Gráficas con ponderación

CEPAL

24/2/2022

Lectura de la base

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

Definir diseño de la muestra con `srvyr`

```
library(srvyr)

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

definir nuevas variables

```
diseno <- diseno %>% 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

Extraer sub-grupos de la encuesta.

```
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

```
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 = 0.5),
        ...)
```

Gráficas de variables continuas.

Histogramas

```
library(ggplot2) ; library(patchwork)
plot1_Ponde <- ggplot(encuesta,
                aes(x = Income, weight = wk))+
  geom_histogram(aes(y = ..density..)) + ylab("") +
  ggtitle("Ponderado") + theme_cepal()
```

Histogramas

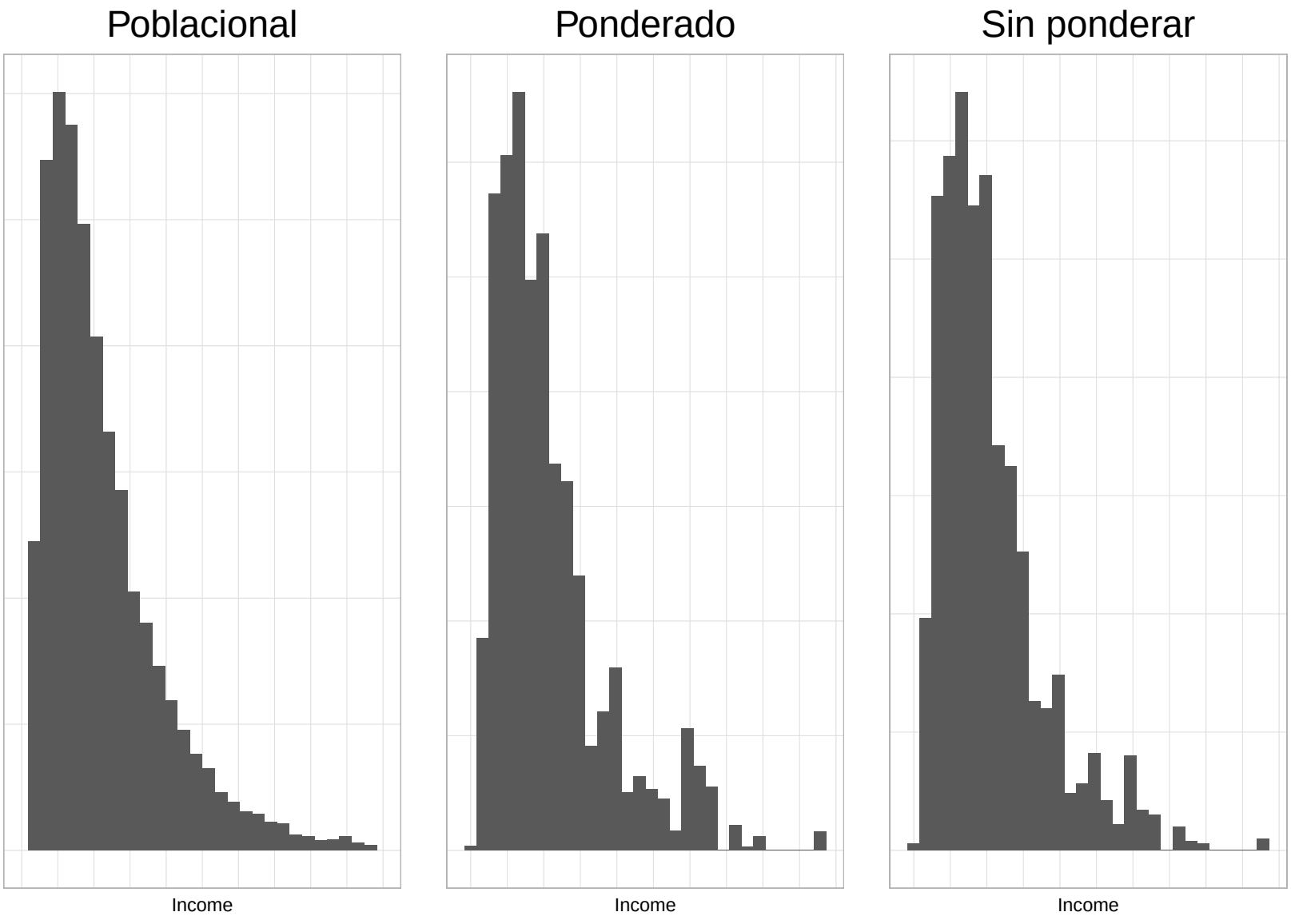
```
plot1_SinPonde <-
  ggplot(encuesta, aes(x = Income))+
  geom_histogram(aes(y = ..density..)) + ylab("") +
    ggtitle("Sin ponderar") + theme_cepal()
```

Histogramas

```
plot1_censo <- ggplot(BigCity, aes(x = Income))+
  geom_histogram(aes(y = ..density..)) + ylab("") +
   ggtitle("Poblacional") + theme_cepal() +
   xlim(0, 2500)

plot1_censo | plot1_Ponde | plot1_SinPonde
```

Histograma



Histogramas

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

Histogramas

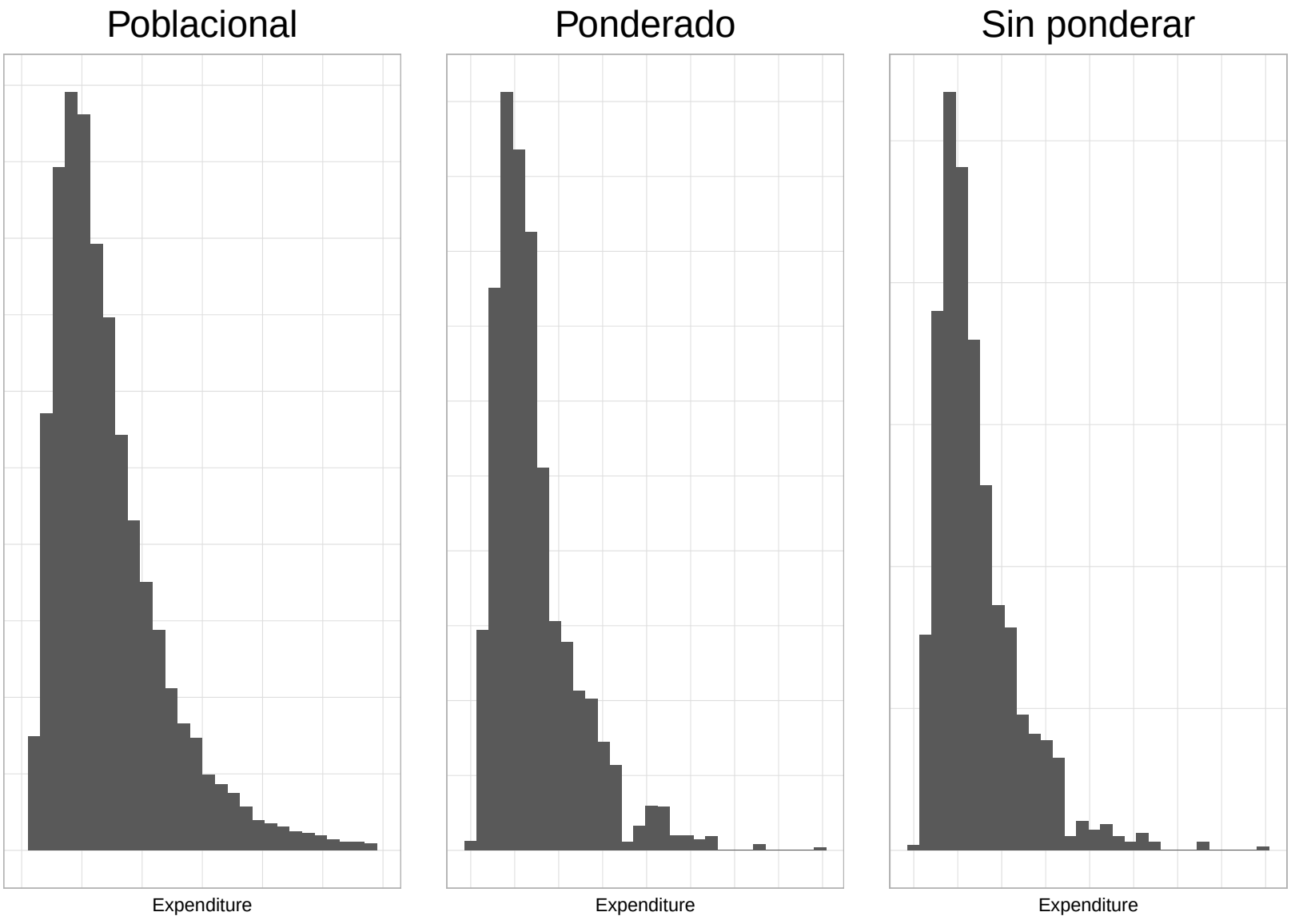
```
plot2_SinPonde <- ggplot(encuesta, aes(x = Expenditure))+
  geom_histogram(aes(y = ..density..)) + ylab("") +
    ggtitle("Sin ponderar") + theme_cepal()
```

Histogramas

```
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



Histogramas por sub-grupos

```
plot3_Ponde <- ggplot(encuesta,
                      aes(x = Income, weight = wk)) +
  geom_histogram(aes(y = ..density.., fill = Zone) ,
                 alpha = 0.5, position = "identity") +
  ylab("") +  ggtitle("Ponderado") + theme_cepal()
```

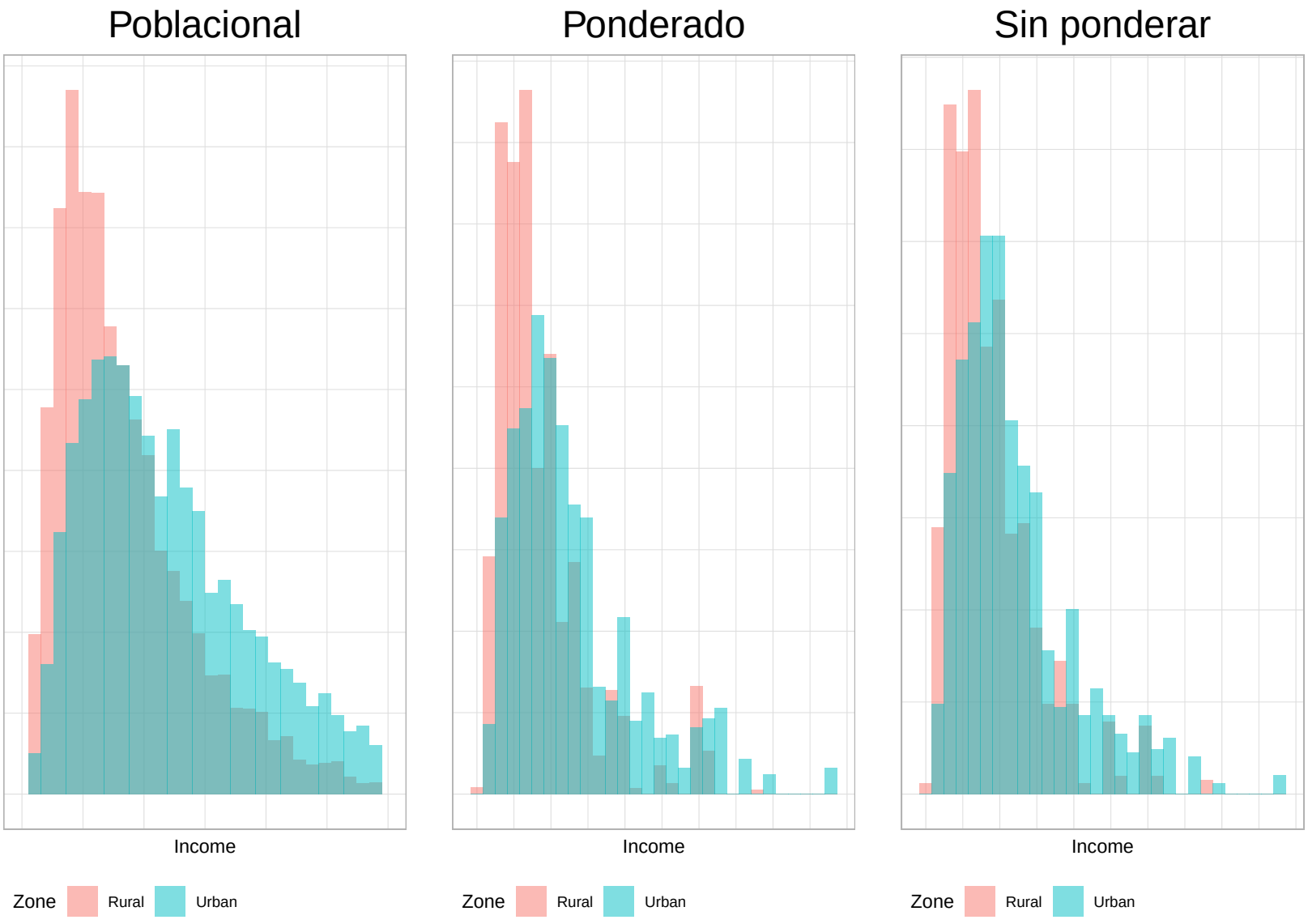
Histogramas por sub-grupos

```
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

```
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



Histogramas por sub-grupos

```
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()
```

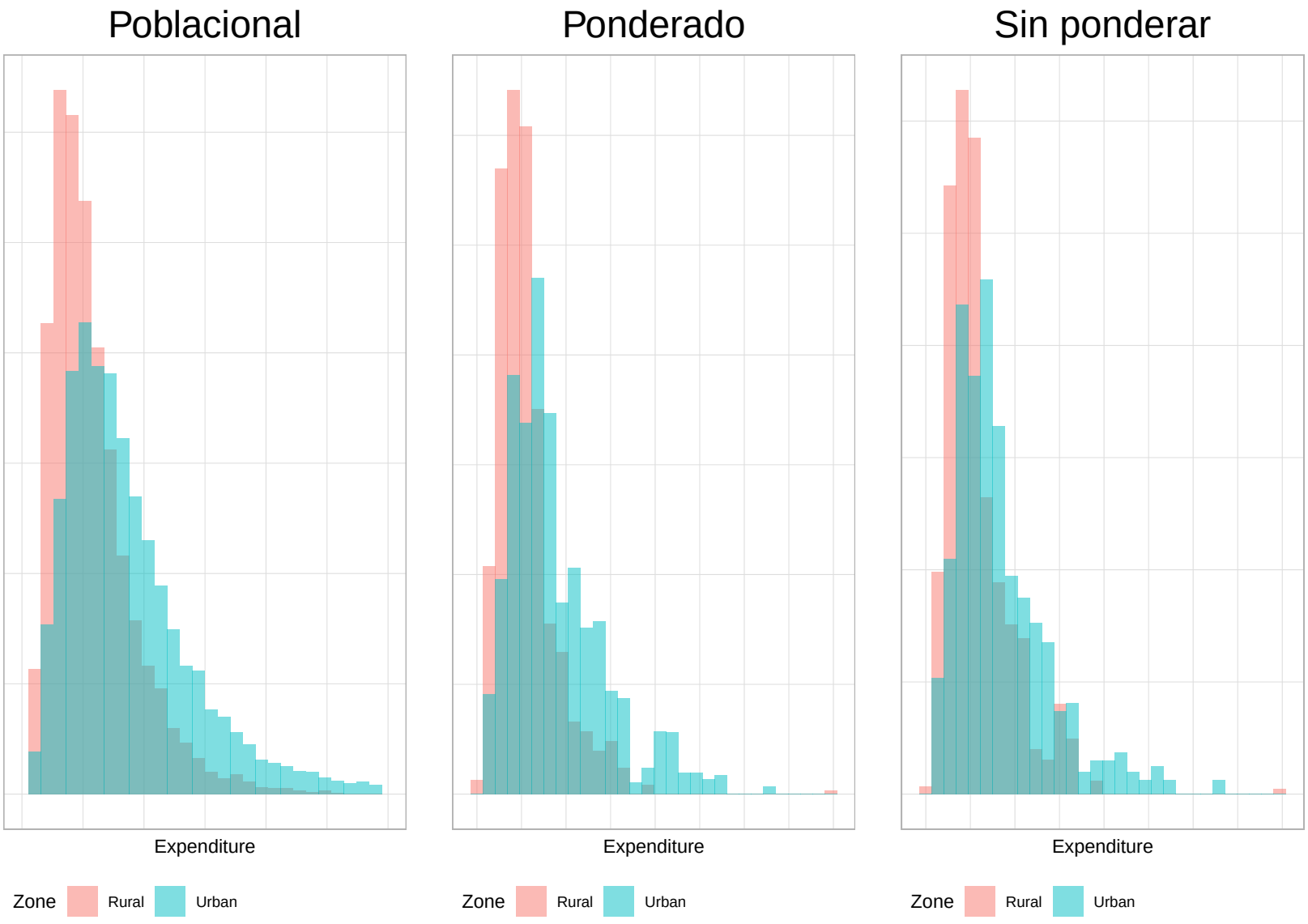
Histogramas por sub-grupos

```
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

```
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



Histogramas por sub-grupos

```
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

```
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

```
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

Histogramas por sub-grupos

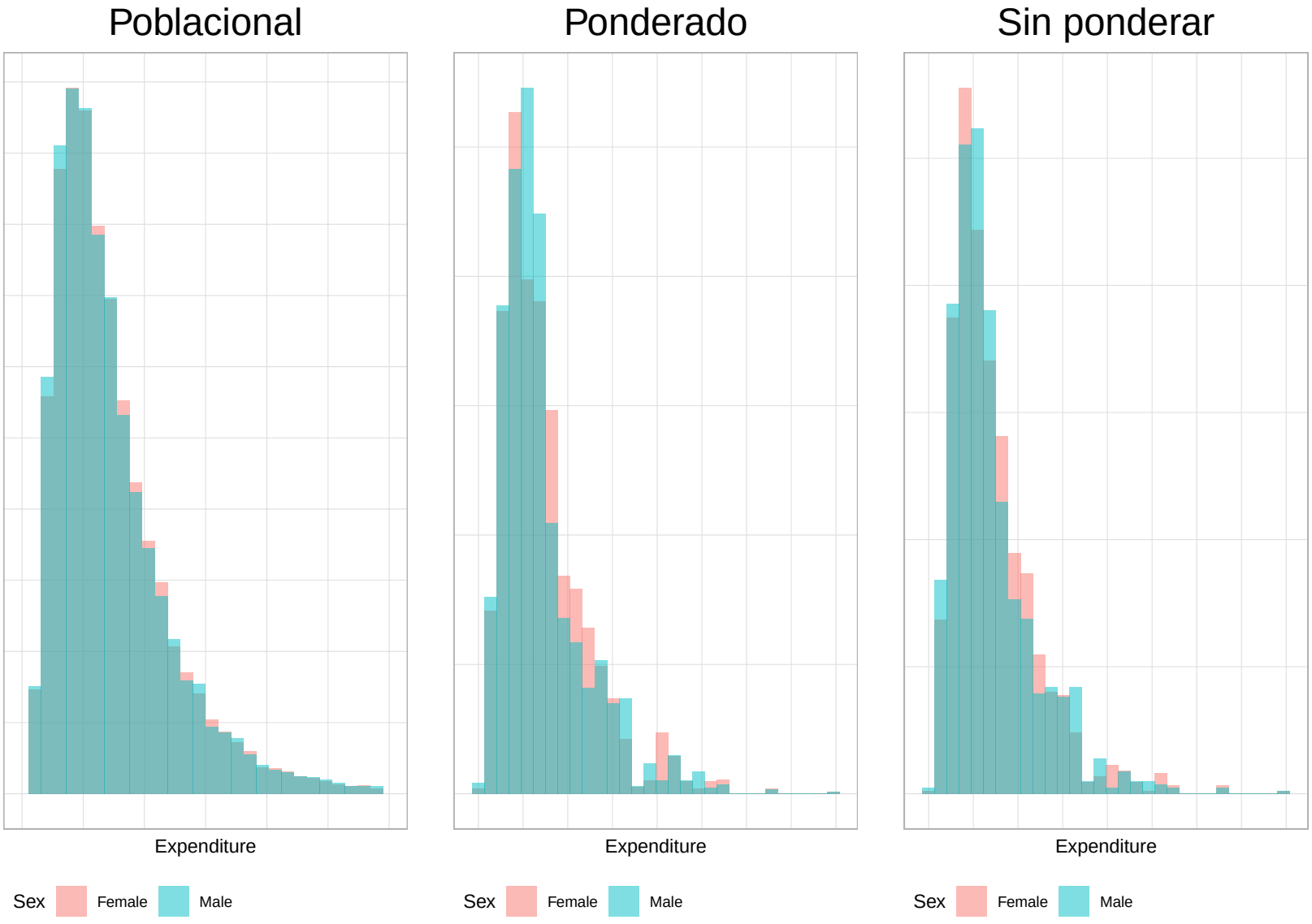
```
plot6_Ponde <- ggplot(encuesta,
                aes(x = Expenditure, weight = wk)) +
  geom_histogram(aes(y = ..density.., fill = Sex) ,
                   alpha = 0.5, position = "identity") +
  ylab("") +  ggtitle("Ponderado") + theme_cepal()
```

Histogramas por sub-grupos

```
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("")
```

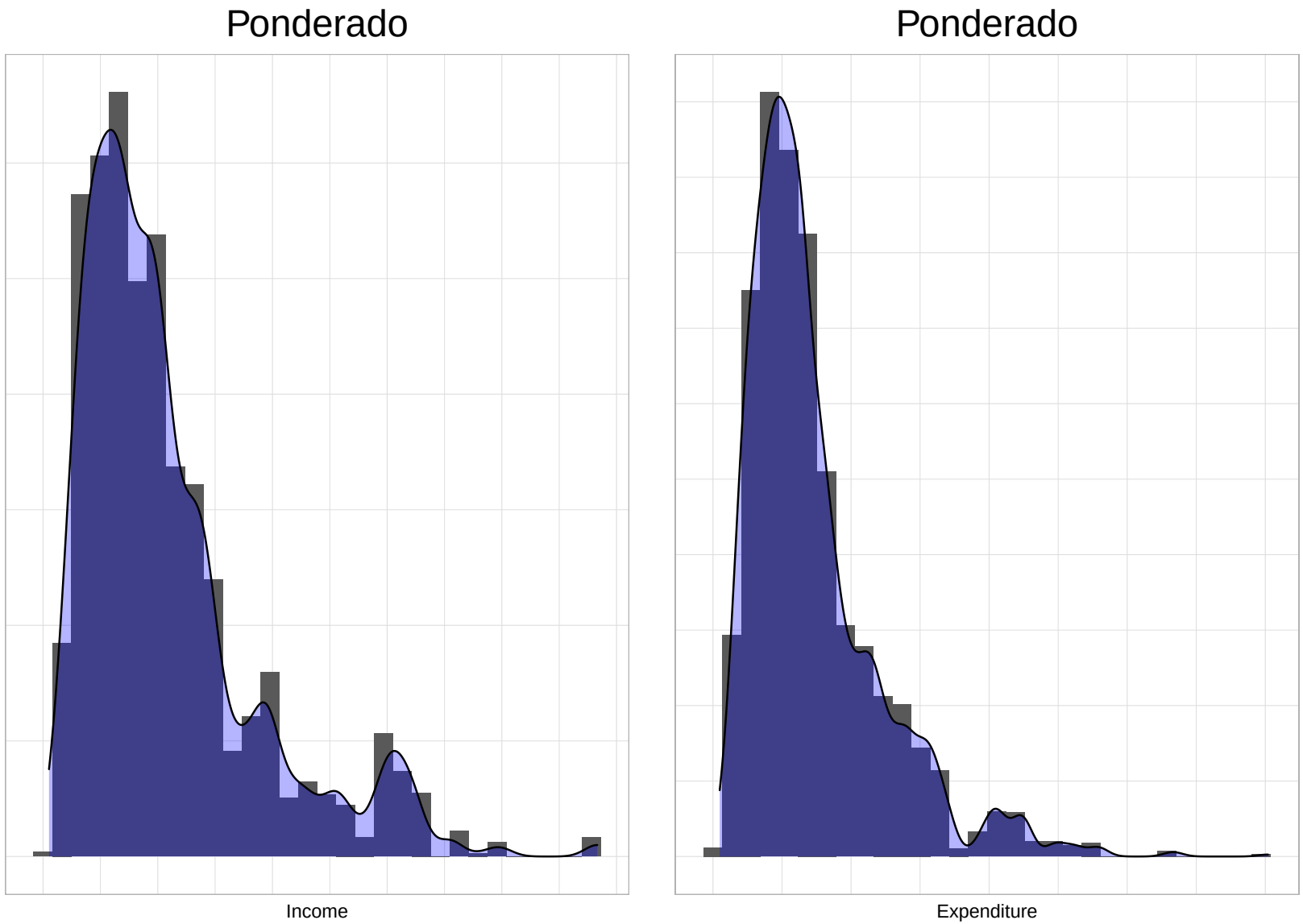
```
plot6_censo <- ggplot(BigCity, aes(x = Expenditure))+
  geom_histogram(aes(y = ..density.., fill = Sex) ,
                 alpha = 0.5, position = "identity") +
   ggtitle("Poblacional") + theme_cepal() +
   xlim(0, 1500)  + ylab("")
plot6_censo | plot6_Ponde | plot6_SinPonde
```

Histogramas por sub-grupos



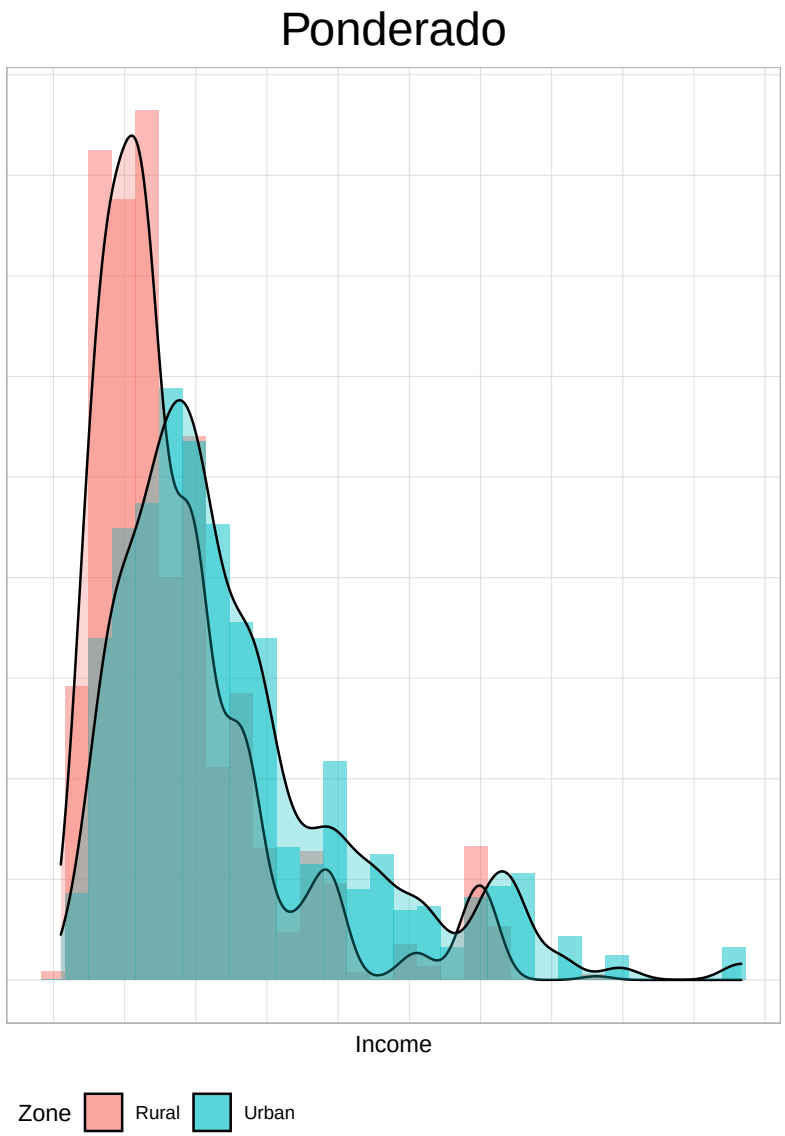
Agregando densidad

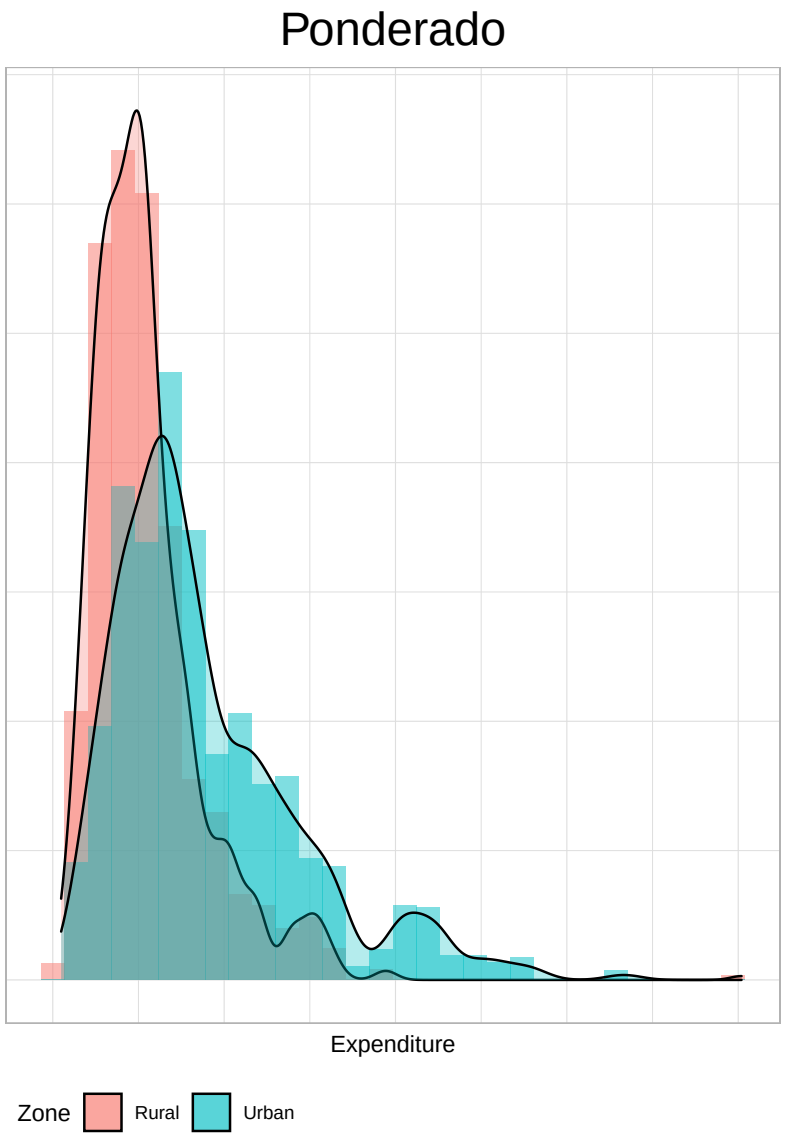
```
plot1_Ponde + geom_density(fill = "blue", alpha = 0.3 ) |
plot2_Ponde + geom_density(fill = "blue", alpha = 0.3)
```



Agregando densidad

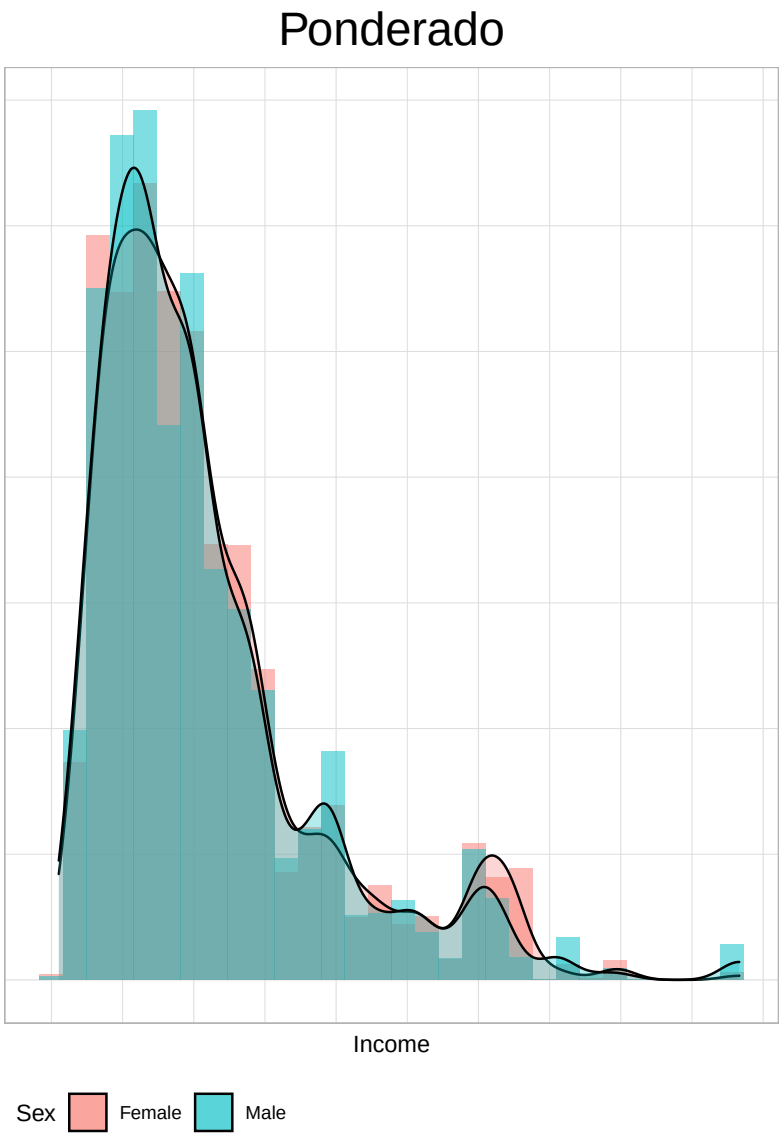
```
plot3_Ponde + geom_density(aes(fill = Zone), alpha = 0.3 )
plot4_Ponde + geom_density(aes(fill = Zone), alpha = 0.3)
```

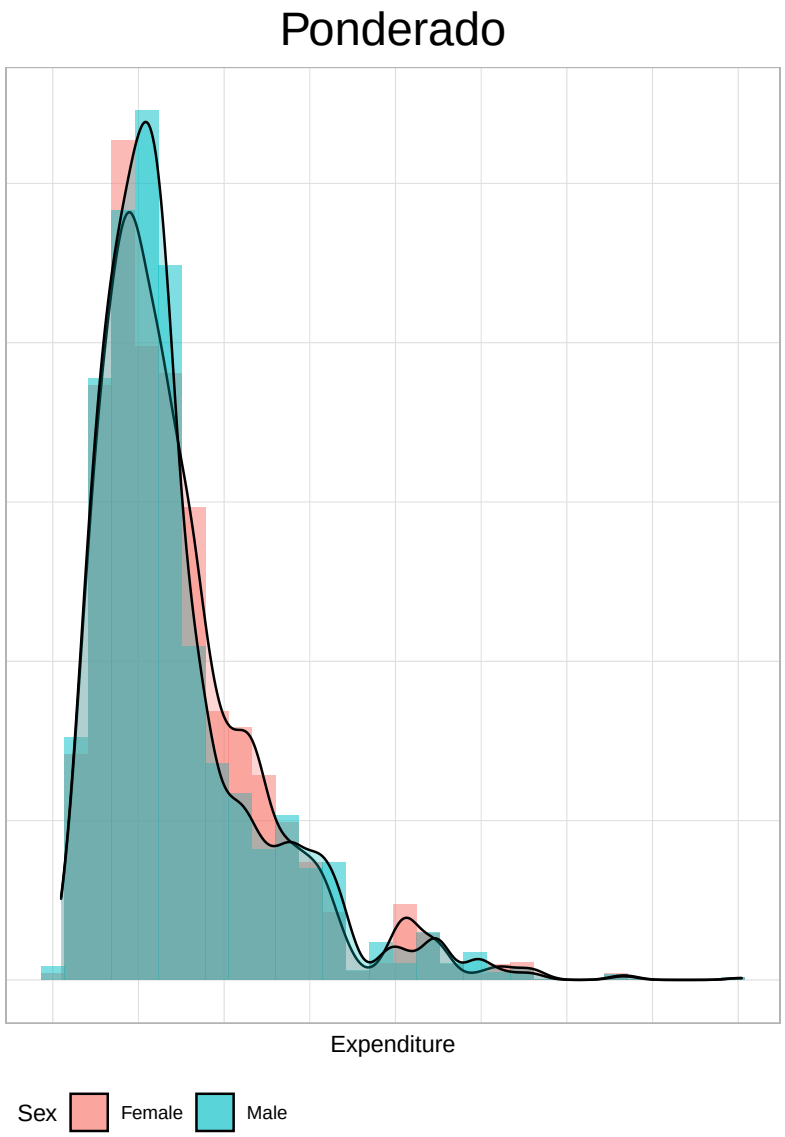




Agregando densidad

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





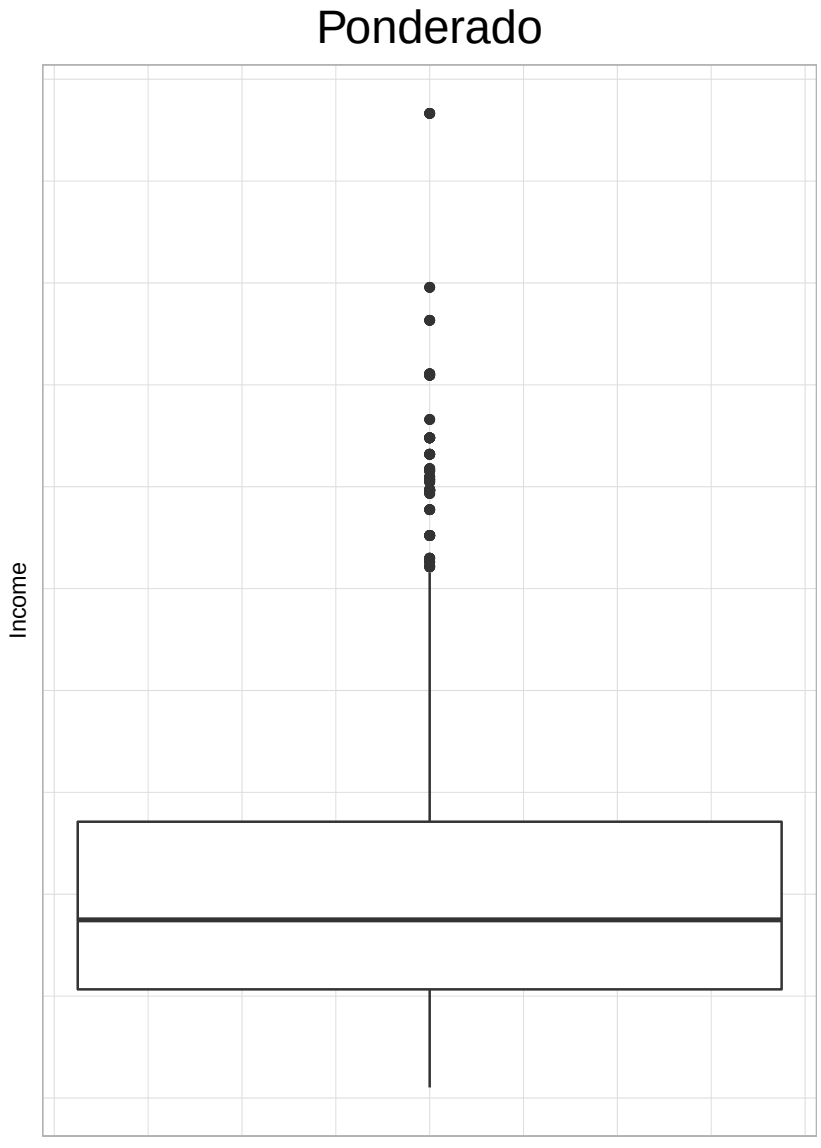
Boxplot

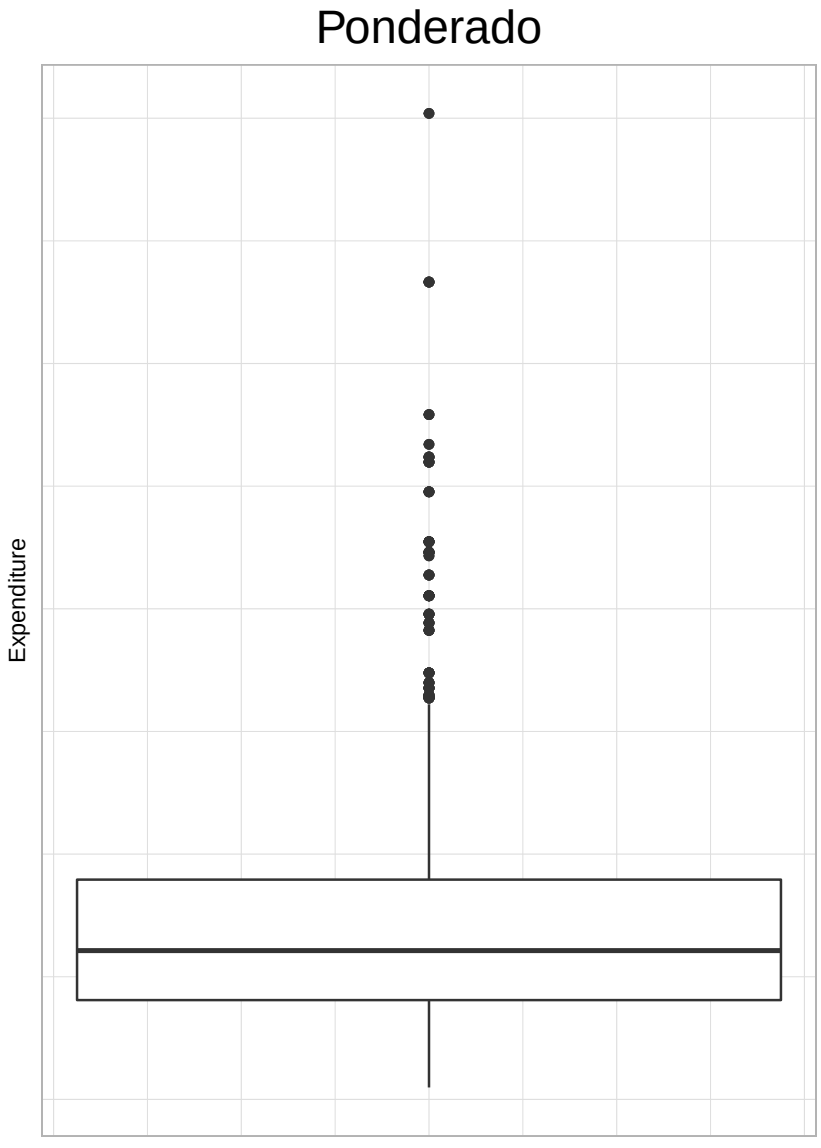
```
plot7_Ponde <- ggplot(encuesta,
                      aes(x = Income, weight = wk)) +
  geom_boxplot() + ggtitle("Ponderado") +
  coord_flip() + theme_cepal()

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

plot7_Ponde | plot8_Ponde
```

Boxplot





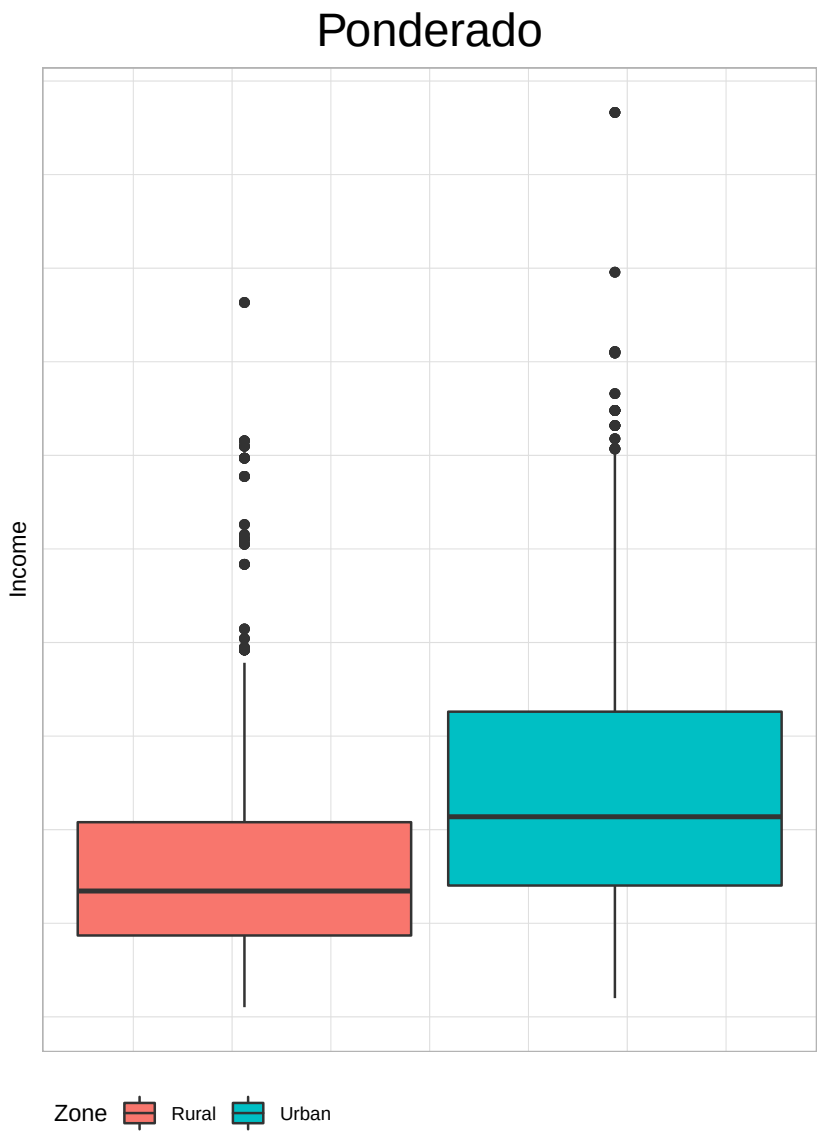
Boxplot

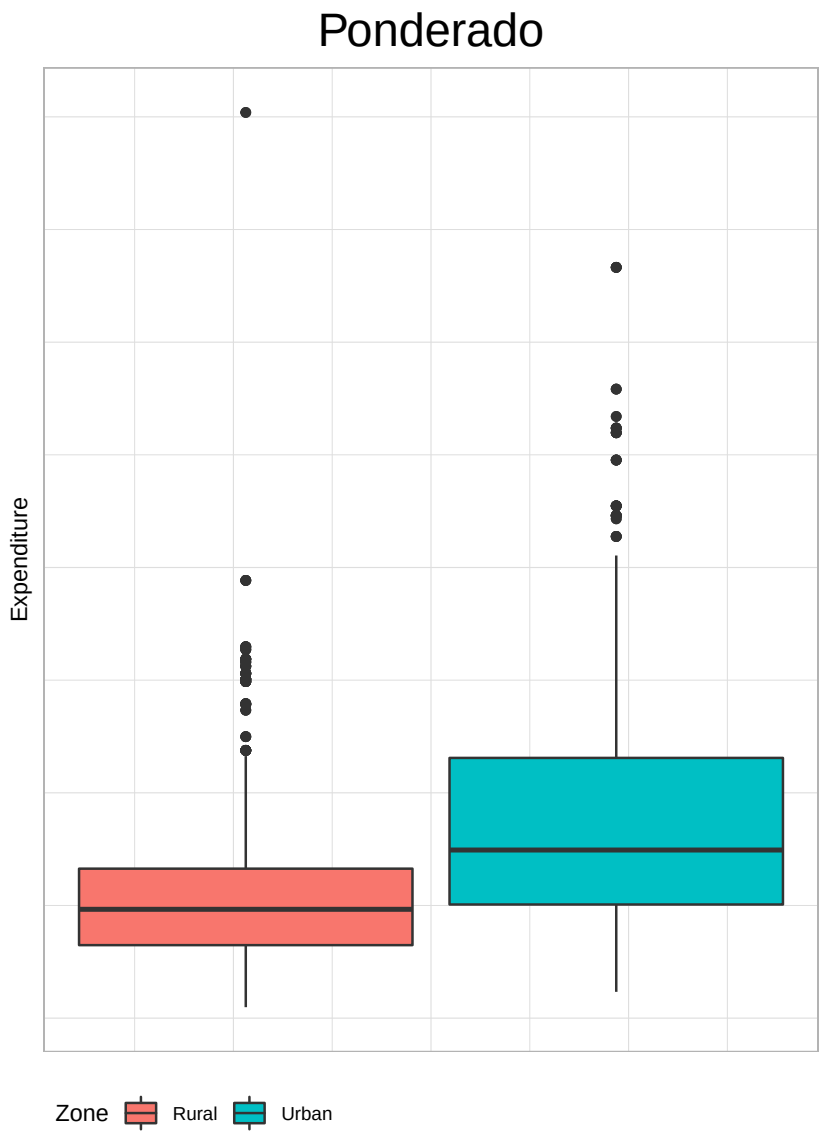
```
plot9_Ponde <- ggplot(encuesta,
                      aes(x = Income, weight = wk)) +
  geom_boxplot(aes(fill = Zone)) +
  ggtitle("Ponderado") +
  coord_flip() + theme_cepal()

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

plot9_Ponde | plot10_Ponde
```

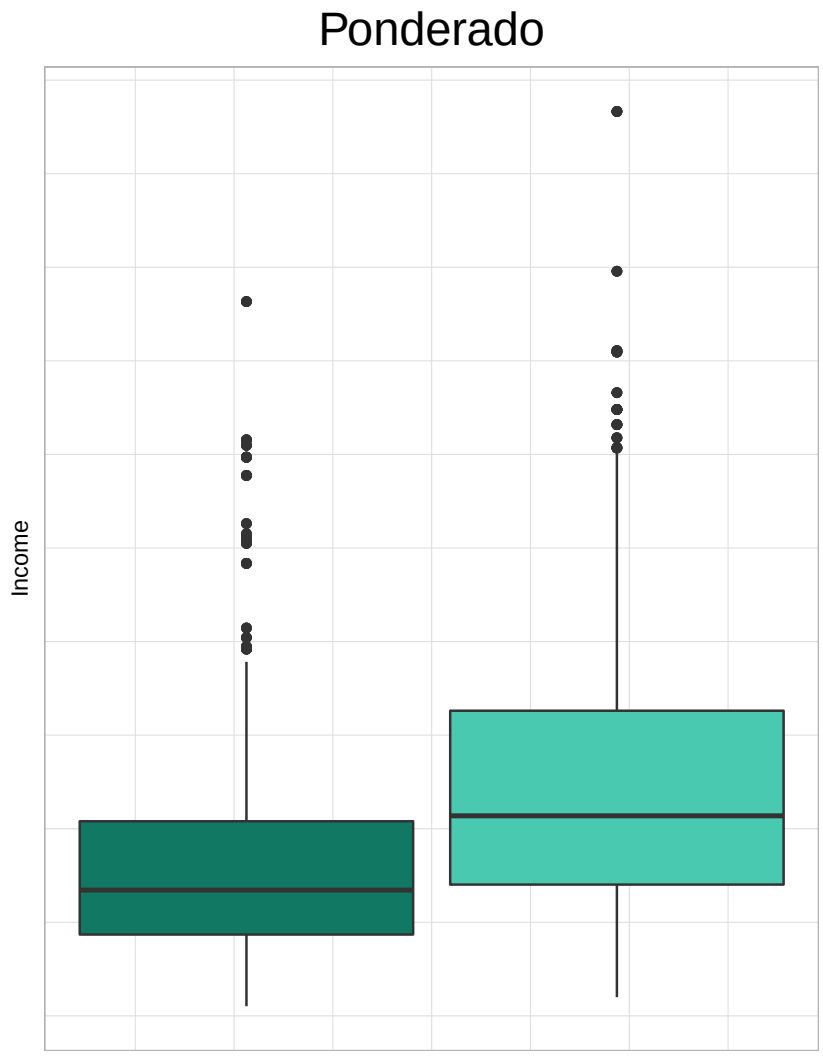
Boxplot

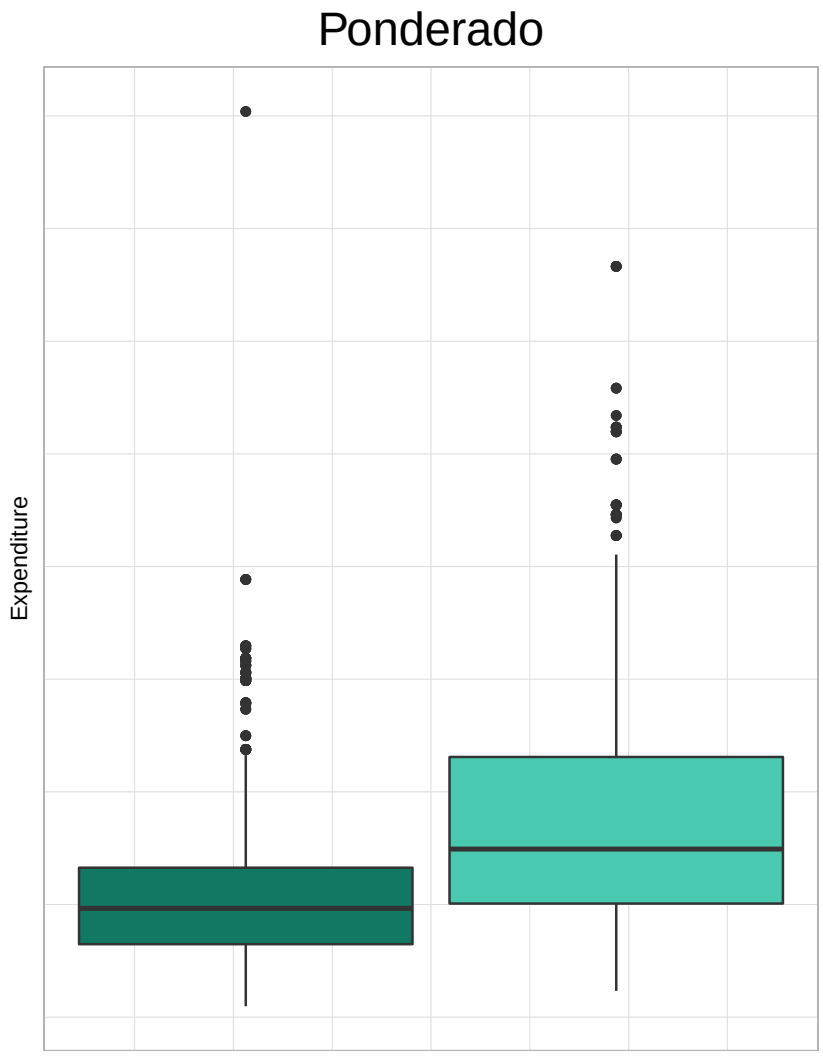

Ponderado
Income
Zone
Rural
Urban


Ponderado
Expenditure
Zone
Rural
Urban

Boxplot

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





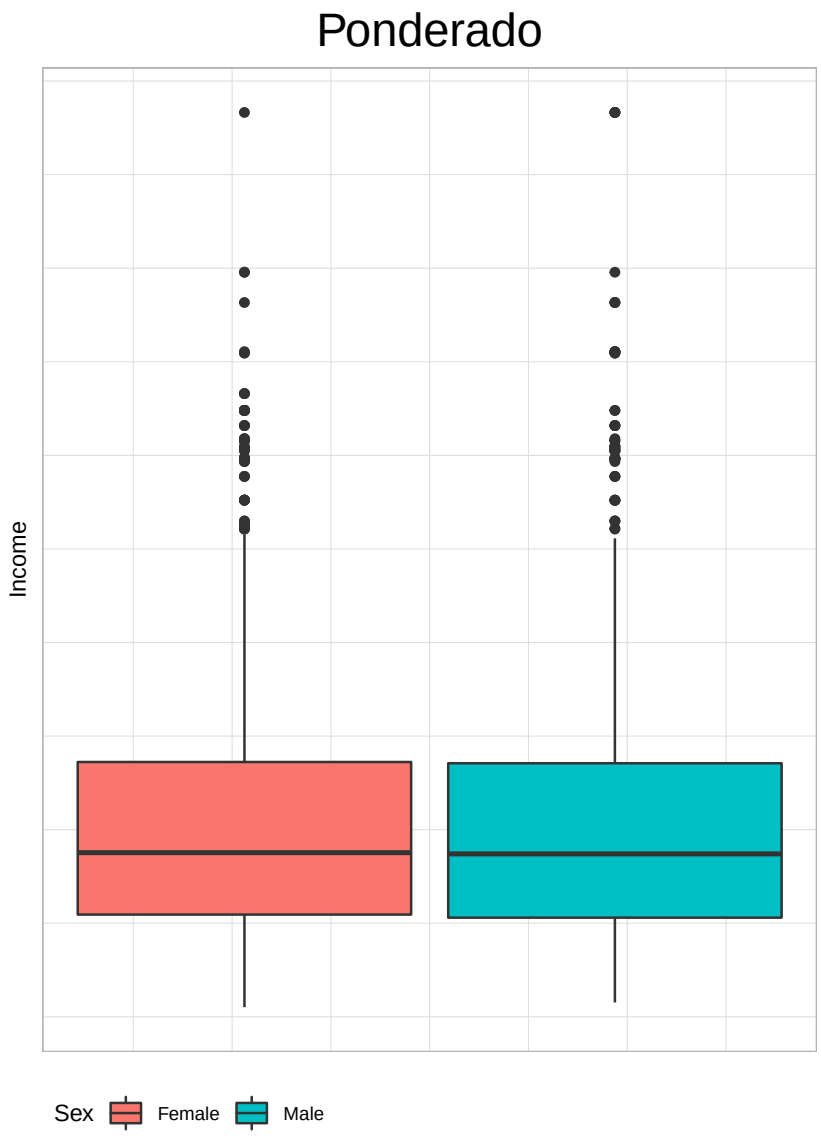
Boxplot

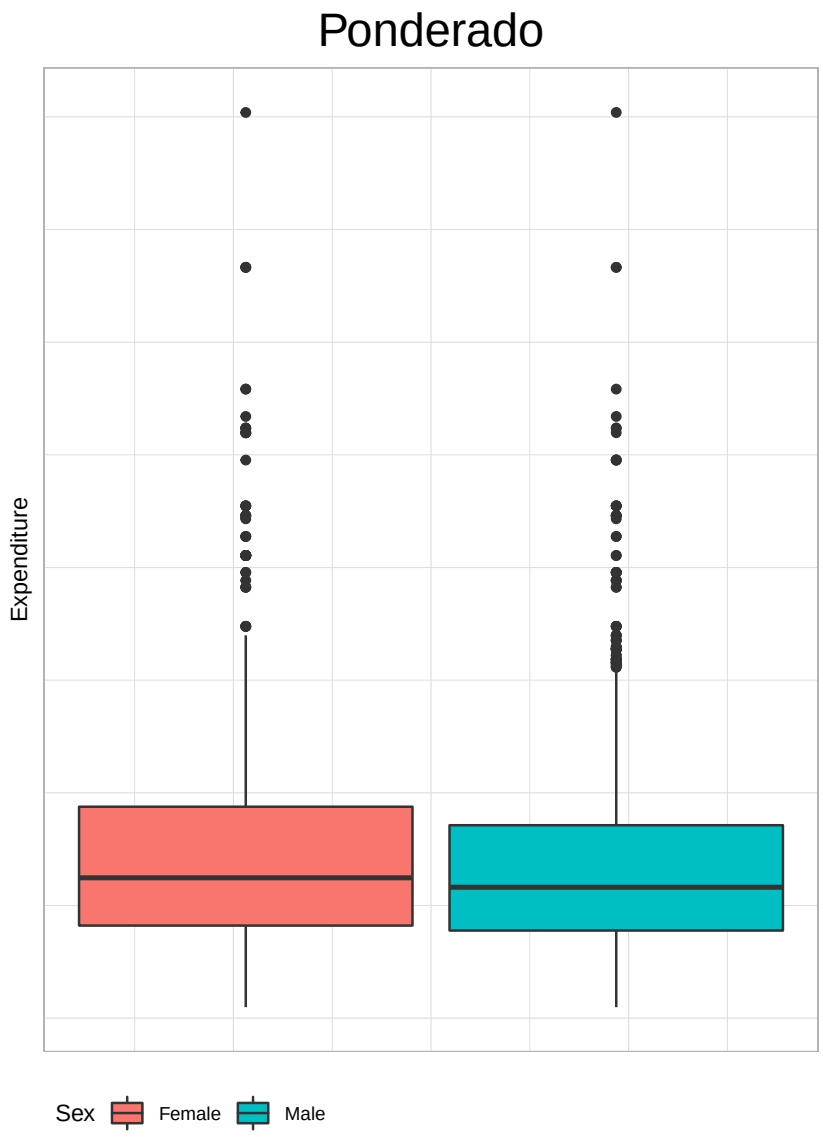
```
plot11_Ponde <- ggplot(encuesta,
                       aes(x = Income, weight = wk)) +
  geom_boxplot(aes(fill = Sex)) +
  ggtitle("Ponderado") +
  coord_flip() + theme_cepal()

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

plot11_Ponde | plot12_Ponde
```

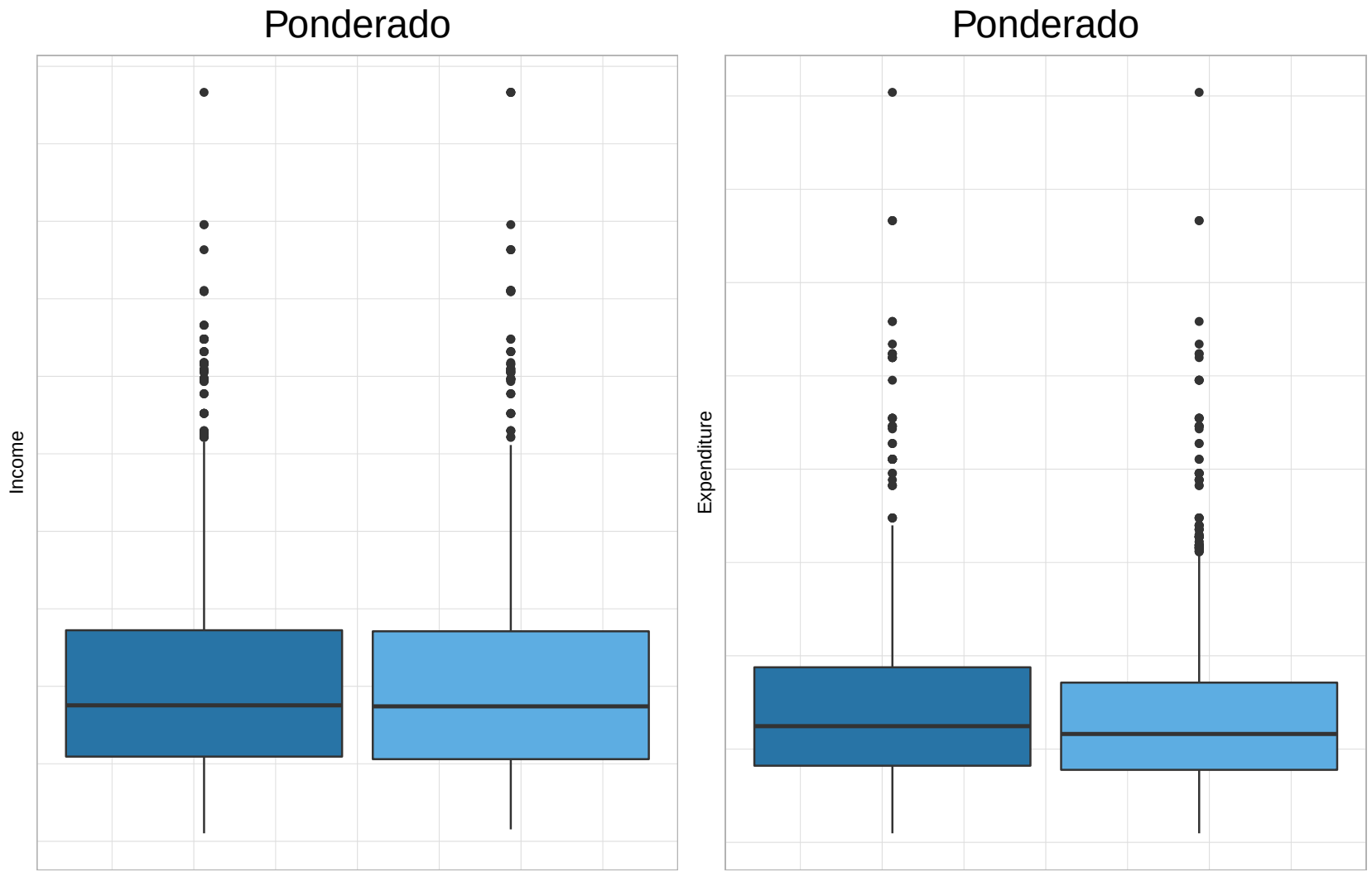
Boxplot


Ponderado
Income
Sex
Female
Male


Ponderado
Expenditure
Sex
Female
Male

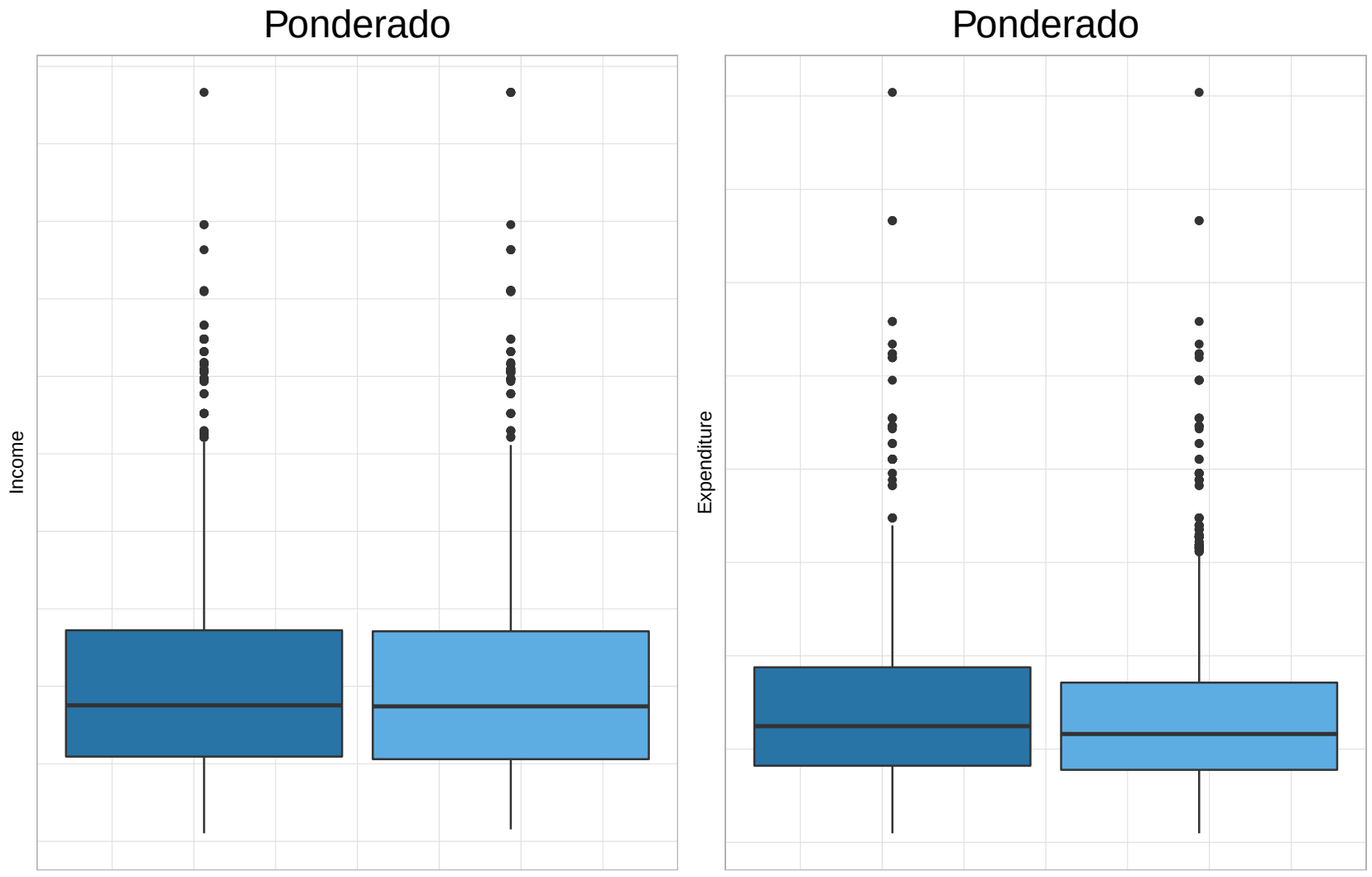
Boxplot

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



Boxplot

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



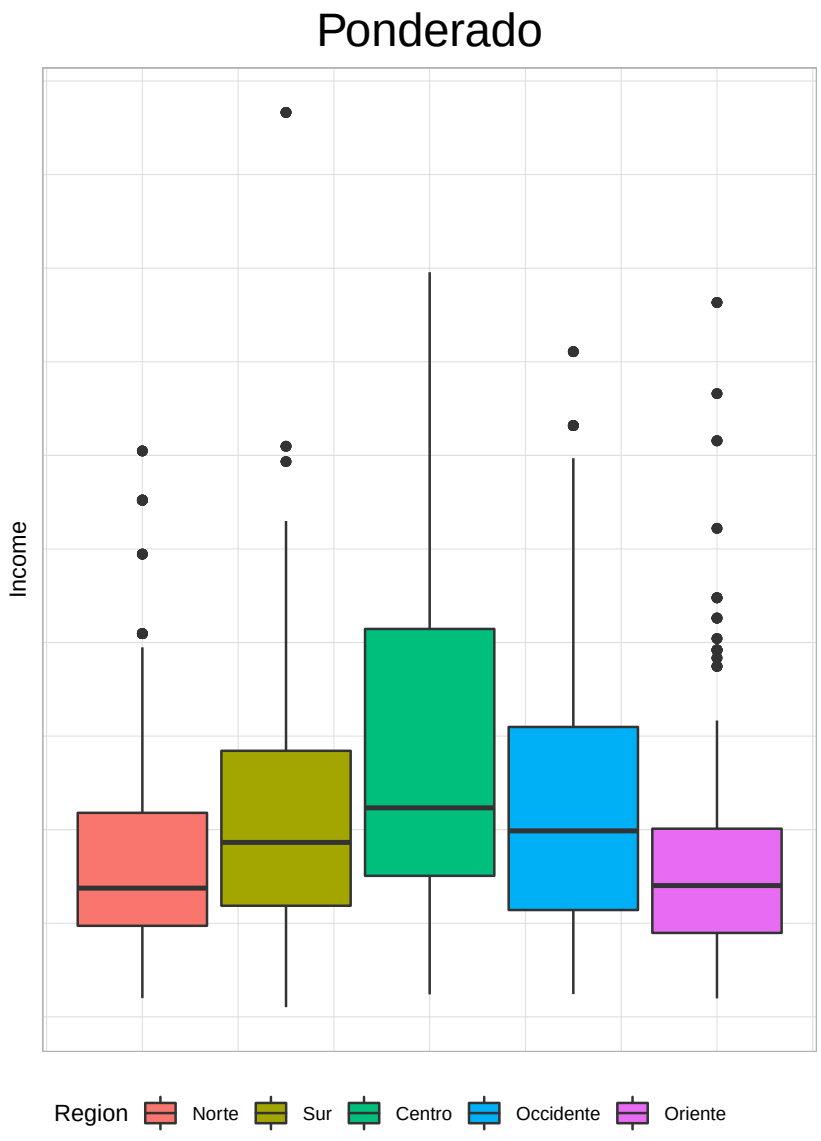
Boxplot

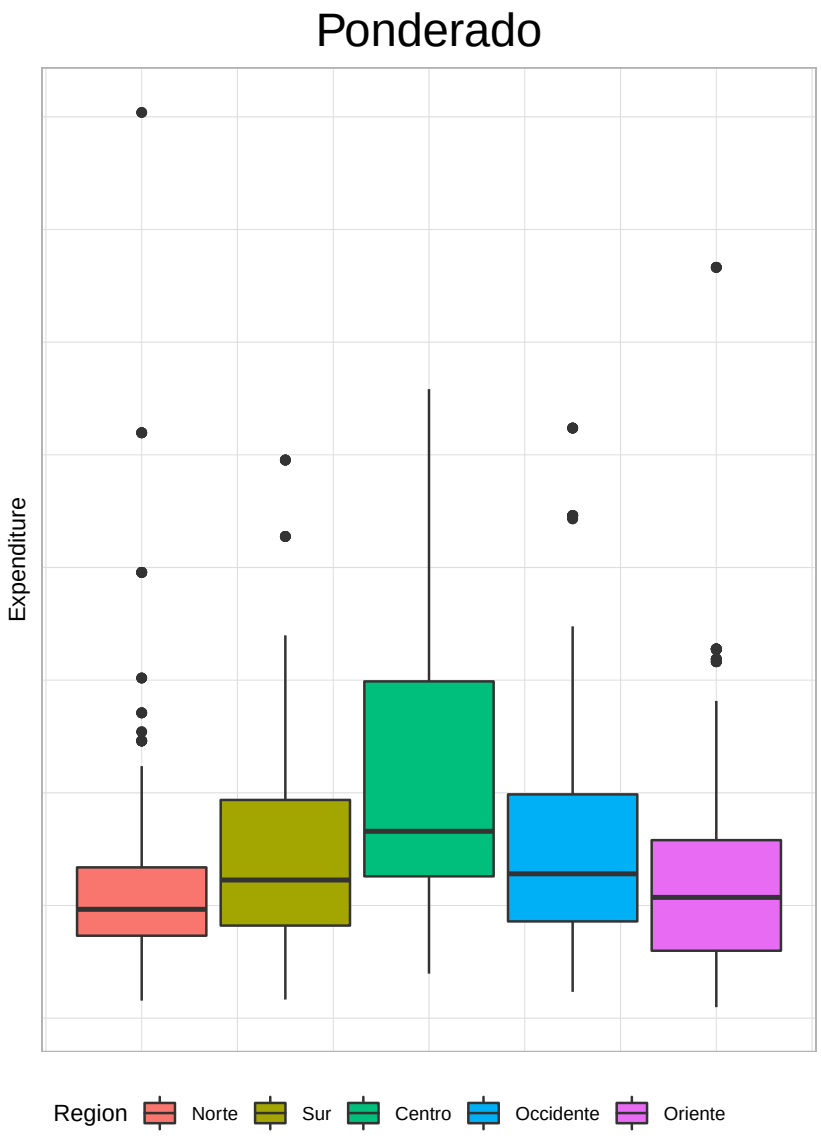
```
plot13_Ponde <- ggplot(encuesta,
                  aes(x = Income, weight = wk)) +
  geom_boxplot(aes(fill = Region)) +
  ggtitle("Ponderado") +
  coord_flip() + theme_cepal()

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

plot13_Ponde | plot14_Ponde
```

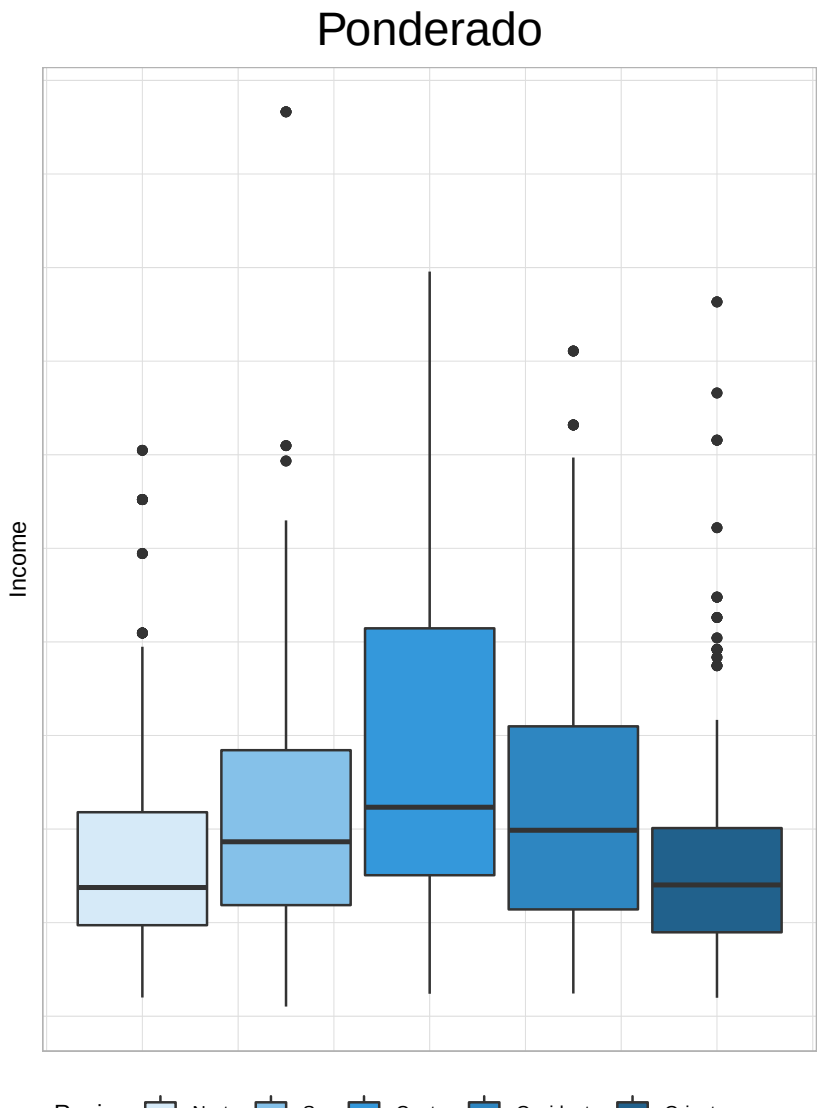
Boxplot

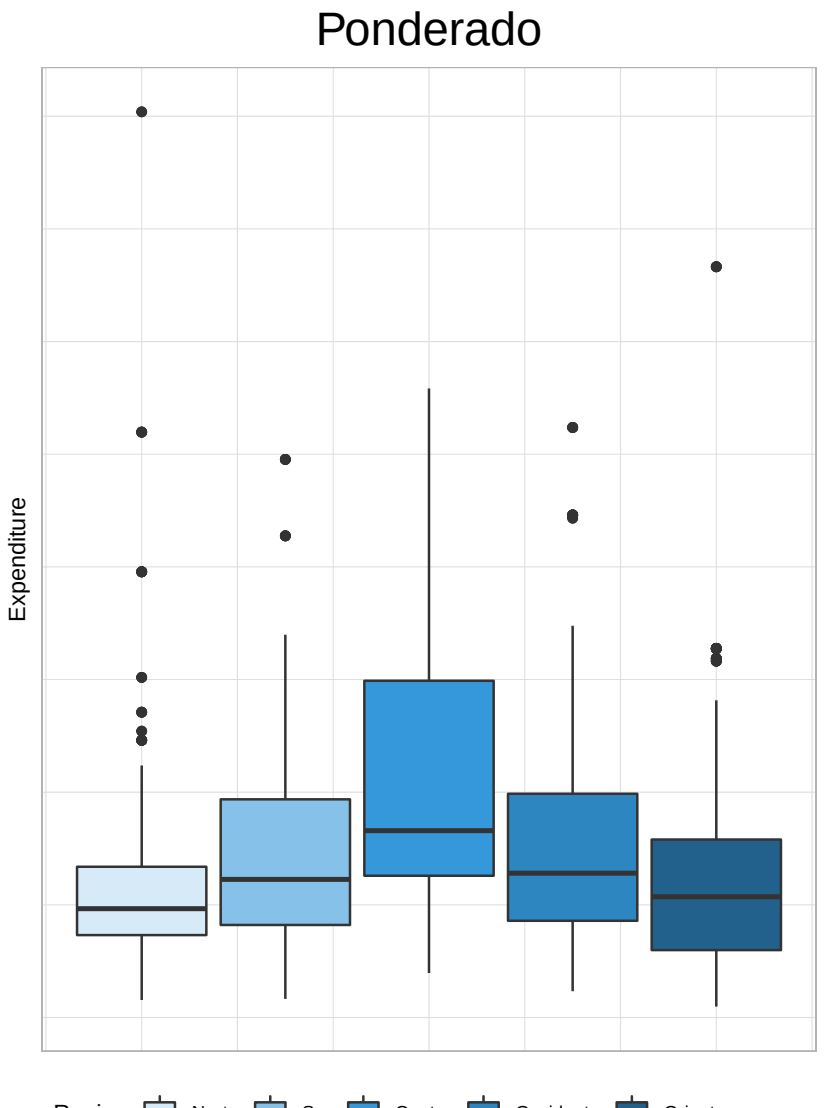

Ponderado
Income
Region
Norte
Sur
Centro
Occidente
Oriente


Ponderado
Expenditure
Region
Norte
Sur
Centro
Occidente
Oriente

Boxplot

```
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)
```

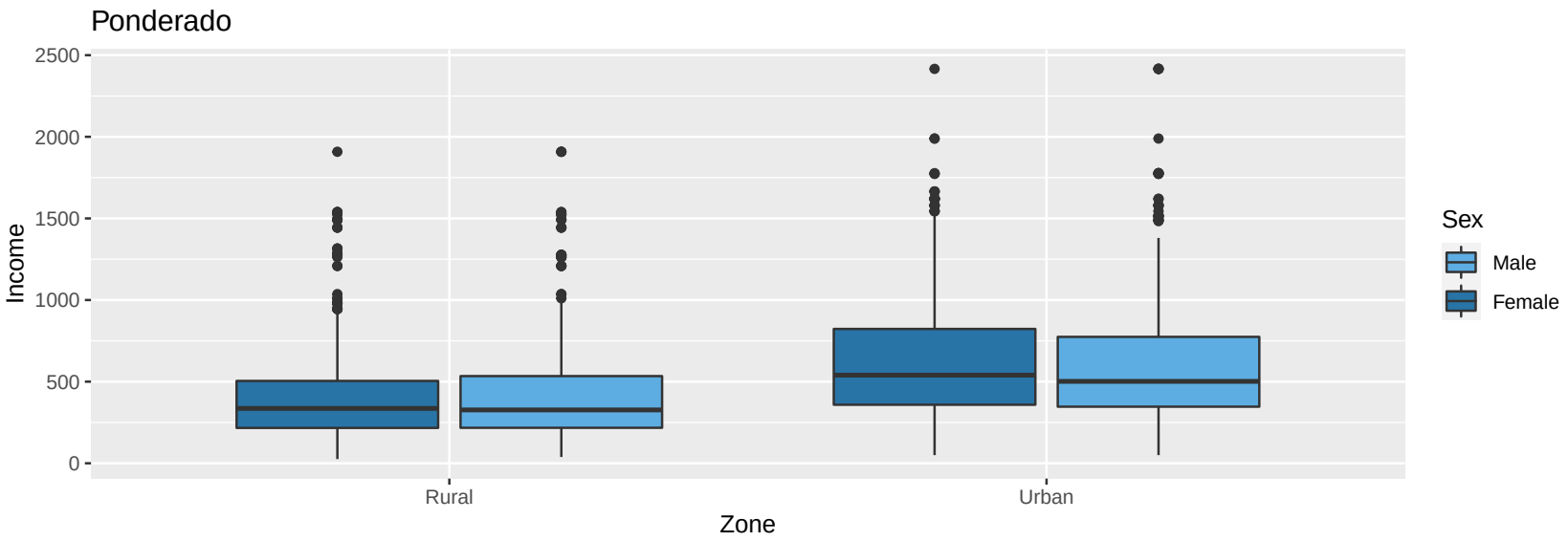


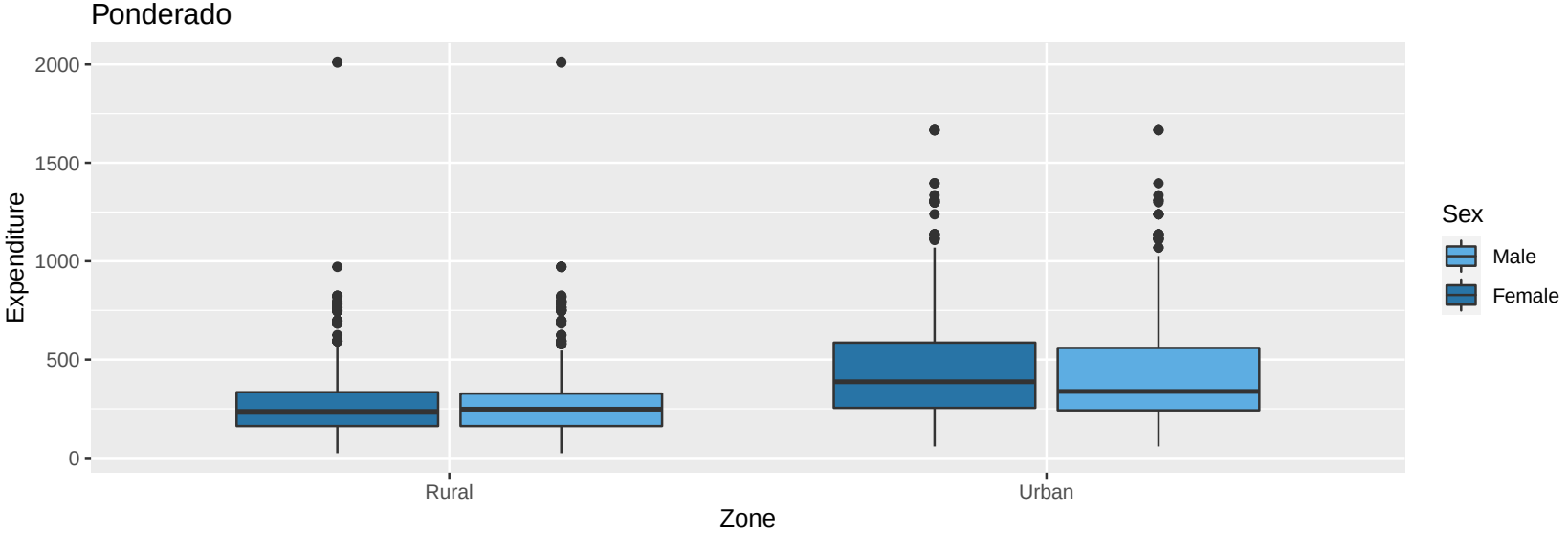


Boxplot

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

Boxplot


Ponderado
2500
2000
1500
1000
500
0
Income
Rural
Urban
Zone
Sex
Male
Female

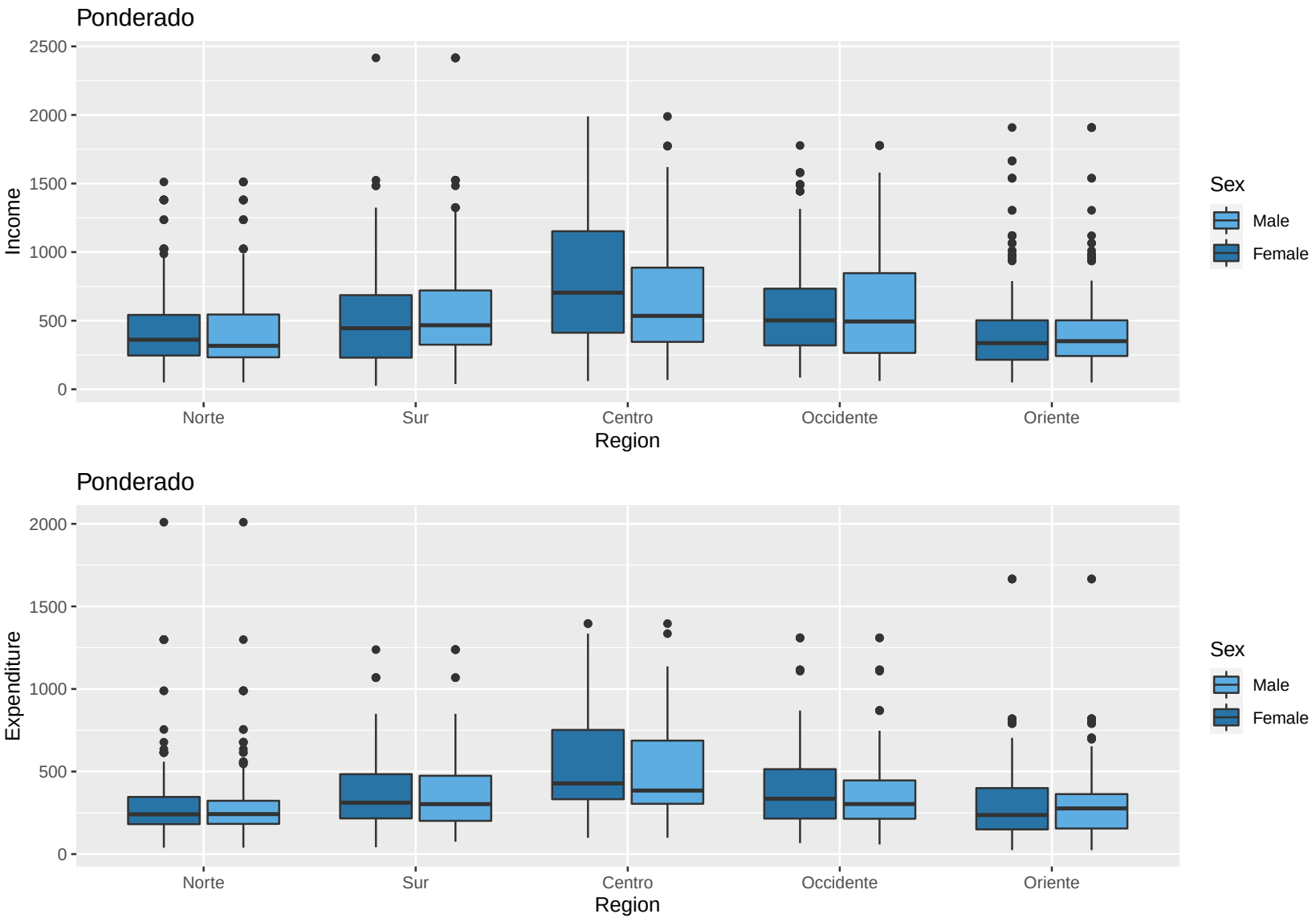

Ponderado
2000
1500
1000
500
0
Expenditure
Rural
Urban
Zone
Sex
Male
Female

Boxplot

```
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()
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
```

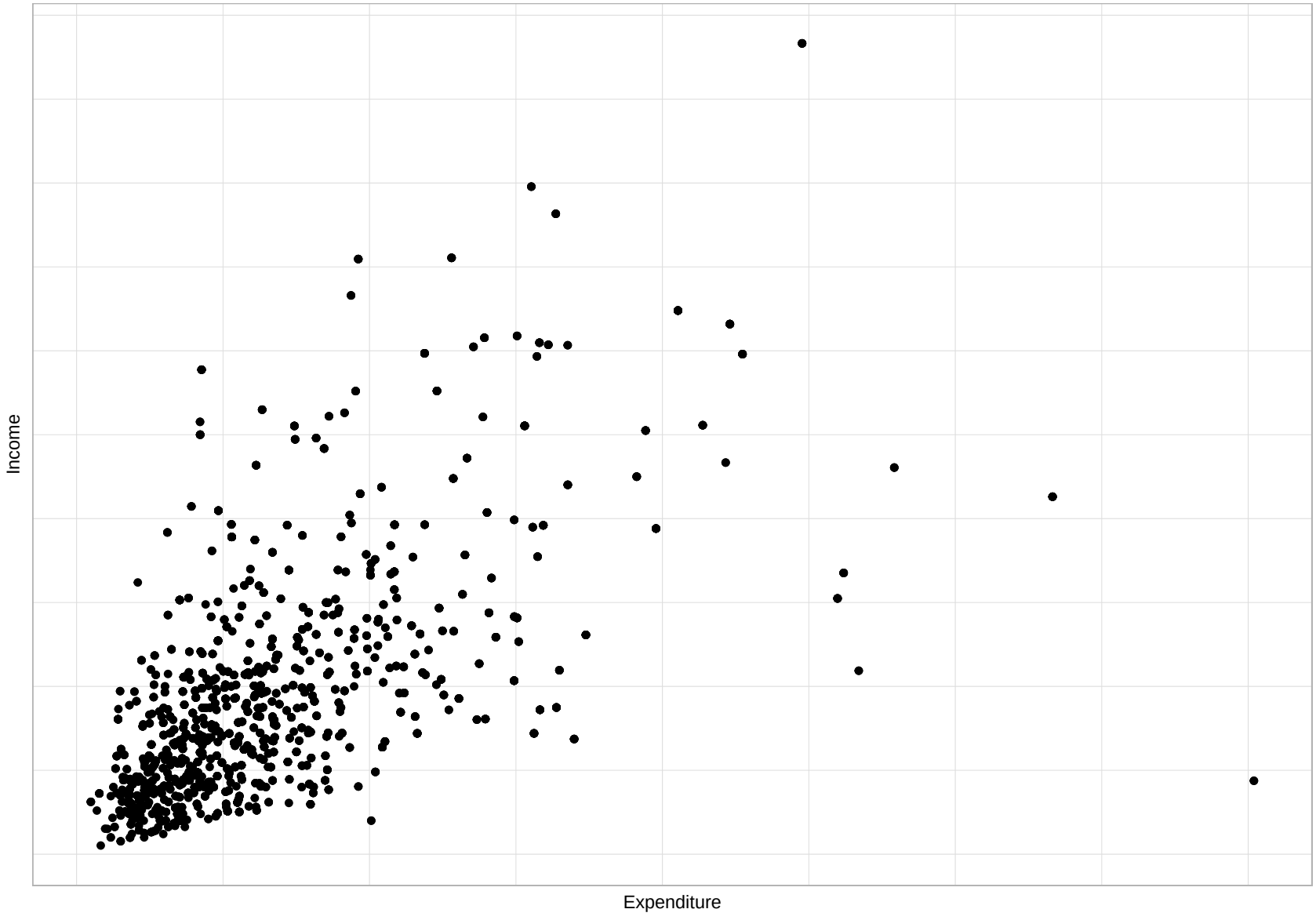
Boxplot


Ponderado
Income
2500
2000
1500
1000
500
0
Norte
Sur
Centro
Occidente
Oriente
Region
Sex
Male
Female
Ponderado
Expenditure
2000
1500
1000
500
0
Norte
Sur
Centro
Occidente
Oriente
Region
Sex
Male
Female

Scaterplot

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

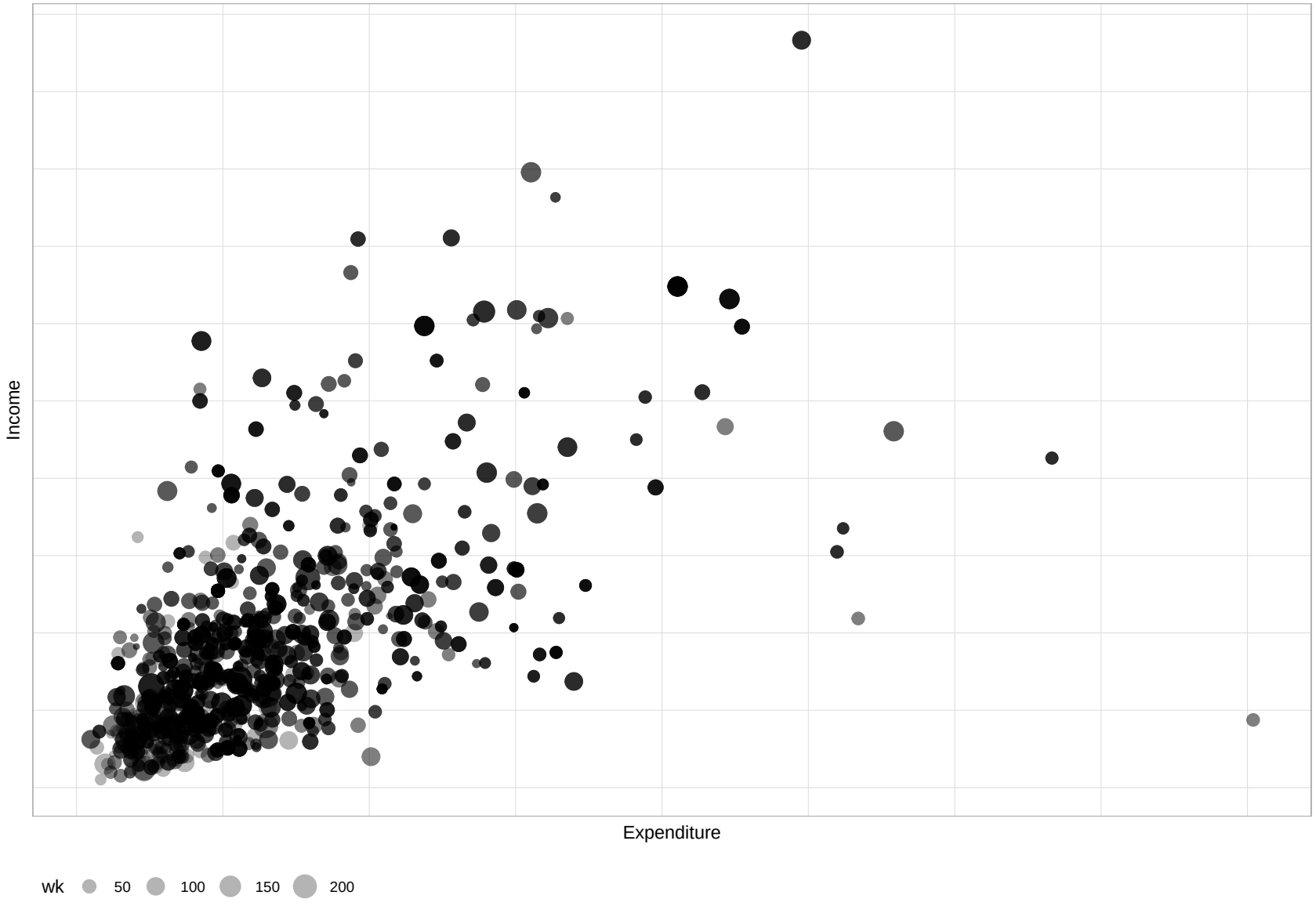
Scaterplot



Scaterplot

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

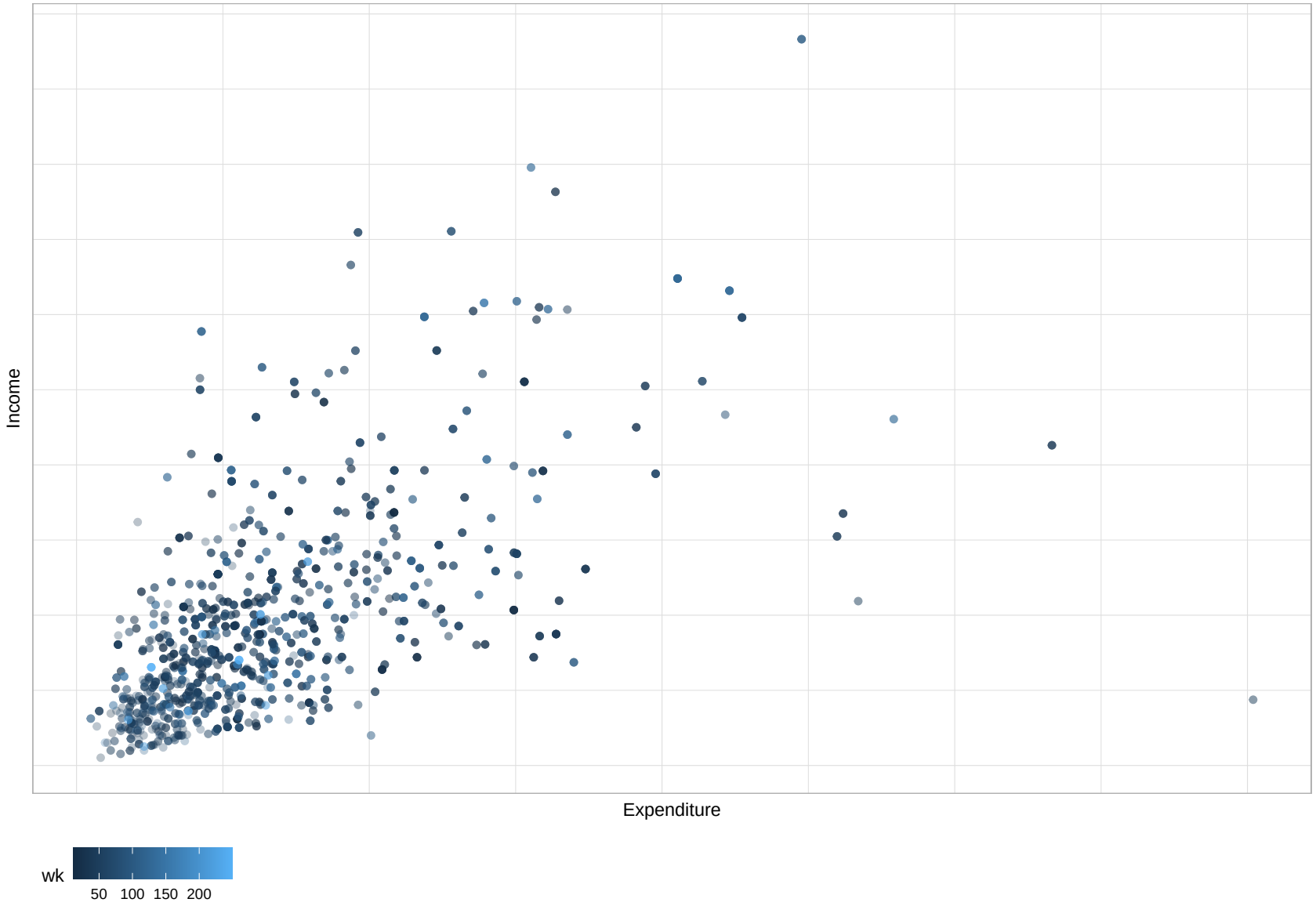
Scaterplot



Scaterplot

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

Scaterplot



Scaterplot

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

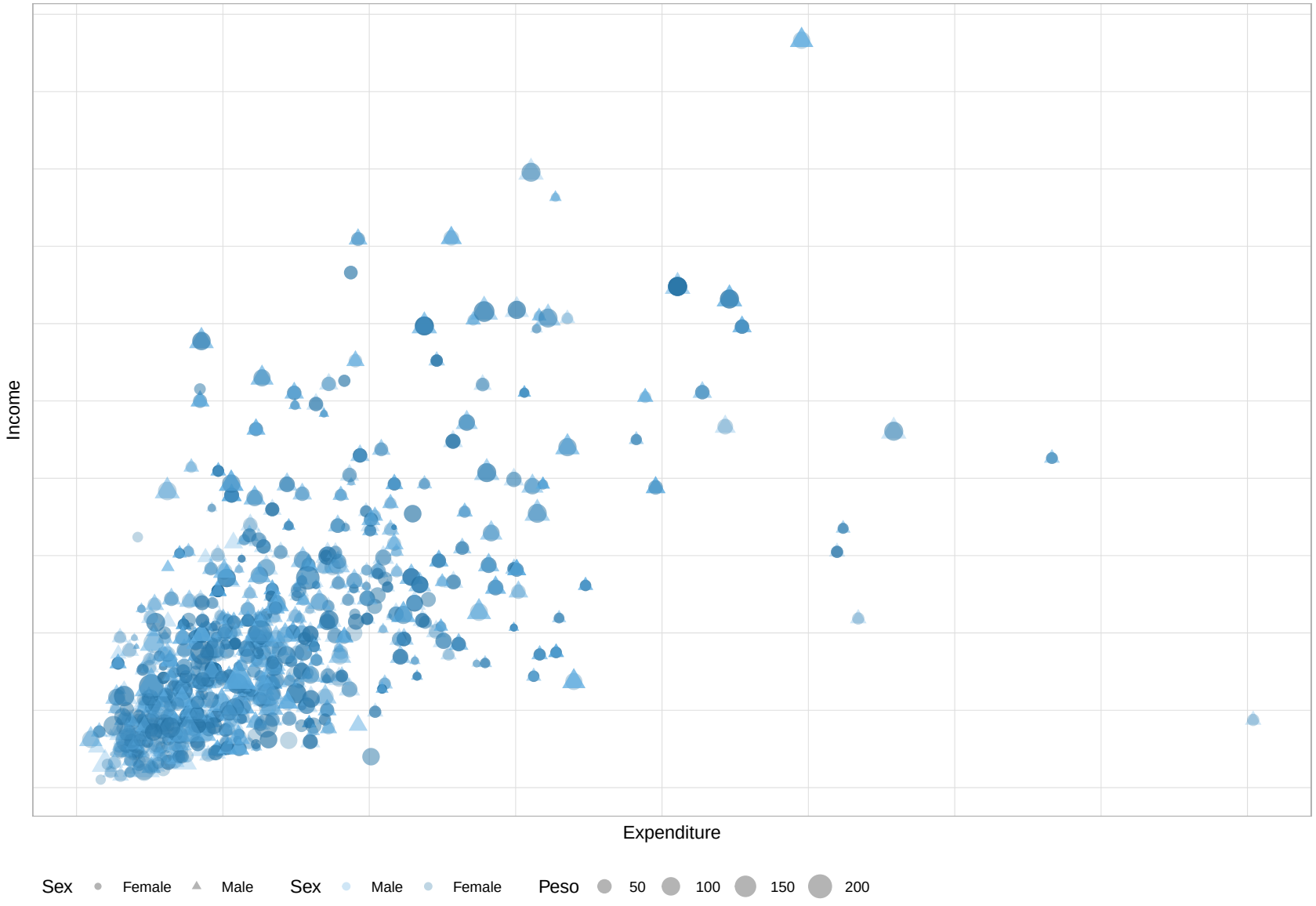
Scaterplot


Income
Expenditure
Peso
50
100
150
200
Zone
Urban
Rural
Zone
Rural
Urban

Scaterplot

```
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
```

Scaterplot


Income
Expenditure
Sex
Female
Male
Sex
Male
Female
Peso
50
100
150
200

Scaterplot

```
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
```

Scaterplot

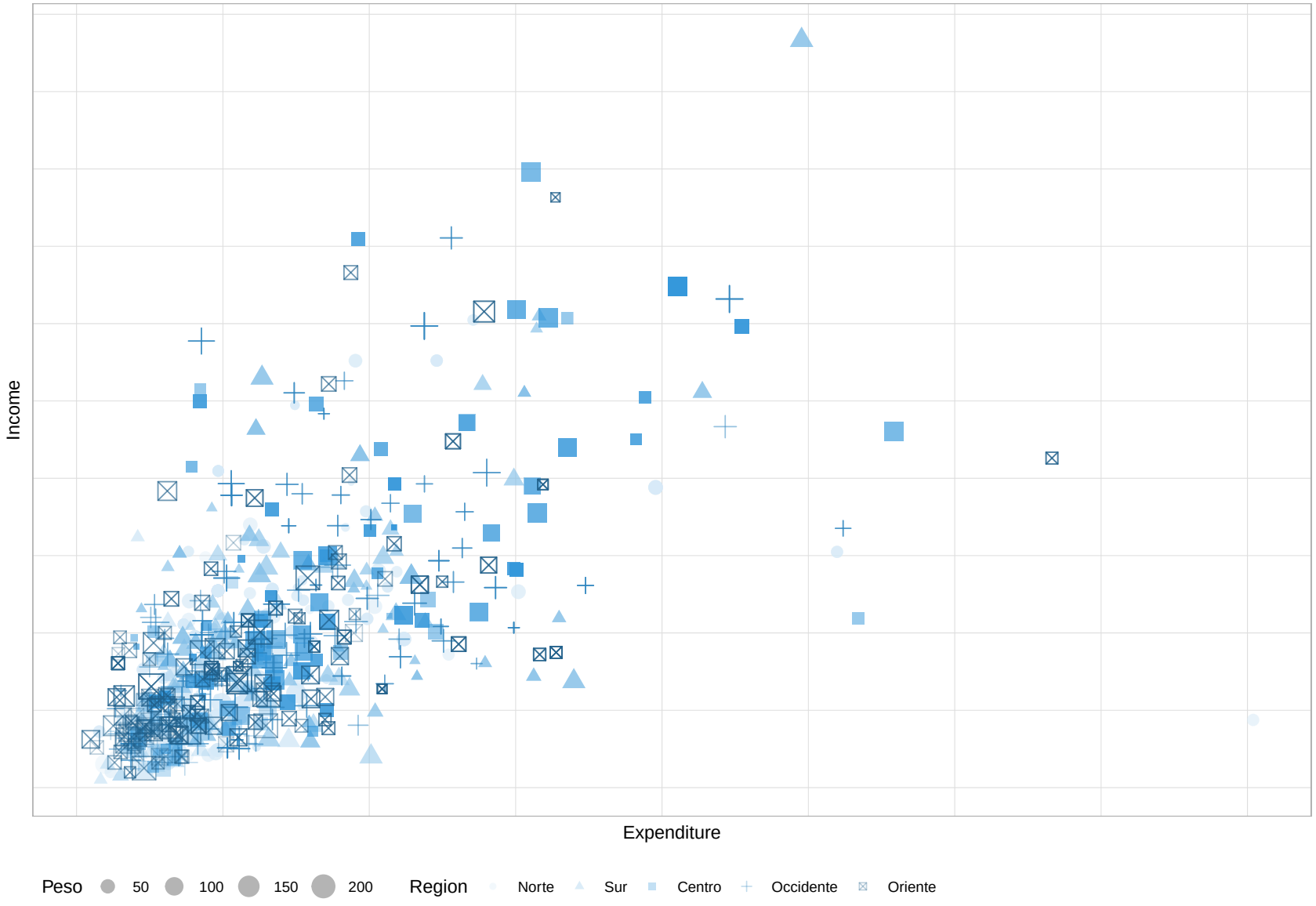

Income
Expenditure
Peso
50
100
150
200
Region
Norte
Sur
Centro
Occidente
Oriente

Diagrama de barras para variables categoricas

Diagrama de barras

```
tamano_zona <- diseno %>% group_by(Zone) %>%
   summarise(
    Nd = survey_total(vartype = c("se","ci")))
plot25_Ponde <- ggplot(data = tamano_zona,
       aes(
         x = Zone,
         y = Nd,
         ymax = Nd_upp,
         ymin = Nd_low,
         fill = Zone
       )) +
  geom_bar(stat = "identity", position = "dodge") +
  geom_errorbar(position = position_dodge(width = 0.9),
                width = 0.3) + theme_bw()
plot25_Ponde
```

Diagrama de barras

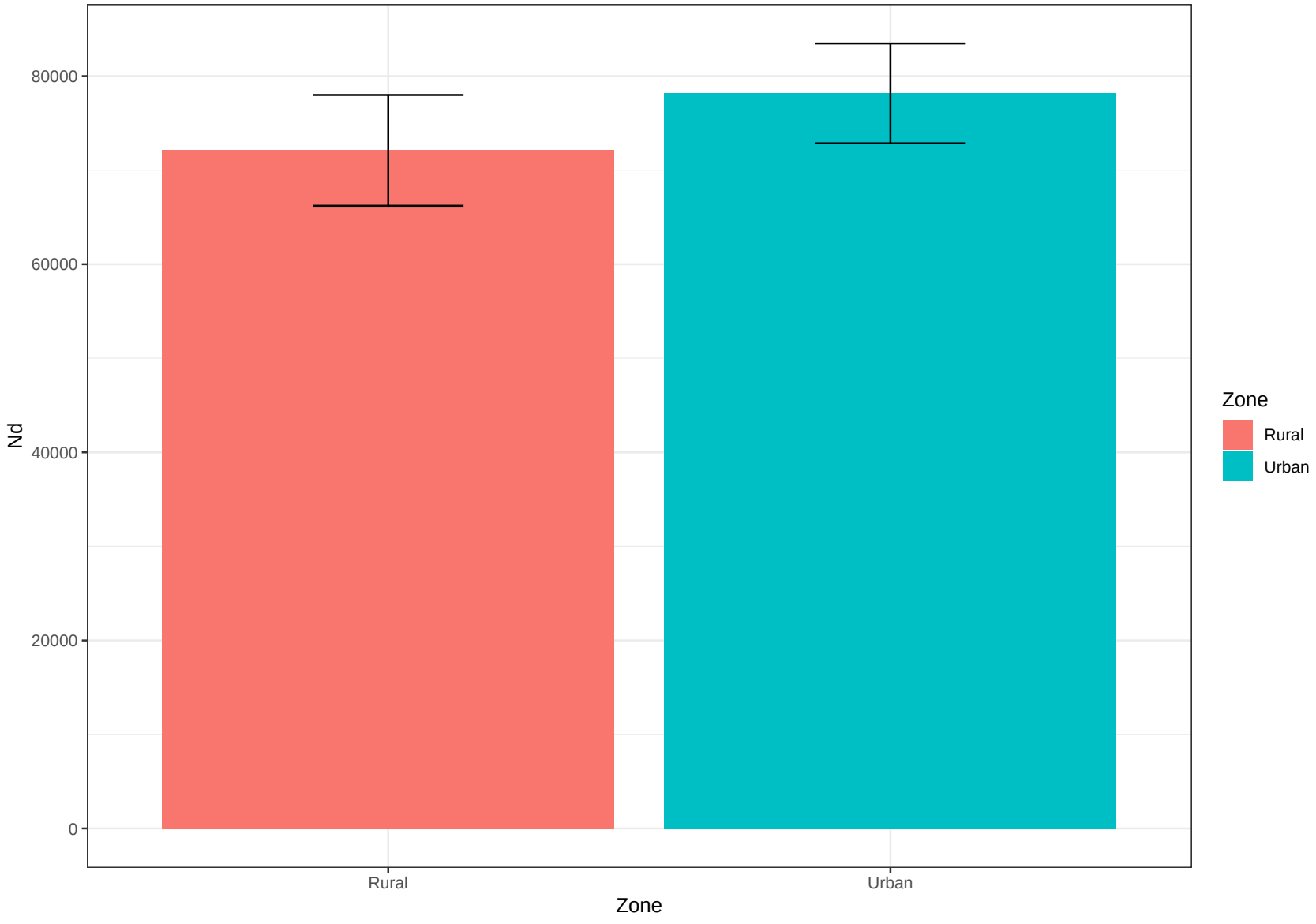


Diagrama de barras

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

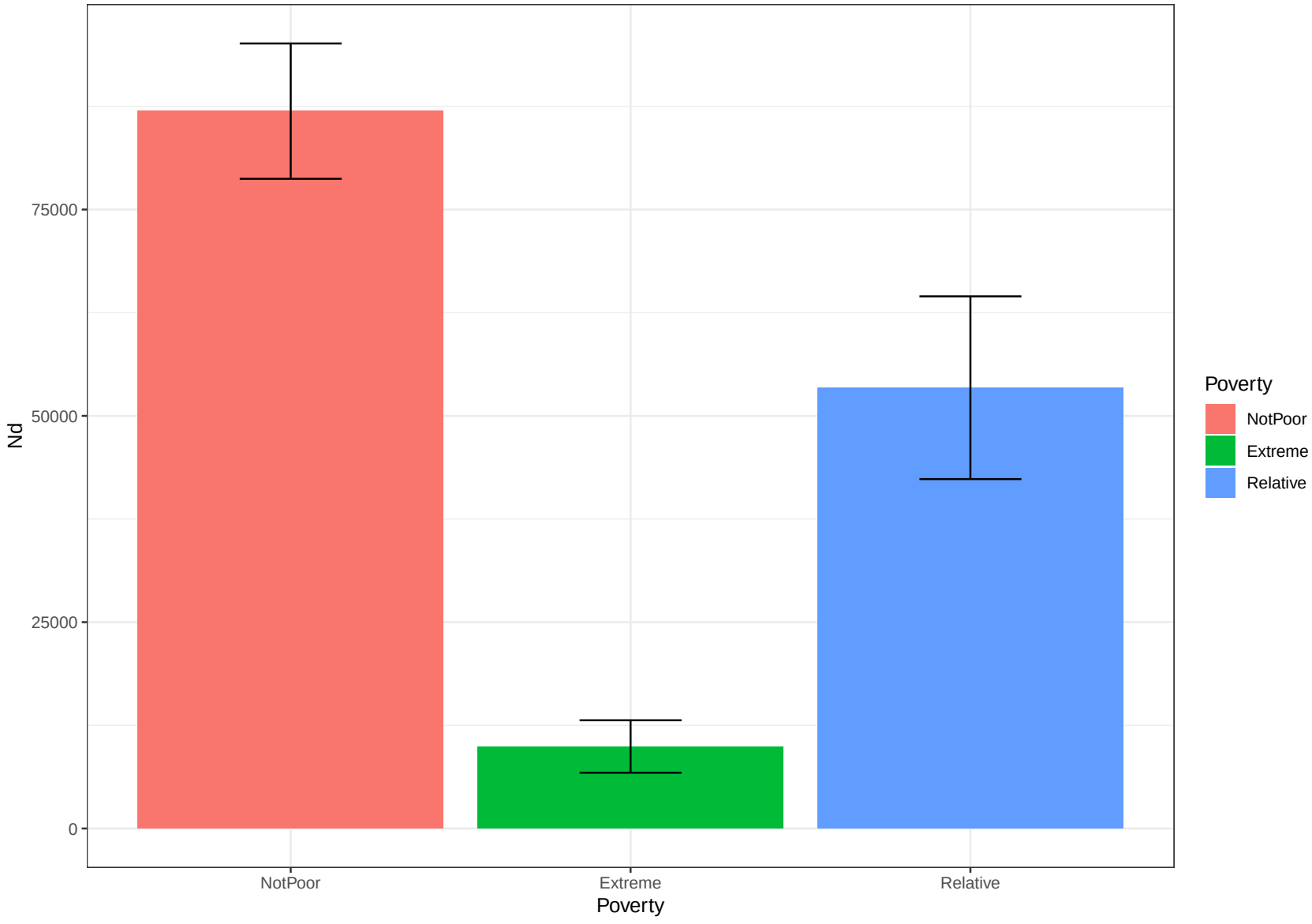


Diagrama de barras

```
tamano_ocupacion_pobreza <- diseno %>%
   group_by(desempleo, Poverty) %>%
   summarise(
       Nd = survey_total(vartype = c("se","ci")),
       .fill = "Total") %>%
   data.frame()
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
```

Diagrama de barras

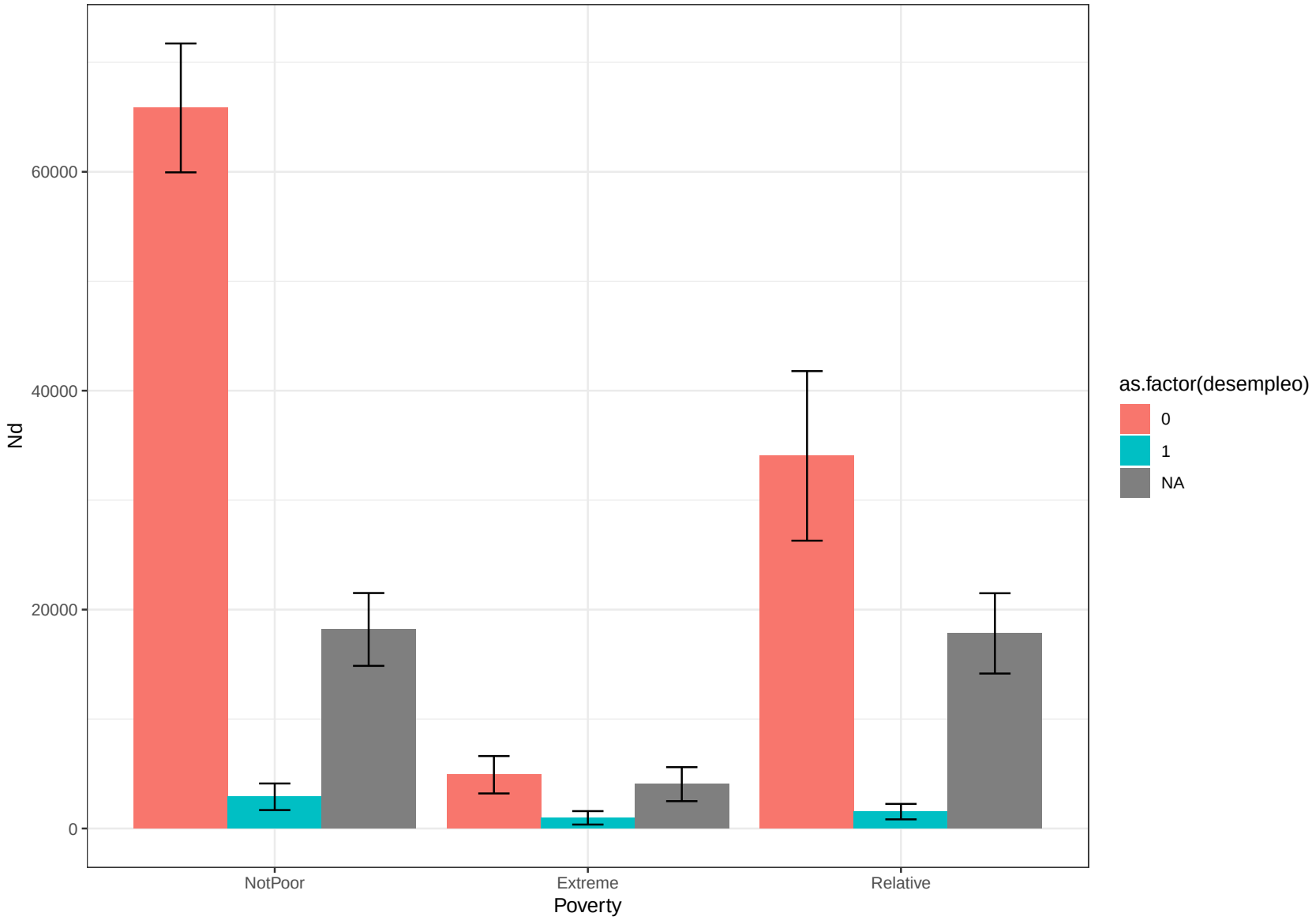


Diagrama de barras

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

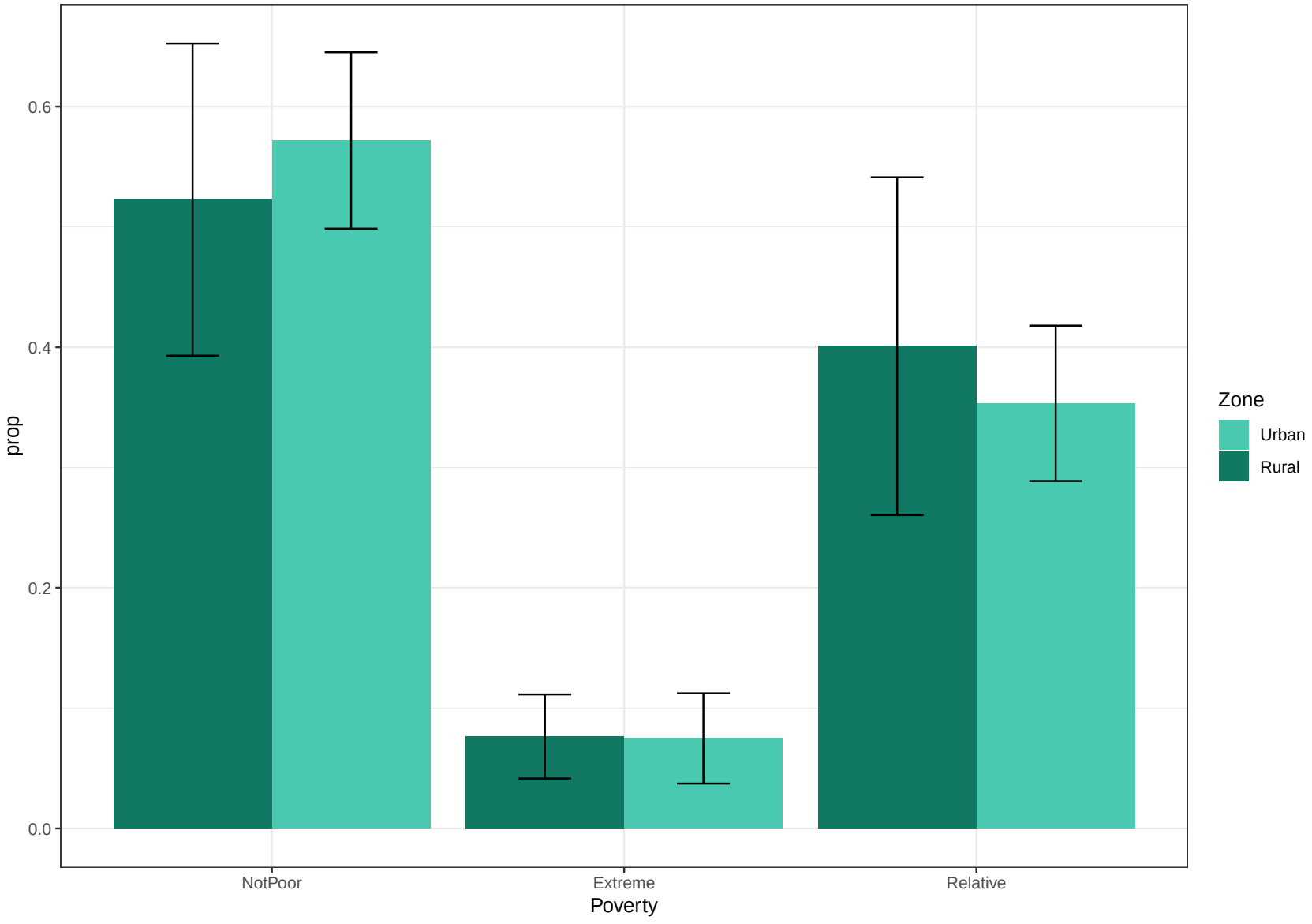
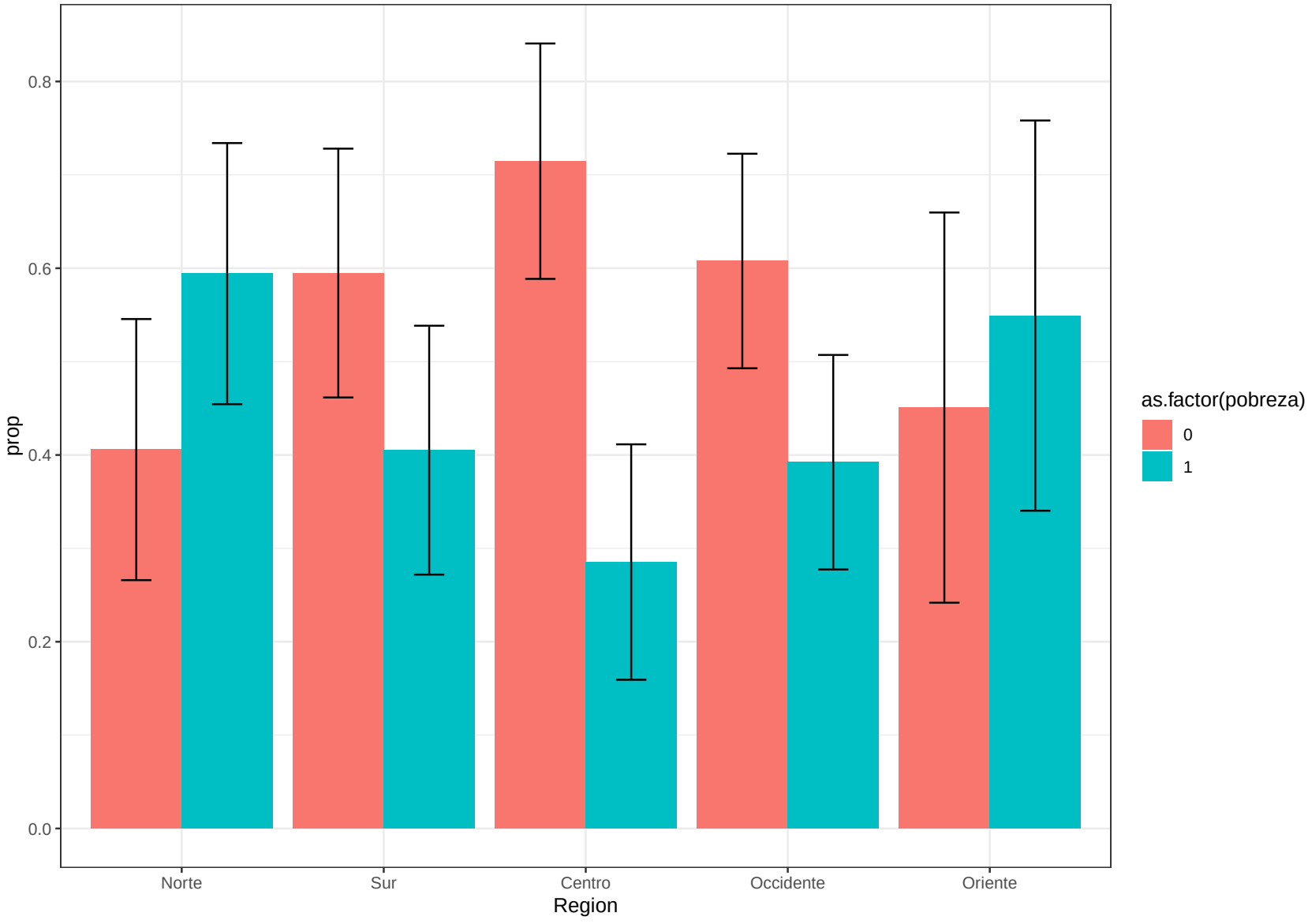


Diagrama de barras

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

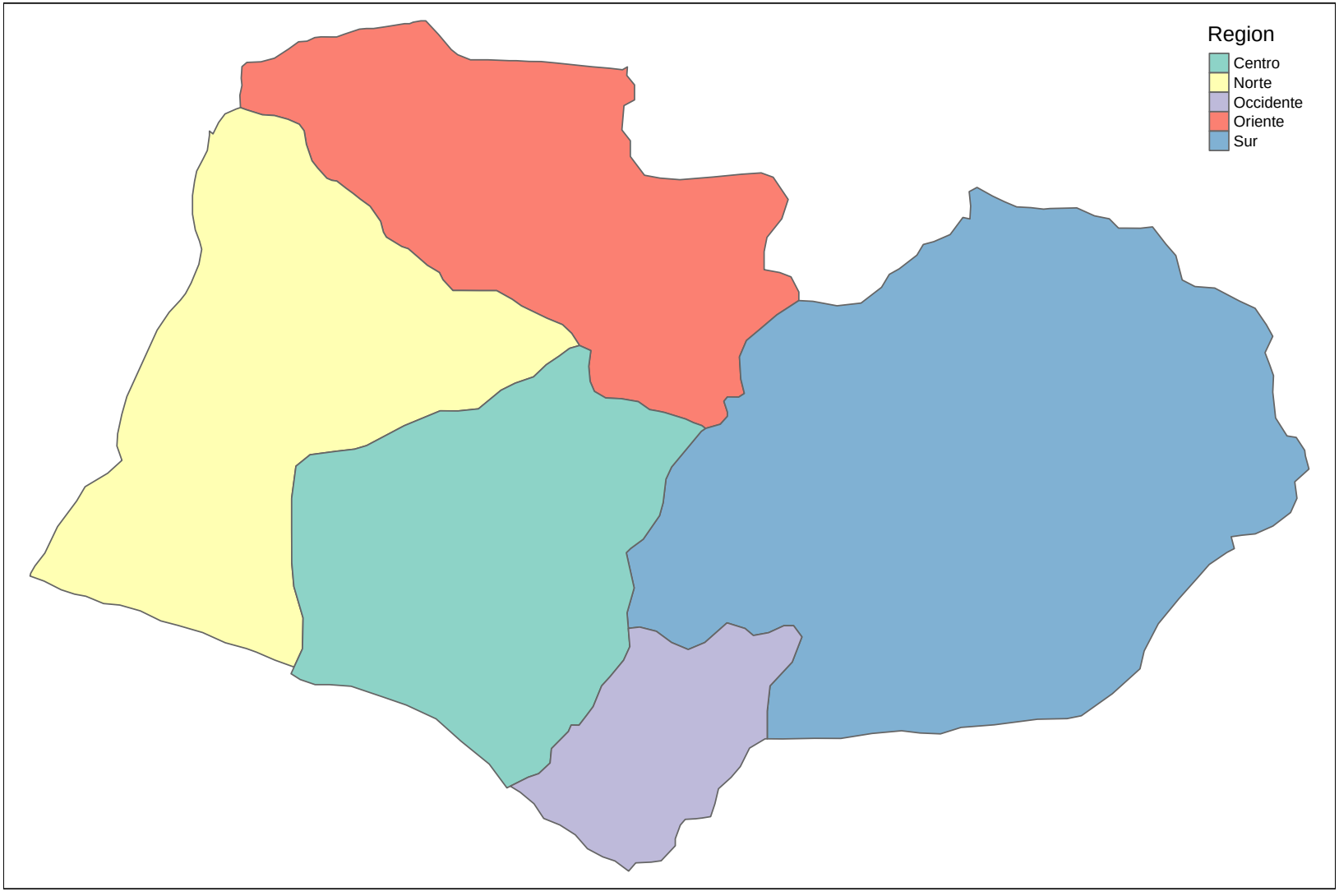


Creando mapas

mapas con tmap.

```
library(sf)
library(tmap)
shapeBigCity <- read_sf( "../Data/shapeBigCity/BigCity.shp"
tm_shape(shapeBigCity) +
tm_polygons(col = "Region")
```

mapas con tmap.

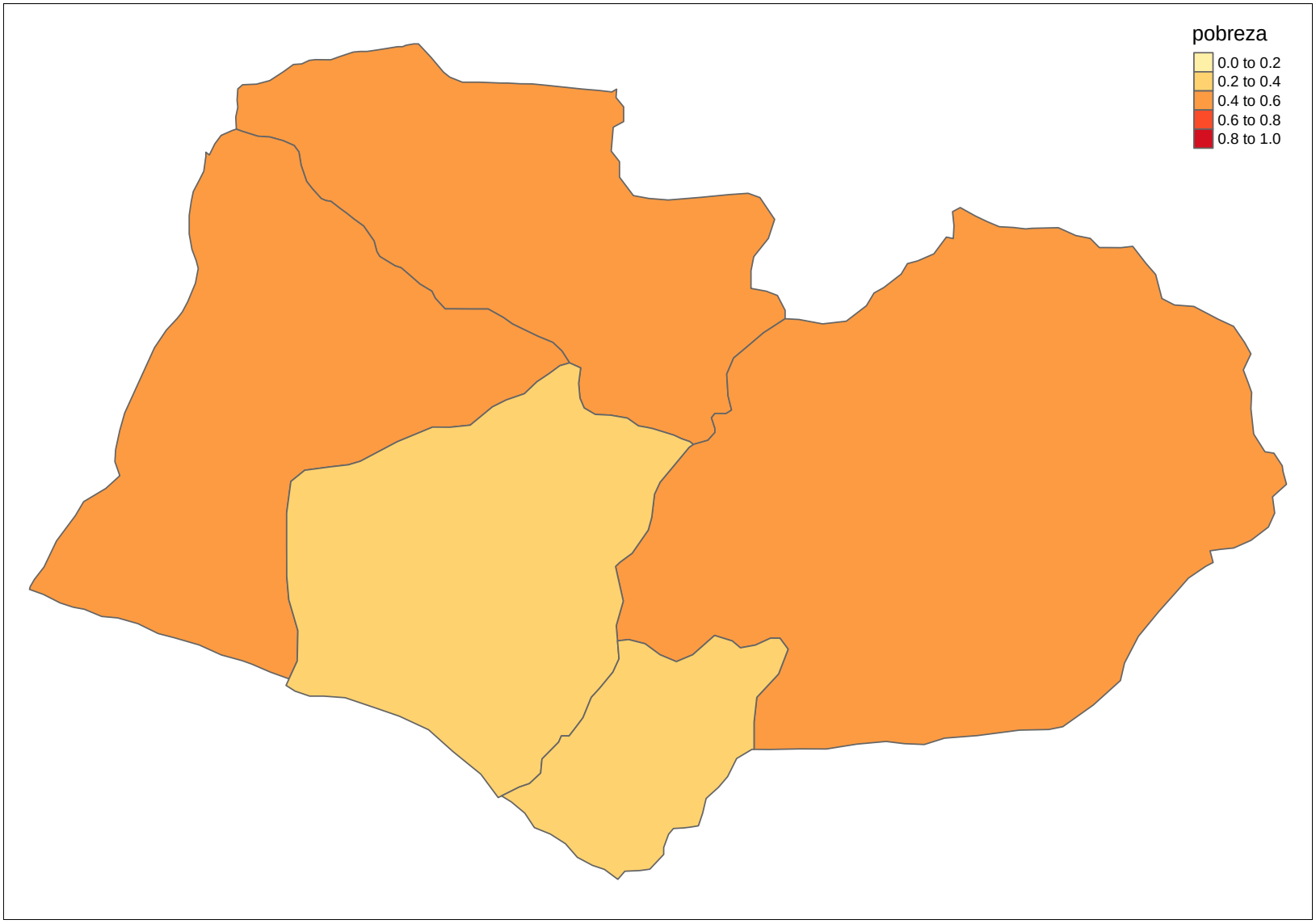


mapas con tmap.

```
brks <- c(0,.2,.4,.6,0.8,1)
shape_temp <- tm_shape(
  shapeBigCity %>%
   left_join(
     prop_RegionH_Pobreza %>%
       filter(pobreza == 1),  by = "Region"))

shape_temp + tm_polygons(
    "prop",
    breaks = brks,
    title = "pobreza",
    palette = "YlOrRd"
  ) + tm_layout(asp = 0)
```

mapas con tmap.



mapas con tmap.

```
(prom_region <- svyby(~Income, ~Region, diseno,
                      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)
```

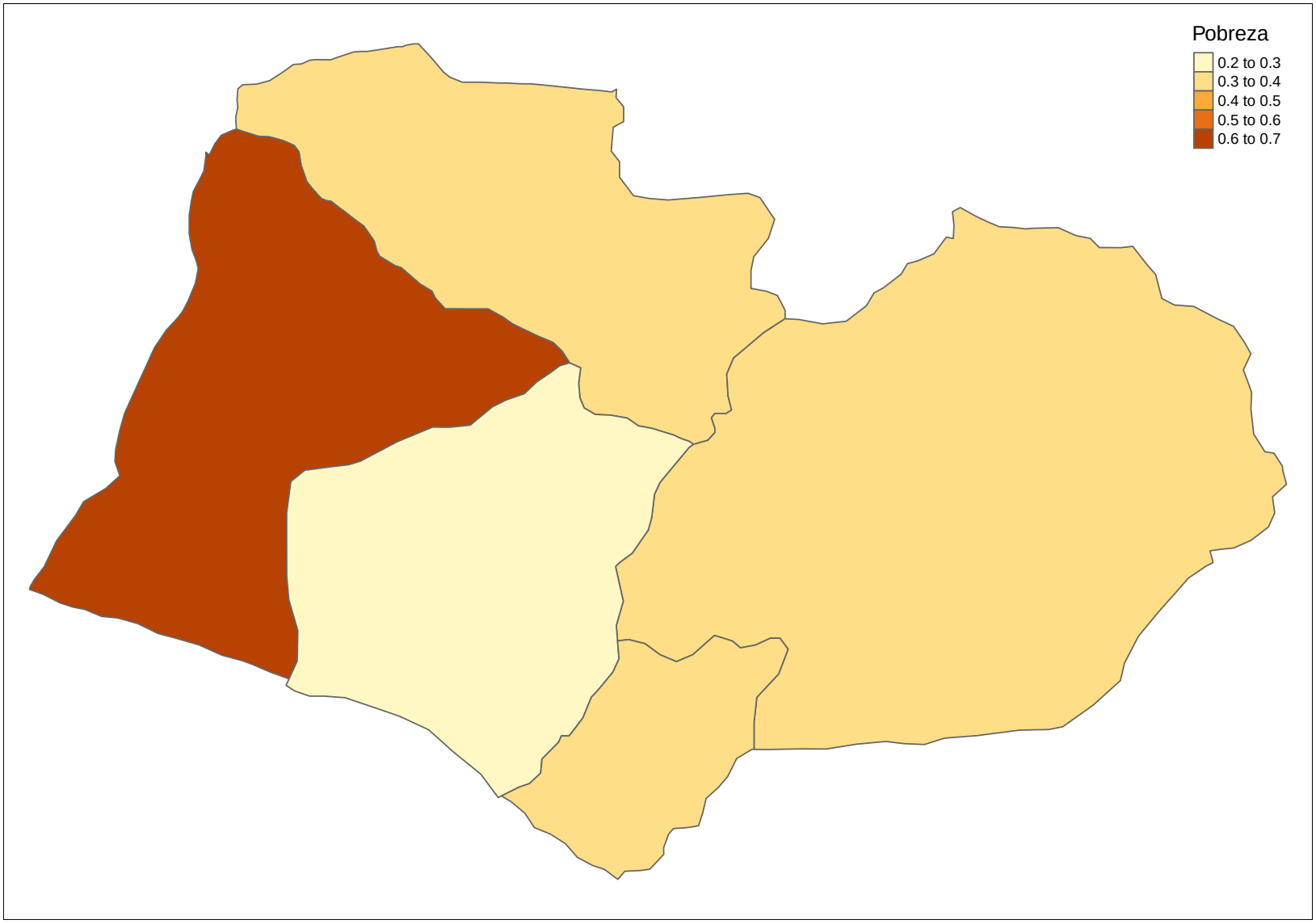
mapas con tmap.

```
prom_region_Sex <- diseno %>% group_by(Region, Zone , Sex,
  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)
```

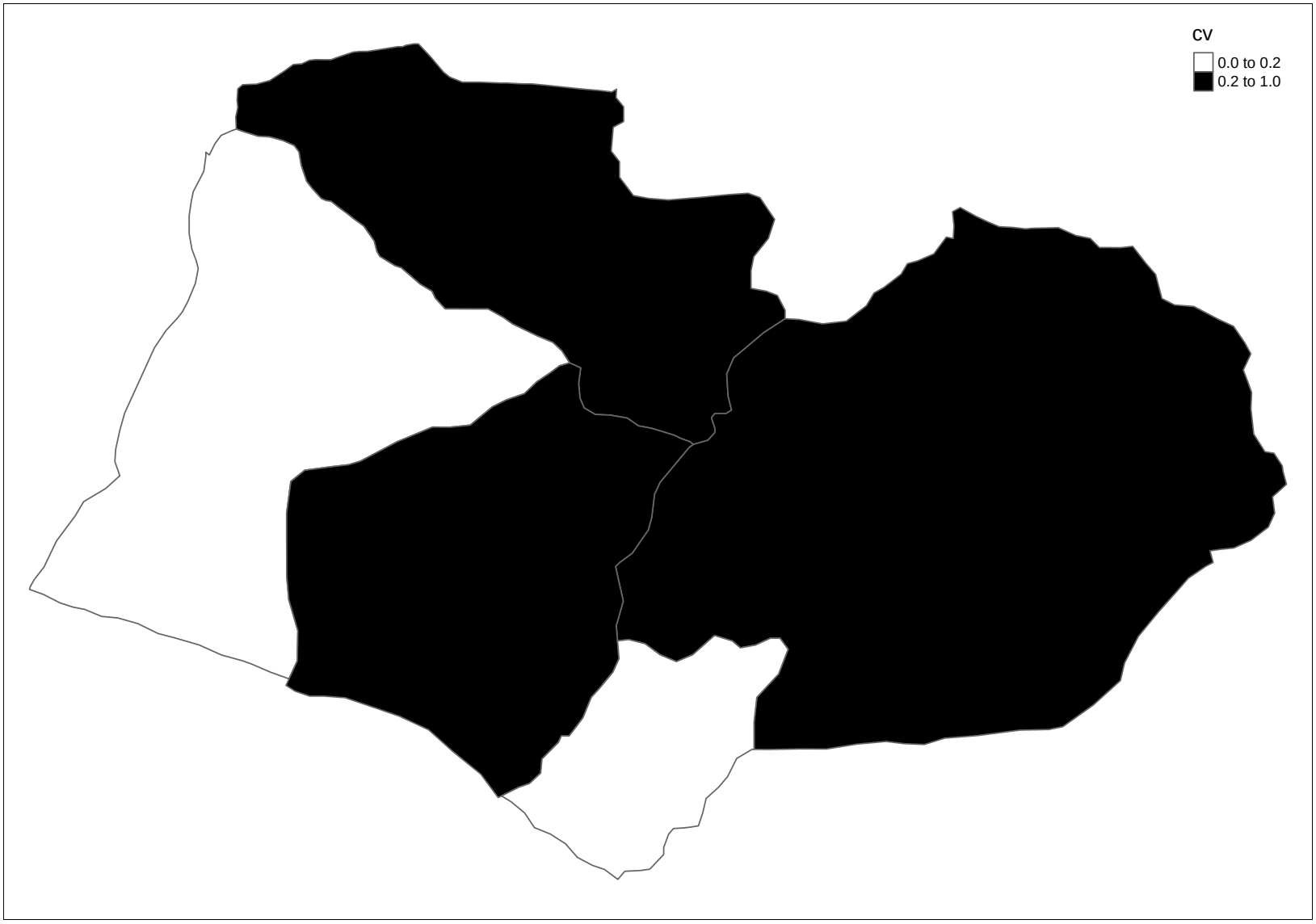
mapas con tmap.



mapas con tmap.

```
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.



mapas con sf

```
pattern <- function(x, pattern) {
  ex = list(
    horizontal = c(1, 2),
    vertical = c(1, 4),
    left2right = c(2, 4),
    right2left = c(1, 3)
  )
  fillgrid = st_make_grid(x)
  endsf = lapply(1:length(fillgrid), function(j)
    sf::st_linestring(sf::st_coordinates(fillgrid[j])[ex[[p
  endsf = sf::st_sfc(endsf, crs = sf::st_crs(x))
  endsf = sf::st_intersection(endsf, x)
  endsf = endsf[sf::st_geometry_type(endsf)
                %in% c("LINESTRING", "MULTILINESTRING")]
  endsf = sf::st_line_merge(sf::st_union(endsf))
  return(endsf)
}
```

mapas con sf

```
temp_shape <- shapeBigCity %>%
   left_join(
     prom_region_Sex ,  by = "Region")

serbgrid = pattern(x = temp_shape[temp_shape$prop_cv>0.2,]

par(mar = c(0, 0, 0, 0))
plot(st_geometry(temp_shape), border = 'red',
                  axes = FALSE)
plot(serbgrid, add = T)
```

mapas con ggplot

```
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)
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

```
# Crear la leyenda para el mapa
legend1 <- bi_legend(pal = "GrPink",  dim = k,
                     xlab = "Coeficiente de variación",
                     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 muje
            vjust = -13, size = 18)

mapa1
```

mapas con ggplot