	83802

# Diagrama de barras

Análisis de encuestas de hogares con R

Gráficas de variables continuas.

Diagrama de barras para variables categoricas

Creando mapas

```
plot25_Ponde <- ggplot(  
  data = tamano_zona, # fuente de los datos  
  aes(  
    x = Zone,          # Valores en el eje x  
    y = Nd,            # Altura de la barras  
    ymax = Nd_upp,     # Limite superior del IC  
    ymin = Nd_low,     # Limite inferior del IC  
    fill = Zone        # Color del relleno  
  )  
) +  
  geom_bar(  
    stat = "identity", # Valor incluido en la tabla  
    position = "dodge") +  
  geom_errorbar(      # Gráfica del IC.  
    position = position_dodge(width = 0.9),  
    width = 0.3  
  ) +  
  theme_bw()  
plot25_Ponde
```

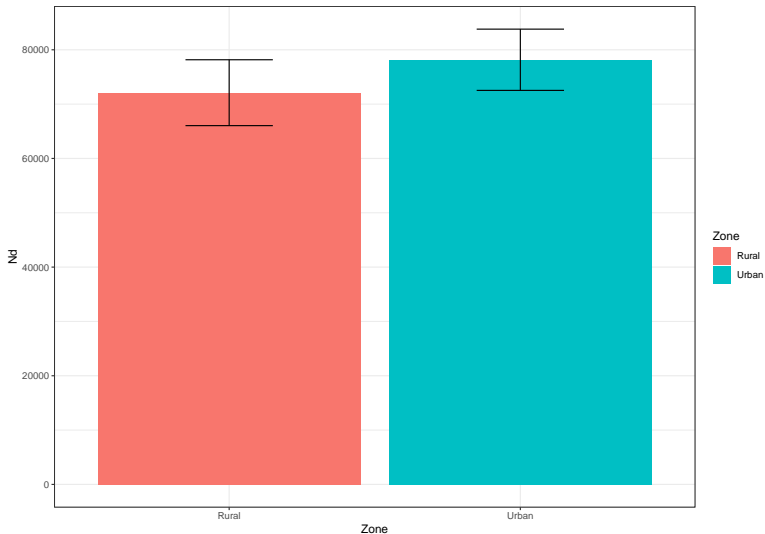
# Diagrama de barras

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas



# Diagrama de barras

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

Como se ha visto en los gráficos anteriores podemos extender a muchas categorías.

```
(tamano_pobreza <- disenio %>%  
  group_by(Poverty) %>%  
  summarise(  
    Nd = survey_total(vartype = c("se", "ci"))  
  ))
```

Poverty	Nd	Nd_se	Nd_low	Nd_upp
NotPoor	91398	4395	82696	100101
Extreme	21519	4949	11719	31319
Relative	37349	3695	30032	44666

# Diagrama de barras

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

El gráfico se obtiene con una sintaxis homologa a la anterior.

```
plot26_Ponde <- ggplot(  
  data = tamano_pobreza,  
  aes(  
    x = Poverty,  
    y = Nd,  
    ymax = Nd_upp,  
    ymin = Nd_low,  
    fill = Poverty  
  )  
) +  
  geom_bar(stat = "identity", position = "dodge") +  
  geom_errorbar(  
    position = position_dodge(width = 0.9),  
    width = 0.3  
  ) +  
  theme_bw()  
plot26_Ponde
```

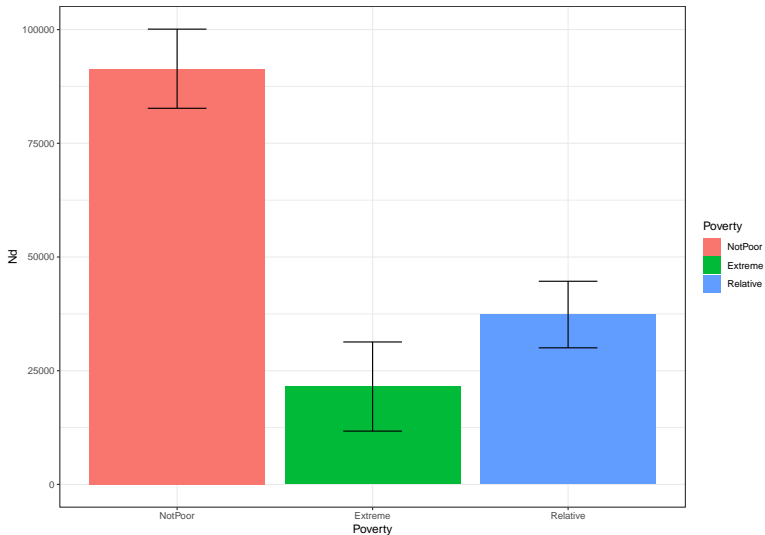
# Diagrama de barras

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas



# Diagrama de barras

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

De forma similar a los gráficos de Caja es posible realizar comparaciones entre más dos variables.

```
tamano_ocupacion_pobreza <- diseno %>%  
  group_by(desempleo, Poverty) %>%  
  summarise(  
    Nd = survey_total(vartype = c("se", "ci"))  
  ) %>% as.data.frame() %>%  
  mutate(  
    desempleo = ifelse(is.na(desempleo),  
                        "Ninos",desempleo))
```



# Diagrama de barras

El gráfico para la tabla anterior queda de la siguiente forma.

```
plot27_Ponde <-  
  ggplot(  
    data = tamano_ocupacion_pobreza,  
    aes(  
      x = Poverty,  
      y = Nd,  
      ymax = Nd_upp,  
      ymin = Nd_low,  
      fill = as.factor(desempleo)  
    )  
  ) +  
  geom_bar(stat = "identity", position = "dodge") +  
  geom_errorbar(  
    position = position_dodge(width = 0.9),  
    width = 0.3  
  ) +  
  theme_bw()  
plot27_Ponde
```

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

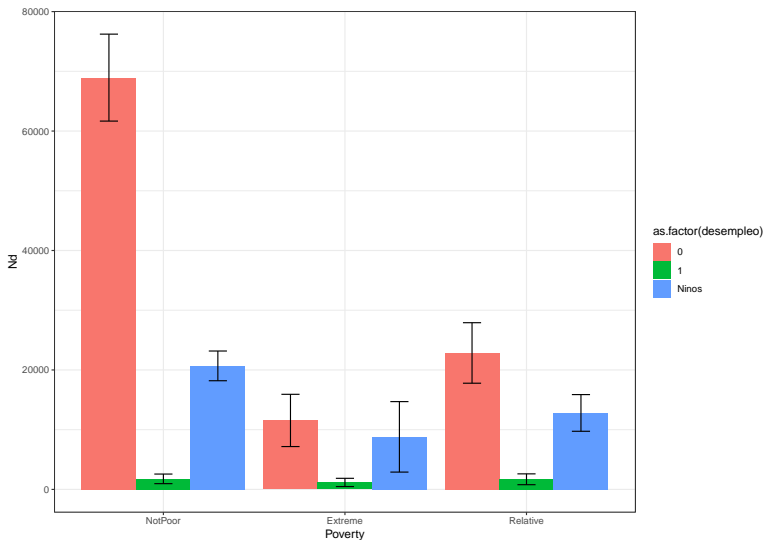
# Diagrama de barras

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas



# Diagrama de barras

En estos gráficos podemos presentar proporciones por variables.

```
(prop_ZonaH_Pobreza <- sub_Hombre %>%  
  group_by(Zone, Poverty) %>%  
  summarise(  
    prop = survey_prop(  
      vartype = c("se", "ci")  
    )  
  ) %>%  
  data.frame())
```

Zone	Poverty	prop	prop_se	prop_low	prop_upp
Rural	NotPoor	0.5488	0.0626	0.4248	0.6729
Rural	Extreme	0.1975	0.0675	0.0640	0.3311
Rural	Relative	0.2536	0.0372	0.1799	0.3274
Urban	NotPoor	0.6599	0.0366	0.5874	0.7324

# Diagrama de barras

Análisis de encuestas de hogares con R

Gráficas de variables continuas.

Diagrama de barras para variables categóricas

Creando mapas

Después de tener la tabla con los valores a presentar el gráfico se realiza con la siguiente sintaxis.

```
plot28_Ponde <- ggplot(  
  data = prop_ZonaH_Pobreza,  
  aes(  
    x = Poverty, y = prop,  
    ymax = prop_upp, ymin = prop_low,  
    fill = Zone  
  )) +  
  geom_bar(stat = "identity", position = "dodge") +  
  geom_errorbar(  
    position = position_dodge(width = 0.9),  
    width = 0.3  
  ) + scale_fill_manual(values = colorZona) +  
  theme_bw();plot28_Ponde
```

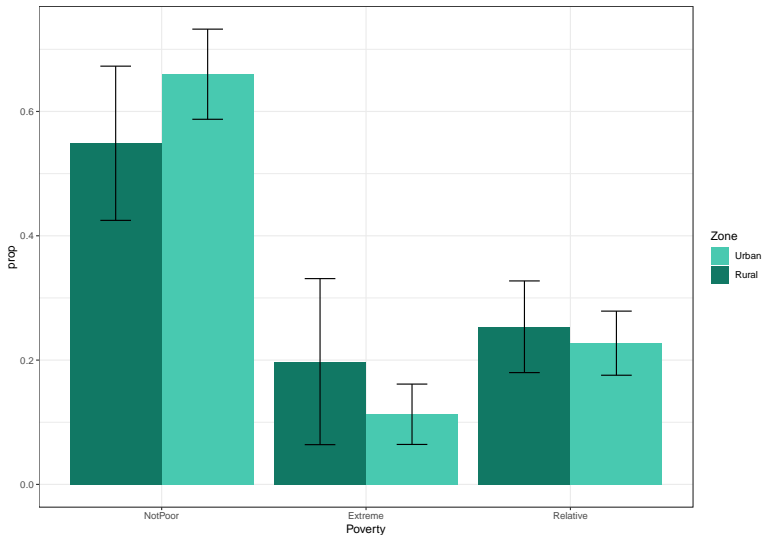
# Diagrama de barras

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas



# Diagrama de barras

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

Proporción de hombres en condición de pobreza por región

```
prop_RegionH_Pobreza <- sub_Hombre %>%  
  group_by(Region, pobreza) %>%  
  summarise(  
    prop = survey_prop(vartype = c("se", "ci"))  
  ) %>%  
  data.frame()
```

# Diagrama de barras

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

```
plot29_Ponde <- ggplot(  
  data = prop_RegionH_Pobreza,  
  aes(  
    x = Region, y = prop,  
    ymax = prop_upp, ymin = prop_low,  
    fill = as.factor(pobreza)  
  )  
) +  
  geom_bar(stat = "identity", position = "dodge") +  
  geom_errorbar(  
    position = position_dodge(width = 0.9),  
    width = 0.3  
  ) +  
  theme_bw()  
plot29_Ponde
```

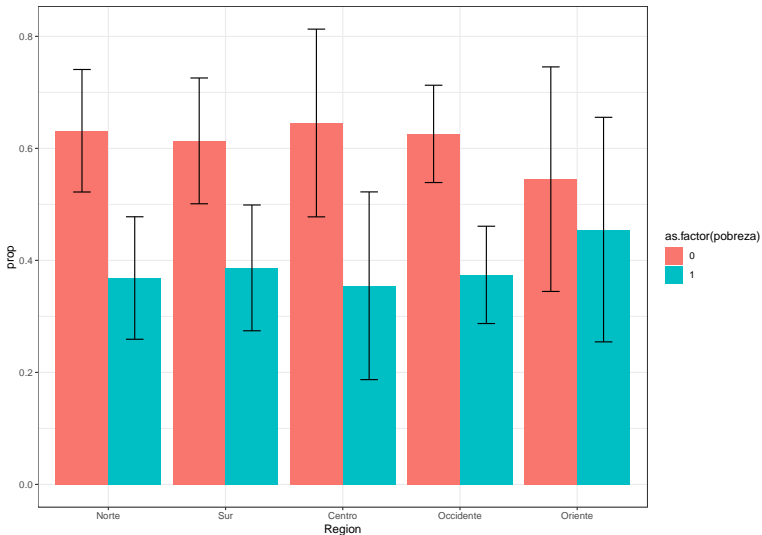
# Diagrama de barras

Análisis de encuestas de hogares con R

Gráficas de variables continuas.

Diagrama de barras para variables categóricas

Creando mapas





Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

## Creando mapas

# Mapas con tmap.

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

Para realizar el mapa hay que contar con el archivo de *shepefile*

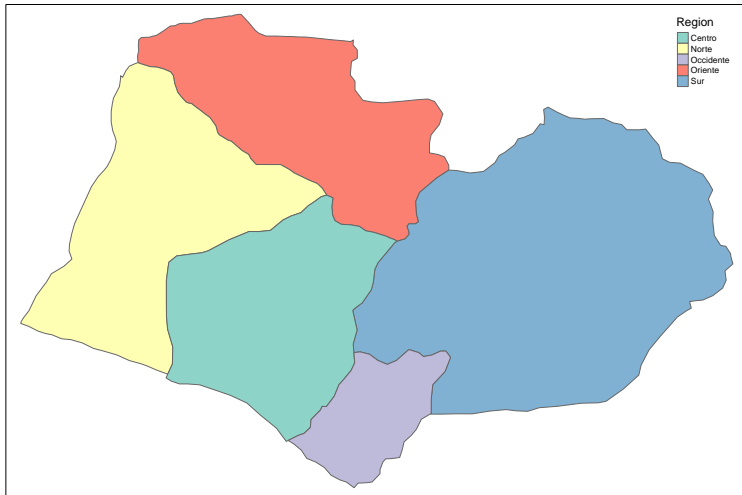
```
library(sf)
library(tmap)
shapeBigCity <- read_sf("/Data/shapeBigCity/BigCity.s
```

La forma más simple de crear el mapa es con la sintaxis.

```
tm_shape(shapeBigCity) + # shepefile
  tm_polygons(col = "Region") # Variable de interés.
```

# Mapas con tmap.

El mapa resultante es:



Análisis de encuestas de hogares con R

Gráficas de variables continuas.

Diagrama de barras para variables categoricas

Creando mapas

# Mapas con tmap.

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

```
brks <- c(0, .2, .4, .6, 0.8, 1)
shape_temp <- tm_shape(
  shapeBigCity %>%          # shapefile
  left_join(                # Agregando una variable
    prop_RegionH_Pobreza %>%
      filter(pobreza == 1), # Filtrando el nivel de interés.
    by = "Region"
  )
)

shape_temp + tm_polygons(
  "prop",                  # Nombre de la columna
  breaks = brks,           # Puntos de corte
  title = "pobreza",      # Titulo del labels.
  palette = "YlOrRd"       # Paleta y dirección de colores
)
```

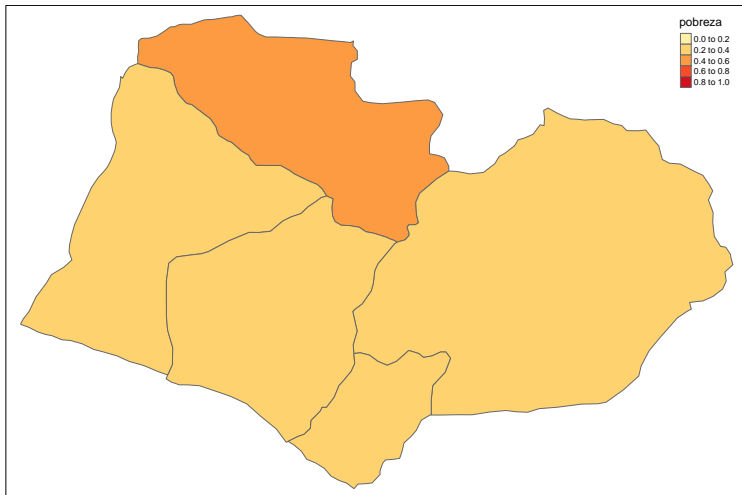
# Mapas con tmap.

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas



# Mapas con tmap.

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

```
(prom_region <- svyby(~Income, ~Region, disen,
  svymean,
  na.rm = T, covmat = TRUE,
  vartype = c("cv")
))
brks <- c(0, 0.2, 1)
shape_temp <- tm_shape(
  shapeBigCity %>%
    left_join(
      prom_region,
      by = "Region"
    )
)

shape_temp + tm_polygons(
  "cv",
  breaks = brks,
  title = "cv",
  palette = c("#FFFFFF", "#000000"),
) + tm_layout(asp = 0)
```

# Mapas con tmap.

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

	Region	Income	cv
Norte	Norte	552.4	0.1002
Sur	Sur	625.8	0.0997
Centro	Centro	650.8	0.0945
Occidente	Occidente	517.0	0.0894
Oriente	Oriente	541.8	0.1323

# Mapas con tmap.

Análisis de encuestas de hogares con R

Gráficas de variables continuas.

Diagrama de barras para variables categoricas

Creando mapas

```
prom_region_Sex <- diseno %>%
  group_by(Region, Zone, Sex, pobreza) %>%
  summarise(prop = survey_mean(vartype = "cv")) %>%
  filter(pobreza == 1, Zone == "Rural", Sex == "Female")

shape_temp <- tm_shape(
  shapeBigCity %>%
    left_join(
      prom_region_Sex,
      by = "Region"
    )
)

shape_temp + tm_polygons(
  "prop",
  title = "Pobreza",
) + tm_layout(asp = 0)
```



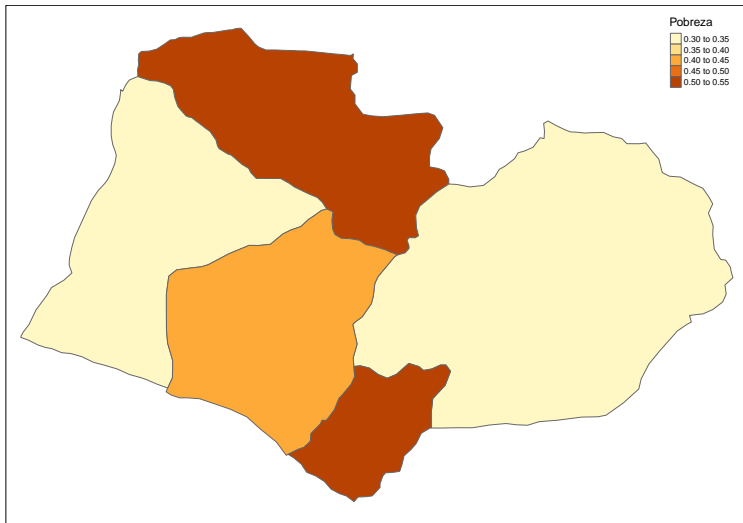
# Mapas con tmap.

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas



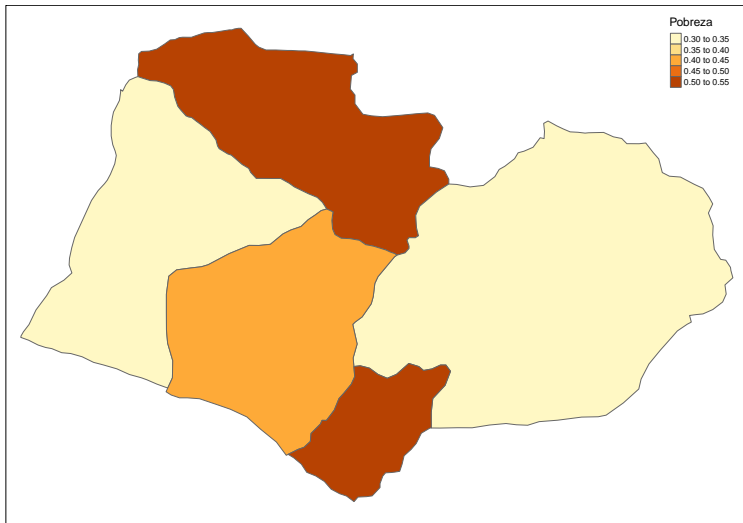
# Mapas con tmap.

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas



# Mapas con tmap.

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

```
shape_temp + tm_polygons(  
  "prop_cv",  
  title = "cv",  
  palette = c("#FFFFFF", "#000000"),  
  breaks = c(0, 0.2, 1)  
) + tm_layout(asp = 0)
```

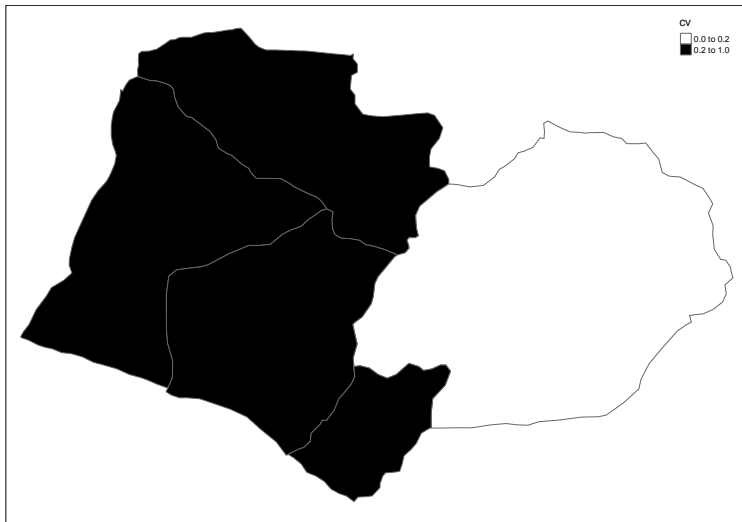
# Mapas con tmap.

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas



# Mapas con ggplot

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

```
library(biscale)
library(cowplot)
temp_shape <- shapeBigCity %>%
  left_join(
    prom_region_Sex,
    by = "Region"
  )
k <- 3
datos.RM.bi <- bi_class(temp_shape,
  y = prop, x = prop_cv, dim = k,
  style = "fisher"
)
map.RM <- ggplot() +
  geom_sf(
    data = datos.RM.bi,
    aes(fill = bi_class, geometry = geometry),
    colour = "white", size = 0.1
  ) +
  bi_scale_fill(pal = "GrPink", dim = k) +
  bi_theme() +
  theme(legend.position = "none")
```

# mapas con ggplot

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

```
# Crear la leyenda para el mapa
legend1 <- bi_legend(
  pal = "GrPink", dim = k,
  xlab = "Coeficiente de variaci<U+00F3>n",
  ylab = "Pobreza", size = 8
)

mapa1 <- ggdraw() +
  draw_plot(map.RM, 0, 0, 1, scale = 0.7) +
  draw_plot(legend1, 0.75, 0.4, 0.2, 0.2, scale = 1) +
  draw_text("Estimaciones directas de la pobreza en la mujer rural",
    vjust = -13, size = 18
  )

mapa1
```

# mapas con ggplot

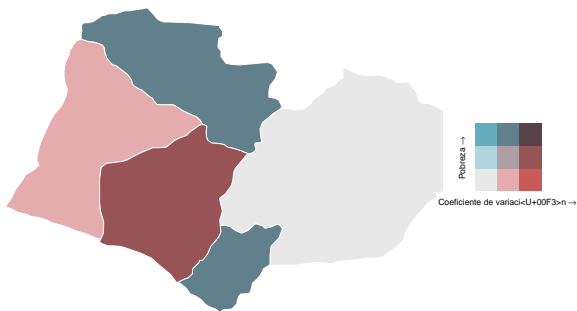
Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

## Estimaciones directas de la pobreza en la mujer rural



# ¡Gracias!

Análisis de  
encuestas de  
hogares con R

Gráficas de  
variables  
continuas.

Diagrama de  
barras para  
variables  
categoricas

Creando  
mapas

*Email:* [andres.gutierrez@cepal.org](mailto:andres.gutierrez@cepal.org)