Anexo 1

Proceso de análisis de la encuesta Einsfoessa realizado con R

# Introducción

La encuesta **EINSFOESSA** se realizó en 2017 por la Fundación FOESSA, publicando los resultados de la misma en el informe de 2019 **VIII Informe sobre Exclusión y Desarrollo Social en España (Fundación FOESSA)**. En este documento trabajaremos sobre los datos de esta encuesta para el análisis y caracterización de la población general y vulnerable en situación de Vulnerabilidad energética (de acuerdo con los indicadores oficiales).

Aquí no entraremos en las especificidades metodológicas de diseño de encuesta de EINSFOESSA, dado que para esta información nos remitimos al capítulo metodológico del VIII Informe.

Sólo destacaremos algunos aspectos por los cuales hemos optado por utilizar los microdatos de esta encuesta (y no las encuestas poblacionales oficiales, como la Encuesta de Condiciones de Vida o la Encuesta de Presupuestos Familiares) para la consecución de los objetivos de investigación.

El **objetivo inicial de la encuesta** es cuantificar y analizar las condiciones de vida y la exclusión social en España, a través de datos representativos a nivel nacional y a nivel de comunidades autónomas.

La justificación para la elección de la base de datos de la encuesta de EINSFOESSA para la caracterización de la población española en situación de vulnerabilidad energética, es que la metodología usada por esta encuesta asegura la inclusión y representación de los hogares con exclusión social. Esta circunstancia nos es de utilidad para poder analizar la prevalencia, características y perfil de la población vulnerable en situación de precariedad energética.

Entre los motivos para el uso de esta base de datos para la caracterización de la población en situación de precariedad energética, apuntamos: 1. La Encuesta EINSFOESSA nos permite comparar los cuatro indicadores oficiales, adoptados por el gobierno español a través de la Estrategia Nacional Contra la Pobreza Energética, en una sola muestra. 2. La Encuesta EINSFOESSA, además, proporciona un amplio abanico de análisis y caracterización de esta población, permitiendo un análisis de las condiciones sociales, demográficas y territoriales de la población en situación de vulnerabilidad energética. 3. El tipo de muestreo utilizado en la encuesta EINSFOESSA asegura la representatividad y presencia en la muestra de hogares en situación de vulnerabilidad que, a menudo, resultan infrarepresentados en las estadísticas oficiales, y en particular, en el caso de la vulnerabilidad y precariedad energética se documenta en los casos de pobreza energética escondida.

La fundación FOESSA, en 2020, también publica un informe específico sobre pobreza energética a partir de los datos de EINSFOESSA. Esta investigación reproduce parcialmente los resultados ya publicados, y los amplía, para profundizar en el conocimiento y caracterización de la población y hogares en situación de vulnerabilidad energética.

*Para el seguimiento del código mostrado en este anexo metodológico se ha creado un proyecto específico de R con los datos necesarios para la ejecución del código. En caso que no se utilize un proyecto de R, deberá modificarse el código debidamente para adaptar las referencias al directorio de trabajo que corresponda.*

## Preparación de los datos a partir de los microdatos de la Encuesta EINSFOESSA 2018

En primer lugar generamos una base de datos reducida en la que sólo tomamos las variables de interés en relación con nuestro objetivo de investigación. Dado que las variables usadas para el cálculo de los indicadores oficiales de pobreza energética que nos servirán como base para nuestro indicador se aplican a nivel de hogar, aplicaremos el filtro hogares de los microdatos proporcionados por la fundación EINSFOESSA. Con este paso, conseguimos una muestra de hogares correspondiente a 11655 hogares.

foessa <- foessa %>% filter(SELEC\_HOGAR == "Selected") #Filtramos y seleccionamos las observaciones correspondientes a hogares  
  
foessa2 <- foessa %>% select(CCAA = CCAA,   
 PROVINCIA = PROVINCIA,   
 tamano\_hogar = PERSONAS,   
 edad = A1,   
 sexo = A2,   
 educacion = B12,   
 salud = C21,   
 dependencia = C26,   
 ingresos\_UC = INGR\_UC\_FIABLE,  
 ingresos\_hogar = INGR\_SUMAFIABLE\_HOGAR,  
 uc\_hogar = UC\_HOGAR,   
 excl4grupos = excl4grupos,   
 exclusion = exclusion,   
 etnia = etnia,   
 monoparental = hmp,   
 anciano = anciano,   
 menor = menor,   
 joven = joven,   
 discapacidad = discapacidad,   
 barrio\_dummy = u1agrup2,   
 barrio = U2,   
 alojamiento = U3,   
 tamano\_municipio = tam\_habitat,   
 ocupado = ocupado,   
 parado = parado,   
 gasto\_energia = E64\_3,   
 gasto\_agua = E64\_4,   
 retrasos = E65,   
 clase\_subjetiva = E70,   
 evolucion12 = E73,   
 avisos\_cortes = E75\_1\_1,   
 dinero\_gastoscasa = E75\_1\_2,   
 reducir\_gfijos = E75\_1\_4,   
 reducir\_galim = E75\_2\_1,   
 dieta\_inadec = E75\_2\_2,   
 reducir\_ocio = E75\_4\_1,   
 perdida\_relaciones =E75\_4\_3,   
 insalubridad = F78\_2,   
 entorno\_degradado = F78\_4,   
 barrio\_conflictivo = F78\_5,   
 necesidad\_vivienda = F79,   
 dispone\_agua = F82\_B\_1,   
 dispone\_agua\_cal = F82\_B\_2,   
 dispone\_elect = F82\_B\_3,   
 dispone\_calef = F82\_B\_6,   
 temp\_adecuada = F82\_B\_18,   
 ppalocupacion = ppalocupacion,  
 tenencia = F77,   
 rehab\_cocina = F80\_1,   
 rehab\_baño = F80\_2,   
 rehab\_instal = F80\_5,   
 rehab\_calef = F80\_6,   
 rehab\_ventana = F80\_7,   
 rehab\_tabiques = F80\_3,   
 rehab\_suelo = F80\_4,   
 rehab\_barreras = F80\_8,   
 dispone\_baño = F82\_B\_5,  
 dispone\_cocina = F82\_B\_8,   
 dispone\_frigo = F82\_B\_9,   
 dispone\_lava = F82\_B\_10,   
 dispone\_pc = F82\_B\_14,   
 dispone\_internet = F82\_B\_15,   
 dispone\_tv = F82\_B\_12,   
 ingresos\_calidad = INGR\_CALIDAD,   
 ruina = F78\_1,   
 discrim\_etnia = I115\_1,   
 discrim\_mujer = I115\_2,   
 discrim\_fisico= I115\_3,   
 discrim\_sexual = I115\_4,   
 discrim\_nunca = I115\_5,   
 discrim\_nosabe = I115\_9,   
 discrim\_nocont = I115\_0,   
 proteina = F82\_B\_17,   
 peso = peso\_corregido5  
 )  
  
  
View(foessa2) # Comprobamos que la base de datos reducida se ha creado correctamente.   
  
# Eliminamos las etiquetas de datos   
  
foessa2 <- remove\_all\_labels(foessa2)

En una primera exploración de la base de datos reducida vemos que esta tiene 29,953 observaciones y un total de 73 variables. Podemos ver que la base de datos está preparada para su análisis partiendo del individuo como unidad de análisis de la encuesta (a pesar que tanto el muestreo como la recogida de datos parten del hogar como unidad de encuesta).

*A continuación adjuntamos un cuadro resumen de las variables creadas:*

## Herramientas metodológicas: funciones propias y preparación de código

### Construcción de tablas de frecuencia para una variable (descriptiva univariante)

En primer lugar, definimos la función que se utilizará para la creación de tablas de frecuencia de los indicadores de pobreza energética (tablas de frecuencia para una sola variable) y que se utilizarán a lo largo de la tesis:

#Función para crear tablas de indicadores (una variable)  
  
tabla\_indicador <- function(x,t){  
   
 # Creamos tabla de frecuencias ponderada y convertimos a objeto flextable   
   
 z <- wtable(x = x,   
 weights = foessa2$peso,   
 stat = "prop",   
 mar = FALSE)  
   
 z <- as.data.frame(z)  
   
 z <- qflextable(z)  
   
 # Definimos las propiedades generales de la tabla   
 set\_table\_properties(z, width = 1, layout = "autofit")  
   
 # Definimos y creamos objetos de texto con formato   
 # Formato general  
 def\_par <- fp\_par(text.align = "center", padding = 5)   
 # Formato texto   
 def\_text <- fp\_text(font.size = 9,   
 italic = FALSE,   
 font.family = "Arial")  
 #Formato de la cabecera   
 def\_text\_header <- fp\_text(font.size = 10,   
 italic = FALSE,   
 font.family = "Arial",  
 color="dodgerblue4",  
 bold = TRUE)  
 # Creamos el título   
 z <- add\_header\_lines(x = z,   
 values = t)  
  
 # Aplicamos los estilos de texto diseñados anteriormente   
 z <- style(x = z, pr\_p = def\_par, pr\_t = def\_text, part = "all")   
 z <- style(x = z, pr\_t = def\_text\_header, part = "header")  
 z <- set\_header\_labels(x = z,   
 Var1 = "Categoría",   
 Freq = "Porcentaje")  
 z <- fontsize(z, i=2, size = 9, part = "header")  
 z <- bold(z, j =1, bold = TRUE, part = "body")  
 z <- color(z, j = 1, color= "dodgerblue4", part = "body")  
   
 # Bordes   
 z <- border\_remove(z) # Primero eliminamos los bordes  
  
 big\_border = fp\_border(color ="dodgerblue4",   
 style = "dotted",   
 width = 2)  
 std\_border = fp\_border(color="dodgerblue4", # color  
 style = "solid", # estilo   
 width = 1) #anchura borde  
 z <- hline(z, i =1, border = big\_border, part = "header")  
 z <- hline(z, i =2, border = std\_border, part = "header")  
 z <- hline\_bottom(z, part="body", border = std\_border )  
   
 # Imprimimos la tabla  
 z  
}  
  
tabla\_indicador(x = foessa2$temp\_adecuada, #variable con marcador $   
 t = "Temperatura adecuada") # Título de la tabla (nombre del indicador)

| **Temperatura adecuada** | |
| --- | --- |
| **Categoría** | **Porcentaje** |
| **Si** | 83.4 |
| **No** | 16.6 |

### Construcción de tablas de contingencia para dos o más variables (descriptiva multivariante)

A continuación se define la función utilizada para la creación de tablas de contingencia presentadas en esta tesis doctoral. Crearemos una función propia para la creación de tablas de contingencia con los datos que iremos obteniendo. La creación de esta función facilita y simplifica la ejecución del código posterior.

crear\_tabla <- function(var1,var2,title,varname,numcol){  
 # Creamos la tabla de frecuencias   
 z <- proc\_freq(foessa2,  
 row = var1, # Variable 1  
 col = var2, # Variable 2  
 include.row\_percent = TRUE,  
 include.column\_percent = TRUE,  
 include.table\_percent = FALSE,  
 include.column\_total = FALSE,  
 include.row\_total = FALSE,  
 include.header\_row = FALSE,   
 weight = "peso")   
   
 # Empezamos a definir las características de formato de la tabla   
 set\_table\_properties(z, width = 1, layout = "autofit")  
   
 # Creamos objetos específicos de definición de formato de texto   
 def\_par <- fp\_par(text.align = "center", padding = 5)  
 def\_text <- fp\_text(font.size = 8, italic = FALSE, font.family = "Arial")  
 def\_text\_header <- fp\_text(font.size = 8, italic = FALSE, bold = FALSE, font.family = "Arial", color="dodgerblue4", underlined = FALSE)  
  
 # Establecemos las cabeceras de la tabla  
 z <- set\_header\_labels(x = z,   
 label = "")  
 z <- add\_header\_row( x = z,   
 values = c(" ","Cálculo", varname),  
 colwidths = c(1,1,numcol))  
 z <- add\_header\_lines(x = z,   
 values = title)  
   
 # Aplicamos los diseños da toda la tabla   
 z <- style(x = z,   
 pr\_p = def\_par,   
 pr\_t = def\_text,   
 part = "all")   
 # Aplicamos los diseños de texto a las cabeceras de la tabla de la tabla  
 z <- style(x = z,   
 pr\_t = def\_text\_header,   
 part = "header")  
 z <- fontsize(z, i = 2, size = 8, part = "header")  
 z <- bold(z, i = 2, part = "header")  
 z <- fontsize(z, i = 1, size = 10, part = "header")  
 z <- bold(z, i = 1, part = "header")  
 #z <- bg(z, i = 2,bg = "aliceblue", part = "header")  
 z <- bg(z, i = 3,bg = "aliceblue", part = "header")  
   
 # Modificamos características específicas del cuerpo de la tabla   
 z <- bg(z, j = 1,bg = "dodgerblue4", part = "body")  
 z <- color(z, j=1, color = "white", part = "body")  
 z <- bold(z, j = 1, part = "body")  
 z <- italic(z, j = 2, part = "body")  
   
 #Definimos los bordes de la tabla   
 z <- border\_remove(z) # Primero eliminamos los bordes  
  
 big\_border = fp\_border(color ="dodgerblue4",   
 style = "dotted",   
 width = 2)  
 std\_border = fp\_border(color="dodgerblue4", # color  
 style = "solid", # estilo   
 width = 1) #anchura borde  
 z <- hline(z, i =1, border = big\_border, part = "header")  
 z <- hline(z, i =2, border = std\_border, part = "header")  
 z <- hline\_bottom(z, part="header", border = std\_border )  
 z <- hline(z, i =3, border = std\_border, part = "body")  
 z <- hline(z, i =6, border = std\_border, part = "body")  
 z <- hline\_bottom(z, part="body", border = std\_border )  
   
 # Refinamos los ajustes y los espaciados de la tabla y cabecera   
 z <- line\_spacing(z, i =2, space = 0.7, part = "header")  
 z <- padding(z, i =3, padding = 2, part = "header")  
  
 # Imprimimos la tabla   
 z  
}  
  
  
  
  
crear\_tabla(var1 ="temp\_adecuada", #variable dependiente  
 var2 ="sexo", # variable independiente  
 title = "Temperatura adecuada y sexo", # título de la tabla   
 varname = "Sexo", # Variable independiente  
 numcol = 2) # Número de valores de la variable independiente

| **Temperatura adecuada y sexo** | | | |
| --- | --- | --- | --- |
|  | **Cálculo** | **Sexo** | |
| temp\_adecuada |  | Varón | Mujer |
| **Si** | *Frequency* | 4084 | 5603 |
| *Row Pct* | 42.16% | 57.84% |
| *Col Pct* | 85.93% | 81.65% |
| **No** | *Frequency* | 669 | 1259 |
| *Row Pct* | 34.68% | 65.32% |
| *Col Pct* | 14.07% | 18.35% |

### Graficación de tablas de contingencia para dos o más variables (descriptiva multivariante)

También definiremos una función propia para el diseño de gráficos de barra con el objetivo de representar las tablas de contingencia más relevantes que creemos a lo largo de esta sección cuantitativa de la tesis:

crear\_grafico <- function(x,y,title,xtitle){  
 df <- foessa2 %>%   
 drop\_na({{x}}) %>%  
 drop\_na({{y}}) %>%  
 group\_by({{y}},{{x}}) %>%   
 tally(wt=peso) %>%   
 complete({{x}}, fill = list(n = 0)) %>%   
 mutate(percentage = n / sum(n) \* 100)  
   
 ggplot(df, aes({{x}}, percentage, fill = {{y}})) +   
 geom\_bar(stat = 'identity', position = 'dodge') +  
 scale\_y\_continuous("Porcentaje de población", expand = c(0,0))+   
 scale\_x\_discrete(xtitle)+   
 scale\_fill\_manual("Vulnerabilidad energética", values = c("#15607a",   
 "#18a1cd",  
 "cyan")) +  
 ggtitle(title) +  
 theme\_classic(base\_size=10) +   
 theme(plot.title = element\_text(face = "bold",   
 size = 10,   
 color = "#00344c",   
 vjust = 3,   
 hjust = 0.5),  
 axis.text.x = element\_text(#angle = 30,   
 hjust = 0.5,   
 vjust = 2,   
 colour = "#00344c",   
 size = rel(0.9)),  
 axis.title.x = element\_text(face = "bold",   
 size = 9),  
 axis.title.y = element\_text(face = "bold",   
 size = 9,   
 vjust = 3),  
 axis.text.y = element\_text(size=8,   
 hjust = 0),  
 axis.line = element\_blank(),  
 axis.ticks.x = element\_blank(),   
 legend.title = element\_text(colour = "#00344c", size = 8, face = "bold"))   
   
}

# Indicadores básicos de vulnerabilidad energética

## Indicador consensual: temperatura adecuada en el hogar

El primer indicador analizado es el indicador consensual de **Temperatura adecuada en el hogar**. Este indicador se construye a partir de la variable temp\_adecuada de la encuesta, que se extrae de la pregunta:

Pregunta F.82: De este listado de cosas, dígame las que considera necesario y de cuales dispone: (…) p) Mantener la vivienda a una temperatura adecuada.

A partir de los resultados, obtenemos la siguiente tabla:

# Si nos fijamos, en el caso del indicador de temperatura adecuada, la persona estará afectada cuando la respuesta es 'No'. Para evitar problemáticas, recodificaremos esta variable para indicar que la respuesta 'Sí' indica que la persona está afectada por este indicador  
  
foessa2$temp\_adecuada\_r <- fct\_recode(foessa2$temp\_adecuada, Sí ="No", No ="Si")  
  
  
tabla\_indicador(x = foessa2$temp\_adecuada\_r,   
 t = "Temperatura inadecuada en el hogar")

| **Temperatura inadecuada en el hogar** | |
| --- | --- |
| **Categoría** | **Porcentaje** |
| **No** | 83.4 |
| **Sí** | 16.6 |

De acuerdo con estos resultados, un 16,2% de la población no puede mantener su hogar a una temperatura adecuada.

## Indicador consensual de retrasos en el pago de facturas de suministros

El segundo indiciador, también de tipo consensual, será el de **Retrasos en el pago de facturas**. A partir de la variable retrasos de la base de datos

Pregunta E.65: ¿Tuvo, durante el año 2017, algún retraso en el pago de recibos (agua, gas, calefacción, electricidad)?

De los resultados obtenidos para la población, extraemos los siguientes resultados:

#Limpiaremos la variable para su posterior recodificación   
foessa2$retrasos <- as.character(foessa2$retrasos) # Convertimos a carácter (character string)  
foessa2$retrasos <- str\_trim(foessa2$retrasos) # Eliminamos los espacios en blanco del string   
foessa2$retrasos <- as.factor(foessa2$retrasos) # Reconvertimos en factor para su posterior manipulación   
  
#Recodificación de la variable (simplificación de niveles de respuesta)  
foessa2$retrasos1 <- fct\_collapse(foessa2$retrasos,   
 "Sí" = c("Sí, dos veces o más","Sí, solamente una vez"),   
 "No" = "No",   
 "NSNC" = c("No contesta", "No sabe"))  
  
foessa2$retrasos2<- fct\_collapse(foessa2$retrasos,   
 "Sí" = "Sí, dos veces o más",   
 "No" = c("No", "Sí, solamente una vez"),  
 "NSNC" = c("No contesta", "No sabe"))  
  
# Eliminar el valor NSNC a NA   
foessa2 <- foessa2 %>% replace\_with\_na(replace = list(retrasos1 = "NSNC"))  
foessa2$retrasos1 <- fct\_drop(foessa2$retrasos1)  
  
# Eliminar el valor NSNC a NA   
foessa2 <- foessa2 %>% replace\_with\_na(replace = list(retrasos2 = "NSNC"))  
  
foessa2$retrasos2 <- fct\_drop(foessa2$retrasos2)  
  
# Indicador de uno o más retrasos  
  
tabla\_indicador(x = foessa2$retrasos1,   
 t = "Uno o más retrasos en el pago de facturas")

| **Uno o más retrasos en el pago de facturas** | |
| --- | --- |
| **Categoría** | **Porcentaje** |
| **No** | 87 |
| **Sí** | 13 |

tabla\_indicador(x = foessa2$retrasos2,   
 t = "Dos o más retrasos en el pago de facturas")

| **Dos o más retrasos en el pago de facturas** | |
| --- | --- |
| **Categoría** | **Porcentaje** |
| **No** | 91.9 |
| **Sí** | 8.1 |

Tal como nos indican los resultados, el 13% de los hogares tuvieron uno o más retrasos en el pago de facturas de suministros durante el año 2017, y un 8.1% tuvo dos o más retrasos.

## Indicador de Gasto desproporcionado (2M)

El tercer indicador analizado es un indicador basado en el gasto. El indicador de Gasto Desproporcionado (2M) se define como el porcentaje de hogares cuya participación del gasto energético en los ingresos es más del doble de la mediana nacional

Para calcular este indicador hemos partido de diversas variables presentes en la encuesta EINSFOESSA, principalmente la variable gasto\_energia e ingresos\_hogar. Estas variables se extraen a partir de las preguntas del cuestionario siguientes:

Pregunta E64\_3 sobre tipos de gastos del hogar, en que se contabilizan los gastos relativos a electricidad y gas.

La variable INGR\_SUMAFIABLE\_HOGAR de los microdatos proporcionados por la Fundación FOESSA, construida a partir de los datos de ingresos familiares obtenidos en la sección de Economía de la encuesta.

# Recodificación de variables   
  
foessa2$gasto\_energia <- fct\_recode(foessa2$gasto\_energia, "0" = "No tiene gasto")#Recodificamos la variable sobre gasto de energia - que, inicialmente, es un factor - para convertir el nivel "No tiene gasto" en O  
  
#Transformamos la variable gasto\_energia en numérica   
foessa2$gasto\_energia\_num <- as.numeric(as.character(foessa2$gasto\_energia))  
  
weightedMedian(foessa2$gasto\_energia\_num, w= foessa2$peso, na.rm = TRUE) #Con la siguiente función calculamos la mediana ponderada de la variable de gasto energético

## [1] 780

A continuación, calcularemos la proporción del gasto energético en relación a los ingresos disponibles de cada uno de los hogares de la base de datos:

#foessa2$gasto\_energia\_eq <- foessa2$gasto\_energia\_num/foessa2$uc\_hogar  
  
foessa2 <- foessa2 %>% mutate(share\_energy = gasto\_energia\_num/ingresos\_hogar\*100)  
  
weightedMedian(foessa2$share\_energy, w= foessa2$peso, na.rm = TRUE) #Con la siguiente función calculamos la mediana ponderada de la variable de gasto energético

## [1] 4

A continuación creamos un indicador que indique si el porcentaje de gasto energético sobre los ingresos del hogar se situan por encima del doble de la mediana.

foessa2 <- foessa2 %>% mutate("TWO\_M" = case\_when(  
 share\_energy > 3.979\*2 ~ "Sí",   
 TRUE ~ "No"))  
  
foessa2$TWO\_M <- as.factor(foessa2$TWO\_M)#Convertir a factor

Finalmente creamos una tabla con los resultados:

tabla\_indicador(x = foessa2$TWO\_M,   
 t = "Gasto desproporcionado")

| **Gasto desproporcionado** | |
| --- | --- |
| **Categoría** | **Porcentaje** |
| **No** | 83.9 |
| **Sí** | 16.1 |

De acuerdo con los cálculos realizados, un 16.1% de los hogares de la muestran tienen un gasto energético desproporcionado, en tanto la proporción de gasto energético sobre sus ingresos se sitúa por encima del doble de la mediana de la muestra.

## Indicador de Pobreza Energética Escondida (M/2)

La pobreza energética escondida (HEP) se define como el porcentaje de los hogares cuyo gasto energético es inferior a la mitad de la mediana. Para el cálculo de este indicador con los datos de EINSFOESSA se ha partido de la variable gasto\_energia.

#Creamos la nueva variable para el indicador "HEP" (Pobreza energética escondida)  
  
foessa2 <- foessa2 %>% mutate("HEP" = case\_when(  
 gasto\_energia\_num < 780/2 ~ "Sí",   
 TRUE ~ "No"  
))  
  
foessa2$HEP <- as.factor(foessa2$HEP)#Convertir a factor  
  
#Ejecutamos la tabla   
tabla\_indicador(x = foessa2$HEP,   
 t = "Pobreza Energética Escondida")

| **Pobreza Energética Escondida** | |
| --- | --- |
| **Categoría** | **Porcentaje** |
| **No** | 85.7 |
| **Sí** | 14.3 |

Los resultados nos muestran que un 14.3% de la población tiene un gasto energético inferior a la mitad de la mediana de gasto.

## Indicador de deficiencias constructivas en el hogar

# Creamos una primera variable que nos indique si el hogar necesita o no rehabilitación  
foessa2 <- foessa2 %>% mutate(rehab = case\_when(  
 necesidad\_vivienda == "Necesitan rehabilitar la vivienda actual" ~ "Sí",   
 TRUE ~ "No"  
))

# Variable sobre deficiencias del hogar:   
  
levels(foessa2$insalubridad) # Variable sobre insalubridad en el hogar

## [1] "No" "Sí" "No Consta"

levels(foessa2$ruina)# Variable sobre si el hogar se encuentra en estado de ruina

## [1] "No" "Sí"

levels(foessa2$rehab) # Variable sobre la necesidad de rehabilitación

## NULL

# Construcción de la variable sobre deficiencias del hogar   
foessa2 <- foessa2 %>% mutate("def\_hogar" = case\_when(  
 insalubridad == "Sí" ~ "Sí",  
 insalubridad == "No" ~ "No",   
 ruina == "Sí" ~ "Sí",   
 ruina == "No" ~ "No",   
 rehab == "Sí" ~ "Sí",  
 rehab == "No" ~ "No",   
 TRUE ~ "No"  
))  
  
  
tabla\_indicador(foessa2$def\_hogar, "Deficiencias en la viviendar")

| **Deficiencias en la viviendar** | |
| --- | --- |
| **Categoría** | **Porcentaje** |
| **No** | 93.2 |
| **Sí** | 6.8 |

### Tipo de rehabilitación necesaria

Una de las variables que componen este indicador sobre deficiencias en la vivienda es si los hogares han expresado la necesidad de realizar algun tipo de rehabilitación.

foessa2$rehab <- as.factor(foessa2$rehab)  
tabla\_indicador(foessa2$rehab, "Necesidad de rehabilitación")

| **Necesidad de rehabilitación** | |
| --- | --- |
| **Categoría** | **Porcentaje** |
| **No** | 94.8 |
| **Sí** | 5.2 |

De los resultados obtenidos en la tabla, podemos apreciar que el porcentaje total de viviendas que necesitan rehabilitación es bajo (sólo un 5.2% de los hogares de la muestra).

En este punto, nos interesa analizar qué tipo de intervenciones de rehabilitación son necesarias en los hogares que así lo han expresado. Para ello crearemos una nueva variable:

#Creación de una nueva variable sobre el tipo de rehabilitción (relevante en términos de precariedad energética)  
foessa2 <- foessa2 %>% mutate(rehab\_tipo =   
 case\_when(rehab == "Sí" & rehab\_cocina == "Sí" ~ "Cocina",   
 rehab == "Sí" & rehab\_baño == "Sí" ~ "Baño",  
 rehab == "Sí" & rehab\_instal == "Sí" ~ "Instalaciones",  
 rehab == "Sí" & rehab\_calef == "Sí" ~ "Calefacción",  
 rehab == "Sí" & rehab\_ventana == "Sí" ~ "Ventana",  
 rehab == "Sí" & rehab\_tabiques == "Sí" ~ "Tabiques",  
 rehab == "Sí" & rehab\_suelo == "Sí" ~ "Suelo",  
 rehab == "Sí" & rehab\_barreras == "Sí" ~ "Barreras arquitectónicas",  
 rehab == "No" ~ "No",  
 TRUE ~ "Otros"))  
  
  
foessa2$rehab\_tipo <- as.factor(foessa2$rehab\_tipo)  
summary(foessa2$rehab\_tipo)

## Baño Barreras arquitectónicas Calefacción   
## 61 10 11   
## Cocina Instalaciones No   
## 101 42 10814   
## Otros Suelo Tabiques   
## 574 8 5   
## Ventana   
## 29

tabla\_indicador(foessa2$rehab\_tipo, "Tipo de rehabilitación")

| **Tipo de rehabilitación** | |
| --- | --- |
| **Categoría** | **Porcentaje** |
| **Baño** | 0.4 |
| **Barreras arquitectónicas** | 0.1 |
| **Calefacción** | 0.0 |
| **Cocina** | 0.5 |
| **Instalaciones** | 0.3 |
| **No** | 94.8 |
| **Otros** | 3.7 |
| **Suelo** | 0.0 |
| **Tabiques** | 0.0 |
| **Ventana** | 0.1 |

## Avisos por impagos

Por último, analizamos también la variable de avisos por impagos, esto es, las personas que han recibido uno o más avisos de corte de suministros por impagos en los últimos doce meses.

tabla\_indicador(foessa2$avisos\_cortes, "Aviso de corte de suministros básicos por impago" )

| **Aviso de corte de suministros básicos por impago** | |
| --- | --- |
| **Categoría** | **Porcentaje** |
| **Sí** | 7.2 |
| **No** | 92.8 |

## Superposición de indicadores: afectación de dos o más indicadores

Como hemos mencionado anteriormente, los indicadores oficiales de pobreza y vulnerabilidad energética se extraren de dos encuestas oficiales distintas: la [Encuesta sobre condiciones de vida](url) y la [Encuesta de Presupuestos Familiares](url). Este hecho ha dificultado el análisis de superposición o coincidencia de distintos factores de vulnerabilidad energética sobre una misma población. El primer análisis de este tipo en España fue llevado a cabo en su último informe sobre pobreza energética de la Asociación Española de Ciencias Ambientales.

En este caso, gracias a la posibilidad de construir todos los indicadores sobre una misma muestra, y además asegurar la representación de hogares vulnerables a través del tipo de muestreo empleado en la encuesta EINSFOESSA, nos es posible analizar la superposición de indicadores y la afectación de los mismos en sus distintos grados a una población específica.

En primer lugar, podemos ver un resumen de los cuatro indicadores construidos hasta el momento:

#Overlapping entre indicadores   
  
## Retrasos en el pago de facturas   
wtable(x = foessa2$retrasos1,   
 weights = foessa2$peso,   
 stat = "prop",   
 mar = FALSE)

## No Sí   
## 87 13

## Temperatura adecuada  
wtable(x = foessa2$temp\_adecuada,   
 weights = foessa2$peso,   
 stat = "prop",   
 mar = FALSE)

## Si No   
## 83.4 16.6

## Gasto desproporcionado  
wtable(x = foessa2$TWO\_M,   
 weights = foessa2$peso,   
 stat = "prop",   
 mar = FALSE)

## No Sí   
## 83.9 16.1

## Pobreza energética escondida   
wtable(x = foessa2$HEP,   
 weights = foessa2$peso,   
 stat = "prop",   
 mar = FALSE)

## No Sí   
## 85.7 14.3

## Deficiencias en la vivienda  
wtable(x = foessa2$def\_hogar,   
 weights = foessa2$peso,   
 stat = "prop",   
 mar = FALSE)

## No Sí   
## 93.2 6.8

A continuación, crearemos una nueva variable para el análisis de la coincidencia de afectación de los distintos indicadores de pobreza energética sobre una misma muestra de hogares. En primer lugar creamos una variable que nos indique si un hogar está afectado por al menos uno de los cuatro indicadores:

#Creamos una nueva variable llamada 'PE' que nos indicará si un individuo está afectado por al menos un indicador:   
  
foessa2 <- foessa2 %>% mutate(PE = case\_when(  
 retrasos1 == "Sí" ~ "PE",   
 temp\_adecuada == "No" ~ "PE",   
 TWO\_M == "Sí" ~ "PE",   
 HEP == "Sí" ~ "PE",   
 def\_hogar == "Sí" ~ "PE",   
 TRUE ~ "NO PE"  
))  
  
foessa2$PE <- as.factor(foessa2$PE) #Transformamos a factor para su interpretación   
  
tabla\_indicador(x = foessa2$PE,   
 t = "Hogares afectados por alguno de los indicadores PE")

| **Hogares afectados por alguno de los indicadores PE** | |
| --- | --- |
| **Categoría** | **Porcentaje** |
| **NO PE** | 54.9 |
| **PE** | 45.1 |

A partir de estos datos podemos ver que un 45.1% de los hogares están afectada por alguno de los indicadores seleccionados para nuestro indicador multidimensional.

A continuación nos interesará evaluar qué proporción de la población de la muestra está afectada por dos o más indicadores:

# Recodificamos las variables afectadas para convertirlas en variables 'dummy'   
df <- dummy\_cols(foessa2,   
 select\_columns = c("retrasos1", "temp\_adecuada", "TWO\_M", "HEP", "def\_hogar"),   
 ignore\_na = TRUE)   
  
df <- df %>% mutate(PE\_Suma = retrasos1\_Sí + temp\_adecuada\_No + TWO\_M\_Sí + HEP\_Sí + def\_hogar\_Sí) #Creamos una variable para saber por cuantos de los indicadores una persona está afectada   
  
df$PE\_Suma <- as.factor(df$PE\_Suma) # Convertimos la nueva variable a factor para su mejor interpretación

También podemos obtener estos datos en una tabla:

tabla\_indicador(x = df$PE\_Suma,   
 t = "Hogares afectados por 1, 2 y 3 indicadores")

| **Hogares afectados por 1, 2 y 3 indicadores** | |
| --- | --- |
| **Categoría** | **Porcentaje** |
| **0** | 55.2 |
| **1** | 29.6 |
| **2** | 9.8 |
| **3** | 4.7 |
| **4** | 0.7 |
| **5** | 0.0 |

A partir de este nuevo indicador, podemos ver: \* Un 29.6% de los hogares están afectados por uno de los indicadores \* Un 9.8% de los hogares están afectados por dos de los indicadores  
\* Un 4.7% de los hogares están afectados por tres de los indicadores  
\* Un 0.7% de los hogares están afectados por cuatro de los indicadores \* Ningun hogar está afectado por todos los indicadores

## Indicador multidimensional de pobreza energética

En este apartado construiremos un indicador multidimensional de vulnerabilidad energética que nos permite, con un solo indicador, contemplar e incluir las distintas dimensiones de vulnerabilidad y precariedad energética.

El objetivo de construir un indicador multidimensional de vulnerabilidad energética es poder analizar las características poblacionales y de contexto, a partir de los datos disponibles, teniendo en cuenta la población afectada

### Selección de variables

Para hacerlo, hemos seleccionado una serie de variables, a partir de la literatura previa y de la selección de indicadores oficiales del estado español.

Los indicadores incorporados a nuestro indicador son los siguientes:

* Temperatura inadecuada en el hogar
* Uno o más atrasos en el pago de facturas
* Gasto excesivo (2M)
* Consumo energético excesivamente bajo (HEP)
* Deficiencias en el hogar (construido a partir de variables existentes relativas al estado de las viviendas)

Los cuatro primeros indicadores, que se corresponden con los indicadores oficiales de pobreza energética, hacen referencia a indicadores relativos a la situación económica y el gasto energético (indicador de retrasos y el indicador de gasto desproporcional), consumo energético (indicador de consumo energético excesivamente bajo) y de confort térmico (mantener la vivienda a una temperatura adecuada). El último indicador, sobre deficiencias del hogar, se ha creado a partir de variables existentes en la base de datos.

### Cálculo de redundancias

En cuanto al método de construcción de este indicador multidimensional utilizaremos el método descrito por Sokolowski et al. (2020) basado en la propuesta de Alkire et al. (2015). Este método se basa en la construcción de un indicador multidimensional, a partir de la previa selección de indicadores, teniendo en cuenta el cálculo de redundancia entre las variables seleccionadas.

# CREAMOS UNA FUNCIÓN PARA CALCULAR LA REDUNDANCIA   
  
redundancia <- function(var1,var2){  
 var1 <- enquo(var1)  
 var2 <- enquo(var2)  
   
 #Total población  
 total <- nrow(foessa2)  
   
 #Porcentaje de población afectada por el indicador 1  
 ind\_1\_p <- nrow(foessa2 %>% filter(!!var1 == "Sí"))  
 ind\_1\_p <- (ind\_1\_p/total)\*100  
   
 #Porcentaje de población afectada por el indicador 2  
 ind\_2\_p <- nrow(foessa2 %>% filter(!!var2 == "Sí"))  
 ind\_2\_p <- (ind\_2\_p/total)\*100  
   
 # Porcentaje de población afectada por ambos indicadores   
 ind\_mix <- nrow(foessa2 %>% filter(!!var1 == "Sí" & !!var2 == "Sí"))  
 ind\_mix <- (ind\_mix/total)\*100  
   
 dos <- c(ind\_1\_p, ind\_2\_p)  
   
 #Cálculo de la redundancia  
 red <- ind\_mix/min(dos)  
 red  
}

A partir de esta función, calcularemos la redundancia entre las variables seleccionadas para el indicador multidimensional que construiremos:

#Redundancia entre temperatura inadecuada y retrasos - 0.56  
redundancia(var1 = temp\_adecuada\_r,   
 var2 = retrasos1)

## [1] 0.5639936

#Redundancia entre temperatura inadecuada y 2M - 0.37  
redundancia(var1 = temp\_adecuada\_r,   
 var2 = TWO\_M)

## [1] 0.3516833

#Redundancia entre temperatura inadecuada y 2M - 0.16  
redundancia(var1 = temp\_adecuada\_r,   
 var2 = HEP)

## [1] 0.2839572

#Redundancia entre retrasos facturas y 2M - 0.33  
redundancia(var1 = retrasos1,   
 var2 = TWO\_M)

## [1] 0.3159851

#Redundancia entre retrasos facturas y hep - 0.12  
redundancia(var1 = retrasos1,   
 var2 = HEP)

## [1] 0.1935829

#Redundancia entre 2m y hep - 0  
redundancia(var1 = TWO\_M,   
 var2 = HEP)

## [1] 0.04919786

#Redundancia entre temperatura inadecuada y deficiencias en el hogar - 0.41  
redundancia(var1 = temp\_adecuada\_r,   
 var2 = def\_hogar)

## [1] 0.4106796

#Redundancia entre retrasos y deficiencias en el hogar - 0.30  
redundancia(var1 = retrasos1,   
 var2 = def\_hogar)

## [1] 0.3067961

#Redundancia entre 2M y deficiencias en el hogar - 0.24  
redundancia(var1 = TWO\_M,   
 var2 = def\_hogar)

## [1] 0.2427184

#Redundancia entre HEP y deficiencias en el hogar - 0.23  
redundancia(var1 = HEP,   
 var2 = def\_hogar)

## [1] 0.2320388

A continuación podemos ver los valores obtenidos a partir del cálculo de redundancias anterior:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Temperatura inadecuada | Retrasos facturas | 2M | HEP | Deficiencias hogar |
| Temperaturainadecuada | 1 | 0.56 | 0.37 | 0.16 | 0.41 |
| Retrasos facturas | 0.56 | 1 | 0.33 | 0.12 | 0.30 |
| 2M | 0.37 | 0.33 | 1 | 0 | 0.24 |
| HEP | 0.16 | 0.12 | 0 | 1 | 0.23 |
| Deficienciashogar | 0.42 | 0.35 | 0.22 | 0.16 | 1 |

En general, podemos ver que las correlaciones y redundancias encontradas no son altas, lo que nos indica que los indicadores seleccionados están analizando dimensiones del fenómeno diferentes. Esto ocurre en todos los casos, excepto entre las variables subjetivas de temperatura inadecuada en el hogar y retrasos en el pago de facturas de suministros. El índice de redundancia entre estos dos indicadores es de 0.56, mostrando que en un 56% de los casos, la población afectada por el indicador de temperatura inadecada en el hogar también está afectada por el indicador de retrasos en el pago de facturas de suministros. Otro índice que destaca es el índice de redundancia y correlación entre Deficiencias del hogar y el indicador subjetivo de temperatura inadecuada, aún sin superar el umbral de 0.5.

Dada esta circunstancia, se tomarán decisiones en la asignación del peso relativo a estas dimensiones, con el objetivo de no sobredimensionar estos indicadores que nos mostrarían un grado relativamente alto de correlación.

### Construcción del indicador

Como hemos indicado, en este punto construiremos un indicador multidimensional que tendrá en cuenta los cuatro subindicadores anteriores. Para ello se tomarán una serie de decisiones en relación, por una parte, al peso relativo de cada uno de los subindicadores, y por otra parte, la forma de agregación.

En lo que respecta al peso relativo, se aplicará la siguiente distribución:

|  |  |
| --- | --- |
| Indicador | Peso relativo sobre 1 |
| Temperaturainadecuada | 0.150 |
| Retrasosfactura | 0.150 |
| 2M | 0.275 |
| HEP | 0.275 |
| Deficienciasvivienda | 0.150 |

Se ha partido de una asignación igualitaria de peso a los diferentes indicadores, matizando posteriormente el peso asignado a las variables subjetivas de temperatura inadecuada y retrasos en las facturas por haber obtenido índices de redundancia sensiblemente superiores al resto.

En relacióna al modo de agregación, se optará por una agregación additiva:

foessa2 <- dummy\_cols(foessa2,   
 select\_columns = c("retrasos1", "temp\_adecuada", "TWO\_M", "HEP", "def\_hogar"),  
 remove\_first\_dummy = TRUE,  
 ignore\_na = TRUE)   
  
# Asignamos el peso relativo a cada una de las columnas   
  
foessa2$temp\_adecuada\_No\_w <- foessa2$temp\_adecuada\_No\*0.150  
foessa2$retrasos1\_Sí\_w <- foessa2$retrasos1\_Sí\*0.150  
foessa2$TWO\_M\_Sí\_w <- foessa2$TWO\_M\_Sí\*0.275  
foessa2$HEP\_Sí\_w <- foessa2$HEP\_Sí\*0.275  
foessa2$def\_hogar\_Sí\_w <- foessa2$def\_hogar\_Sí\*0.150  
  
foessa2$Vulnerabilidad\_num <- foessa2 %>% select(temp\_adecuada\_No\_w, retrasos1\_Sí\_w, TWO\_M\_Sí\_w, HEP\_Sí\_w) %>% rowSums(na.rm=TRUE)  
  
foessa2$Vulnerabilidad\_num <- as.numeric(foessa2$Vulnerabilidad\_num)  
  
summary(foessa2$Vulnerabilidad\_num)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 0.0000 0.0000 0.0000 0.1523 0.2750 0.8500

Una vez hemos calculado el índice que irá desde 0 a 1, podemos agrupar las observaciones en cuatro grupos:

foessa2 <- foessa2 %>% mutate ("Vulnerabilidad\_Energetica" = case\_when(  
 Vulnerabilidad\_num == 0 ~ "No vulnerable",   
 Vulnerabilidad\_num > 0 & Vulnerabilidad\_num <=0.5 ~ "Vulnerabilidad baja",   
 Vulnerabilidad\_num >0.5 ~ "Vulnerabilidad alta",   
 TRUE ~ "No vulnerable"  
))  
  
# Ordenamos los niveles obtenidos   
foessa2$Vulnerabilidad\_Energetica <- factor(foessa2$Vulnerabilidad\_Energetica,  
 levels = c("No vulnerable", "Vulnerabilidad baja", "Vulnerabilidad alta"),  
 ordered = TRUE)  
  
wtable(x = foessa2$Vulnerabilidad\_Energetica,   
 weights = foessa2$peso, stat = "prop")

## No vulnerable Vulnerabilidad baja Vulnerabilidad alta Sum   
## 57.6 38.0 4.4 100.0

También crearemos una segunda variable de Pobreza Energética que actúe como variable dicotómica, a partir de los datos anteriores:

foessa2 <- foessa2 %>% mutate ("Vulnerabilidad\_dummy" = case\_when(  
 Vulnerabilidad\_num > 0.1 ~ "Vulnerable",   
 TRUE ~ "No vulnerable"  
))  
  
foessa2$Vulnerabilidad\_dummy <- as.factor(foessa2$Vulnerabilidad\_dummy)  
  
wtable(x = foessa2$Vulnerabilidad\_dummy,   
 weights = foessa2$peso, stat = "prop")

## No vulnerable Vulnerable Sum   
## 57.6 42.4 100.0

# Caracterización de la población en situación de precariedad energética

## Análisis descriptivo y tablas de contingencia entre variables

Hasta el momento nos hemos centrado en la construcción de los indicadores de vulnerabilidad energética que nos ayuden a identificar la población en este tipo de situación. Ahora bien, también nos interesará ir más allá de los indicadores y caracterizar esta población. Para ello, combinaremos las variables de caracterización propuestas por el gobierno español en su Estrategia Nacional contra la Pobreza Energética, con el objetivo de facilitar la comparación de datos, y por otro lado, aprovechando la amplitud de la información recogida por la encuesta de EINSFOESSA, incorporaremos nuevas variables de caracterización que nos proporcionan una imagen más concreta y definida de las condiciones de vida de la población afectada por algunos de los indicadores de vulnerabilidad energética.

A continuación especificamos las dimensiones en torno a las cuales hemos agrupado las variables de caracterización analizadas: 1. Dimensión 1: Características familiares 2. Dimensión 2: características económicas 4. Dimensión 3: Características de la vivienda 3. Dimensión 4: Características del entorno urbano 5. Dimensión 5: Impactos sobre la salud y relaciones sociales 6. Dimensión 6: Vulnerabilidad energética y colectivos minoritarios

## Prueba de hipótesis de independencia estadística entre variables y análisis de asociación

El análisis exploratorio de cada una de las variables de caracterización, será complementado con un análisis de asociación entre variables. En el caso concreto que nos ocupa, trabajamos fundamentalmente con tablas de contingencia para poner en relación variables categóricas.

Siguiendo la propuesta de Puente Viedma(2009), nuestro análisis de asociación será de tipo frecuentista y constará de tres partes: (1) la detección de asociación a través del Chi Cuadrado, (2) la fuerza (V de Cramer) y la dirección de la asociación, y, finalmente, (3) a qué celdas se debe la asociación (a través de un análisis de residuos).

##### Chi Cuadrado (x2)

La prueba de Chi cuadrado como test de independencia compara dos o más patrones de frecuencias con el objetivo de comprobar si son diferentes entre ellos, es decir, si son independientes.

Verificaremos si existe algun tipo de relación entre variables comparando los datos observados en nuestra muestra con una tabla ideal que representa la distribución recíproca que tendrían esas variables si estuvieran distribuidas al azar. Esta tabla (que es una creación estadística) la llamamos *modelo de independencia* y nos indica la forma que tendrían los conteos si las variables fueran independientes. Para obtener el estadístico usamos la formula χ2 (Chi cuadrado). Al aplicar una prueba de hipótesis de independencia estadística para una tabla de contingencia estaremos estimando la probabilidad de nuestra tabla, dada la hiótesis de independencia entre filas y columnas.

Para poder interpretar los datos que presentamos a continuación debemos tener en cuenta que la *prueba de hipótesis* sólo nos informa sobre la probabilidad de nuestros datos dada una hipótesis de nulidad. Debe tenerse en cuenta que rechazar la hipótesis de nulidad no afirma automáticamente nuestra hipótesis ni tampoco nos da información sobre la calidad de la hipótesis de nulidad.

Al usar el Chi Cuadrado como test de independencia debemos tener en cuenta una serie de condiciones, entre ellas destacamos (Puente Viedma, 2009; Hinton, 2005):

* La muestra debe ser aleatoria.
* Las categorías deben ser mutuamente excluyentes
* La frecuencia esperada en una celda debe ser mayor a 5. En caso que no lo sea, deberá optarse por no aplicar Chi cuadrado.
* No se recomienda usar Chi cuadrado en las tablas de 2x2 dado que la curva no se ajusta a una curva normal dificultando la aplicación de criterios de probabilidad.

##### Intensidad o fuerza de la relación: V de Cramer y Coeficiente de contingencia

La *V de Cramer* es una extensión del coeficiente Phi que se encuentra normalizada. Esta medida de asociación nos indica la intensidad de la relación entre las dos variables. Los valores que obtenemos de la V de Cramer oscilan entre 0 y 1 (en que los valores prócimos a 0 indican no asociación y los cercanos a 1 fuerte asociación).

Para su interpretación, optaremos por los siguientes niveles:

|  |  |
| --- | --- |
| Resultado | Interpretación |
| Resultado entre 0 y 0,2 | No hay asociación |
| Resultado de 0,2 indica | Asociación debil |
| Resultado entre 0,2 y 0,6 | Asociación moderada |
| Resultado entre 0,6 y 1 | Asociación fuerte |

El *Coeficiente de contingencia* esuna medida de asociación derivada del Chi cuadrado. Este coeficiente nos muestra un valor entre 0 y 1, en que 0 indica que no hay asociación entre las variables y 1 que existe una gran relación. El mayor inconveniente que tiene esta medida es que es proporcional al número de observaciones. Estop implica que el valor nunca alcanzará a ser 1 sinó que solo llegará a un máximo dependiente del número de filas y columnas.

##### Dirección de la relación: Gamma de Goodman and Kruskal

El test de Goodman-Kruskal gamma (γ) (Goodman & Kruskal, 1954) comprueba la relación monotónia entre dos variables ordinales. En este caso disponemos de las variables quintiles y vulnerabilidad energética, ambas variables categóricas ordinales.

Para valorar la intensidad de la asociación usaremos el baremo propuesto por Rea and Parker (1992):

0.00 < 0.10 - No asociación 0.10 < 0.20 - Debil 0.20 < 0.40 - Moderada 0.40 < 0.60 - Relativamente fuerte 0.60 < 0.80 - Fuerte 0.80 < 1.00 - Muy fuerte

##### Análisis de residuos estandarizados de Pearson

Tal como apunta Sharpe (2015), en muchas ocasiones, los científicos sociales sólo llegan hasta el test de independencia pero no van más allá. En esta investigación, complementaremos el test de independencia junto con un *análisis de residuos* que nos permitirá discernir entre qué categorías se da la asociación, en caso que se detecte una relación significativa entre variables. A continuación, podemos visualizar los residuos de Pearson que hemos extraído a través de la prueba de Chi Cuadrado, para ver qué interacciones contribuyen en mayor medida al cálculo del estadístico.

Mientras que el test de Chi cuadrado nos indica si existe evidencia contra la hipótesis nula, los residuos de Pearson - mediante una comparación celda a celda de la tabla de contingencia - nos permite entender mejor esta evidencia (Agresti, 2007, p.38)

En este caso, utilizaremos residuos estandarizados. Tal como indica Agresti (2019) los residuos estandarizados nos muestran el grado de discrepancia entre los valores esperados y los valores observados. Residuos estandarizados cercanos a cero nos indican que no existe una gran desviación, mientras que residuos más altos mostrarán una mayor desviación.

##### Razón de probabilidades (Odds Ratio)

En determinadas ocasiones, en el caso de variables de caracterización binarias, optaremos por el cálculo de la razón de probabilidades.

La razón de momios o razón de probabilidades nos da información sobre la ratio entre las probabilidades de éxito de un suceso respecto a las probabilidades de fracaso. Por lo tanto, establece una correlación entre las variables (no causalidad).

## Dimensión 1: Características familiares

### Vulnerabilidad Energética según tamaño del hogar

Para analizar la distribución de la población en situación de vulnerabilidad y precariedad energética según el tamaño del hogar, primero necesitaremos crear una variable de tamaño del hogar que nos permite su comparabilidad:

foessa2$tamano\_hogarf <- as.factor(foessa2$tamano\_hogar)  
levels(foessa2$tamano\_hogarf) # Verificamos el número de niveles iniciales de la variable tamaño hogar

## [1] "1" "2" "3" "4" "5" "6" "7" "8" "9" "10"

foessa2$tamano\_hogarf <- fct\_collapse(foessa2$tamano\_hogarf,   
 "1 miembro" = c("1"),   
 "2 miembros" = c("2"),   
 "3 miembros" = c("3"),   
 "4 miembros" = c("4"),   
 "5 miembros o más" = c("5", "6", "7", "8", "9", "10"),   
 other\_level = NULL)  
  
levels(foessa2$tamano\_hogarf) # Verificamos el número de niveles finales de la variable recodificada

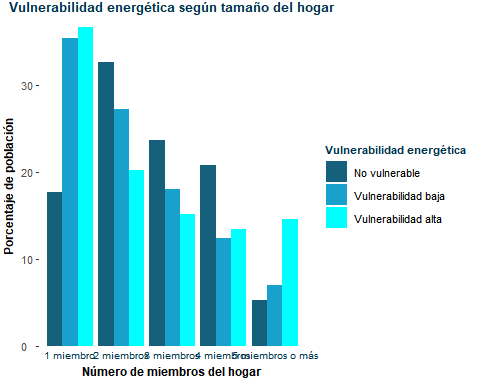
## [1] "1 miembro" "2 miembros" "3 miembros" "4 miembros"   
## [5] "5 miembros o más"

Una vez tenemos creada la nueva variable de tamaño del hogar, podemos generar tablas de contingencia a partir de las variables de interés para los distintos indicadores, empezando con la variable general de PE que nos indica la población afectada por alguno de los indicadores:

crear\_tabla(var1 ="Vulnerabilidad\_Energetica", #variable dependiente  
 var2 ="tamano\_hogarf", # variable independiente  
 title = "vulnerabilidad energética según tamaño del hogar", # título de la tabla   
 varname = "Número de miembros del hogar", # Variable independiente  
 numcol = 5) # Número de valores de la variable independiente

| **vulnerabilidad energética según tamaño del hogar** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Cálculo** | **Número de miembros del hogar** | | | | |
| Vulnerabilidad\_Energetica |  | 1 miembro | 2 miembros | 3 miembros | 4 miembros | 5 miembros o más |
| **No vulnerable** | *Frequency* | 1186 | 2196 | 1591 | 1394 | 351 |
| *Row Pct* | 17.66% | 32.69% | 23.68% | 20.75% | 5.22% |
| *Col Pct* | 40.35% | 62.62% | 64.56% | 69.31% | 47.84% |
| **Vulnerabilidad baja** | *Frequency* | 1567 | 1208 | 796 | 549 | 309 |
| *Row Pct* | 35.38% | 27.27% | 17.98% | 12.4% | 6.97% |
| *Col Pct* | 53.3% | 34.44% | 32.31% | 27.31% | 42.07% |
| **Vulnerabilidad alta** | *Frequency* | 187 | 103 | 77 | 68 | 74 |
| *Row Pct* | 36.69% | 20.22% | 15.18% | 13.37% | 14.55% |
| *Col Pct* | 6.35% | 2.93% | 3.13% | 3.38% | 10.09% |

crear\_grafico(x = tamano\_hogarf, #variable independiente  
 y = Vulnerabilidad\_Energetica, # variable dependiente   
 title = "Vulnerabilidad energética según tamaño del hogar", # título del gráfico  
 xtitle = "Número de miembros del hogar") # título axis X



#### Indicadores individuales

##### Indicador de temperatura adecuada segun tamaño del hogar

crear\_tabla (var1 = "temp\_adecuada",   
 var2 = "tamano\_hogarf",   
 title = "Temperatura adecuada según tamaño del hogar",   
 varname = "Puede tener el hogar a una temperatura adecuada",   
 numcol = 5)

| **Temperatura adecuada según tamaño del hogar** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Cálculo** | **Puede tener el hogar a una temperatura adecuada** | | | | |
| temp\_adecuada |  | 1 miembro | 2 miembros | 3 miembros | 4 miembros | 5 miembros o más |
| **Si** | *Frequency* | 2317 | 2979 | 2070 | 1782 | 539 |
| *Row Pct* | 23.91% | 30.75% | 21.37% | 18.4% | 5.57% |
| *Col Pct* | 79.13% | 85.38% | 84.24% | 88.7% | 73.7% |
| **No** | *Frequency* | 611 | 510 | 387 | 227 | 192 |
| *Row Pct* | 31.68% | 26.46% | 20.09% | 11.78% | 9.98% |
| *Col Pct* | 20.87% | 14.62% | 15.76% | 11.3% | 26.3% |

##### Indicador de retrasos en el pago de facturas de suministros segun tamaño del hogar

crear\_tabla("retrasos1",  
 "tamano\_hogarf",   
 "Retrasos en el pago de factures según tamaño del hogar",   
 "Número de miembros del hogar",   
 5)

| **Retrasos en el pago de factures según tamaño del hogar** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Cálculo** | **Número de miembros del hogar** | | | | |
| retrasos1 |  | 1 miembro | 2 miembros | 3 miembros | 4 miembros | 5 miembros o más |
| **No** | *Frequency* | 2563 | 3189 | 2133 | 1751 | 484 |
| *Row Pct* | 25.33% | 31.51% | 21.08% | 17.3% | 4.78% |
| *Col Pct* | 87.35% | 91% | 86.63% | 87.23% | 66.48% |
| **Sí** | *Frequency* | 371 | 315 | 329 | 256 | 244 |
| *Row Pct* | 24.49% | 20.79% | 21.71% | 16.91% | 16.1% |
| *Col Pct* | 12.65% | 9% | 13.37% | 12.77% | 33.52% |

##### Indicador de gasto desproporcionado (2M) segun tamaño del hogar

crear\_tabla( "TWO\_M",  
 "tamano\_hogarf",   
 "Gasto desproporcionado según tamaño del hogar",   
 "Número de miembros del hogar", 5)

| **Gasto desproporcionado según tamaño del hogar** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Cálculo** | **Número de miembros del hogar** | | | | |
| TWO\_M |  | 1 miembro | 2 miembros | 3 miembros | 4 miembros | 5 miembros o más |
| **No** | *Frequency* | 2252 | 2958 | 2184 | 1783 | 598 |
| *Row Pct* | 23.04% | 30.26% | 22.35% | 18.24% | 6.12% |
| *Col Pct* | 76.6% | 84.35% | 88.66% | 88.64% | 81.52% |
| **Sí** | *Frequency* | 688 | 549 | 279 | 228 | 136 |
| *Row Pct* | 36.6% | 29.18% | 14.86% | 12.15% | 7.21% |
| *Col Pct* | 23.4% | 15.65% | 11.34% | 11.36% | 18.48% |

##### Indicador de Pobreza Energética Escondida (M/2) segun tamaño del hogar

crear\_tabla("HEP", "tamano\_hogarf", "Pobreza Energética Escondida según tamaño del hogar", "Número de miembros del hogar", 5)

| **Pobreza Energética Escondida según tamaño del hogar** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Cálculo** | **Número de miembros del hogar** | | | | |
| HEP |  | 1 miembro | 2 miembros | 3 miembros | 4 miembros | 5 miembros o más |
| **No** | *Frequency* | 2183 | 3100 | 2214 | 1829 | 668 |
| *Row Pct* | 21.85% | 31.02% | 22.15% | 18.3% | 6.68% |
| *Col Pct* | 74.26% | 88.4% | 89.86% | 90.94% | 91.01% |
| **Sí** | *Frequency* | 757 | 407 | 250 | 182 | 66 |
| *Row Pct* | 45.55% | 24.48% | 15.03% | 10.97% | 3.97% |
| *Col Pct* | 25.74% | 11.6% | 10.14% | 9.06% | 8.99% |

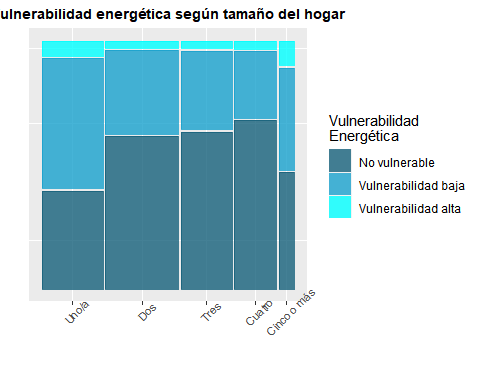
##### Indicador de deficiencias en la vivienda segun tamaño del hogar

crear\_tabla("def\_hogar", "tamano\_hogarf", "Pobreza Energética Escondida según tamaño del hogar", "Número de miembros del hogar", 5)

| **Pobreza Energética Escondida según tamaño del hogar** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Cálculo** | **Número de miembros del hogar** | | | | |
| def\_hogar |  | 1 miembro | 2 miembros | 3 miembros | 4 miembros | 5 miembros o más |
| **No** | *Frequency* | 2716 | 3316 | 2275 | 1906 | 653 |
| *Row Pct* | 25% | 30.52% | 20.94% | 17.54% | 6.01% |
| *Col Pct* | 92.39% | 94.56% | 92.33% | 94.76% | 89.01% |
| **Sí** | *Frequency* | 224 | 191 | 189 | 105 | 81 |
| *Row Pct* | 28.33% | 24.17% | 23.94% | 13.36% | 10.21% |
| *Col Pct* | 7.61% | 5.44% | 7.67% | 5.24% | 10.99% |

### Test de independencia y medidas de asociación entre variables

#Modificamos los niveles para la graficación posterior  
foessa2$tamano\_hogarf <- recode(foessa2$tamano\_hogarf,   
 "1 miembro" = "Uno/a",   
 "2 miembros" = "Dos",   
 "3 miembros" = "Tres",   
 "4 miembros" = "Cuatro",   
 "5 miembros o más" = "Cinco o más")  
  
# Creamos una table de contingencia ponderada  
t1 <- wtable(foessa2$Vulnerabilidad\_Energetica,foessa2$tamano\_hogarf, weights=foessa2$peso, mar = FALSE)  
  
#Establecemos la paleta de colores para el gráfico   
colors <- c("#15607a","#18a1cd","cyan")  
  
# Visualizamos la distribución de datos  
ggplot(data = subset(foessa2, !is.na(tamano\_hogarf))) + # Subset si queremos eliminar NA de facet\_grid  
 geom\_mosaic(aes(weight = peso,   
 x = ggmosaic::product(Vulnerabilidad\_Energetica, tamano\_hogarf),   
 fill= Vulnerabilidad\_Energetica),  
 na.rm = TRUE) +  
 labs(x="",   
 y= "",   
 title='Vulnerabilidad energética según tamaño del hogar') +   
 theme(axis.text.y = element\_blank(),  
 axis.ticks.y = element\_blank(),   
 axis.text.x = element\_text(angle = 45, size = 8),   
 plot.title = element\_text(size = 10.5, hjust = 0.5, face = "bold")) +   
 scale\_fill\_manual(name="Vulnerabilidad\nEnergética",   
 values=colors)



# + facet\_grid(~entorno\_degradado) # Tercera variable

##### Test de independencia y medidas de asociación

A continuación ejecutaremos la prueba de independencia en relación a la hipótesis nula *“No existe relación entre el nivel de vulnerabilidad energética y el tamaño del hogar”* y también analizaremos la medida de asociación a través del test V de Cramer:

chisq1 <- chisq.test(t1) #Relación significativa  
chisq1 # 2.2e-16

##   
## Pearson's Chi-squared test  
##   
## data: t1  
## X-squared = 636.12, df = 8, p-value < 2.2e-16

assocstats(t1)

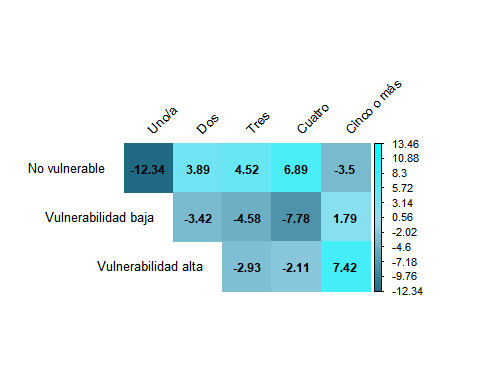
## X^2 df P(> X^2)  
## Likelihood Ratio 623.66 8 0  
## Pearson 636.12 8 0  
##   
## Phi-Coefficient : NA   
## Contingency Coeff.: 0.227   
## Cramer's V : 0.165

Por una parte vemos que la relación entre ambas variables es significativa, por lo que rechazamos la hipótesis nula. Por otra parte, vemos que la prueba de V de Cramer nos indica una medida de asociación de 0.16, tratándose de una medida muy baja, por debajo de 0.2, por lo que no se estimará asociación.

#### Análisis de residuos de Pearson

A continuación, podemos visualizar los residuos de Pearson que hemos extraído a través de la prueba de Chi Cuadrado, para ver qué interacciones contribuyen en mayor medida al cálculo del estadístico.

#install.packages("corrplot")  
library(corrplot)  
corrplot(chisq1$residuals,   
 type = "upper",   
 tl.col= "black",   
 tl.srt = 45,   
 is.cor = FALSE,  
 tl.offset = 1,  
 method = "color",   
 addCoef.col = "black",  
 number.cex=0.75,  
 tl.cex = 0.8,   
 cl.cex = 0.7,   
 cl.align.text = "l",  
 col=colorRampPalette(c("#15607a","#9ddaed","cyan"))(200))



c("turquoise","white","blue")

## [1] "turquoise" "white" "blue"

En este test en el que analizamos los residuos estandarizados en cada una de las celdas, vemos que la única que presenta valores significativos negativos es entre la condición de no vulnerable y la condición de hogar con un solo miembro.

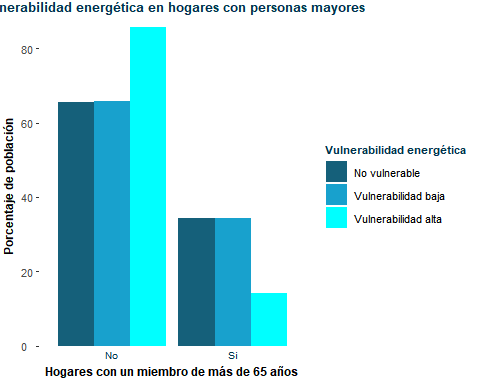
### Hogares con (al menos) una persona de más de 65 años

En primer logar, veremos la distribución de este perfil de hogar en la población afectada por alguno de los indicadores de Vulnerabilidad energética:

# Como primer paso recodificaremos los valores de la variable anciano para simplificar las etiquetas  
  
foessa2$anciano <- fct\_recode(foessa2$anciano,   
 No ="Hogar sin personas de 65 o más años",   
 Si = "Hogar con personas de 65 o más años")  
  
# Creamos tabla de contingencia entre la variable agregada PE y anciano   
  
crear\_tabla ("Vulnerabilidad\_Energetica", "anciano", "Vulnerabilidad energética en hogares con personas mayores", "Hogares con un miembro de más de 65 años", 2)

| **Vulnerabilidad energética en hogares con personas mayores** | | | |
| --- | --- | --- | --- |
|  | **Cálculo** | **Hogares con un miembro de más de 65 años** | |
| Vulnerabilidad\_Energetica |  | No | Si |
| **No vulnerable** | *Frequency* | 4406 | 2312 |
| *Row Pct* | 65.59% | 34.41% |
| *Col Pct* | 56.82% | 59.26% |
| **Vulnerabilidad baja** | *Frequency* | 2911 | 1517 |
| *Row Pct* | 65.74% | 34.26% |
| *Col Pct* | 37.55% | 38.89% |
| **Vulnerabilidad alta** | *Frequency* | 436 | 72 |
| *Row Pct* | 85.76% | 14.24% |
| *Col Pct* | 5.63% | 1.86% |

crear\_grafico(x = anciano,   
 y = Vulnerabilidad\_Energetica,  
 title = "Vulnerabilidad energética en hogares con personas mayores",   
 xtitle = "Hogares con un miembro de más de 65 años")



#### Indicadores individuales

##### Indicador de temperatura inadecuada en el hogar en hogares con un miembro mayor de 65 años

crear\_tabla ("temp\_adecuada", "anciano", "Temperatura adecuada en hogares con personas mayores", "Hogares con un miembro de más de 65 años", 2)

| **Temperatura adecuada en hogares con personas mayores** | | | |
| --- | --- | --- | --- |
|  | **Cálculo** | **Hogares con un miembro de más de 65 años** | |
| temp\_adecuada |  | No | Si |
| **Si** | *Frequency* | 6443 | 3245 |
| *Row Pct* | 66.51% | 33.49% |
| *Col Pct* | 83.35% | 83.5% |
| **No** | *Frequency* | 1287 | 641 |
| *Row Pct* | 66.74% | 33.26% |
| *Col Pct* | 16.65% | 16.5% |

##### Indicador de retrasos en el pago de suministros en hogares con un miembro mayor de 65 años

crear\_tabla ("retrasos1", "anciano", "Retrasos en el pago de suministros en hogares con personas mayores", "Hogares con un miembro de más de 65 años", 2)

| **Retrasos en el pago de suministros en hogares con personas mayores** | | | |
| --- | --- | --- | --- |
|  | **Cálculo** | **Hogares con un miembro de más de 65 años** | |
| retrasos1 |  | No | Si |
| **No** | *Frequency* | 6502 | 3618 |
| *Row Pct* | 64.25% | 35.75% |
| *Col Pct* | 83.99% | 92.89% |
| **Sí** | *Frequency* | 1239 | 277 |
| *Row Pct* | 81.74% | 18.26% |
| *Col Pct* | 16.01% | 7.11% |

##### Indicador de gasto desproporcionado energético (2M) en hogares con un miembro mayor de 65 años

crear\_tabla( "TWO\_M", "anciano", "Gasto desproporcionado en hogares con personas mayores", "Hogares con un miembro de más de 65 años", 2)

| **Gasto desproporcionado en hogares con personas mayores** | | | |
| --- | --- | --- | --- |
|  | **Cálculo** | **Hogares con un miembro de más de 65 años** | |
| TWO\_M |  | No | Si |
| **No** | *Frequency* | 6514 | 3261 |
| *Row Pct* | 66.64% | 33.36% |
| *Col Pct* | 84.01% | 83.58% |
| **Sí** | *Frequency* | 1239 | 641 |
| *Row Pct* | 65.93% | 34.07% |
| *Col Pct* | 15.99% | 16.42% |

##### Indicador de pobreza energética escondida (M/2) en hogares con un miembro mayor de 65 años

crear\_tabla("HEP", "anciano", "Pobreza Energética Escondida en hogares con personas mayores", "Hogares con un miembro de más de 65 años", 2)

| **Pobreza Energética Escondida en hogares con personas mayores** | | | |
| --- | --- | --- | --- |
|  | **Cálculo** | **Hogares con un miembro de más de 65 años** | |
| HEP |  | No | Si |
| **No** | *Frequency* | 6583 | 3410 |
| *Row Pct* | 65.87% | 34.13% |
| *Col Pct* | 84.91% | 87.41% |
| **Sí** | *Frequency* | 1170 | 491 |
| *Row Pct* | 70.44% | 29.56% |
| *Col Pct* | 15.09% | 12.59% |

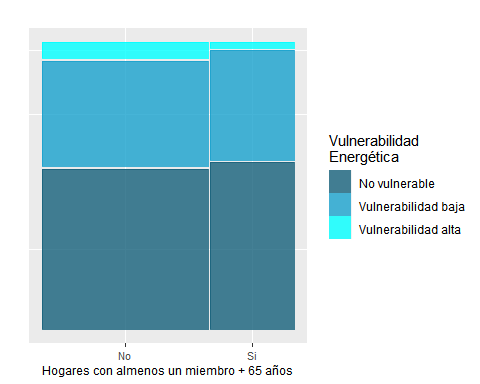
##### Indicador de deficiencias de la vivienda en hogares con un miembro mayor de 65 años

crear\_tabla ("def\_hogar", "anciano", "Temperatura adecuada en hogares con personas mayores", "Hogares con un miembro de más de 65 años", 2)

| **Temperatura adecuada en hogares con personas mayores** | | | |
| --- | --- | --- | --- |
|  | **Cálculo** | **Hogares con un miembro de más de 65 años** | |
| def\_hogar |  | No | Si |
| **No** | *Frequency* | 7204 | 3662 |
| *Row Pct* | 66.3% | 33.7% |
| *Col Pct* | 92.91% | 93.86% |
| **Sí** | *Frequency* | 550 | 240 |
| *Row Pct* | 69.65% | 30.35% |
| *Col Pct* | 7.09% | 6.14% |

#### Test de independencia y medidas de asociación entre variables

# Creamos una table de contingencia ponderada  
t1 <- wtable(foessa2$Vulnerabilidad\_Energetica,foessa2$anciano, weights=foessa2$peso, mar = FALSE)  
  
#Establecemos la paleta de colores para el gráfico   
colors <- c("#15607a","#18a1cd","cyan")  
  
# Visualizamos la distribución de datos  
ggplot(data = subset(foessa2, !is.na(anciano))) + # Subset si queremos eliminar NA de facet\_grid  
 geom\_mosaic(aes(weight = peso,   
 x = ggmosaic::product(Vulnerabilidad\_Energetica, anciano),   
 fill= Vulnerabilidad\_Energetica),  
 na.rm = TRUE) +  
 labs(x="Hogares con almenos un miembro + 65 años",   
 y= "",   
 title='') +   
 theme(axis.text.y = element\_blank(),  
 axis.ticks.y = element\_blank(),   
 axis.text.x = element\_text(size = 8),   
 plot.title = element\_text(size = 10.5, hjust = 0.5, face = "bold"),  
 axis.title.x = element\_text(size = 9)) +   
 scale\_fill\_manual(name="Vulnerabilidad\nEnergética",values=colors)



# + facet\_grid(~entorno\_degradado) # Tercera variable

##### Chi Cuadrado (x2), V de Cramer y Coeficiente de contingencia

A continuación ejecutaremos la prueba de independencia en relación a la hipótesis nula *“No existe relación entre el nivel de vulnerabilidad energética y los hogares al menos un miembro mayor de 65 años”* y también analizaremos la medida de asociación a través del test V de Cramer:

chisq1 <- chisq.test(t1)   
chisq1 # 2.2e-16

##   
## Pearson's Chi-squared test  
##   
## data: t1  
## X-squared = 88.329, df = 2, p-value < 2.2e-16

assocstats(t1)

## X^2 df P(> X^2)  
## Likelihood Ratio 101.540 2 0  
## Pearson 88.329 2 0  
##   
## Phi-Coefficient : NA   
## Contingency Coeff.: 0.087   
## Cramer's V : 0.087

Por una parte vemos que la relación entre ambas variables es significativa, por lo que rechazamos la hipótesis nula. Por otra parte, vemos que la prueba de V de Cramer nos indica una medida de asociación de 0.08, tratándose de una medida extremadamente baja, por debajo de 0.2, por lo que no se estimará asociación.

En este caso, dado la baja relación no analizaremos los residuos de Pearson.

### Hogares con dos adultos (sin niños) y uno de ellos de más de 65 años

Para el cálculo de estos indicadores, primero crearemos una nueva variable de referencia para identificar esta tipología específica de hogares:

levels(foessa2$anciano)

## [1] "No" "Si"

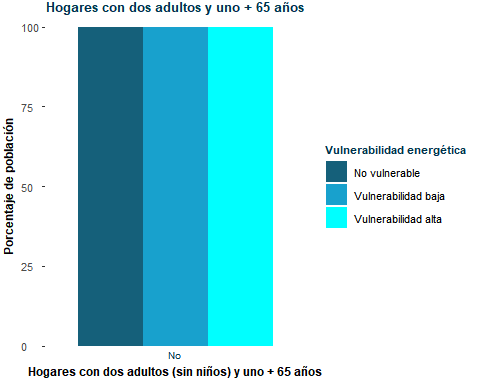
foessa2 <- foessa2 %>% mutate(dos\_adult\_65 = case\_when(  
 tamano\_hogarf == "2 miembros" &   
 menor == "Hogar sin menores de 18 años" &   
 anciano == "Si" ~ "Sí",  
 TRUE ~ "No"))  
  
foessa2$dos\_adult\_65 <- as.factor(foessa2$dos\_adult\_65)

Aplicamos esta variable al indicador general de Vulnerabilidad energética:

crear\_tabla ("Vulnerabilidad\_Energetica", "dos\_adult\_65", "Vulnerabilidad energética en hogares con dos adultos y uno + 65 años", "Hogares con dos adultos (sin niños) y uno + 65 años", 2)

## Error in add\_header\_row(x = z, values = c(" ", "Cálculo", varname), colwidths = c(1, : colwidths' sum must be equal to the number of col\_keys (3)

crear\_grafico(x = dos\_adult\_65,   
 y = Vulnerabilidad\_Energetica,  
 title = "Hogares con dos adultos y uno + 65 años",   
 xtitle = "Hogares con dos adultos (sin niños) y uno + 65 años" )



A partir de esta nueva variable, ya podremos calcular las correspondientes tablas de población afectada por alguno de los indicadores de pobreza energética en este tipo de hogares.

#### Indicadores individuales

##### Indicador de temperatura inadecuada en el hogar en hogares con dos adultos (sin niños) y uno de ellos de más de 65 años

crear\_tabla ("temp\_adecuada", "dos\_adult\_65", "Temperatura inadecuada en hogares con dos adultos y uno + 65 años", "Hogares con dos adultos (sin niños) y uno + 65 años", 2)

## Error in add\_header\_row(x = z, values = c(" ", "Cálculo", varname), colwidths = c(1, : colwidths' sum must be equal to the number of col\_keys (3)

##### Indicador de retrasos en el pago de suministros en hogares con dos adultos (sin niños) y uno de ellos de más de 65 años

crear\_tabla ("retrasos1", "dos\_adult\_65", "Retrasos en el pago de suministros en hogares con dos adultos y uno + 65 años", "Hogares con dos adultos (sin niños) y uno + 65 años", 2)

## Error in add\_header\_row(x = z, values = c(" ", "Cálculo", varname), colwidths = c(1, : colwidths' sum must be equal to the number of col\_keys (3)

##### Indicador de gasto desproporcionado energético (2M) en hogares con dos adultos (sin niños) y uno de ellos de más de 65 años

crear\_tabla ("TWO\_M", "dos\_adult\_65", "Gasto energético desproporcionado en hogares con dos adultos y uno + 65 años", "Hogares con dos adultos (sin niños) y uno + 65 años", 2)

## Error in add\_header\_row(x = z, values = c(" ", "Cálculo", varname), colwidths = c(1, : colwidths' sum must be equal to the number of col\_keys (3)

##### Indicador de pobreza energética escondida (M/2) en hogares con dos adultos (sin niños) y uno de ellos de más de 65 años

crear\_tabla("HEP", "dos\_adult\_65", "Pobreza Energética escondida en hogares con dos adultos y uno + 65 años", "Hogares con dos adultos (sin niños) y uno + 65 años", 2)

## Error in add\_header\_row(x = z, values = c(" ", "Cálculo", varname), colwidths = c(1, : colwidths' sum must be equal to the number of col\_keys (3)

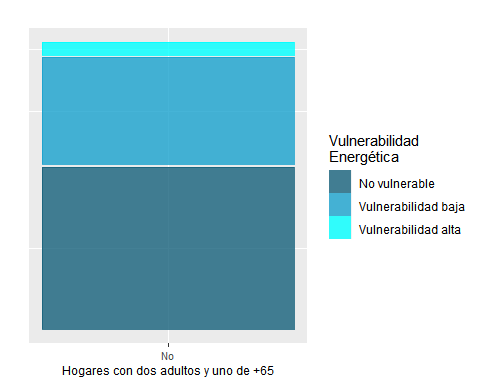
##### Indicador de deficiencias de la vivienda en hogares con dos adultos (sin niños) y uno de ellos de más de 65 años

crear\_tabla("def\_hogar", "dos\_adult\_65", "Pobreza Energética escondida en hogares con dos adultos y uno + 65 años", "Hogares con dos adultos (sin niños) y uno + 65 años", 2)

## Error in add\_header\_row(x = z, values = c(" ", "Cálculo", varname), colwidths = c(1, : colwidths' sum must be equal to the number of col\_keys (3)

#### Test de independencia y medidas de asociación entre variables

# Creamos una table de contingencia ponderada  
t1 <- wtable(foessa2$Vulnerabilidad\_Energetica,foessa2$dos\_adult\_65, weights=foessa2$peso, mar = FALSE)  
  
#Establecemos la paleta de colores para el gráfico   
colors <- c("#15607a","#18a1cd","cyan")  
  
# Visualizamos la distribución de datos  
ggplot(data = subset(foessa2, !is.na(dos\_adult\_65))) + # Subset si queremos eliminar NA de facet\_grid  
 geom\_mosaic(aes(weight = peso,   
 x = ggmosaic::product(Vulnerabilidad\_Energetica, dos\_adult\_65),   
 fill= Vulnerabilidad\_Energetica),  
 na.rm = TRUE) +  
 labs(x="Hogares con dos adultos y uno de +65",   
 y= "",   
 title='') +   
 theme(axis.text.y = element\_blank(),  
 axis.ticks.y = element\_blank(),   
 axis.text.x = element\_text(size = 8),   
 plot.title = element\_text(size = 10.5, hjust = 0.5, face = "bold"),   
 axis.title.x = element\_text(size = 9)) +   
 scale\_fill\_manual(name="Vulnerabilidad\nEnergética",   
 values=colors)



# + facet\_grid(~entorno\_degradado) # Tercera variable

##### Chi Cuadrado (x2), V de Cramer y Coeficiente de contingencia

A continuación ejecutaremos la prueba de independencia en relación a la hipótesis nula *“No existe relación entre el nivel de vulnerabilidad energética y los hogares compuestos por dos adultos sin hijos en que al menos un miembro es mayor de 65 años”* y también analizaremos la medida de asociación a través del test V de Cramer:

chisq1 <- chisq.test(t1) #Relación significativa  
chisq1 # 2.2e-16

##   
## Chi-squared test for given probabilities  
##   
## data: t1  
## X-squared = 5075.6, df = 2, p-value < 2.2e-16

assocstats(t1)

## X^2 df P(> X^2)  
## Likelihood Ratio 0 0 1  
## Pearson 0 0 1  
##   
## Phi-Coefficient : NA   
## Contingency Coeff.: 0   
## Cramer's V : NaN

Por una parte vemos que la relación entre ambas variables es significativa, por lo que debemos rechazar la hipótesis nula. Por otra parte, vemos que la prueba de V de Cramer nos indica una medida de asociación de 0.073, tratándose de una medida extremadamente baja, por debajo de 0.2, por lo que no se estimará asociación. El Coeficiente de contingencia también es igualmente bajo.

En este caso, dado la baja relación no analizaremos los residuos de Pearson.

### Hogares Monoparentales y Monomarentales

En primer lugar, en relación a los hogares monomarentales y monoparentales, lo que haremos será distinguir entre dos variables distintas: \* Hogares monoparentales, a partir de la variable monoparental para hogares con niños y un solo adulto. En este caso, se consideran tanto los hogares en que el único adulto es un hombre o una mujer. \* Hogares monomarentales, en que el único adulto es una mujer

# Primero cambiaremos el orden y el nombre de los niveles en la variable monoparental general para facilitar la comprensión y comparación  
  
foessa2$monoparental <- relevel(foessa2$monoparental, "Hay nucleo monoparental")  
  
levels(foessa2$monoparental)[levels(foessa2$monoparental)=="Hay nucleo monoparental"] <- "Hogar monoparental"  
  
levels(foessa2$monoparental)[levels(foessa2$monoparental)=="No hay nucleo monoparental"] <- "Hogar no monoparental"  
  
  
# Primero crearemos una variable para identificar los hogares monomarentales en que el adulto es una mujer   
  
foessa2 <- foessa2 %>% mutate(monomarental = case\_when(  
 monoparental == "Hogar monoparental" & edad >= 18 & sexo == "Mujer" ~ "Hogar monomarental",   
 TRUE ~ "Hogar no monomarental"  
))  
  
foessa2$monomarental <- as.factor(foessa2$monomarental)  
  
summary(foessa2$monomarental)

## Hogar monomarental Hogar no monomarental   
## 931 10724

# Podemos crear una variable para analizar los hogares monoparentales en que el adulto es un hombre  
  
foessa2 <- foessa2 %>% mutate(monoparental\_h = case\_when(  
 monoparental == "Hogar monoparental" & edad >= 18 & sexo == "Varón" ~ "Hogar monoparental (h)",   
 TRUE ~ "Hogar no monoparental (h)"  
))  
  
foessa2$monoparental\_h <- as.factor(foessa2$monoparental\_h)  
  
summary(foessa2$monoparental\_h)

## Hogar monoparental (h) Hogar no monoparental (h)   
## 317 11338

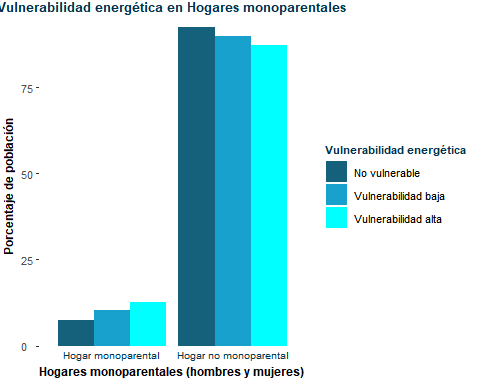
Una vez tenemos creada la nueva variable para hogares monomarentales, podemos analizar el impacto de la vulnerabilidad energética en estos hogares con la variable Vulnerabilidad\_Energetica:

**Prevalencia de la Vulnerabilidad energética en hogares monoparentales (hombres y mujeres):**

crear\_tabla("Vulnerabilidad\_Energetica", "monoparental", "Pobreza Energética en hogares monoparentales", "Hogares monoparentales", 2)

| **Pobreza Energética en hogares monoparentales** | | | |
| --- | --- | --- | --- |
|  | **Cálculo** | **Hogares monoparentales** | |
| Vulnerabilidad\_Energetica |  | Hogar monoparental | Hogar no monoparental |
| **No vulnerable** | *Frequency* | 504 | 6214 |
| *Row Pct* | 7.5% | 92.5% |
| *Col Pct* | 49.44% | 58.42% |
| **Vulnerabilidad baja** | *Frequency* | 451 | 3978 |
| *Row Pct* | 10.17% | 89.83% |
| *Col Pct* | 44.19% | 37.4% |
| **Vulnerabilidad alta** | *Frequency* | 65 | 444 |
| *Row Pct* | 12.76% | 87.24% |
| *Col Pct* | 6.37% | 4.17% |

crear\_grafico(x = monoparental,   
 y = Vulnerabilidad\_Energetica,  
 title = "Vulnerabilidad energética en Hogares monoparentales",   
 xtitle = "Hogares monoparentales (hombres y mujeres)")

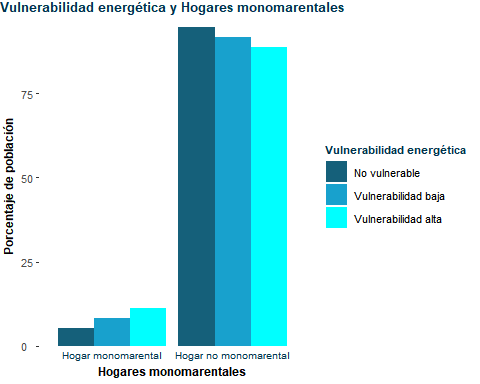


**Impacto de la Vulnerabilidad energética en hogares monomarentales:**

crear\_tabla("Vulnerabilidad\_Energetica", "monomarental", "Vulnerabilidad Energética en hogares monomarentales", "Hogares monomarentales", 2)

| **Vulnerabilidad Energética en hogares monomarentales** | | | |
| --- | --- | --- | --- |
|  | **Cálculo** | **Hogares monomarentales** | |
| Vulnerabilidad\_Energetica |  | Hogar monomarental | Hogar no monomarental |
| **No vulnerable** | *Frequency* | 348 | 6369 |
| *Row Pct* | 5.19% | 94.81% |
| *Col Pct* | 45.37% | 58.5% |
| **Vulnerabilidad baja** | *Frequency* | 362 | 4067 |
| *Row Pct* | 8.18% | 91.82% |
| *Col Pct* | 47.14% | 37.35% |
| **Vulnerabilidad alta** | *Frequency* | 58 | 451 |
| *Row Pct* | 11.32% | 88.68% |
| *Col Pct* | 7.5% | 4.14% |

crear\_grafico(x = monomarental,   
 y = Vulnerabilidad\_Energetica,  
 title = "Vulnerabilidad energética y Hogares monomarentales",   
 xtitle = "Hogares monomarentales")

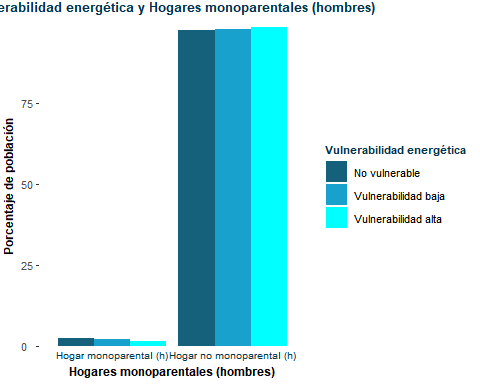


**Impacto de la Vulnerabilidad energética en hogares monoparentales (hombres):**

crear\_tabla("Vulnerabilidad\_Energetica", "monoparental\_h", "Vulnerabilidad Energética en hogares monoparentales (hombres)", "Hogares monoparentales(hombres)", 2)

| **Vulnerabilidad Energética en hogares monoparentales (hombres)** | | | |
| --- | --- | --- | --- |
|  | **Cálculo** | **Hogares monoparentales(hombres)** | |
| Vulnerabilidad\_Energetica |  | Hogar monoparental (h) | Hogar no monoparental (h) |
| **No vulnerable** | *Frequency* | 156 | 6562 |
| *Row Pct* | 2.32% | 97.68% |
| *Col Pct* | 61.88% | 57.54% |
| **Vulnerabilidad baja** | *Frequency* | 89 | 4340 |
| *Row Pct* | 2% | 98% |
| *Col Pct* | 35.2% | 38.06% |
| **Vulnerabilidad alta** | *Frequency* | 7 | 501 |
| *Row Pct* | 1.44% | 98.56% |
| *Col Pct* | 2.92% | 4.4% |

crear\_grafico(x = monoparental\_h,   
 y = Vulnerabilidad\_Energetica,  
 title = "Vulnerabilidad energética y Hogares monoparentales (hombres)",   
 xtitle = "Hogares monoparentales (hombres)")



De los resultados se desprende que, aunque existe una ligera diferencia entre el impacto de la pobreza energética entre hogares monoparentales (hombres) y hogares monomarentales, la diferencia no es significativa.

A continuación, se analizarán los distintos indicadores de pobreza energética pero únicamente para la variable de hogares monoparentales generales.

#### Indicadores individuales

##### Indicador de temperatura inadecuada en hogares monoparentales

crear\_tabla ("temp\_adecuada", "monoparental", "Temperatura adecuada en el hogar en hogares monoparentales", "Hogares monoparentales", 2)

| **Temperatura adecuada en el hogar en hogares monoparentales** | | | |
| --- | --- | --- | --- |
|  | **Cálculo** | **Hogares monoparentales** | |
| temp\_adecuada |  | Hogar monoparental | Hogar no monoparental |
| **Si** | *Frequency* | 753 | 8934 |
| *Row Pct* | 7.78% | 92.22% |
| *Col Pct* | 74.43% | 84.26% |
| **No** | *Frequency* | 259 | 1669 |
| *Row Pct* | 13.42% | 86.58% |
| *Col Pct* | 25.57% | 15.74% |

##### Indicador de retrasos en el pago de suministros en hogares monoparentales

crear\_tabla ("retrasos1", "monoparental", "Retrasos en el pago de suministros en hogares monoparentales", "Hogares monoparentales", 2)

| **Retrasos en el pago de suministros en hogares monoparentales** | | | |
| --- | --- | --- | --- |
|  | **Cálculo** | **Hogares monoparentales** | |
| retrasos1 |  | Hogar monoparental | Hogar no monoparental |
| **No** | *Frequency* | 821 | 9298 |
| *Row Pct* | 8.11% | 91.89% |
| *Col Pct* | 80.62% | 87.58% |
| **Sí** | *Frequency* | 197 | 1319 |
| *Row Pct* | 13.02% | 86.98% |
| *Col Pct* | 19.38% | 12.42% |

##### Indicador de gasto desproporcionado energético (2M) en hogares monoparentales

crear\_tabla ("TWO\_M", "monoparental", "Gasto energético desproporcionado en hogares monoparentales", "Hogares monoparentales", 2)

| **Gasto energético desproporcionado en hogares monoparentales** | | | |
| --- | --- | --- | --- |
|  | **Cálculo** | **Hogares monoparentales** | |
| TWO\_M |  | Hogar monoparental | Hogar no monoparental |
| **No** | *Frequency* | 788 | 8987 |
| *Row Pct* | 8.06% | 91.94% |
| *Col Pct* | 77.24% | 84.51% |
| **Sí** | *Frequency* | 232 | 1648 |
| *Row Pct* | 12.34% | 87.66% |
| *Col Pct* | 22.76% | 15.49% |

##### Indicador de pobreza energética escondida (M/2) en hogares monoparentales

crear\_tabla ("HEP", "monoparental", "Pobreza energética escondida en hogares monoparentales", "Hogares monoparentales", 2)

| **Pobreza energética escondida en hogares monoparentales** | | | |
| --- | --- | --- | --- |
|  | **Cálculo** | **Hogares monoparentales** | |
| HEP |  | Hogar monoparental | Hogar no monoparental |
| **No** | *Frequency* | 907 | 9087 |
| *Row Pct* | 9.07% | 90.93% |
| *Col Pct* | 88.92% | 85.44% |
| **Sí** | *Frequency* | 113 | 1548 |
| *Row Pct* | 6.8% | 93.2% |
| *Col Pct* | 11.08% | 14.56% |

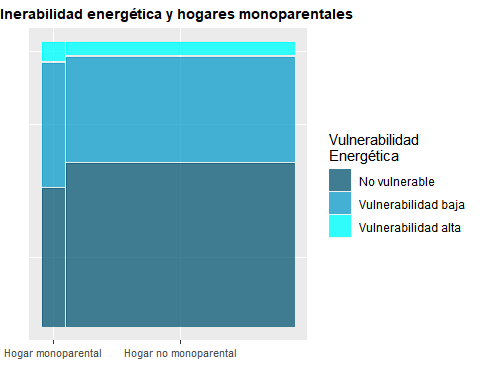
##### Indicador de deficiencias de la vivienda en hogares monoparentales

crear\_tabla ("def\_hogar", "monoparental", "Temperatura adecuada en el hogar en hogares monoparentales", "Hogares monoparentales", 2)

| **Temperatura adecuada en el hogar en hogares monoparentales** | | | |
| --- | --- | --- | --- |
|  | **Cálculo** | **Hogares monoparentales** | |
| def\_hogar |  | Hogar monoparental | Hogar no monoparental |
| **No** | *Frequency* | 922 | 9943 |
| *Row Pct* | 8.49% | 91.51% |
| *Col Pct* | 90.42% | 93.49% |
| **Sí** | *Frequency* | 98 | 692 |
| *Row Pct* | 12.37% | 87.63% |
| *Col Pct* | 9.58% | 6.51% |

#### Test de independencia y medidas de asociación entre variables

# Creamos una table de contingencia ponderada  
t1 <- wtable(foessa2$Vulnerabilidad\_Energetica,foessa2$monoparental, weights=foessa2$peso, mar = FALSE)  
  
#Establecemos la paleta de colores para el gráfico   
colors <- c("#15607a","#18a1cd","cyan")  
  
# Visualizamos la distribución de datos  
ggplot(data = subset(foessa2, !is.na(monoparental))) + # Subset si queremos eliminar NA de facet\_grid  
 geom\_mosaic(aes(weight = peso,   
 x = ggmosaic::product(Vulnerabilidad\_Energetica, monoparental),   
 fill= Vulnerabilidad\_Energetica),  
 na.rm = TRUE) +  
 labs(x="",   
 y= "",   
 title='Vulnerabilidad energética y hogares monoparentales') +   
 theme(axis.text.y = element\_blank(),  
 axis.ticks.y = element\_blank(),   
 axis.text.x = element\_text(size = 8),   
 plot.title = element\_text(size = 10.5, hjust = 0.5, face = "bold")) +   
 scale\_fill\_manual(name="Vulnerabilidad\nEnergética",   
 values=colors)



# + facet\_grid(~entorno\_degradado) # Tercera variable

##### Chi Cuadrado (x2), V de Cramer y Coeficiente de contingencia

A continuación ejecutaremos la prueba de independencia en relación a la hipótesis nula *“No existe relación entre el nivel de vulnerabilidad energética y los hogares de tipo monparental”* y también analizaremos la medida de asociación a través del test V de Cramer:

chisq1 <- chisq.test(t1) #Relación significativa  
chisq1 # 2.2e-16

##   
## Pearson's Chi-squared test  
##   
## data: t1  
## X-squared = 34.559, df = 2, p-value = 3.131e-08

assocstats(t1)

## X^2 df P(> X^2)  
## Likelihood Ratio 33.491 2 5.3399e-08  
## Pearson 34.559 2 3.1306e-08  
##   
## Phi-Coefficient : NA   
## Contingency Coeff.: 0.054   
## Cramer's V : 0.054

Vemos que la relación entre ambas variables es significativa, por lo que rechazamos la hipótesis nula. Por otra parte, vemos que la prueba de V de Cramer nos indica una medida de asociación de 0.05, tratándose de una medida extremadamente baja, por debajo de 0.2, por lo que no se estimará asociación.El coeficiente de contingencia nos indica igualmente un valor muy bajo.

### Hogares con un miembro con discapacidad

Para el cálculo de esta variable utilizaremos la variable de la base de datos ‘EINSFOESSA’ de discapacidad que nos indica si en el núcleo familiar hay algún miembro con una discapacidad reconocida mediante certificado.

La variable discapacidad se crea a partir de la pregunta de la encuesta:

C25: ¿Tiene certificado de discapacidad/ minusvalía?

En primer lugar analizamos la variable y sus niveles:

levels(foessa2$discapacidad)

## [1] "Hogar sin discapacitados según C25"   
## [2] "Hogar con uno o más discapacitados según C25"

A continuación, recodificamos los nombres de los niveles para favorecer su interpretación:

foessa2$discapacidad <- fct\_recode(foessa2$discapacidad, Sí ="Hogar con uno o más discapacitados según C25", No ="Hogar sin discapacitados según C25")  
  
levels(foessa2$discapacidad)

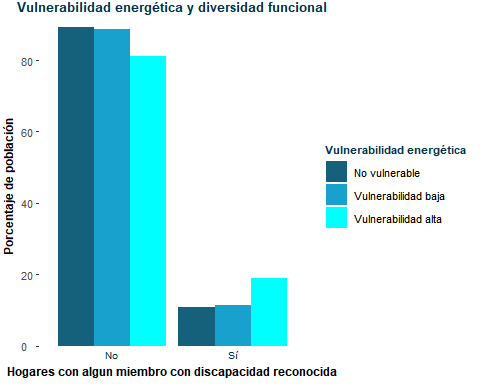
## [1] "No" "Sí"

A partir de esta variable, podemos generar las siguientes tablas de contingencia.

crear\_tabla("Vulnerabilidad\_Energetica", "discapacidad", "Vulnerabilidad energética en hogares con algun miembro con discapacidad reconocida", "Miembros de la unidad familiar con una discapacidad reconocida", 2)

| **Vulnerabilidad energética en hogares con algun miembro con discapacidad reconocida** | | | |
| --- | --- | --- | --- |
|  | **Cálculo** | **Miembros de la unidad familiar con una discapacidad reconocida** | |
| Vulnerabilidad\_Energetica |  | No | Sí |
| **No vulnerable** | *Frequency* | 6000 | 718 |
| *Row Pct* | 89.32% | 10.68% |
| *Col Pct* | 58.04% | 54.48% |
| **Vulnerabilidad baja** | *Frequency* | 3925 | 504 |
| *Row Pct* | 88.63% | 11.37% |
| *Col Pct* | 37.97% | 38.23% |
| **Vulnerabilidad alta** | *Frequency* | 413 | 96 |
| *Row Pct* | 81.13% | 18.87% |
| *Col Pct* | 3.99% | 7.29% |

crear\_grafico(x= discapacidad,   
 y = Vulnerabilidad\_Energetica,   
 title = "Vulnerabilidad energética y diversidad funcional",  
 xtitle = "Hogares con algun miembro con discapacidad reconocida")



Vemos que, de entrada, la prevalencia de Vulnerabilidad energética - con alguno de sus indicadores - es casi 8 puntos superior en los hogares con un miembro con una discapacidad reconocida.

#### Indicadores individuales

##### Temperatura inadecuada en el hogar en hogares con personas con discapacidad reconocida

crear\_tabla("temp\_adecuada", "discapacidad", "Temperatura inadecuada y diversidad funcional","hogares con algún miembro con una discapacidad reconocida",2)

| **Temperatura inadecuada y diversidad funcional** | | | |
| --- | --- | --- | --- |
|  | **Cálculo** | **hogares con algún miembro con una discapacidad reconocida** | |
| temp\_adecuada |  | No | Sí |
| **Si** | *Frequency* | 8688 | 1000 |
| *Row Pct* | 89.68% | 10.32% |
| *Col Pct* | 84.35% | 76% |
| **No** | *Frequency* | 1612 | 316 |
| *Row Pct* | 83.63% | 16.37% |
| *Col Pct* | 15.65% | 24% |

##### Retrasos en el pago de suministros en hogares con personas con discapacidad reconocida

crear\_tabla("retrasos1", "discapacidad", "Retrasos en el pago de suministros y diversidad funcional","hogares con algún miembro con una discapacidad reconocida",2 )

| **Retrasos en el pago de suministros y diversidad funcional** | | | |
| --- | --- | --- | --- |
|  | **Cálculo** | **hogares con algún miembro con una discapacidad reconocida** | |
| retrasos1 |  | No | Sí |
| **No** | *Frequency* | 9065 | 1055 |
| *Row Pct* | 89.58% | 10.42% |
| *Col Pct* | 87.82% | 80.27% |
| **Sí** | *Frequency* | 1257 | 259 |
| *Row Pct* | 82.9% | 17.1% |
| *Col Pct* | 12.18% | 19.73% |

##### Gasto energético desproporcionado en hogares con personas con discapacidad reconocida

crear\_tabla("TWO\_M", "discapacidad", "Gasto energético desproporcionado y diversidad funcional","hogares con algún miembro con una discapacidad reconocida",2 )

| **Gasto energético desproporcionado y diversidad funcional** | | | |
| --- | --- | --- | --- |
|  | **Cálculo** | **hogares con algún miembro con una discapacidad reconocida** | |
| TWO\_M |  | No | Sí |
| **No** | *Frequency* | 8685 | 1090 |
| *Row Pct* | 88.85% | 11.15% |
| *Col Pct* | 84.02% | 82.71% |
| **Sí** | *Frequency* | 1652 | 228 |
| *Row Pct* | 87.89% | 12.11% |
| *Col Pct* | 15.98% | 17.29% |

##### Pobreza energética escondida en hogares con personas con discapacidad reconocida

crear\_tabla("HEP", "discapacidad", "Pobreza energética escondida y diversidad funcional","hogares con algún miembro con una discapacidad reconocida",2 )

| **Pobreza energética escondida y diversidad funcional** | | | |
| --- | --- | --- | --- |
|  | **Cálculo** | **hogares con algún miembro con una discapacidad reconocida** | |
| HEP |  | No | Sí |
| **No** | *Frequency* | 8829 | 1165 |
| *Row Pct* | 88.34% | 11.66% |
| *Col Pct* | 85.4% | 88.43% |
| **Sí** | *Frequency* | 1509 | 152 |
| *Row Pct* | 90.83% | 9.17% |
| *Col Pct* | 14.6% | 11.57% |

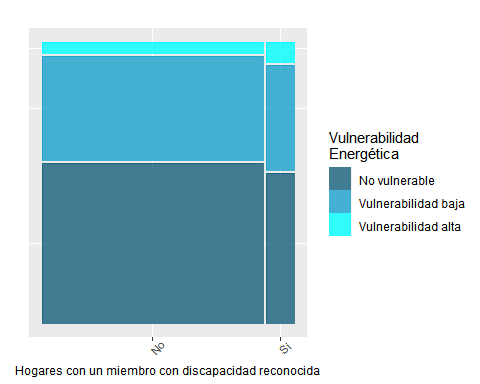
##### Indicador de deficiencias en el hogar en hogares con personas con discapacidad reconocida

crear\_tabla ("def\_hogar", "discapacidad", "Deficiencias en la vivienda y diversidad funcional", "Hogares con algun miembro con una discapacidad reconocida", 2)

| **Deficiencias en la vivienda y diversidad funcional** | | | |
| --- | --- | --- | --- |
|  | **Cálculo** | **Hogares con algun miembro con una discapacidad reconocida** | |
| def\_hogar |  | No | Sí |
| **No** | *Frequency* | 9650 | 1215 |
| *Row Pct* | 88.82% | 11.18% |
| *Col Pct* | 93.35% | 92.25% |
| **Sí** | *Frequency* | 687 | 102 |
| *Row Pct* | 87.06% | 12.94% |
| *Col Pct* | 6.65% | 7.75% |

#### Test de independencia y medidas de asociación entre variables

# Creamos una table de contingencia ponderada  
t1 <- wtable(foessa2$Vulnerabilidad\_Energetica,  
 foessa2$discapacidad,   
 weights=foessa2$peso, mar = FALSE)  
  
#Establecemos la paleta de colores para el gráfico   
colors <- c("#15607a","#18a1cd","cyan")  
  
# Visualizamos la distribución de datos  
ggplot(data = subset(foessa2, !is.na(discapacidad))) + # Subset si queremos eliminar NA de facet\_grid  
 geom\_mosaic(aes(weight = peso,   
 x = ggmosaic::product(Vulnerabilidad\_Energetica, discapacidad),   
 fill= Vulnerabilidad\_Energetica),  
 na.rm = TRUE) +  
 labs(x="Hogares con un miembro con discapacidad reconocida",   
 y= "",   
 title='') +   
 theme(axis.text.y = element\_blank(),  
 axis.ticks.y = element\_blank(),   
 axis.text.x = element\_text(angle = 45, size = 8),   
 plot.title = element\_text(size = 10.5, hjust = 0.5, face = "bold"),   
 axis.title.x = element\_text(size = 9)) +   
 scale\_fill\_manual(name="Vulnerabilidad\nEnergética",   
 values=colors)



# + facet\_grid(~entorno\_degradado) # Tercera variable

##### Chi Cuadrado (x2), V de Cramer y Coeficiente de contingencia

A continuación ejecutaremos la prueba de independencia en relación a la hipótesis nula *“No existe relación entre el nivel de vulnerabilidad energética y el tamaño del hogar”* y también analizaremos la medida de asociación a través del test V de Cramer:

chisq1 <- chisq.test(t1) #Relación significativa  
chisq1 # 2.2e-16

##   
## Pearson's Chi-squared test  
##   
## data: t1  
## X-squared = 31.657, df = 2, p-value = 1.336e-07

assocstats(t1)

## X^2 df P(> X^2)  
## Likelihood Ratio 27.429 2 1.1065e-06  
## Pearson 31.657 2 1.3360e-07  
##   
## Phi-Coefficient : NA   
## Contingency Coeff.: 0.052   
## Cramer's V : 0.052

Vemos que la relación entre ambas variables es significativa, por lo que rechazamos la hipótesis nula. Por otra parte, vemos que la prueba de V de Cramer nos indica una medida de asociación de 0.05, tratándose de una medida muy baja, por debajo de 0.2, por lo que no se estimará asociación. El coeficiente de contingencia nos indica igualmente un valor muy bajo.

### Vulnerabilidad energética y colectivos minoritarios

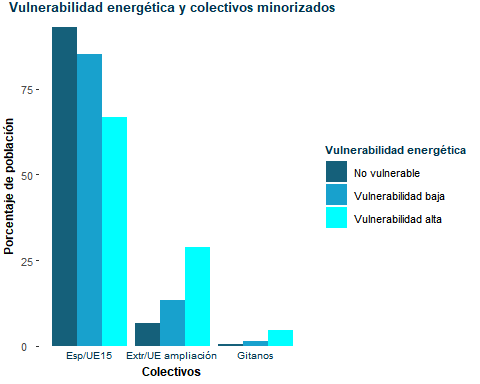
En este apartado analizaremos la tabla de contingencia entre nuestro índice de vulnerabilidad energética y la pertenencia a algun grupo minoritario y, en concreto, los colectivos de población migrante extracomunitaria y la población gitana. Para hacerlo, utilizaremos la variable ‘etnia’ incorporada a la base de datos de EINSFOESSA que distribuye la muestra en tres categorías: población nacional o de la Unión Europea, población de origen extracomunitario y población gitana (sin diferenciar su nacionalidad).

foessa2$etnia <- recode(foessa2$etnia,   
 "Todos españoles o UE15" = "Esp/UE15",   
 "Algún extracomunitario o UE ampliacion" = "Extr/UE ampliación",   
 "Gitanos" = "Gitanos")  
  
crear\_tabla("Vulnerabilidad\_Energetica", "etnia", "Vulnerabilidad energética en hogares y colectivos minorizados", "Colectivos", 3)

| **Vulnerabilidad energética en hogares y colectivos minorizados** | | | | |
| --- | --- | --- | --- | --- |
|  | **Cálculo** | **Colectivos** | | |
| Vulnerabilidad\_Energetica |  | Esp/UE15 | Extr/UE ampliación | Gitanos |
| **No vulnerable** | *Frequency* | 6257 | 439 | 21 |
| *Row Pct* | 93.15% | 6.53% | 0.32% |
| *Col Pct* | 60.31% | 37.2% | 21.41% |
| **Vulnerabilidad baja** | *Frequency* | 3778 | 595 | 56 |
| *Row Pct* | 85.31% | 13.42% | 1.26% |
| *Col Pct* | 36.42% | 50.4% | 55.94% |
| **Vulnerabilidad alta** | *Frequency* | 340 | 146 | 23 |
| *Row Pct* | 66.8% | 28.75% | 4.45% |
| *Col Pct* | 3.28% | 12.4% | 22.65% |

A continuación, crearemos un gráfico para visualizar estos datos:

crear\_grafico(x= etnia,   
 y = Vulnerabilidad\_Energetica,   
 title = "Vulnerabilidad energética y colectivos minorizados",  
 xtitle = "Colectivos")

 #### Indicadores individualizados

##### Temperatura inadecuada en el hogar y colectivos minorizados

crear\_tabla("temp\_adecuada", "etnia", "Temperatura inadecuada y colectivos minorizados","Colectivos",3)

| **Temperatura inadecuada y colectivos minorizados** | | | | |
| --- | --- | --- | --- | --- |
|  | **Cálculo** | **Colectivos** | | |
| temp\_adecuada |  | Esp/UE15 | Extr/UE ampliación | Gitanos |
| **Si** | *Frequency* | 8858 | 788 | 41 |
| *Row Pct* | 91.44% | 8.13% | 0.43% |
| *Col Pct* | 85.67% | 66.98% | 41.44% |
| **No** | *Frequency* | 1481 | 389 | 58 |
| *Row Pct* | 76.83% | 20.15% | 3.02% |
| *Col Pct* | 14.33% | 33.02% | 58.56% |

##### Retrasos en el pago de facturas y colectivos minorizados

crear\_tabla("retrasos1", "etnia", "Retrasos en el pago de facturas y colectivos minorizados","Colectivos",3)

| **Retrasos en el pago de facturas y colectivos minorizados** | | | | |
| --- | --- | --- | --- | --- |
|  | **Cálculo** | **Colectivos** | | |
| retrasos1 |  | Esp/UE15 | Extr/UE ampliación | Gitanos |
| **No** | *Frequency* | 9225 | 852 | 43 |
| *Row Pct* | 91.16% | 8.42% | 0.42% |
| *Col Pct* | 89.03% | 72.42% | 43.93% |
| **Sí** | *Frequency* | 1137 | 324 | 55 |
| *Row Pct* | 75% | 21.4% | 3.6% |
| *Col Pct* | 10.97% | 27.58% | 56.07% |

##### Gasto energético desproporcionado y colectivos minorizados

crear\_tabla("TWO\_M", "etnia", "Gasto energético desproporcionado y colectivos minorizados","Colectivos",3)

| **Gasto energético desproporcionado y colectivos minorizados** | | | | |
| --- | --- | --- | --- | --- |
|  | **Cálculo** | **Colectivos** | | |
| TWO\_M |  | Esp/UE15 | Extr/UE ampliación | Gitanos |
| **No** | *Frequency* | 8780 | 919 | 76 |
| *Row Pct* | 89.82% | 9.4% | 0.77% |
| *Col Pct* | 84.62% | 77.92% | 75.69% |
| **Sí** | *Frequency* | 1595 | 260 | 24 |
| *Row Pct* | 84.85% | 13.86% | 1.29% |
| *Col Pct* | 15.38% | 22.08% | 24.31% |

##### Pobreza energética escondida y colectivos minorizados

crear\_tabla("HEP", "etnia", "Pobreza energética escondida y colectivos minorizados","Colectivos",3)

| **Pobreza energética escondida y colectivos minorizados** | | | | |
| --- | --- | --- | --- | --- |
|  | **Cálculo** | **Colectivos** | | |
| HEP |  | Esp/UE15 | Extr/UE ampliación | Gitanos |
| **No** | *Frequency* | 8979 | 940 | 75 |
| *Row Pct* | 89.85% | 9.41% | 0.75% |
| *Col Pct* | 86.54% | 79.68% | 74.87% |
| **Sí** | *Frequency* | 1396 | 240 | 25 |
| *Row Pct* | 84.06% | 14.43% | 1.51% |
| *Col Pct* | 13.46% | 20.32% | 25.13% |

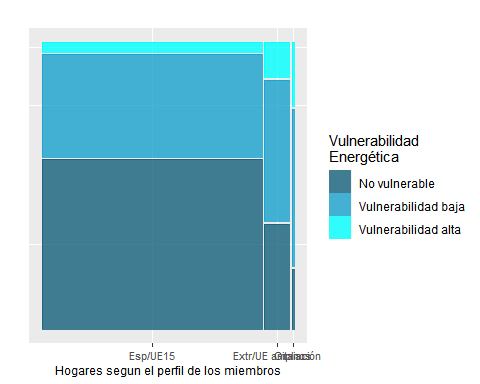
##### Deficiencias en la vivienda y colectivos minorizados

crear\_tabla("def\_hogar", "etnia", "Deficiencias en la vivienda y colectivos minorizados","Colectivos",3)

| **Deficiencias en la vivienda y colectivos minorizados** | | | | |
| --- | --- | --- | --- | --- |
|  | **Cálculo** | **Colectivos** | | |
| def\_hogar |  | Esp/UE15 | Extr/UE ampliación | Gitanos |
| **No** | *Frequency* | 9717 | 1073 | 75 |
| *Row Pct* | 89.43% | 9.88% | 0.69% |
| *Col Pct* | 93.65% | 90.98% | 75.23% |
| **Sí** | *Frequency* | 658 | 106 | 25 |
| *Row Pct* | 83.39% | 13.48% | 3.13% |
| *Col Pct* | 6.35% | 9.02% | 24.77% |

#### Test de independencia y medidas de asociación entre variables

# Creamos una table de contingencia ponderada  
t1 <- wtable(foessa2$Vulnerabilidad\_Energetica,  
 foessa2$etnia,   
 weights=foessa2$peso, mar = FALSE)  
  
#Establecemos la paleta de colores para el gráfico   
colors <- c("#15607a","#18a1cd","cyan")  
  
# Visualizamos la distribución de datos  
ggplot(data = subset(foessa2, !is.na(etnia))) + # Subset si queremos eliminar NA de facet\_grid  
 geom\_mosaic(aes(weight = peso,   
 x = ggmosaic::product(Vulnerabilidad\_Energetica, etnia),   
 fill= Vulnerabilidad\_Energetica),  
 na.rm = TRUE) +  
 labs(x="Hogares segun el perfil de los miembros",   
 y= "",   
 title='') +   
 theme(axis.text.y = element\_blank(),  
 axis.ticks.y = element\_blank(),   
 axis.text.x = element\_text(size = 8),   
 plot.title = element\_text(size = 10.5, hjust = 0.5, face = "bold"),   
 axis.title.x = element\_text(size = 9)) +   
 scale\_fill\_manual(name="Vulnerabilidad\nEnergética",   
 values=colors)



# + facet\_grid(~entorno\_degradado) # Tercera variable

##### Chi Cuadrado (x2), V de Cramer y Coeficiente de contingencia

A continuación ejecutaremos la prueba de independencia en relación a la hipótesis nula *“No existe relación entre el nivel de vulnerabilidad energética y el tamaño del hogar”* y también analizaremos la medida de asociación a través del test V de Cramer:

chisq1 <- chisq.test(t1) #Relación significativa

## Warning in stats::chisq.test(x, y, ...): Chi-squared approximation may be  
## incorrect

chisq1 # 2.2e-16

##   
## Pearson's Chi-squared test  
##   
## data: t1  
## X-squared = 463.17, df = 4, p-value < 2.2e-16

assocstats(t1)

## X^2 df P(> X^2)  
## Likelihood Ratio 382.81 4 0  
## Pearson 463.17 4 0  
##   
## Phi-Coefficient : NA   
## Contingency Coeff.: 0.196   
## Cramer's V : 0.141

Por una parte vemos que la relación entre ambas variables es significativa, por lo que rechazamos la hipótesis nula. Por otra parte, vemos que la prueba de V de Cramer nos indica una medida de asociación de 0.52, tratándose de una medida muy baja, por debajo de 0.2, por lo que no se estimará asociación. El coeficiente de contingencia nos indica igualmente un valor muy bajo.

## Dimensión 2: Características socioeconómicas e indicadores relativos a ingresos familiares

### Vulnerabilidad Energética según quintil de renda

En este punto, a partir de la variable ingresos\_UC correspondiente a los ingresos de la unidad de consumo, crearemos una nueva variable que dividirá a las observaciones de la base de datos en función del quintil de ingresos al cual corresponden:

summary(foessa2$ingresos\_hogar)

## Min. 1st Qu. Median Mean 3rd Qu. Max. NA's   
## 0 10800 17000 21532 27580 229150 1480

foessa2$quintiles <- cut(  
 foessa2$ingresos\_UC,  
 breaks = quantile(foessa2$ingresos\_hogar, c(0, 0.2, 0.4, 0.6, 0.8, 1), na.rm = TRUE),  
 labels = c("Quintil 1", "Quintil 2", "Quintil 3", "Quintil 4", "Quintil 5"),  
 right = FALSE,  
 include.lowest = TRUE  
)

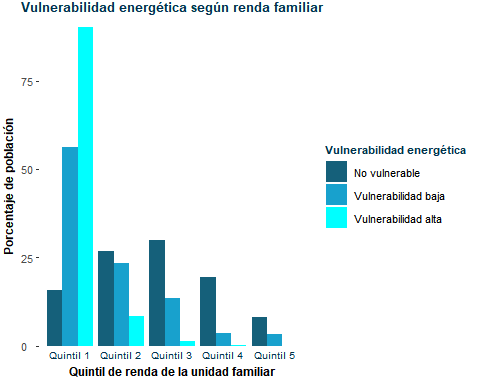
Una vez tenemos creada la nueva variable de quintiles de renda, podemos generar tablas de contingencia a partir de las variables de interés para los distintos indicadores.

En este primer apartado nos centraremos en la variable agregada PE y posteriormente veremos los indicadores desagregados.

crear\_tabla ("Vulnerabilidad\_Energetica", "quintiles", "Vulnerabilidad energética y nivel de renda", "Quintil de renda de la unidad familiar", 5)

| **Vulnerabilidad energética y nivel de renda** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Cálculo** | **Quintil de renda de la unidad familiar** | | | | |
| Vulnerabilidad\_Energetica |  | Quintil 1 | Quintil 2 | Quintil 3 | Quintil 4 | Quintil 5 |
| **No vulnerable** | *Frequency* | 938 | 1591 | 1768 | 1152 | 476 |
| *Row Pct* | 15.83% | 26.86% | 29.84% | 19.44% | 8.04% |
| *Col Pct* | 25.71% | 61.7% | 76.27% | 88.9% | 78.05% |
| **Vulnerabilidad baja** | *Frequency* | 2270 | 947 | 543 | 143 | 134 |
| *Row Pct* | 56.22% | 23.46% | 13.46% | 3.55% | 3.32% |
| *Col Pct* | 62.22% | 36.72% | 23.44% | 11.07% | 21.95% |
| **Vulnerabilidad alta** | *Frequency* | 441 | 41 | 7 | 0 | 0 |
| *Row Pct* | 90.22% | 8.34% | 1.36% | 0.07% | 0% |
| *Col Pct* | 12.08% | 1.58% | 0.29% | 0.03% | 0% |

crear\_grafico(x = quintiles,   
 y = Vulnerabilidad\_Energetica,   
 title = "Vulnerabilidad energética según renda familiar",   
 xtitle = "Quintil de renda de la unidad familiar")



#### Indicadores individuales

##### Indicador de temperatura adecuada segun quintiles de renda

crear\_tabla ("temp\_adecuada", "quintiles", "Temperatura adecuada y nivel de renda", "Quintil de renda de la unidad familiar", 5)

| **Temperatura adecuada y nivel de renda** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Cálculo** | **Quintil de renda de la unidad familiar** | | | | |
| temp\_adecuada |  | Quintil 1 | Quintil 2 | Quintil 3 | Quintil 4 | Quintil 5 |
| **Si** | *Frequency* | 2467 | 2202 | 2181 | 1259 | 578 |
| *Row Pct* | 28.4% | 25.35% | 25.1% | 14.49% | 6.65% |
| *Col Pct* | 67.9% | 85.7% | 94.26% | 97.37% | 94.73% |
| **No** | *Frequency* | 1166 | 368 | 133 | 34 | 32 |
| *Row Pct* | 67.31% | 21.21% | 7.66% | 1.96% | 1.85% |
| *Col Pct* | 32.1% | 14.3% | 5.74% | 2.63% | 5.27% |

##### Indicador de retrasos en el pago de facturas de suministros segun quintiles de renda

crear\_tabla ("retrasos1", "quintiles", "Retrasos en el pago de suministros y nivel de renda", "Quintil de renda de la unidad familiar", 5)

| **Retrasos en el pago de suministros y nivel de renda** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Cálculo** | **Quintil de renda de la unidad familiar** | | | | |
| retrasos1 |  | Quintil 1 | Quintil 2 | Quintil 3 | Quintil 4 | Quintil 5 |
| **No** | *Frequency* | 2728 | 2323 | 2194 | 1272 | 553 |
| *Row Pct* | 30.07% | 25.61% | 24.19% | 14.03% | 6.1% |
| *Col Pct* | 75.02% | 90.07% | 94.67% | 98.23% | 90.68% |
| **Sí** | *Frequency* | 908 | 256 | 124 | 23 | 57 |
| *Row Pct* | 66.41% | 18.73% | 9.03% | 1.68% | 4.16% |
| *Col Pct* | 24.98% | 9.93% | 5.33% | 1.77% | 9.32% |

##### Indicador de gasto desproporcionado (2M) segun quintiles de renda

crear\_tabla ("TWO\_M", "quintiles", "Gasto energético desproporcionado y nivel de renda", "Quintil de renda de la unidad familiar", 5)

| **Gasto energético desproporcionado y nivel de renda** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Cálculo** | **Quintil de renda de la unidad familiar** | | | | |
| TWO\_M |  | Quintil 1 | Quintil 2 | Quintil 3 | Quintil 4 | Quintil 5 |
| **No** | *Frequency* | 2138 | 2326 | 2212 | 1285 | 610 |
| *Row Pct* | 24.95% | 27.13% | 25.81% | 15% | 7.12% |
| *Col Pct* | 58.61% | 90.17% | 95.43% | 99.22% | 100% |
| **Sí** | *Frequency* | 1510 | 254 | 106 | 10 | 0 |
| *Row Pct* | 80.33% | 13.49% | 5.64% | 0.54% | 0% |
| *Col Pct* | 41.39% | 9.83% | 4.57% | 0.78% | 0% |

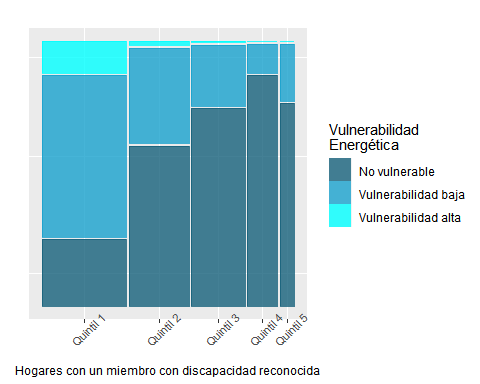
##### Indicador de Pobreza Energética Escondida (M/2) segun quintiles de renda

crear\_tabla ("HEP", "quintiles", "Pobreza energética escondida y nivel de renda", "Quintil de renda de la unidad familiar", 5)

| **Pobreza energética escondida y nivel de renda** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Cálculo** | **Quintil de renda de la unidad familiar** | | | | |
| HEP |  | Quintil 1 | Quintil 2 | Quintil 3 | Quintil 4 | Quintil 5 |
| **No** | *Frequency* | 2957 | 2222 | 2066 | 1202 | 551 |
| *Row Pct* | 32.87% | 24.69% | 22.96% | 13.36% | 6.12% |
| *Col Pct* | 81.04% | 86.13% | 89.12% | 92.78% | 90.31% |
| **Sí** | *Frequency* | 692 | 358 | 252 | 93 | 59 |
| *Row Pct* | 47.56% | 24.6% | 17.34% | 6.43% | 4.07% |
| *Col Pct* | 18.96% | 13.87% | 10.88% | 7.22% | 9.69% |

#### Test de independencia y medidas de asociación entre variables

# Creamos una table de contingencia ponderada  
t1 <- wtable(foessa2$Vulnerabilidad\_Energetica,  
 foessa2$quintiles,   
 weights=foessa2$peso, mar = FALSE)  
  
#Establecemos la paleta de colores para el gráfico   
colors <- c("#15607a","#18a1cd","cyan")  
  
# Visualizamos la distribución de datos  
ggplot(data = subset(foessa2, !is.na(quintiles))) + # Subset si queremos eliminar NA de facet\_grid  
 geom\_mosaic(aes(weight = peso,   
 x = ggmosaic::product(Vulnerabilidad\_Energetica, quintiles),   
 fill= Vulnerabilidad\_Energetica),  
 na.rm = TRUE) +  
 labs(x="Hogares con un miembro con discapacidad reconocida",   
 y= "",   
 title='') +   
 theme(axis.text.y = element\_blank(),  
 axis.ticks.y = element\_blank(),   
 axis.text.x = element\_text(angle = 45, size = 8),   
 plot.title = element\_text(size = 10.5, hjust = 0.5, face = "bold"),   
 axis.title.x = element\_text(size = 9)) +   
 scale\_fill\_manual(name="Vulnerabilidad\nEnergética",   
 values=colors)



# + facet\_grid(~entorno\_degradado) # Tercera variable

##### Chi Cuadrado (x2), V de Cramer y Coeficiente de contingencia

A continuación ejecutaremos la prueba de independencia en relación a la hipótesis nula *“No existe relación entre el nivel de vulnerabilidad energética y el nivel de renda hogar”* y también analizaremos la medida de asociación a través del test V de Cramer:

chisq1 <- chisq.test(t1) #Relación significativa

## Error in stats::chisq.test(x, y, ...): all entries of 'x' must be nonnegative and finite

chisq1 # 2.2e-16

##   
## Pearson's Chi-squared test  
##   
## data: t1  
## X-squared = 463.17, df = 4, p-value < 2.2e-16

assocstats(t1)

## X^2 df P(> X^2)  
## Likelihood Ratio NA 8 NA  
## Pearson NA 8 NA  
##   
## Phi-Coefficient : NA   
## Contingency Coeff.: NA   
## Cramer's V : NA

Por una parte vemos que la relación entre ambas variables es significativa, por lo que rechazamos la hipótesis nula. Por otra parte, vemos que la prueba de V de Cramer nos indica una medida de asociación de 0.49, tratándose de una medida comprendida entre 0,2 y 0,6, por lo que indica una asociación moderada. El coeficiente de contingencia nos indica también un nivel 0.49.

##### Dirección de la relación: Gamma de Goodman and Kruskal

Dado que ambas variables son de tipo ordinal, aplicaremos el text de Gamma de Goodman y Kruskal para medir la intensidad y dirección de la asociación.

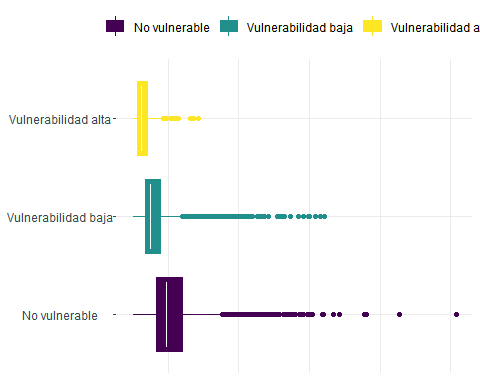
gkgamma(t1)

##   
## Goodman-Kruskal's gamma for ordinal categorical data  
##   
## data: t1  
## Z = NA, p-value = NA  
## 95 percent confidence interval:  
## NA NA  
## sample estimates:  
## Goodman-Kruskal's gamma   
## NA

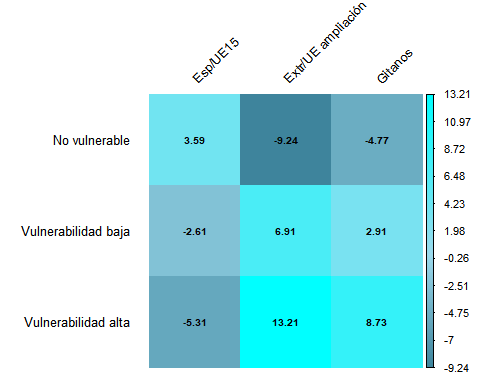
El valor de Gamma es de -0.6635, indicando una asociación fuerte de direccionalidad negativa, es decir, de relación inversa. Esto implicará que valores altos en vulnerabilidad energética nos indican que más bajo será el quintil.

Esto encaja con la tendencia que podemos apreciar cuando ponemos en relación la variable de vulnerabilidad energética y la variable numérica de ingresos en el hogar, como podemos apreciar en el seiguiente gráfico de caja complementario:

# Generate plot  
boxplot = ggplot(foessa2, aes(y = Vulnerabilidad\_Energetica, x = ingresos\_hogar))  
#Stylized Boxplot  
boxplot = boxplot + geom\_boxplot(outlier.colour = NULL, aes(colour=Vulnerabilidad\_Energetica, fill=Vulnerabilidad\_Energetica)) + # geom\_boxplot(notch=T) to compare groups  
 stat\_summary(geom = "crossbar", width=0.65, fatten=0, color="white", fun.data = function(x){ return(c(y=median(x), ymin=median(x), ymax=median(x))) })   
#Theme   
theme = theme\_set(theme\_minimal())  
theme = theme\_update(legend.position="top", legend.title=element\_blank(), panel.grid.major.x=element\_blank())  
#No X Axis  
theme = theme\_update(axis.text.x=element\_blank(), axis.ticks.x = element\_blank(), axis.line.x = element\_blank(), axis.title.x=element\_blank())  
#No Y Axis  
theme = theme\_update(axis.text.y=element\_blank(), axis.ticks.y = element\_blank(), axis.line.y = element\_blank(), axis.title.y=element\_blank())  
#No Y Axis Label + Grey Axis Numbers  
theme = theme\_update(axis.line.y = element\_blank(), axis.title.y=element\_blank(), axis.text.y = element\_text(colour="#334756"), axis.ticks.y= element\_line(colour="#334756"))  
#Imprimir gráfico  
boxplot

 ##### Análisis de residuos de Pearson A continuación, podemos visualizar los residuos de Pearson que hemos extraído a través de la prueba de Chi Cuadrado, para ver qué interacciones contribuyen en mayor medida al cálculo del estadístico.

corrplot(chisq1$residuals,   
 type = "full",   
 tl.col= "black",   
 tl.srt = 45,   
 is.cor = FALSE,  
 tl.offset = 1,  
 method = "color",   
 addCoef.col = "black",  
 addCoefasPercent = FALSE,  
 number.cex=0.65,  
 tl.cex = 0.8,   
 cl.cex = 0.7,   
 cl.align.text = "l",  
 col=colorRampPalette(c("#15607a","#9ddaed","cyan"))(200))

 El análisis de residuos nos confirma lo que ya habíamos detectado previamente: \* Existe una relación negativa entre pertenecer al primer y segundo quintil (quintiles con más bajos ingresos) y no ser vulnerable. De la misma forma, existe una notable asociación positiva entre pertenecer al primer quintil y estar en situación de vulnerabilidad. \* Pertenecer al segundo quintil muestra asociación positiva con estar en sitaución de vulnerabilidad baja, pero a partir del segundo quintil la asociación con la situación de vulnerabilidad alta es de tipo negativo. \* Existe una relación de asociación positiva entre pertenecer al quintil 4 y 5 (quintiles más rico) y no ser vulnerable.

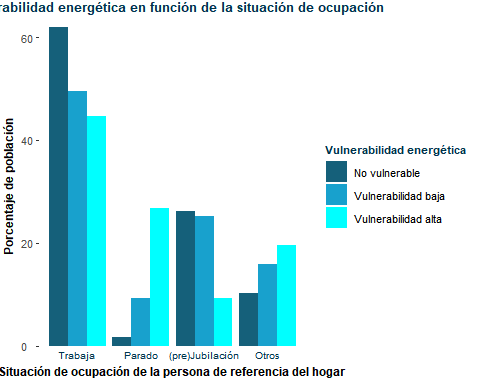
### Vulnerabilidad y precariedad energética en función de la situación de ocupación de la persona de referencia del hogar

Un segundo indicador relevante en relación con la dimensión económica es el impacto de pobreza energética en función de la situación de ocupación de la persona de referencia del hogar. Para analizar esta relación, usaremos la variable ppalocupacion.

foessa2$ppalocupacion <- fct\_recode(foessa2$ppalocupacion,   
 "Trabaja" ="Trabajando",   
 "(pre)Jubilación" = "Percibía p. jubilacion o ingresos prejubilacion",   
 "Parado" = "Buscando empleo",   
 "Otros" = "Otras situaciones")  
  
  
crear\_tabla("Vulnerabilidad\_Energetica", "ppalocupacion", "Vulnerabilidad energética en función de la situación de ocupación", "Situación de ocupación de la persona de referencia del hogar", 4)

| **Vulnerabilidad energética en función de la situación de ocupación** | | | | | |
| --- | --- | --- | --- | --- | --- |
|  | **Cálculo** | **Situación de ocupación de la persona de referencia del hogar** | | | |
| Vulnerabilidad\_Energetica |  | Trabaja | Parado | (pre)Jubilación | Otros |
| **No vulnerable** | *Frequency* | 4167 | 106 | 1759 | 685 |
| *Row Pct* | 62.03% | 1.58% | 26.19% | 10.2% |
| *Col Pct* | 63.23% | 16.33% | 60.21% | 45.95% |
| **Vulnerabilidad baja** | *Frequency* | 2196 | 410 | 1116 | 707 |
| *Row Pct* | 49.59% | 9.25% | 25.2% | 15.96% |
| *Col Pct* | 33.33% | 62.85% | 38.19% | 47.41% |
| **Vulnerabilidad alta** | *Frequency* | 227 | 136 | 47 | 99 |
| *Row Pct* | 44.63% | 26.67% | 9.21% | 19.49% |
| *Col Pct* | 3.45% | 20.82% | 1.6% | 6.65% |

crear\_grafico(ppalocupacion,   
 Vulnerabilidad\_Energetica,   
 "Vulnerabilidad energética en función de la situación de ocupación",   
 "Situación de ocupación de la persona de referencia del hogar")



#### Indicadores individualizados

##### Indicador de temperatura inadecuada en el hogar en función de la situación de ocupación de la persona de referencia

crear\_tabla("temp\_adecuada", "ppalocupacion", "Temperatura adecuada en el hogar en función de la situación de ocupación", "Situación de ocupación de la persona de referencia del hogar", 4)

| **Temperatura adecuada en el hogar en función de la situación de ocupación** | | | | | |
| --- | --- | --- | --- | --- | --- |
|  | **Cálculo** | **Situación de ocupación de la persona de referencia del hogar** | | | |
| temp\_adecuada |  | Trabaja | Parado | (pre)Jubilación | Otros |
| **Si** | *Frequency* | 5672 | 381 | 2455 | 1179 |
| *Row Pct* | 58.55% | 3.94% | 25.34% | 12.17% |
| *Col Pct* | 86.46% | 58.65% | 84.13% | 79.3% |
| **No** | *Frequency* | 889 | 269 | 463 | 308 |
| *Row Pct* | 46.09% | 13.94% | 24.02% | 15.95% |
| *Col Pct* | 13.54% | 41.35% | 15.87% | 20.7% |

##### Indicador de retrasos en el pago de facturas en función de la situación de ocupación de la persona de referencia

crear\_tabla("retrasos1", "ppalocupacion", "Retrasos en el pago de suministros en función de la situación de ocupación", "Situación de ocupación de la persona de referencia del hogar", 4)

| **Retrasos en el pago de suministros en función de la situación de ocupación** | | | | | |
| --- | --- | --- | --- | --- | --- |
|  | **Cálculo** | **Situación de ocupación de la persona de referencia del hogar** | | | |
| retrasos1 |  | Trabaja | Parado | (pre)Jubilación | Otros |
| **No** | *Frequency* | 5776 | 362 | 2708 | 1274 |
| *Row Pct* | 57.08% | 3.57% | 26.76% | 12.59% |
| *Col Pct* | 87.78% | 55.81% | 92.74% | 85.65% |
| **Sí** | *Frequency* | 804 | 286 | 212 | 214 |
| *Row Pct* | 53.05% | 18.89% | 13.98% | 14.08% |
| *Col Pct* | 12.22% | 44.19% | 7.26% | 14.35% |

##### Indicador de gasto desproporcionado (2M) en función de la situación de ocupación de la persona de referencia

crear\_tabla("TWO\_M", "ppalocupacion", "Gasto energético desproporcionado en función de la situación de ocupación", "Situación de ocupación de la persona de referencia del hogar", 4)

| **Gasto energético desproporcionado en función de la situación de ocupación** | | | | | |
| --- | --- | --- | --- | --- | --- |
|  | **Cálculo** | **Situación de ocupación de la persona de referencia del hogar** | | | |
| TWO\_M |  | Trabaja | Parado | (pre)Jubilación | Otros |
| **No** | *Frequency* | 5815 | 378 | 2465 | 1116 |
| *Row Pct* | 59.49% | 3.87% | 25.22% | 11.42% |
| *Col Pct* | 88.25% | 58.08% | 84.36% | 74.84% |
| **Sí** | *Frequency* | 775 | 273 | 457 | 375 |
| *Row Pct* | 41.2% | 14.53% | 24.32% | 19.96% |
| *Col Pct* | 11.75% | 41.92% | 15.64% | 25.16% |

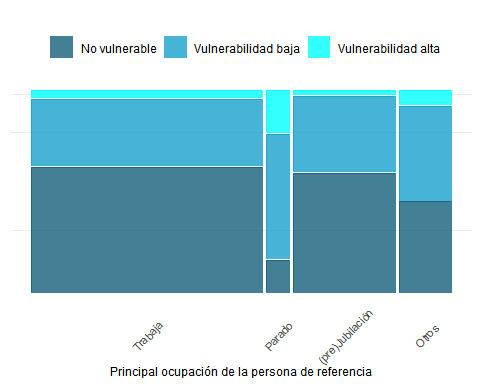
##### Indicador de pobreza energética escondida (2/M) en función de la situación de ocupación de la persona de referencia

crear\_tabla("HEP", "ppalocupacion", "Pobreza energética escondida en función de la situación de ocupación", "Situación de ocupación de la persona de referencia del hogar", 4)

| **Pobreza energética escondida en función de la situación de ocupación** | | | | | |
| --- | --- | --- | --- | --- | --- |
|  | **Cálculo** | **Situación de ocupación de la persona de referencia del hogar** | | | |
| HEP |  | Trabaja | Parado | (pre)Jubilación | Otros |
| **No** | *Frequency* | 5711 | 498 | 2573 | 1213 |
| *Row Pct* | 57.14% | 4.98% | 25.74% | 12.13% |
| *Col Pct* | 86.66% | 76.35% | 88.04% | 81.31% |
| **Sí** | *Frequency* | 879 | 154 | 349 | 279 |
| *Row Pct* | 52.92% | 9.27% | 21.03% | 16.78% |
| *Col Pct* | 13.34% | 23.65% | 11.96% | 18.69% |

#### Test de independencia y medidas de asociación entre variables

# Creamos una table de contingencia ponderada  
t1 <- wtable(foessa2$Vulnerabilidad\_Energetica,  
 foessa2$ppalocupacion,   
 weights=foessa2$peso, mar = FALSE)  
  
#Establecemos la paleta de colores para el gráfico   
colors <- c("#15607a","#18a1cd","cyan")  
  
# Visualizamos la distribución de datos  
ggplot(data = subset(foessa2, !is.na(ppalocupacion))) + # Subset si queremos eliminar NA de facet\_grid  
 geom\_mosaic(aes(weight = peso,   
 x = ggmosaic::product(Vulnerabilidad\_Energetica, ppalocupacion),   
 fill= Vulnerabilidad\_Energetica),  
 na.rm = TRUE) +  
 labs(x="Principal ocupación de la persona de referencia",   
 y= "",   
 title='') +   
 theme(axis.text.y = element\_blank(),  
 axis.ticks.y = element\_blank(),   
 axis.text.x = element\_text(angle = 45, size = 8),   
 plot.title = element\_text(size = 10.5, hjust = 0.5, face = "bold"),   
 axis.title.x = element\_text(size = 9)) +   
 scale\_fill\_manual(name="Vulnerabilidad\nEnergética",   
 values=colors)



# + facet\_grid(~entorno\_degradado) # Tercera variable

##### Chi Cuadrado (x2), V de Cramer y Coeficiente de contingencia

A continuación ejecutaremos la prueba de independencia en relación a la hipótesis nula *“No existe relación entre el nivel de vulnerabilidad energética y el tamaño del hogar”* y también analizaremos la medida de asociación a través del test V de Cramer:

chisq1 <- chisq.test(t1) #Relación significativa  
chisq1 # 2.2e-16

##   
## Pearson's Chi-squared test  
##   
## data: t1  
## X-squared = 932.03, df = 6, p-value < 2.2e-16

assocstats(t1)

## X^2 df P(> X^2)  
## Likelihood Ratio 812.19 6 0  
## Pearson 932.03 6 0  
##   
## Phi-Coefficient : NA   
## Contingency Coeff.: 0.272   
## Cramer's V : 0.2

La relación entre variables es significativa (p-value < 0.05). Vemos que la prueba de V de Cramer es de 0.2, indicando una medida de 0.2 que se sitúa en el límite para considerar la asociación, que se considera muy debil, por lo que bajo una posición conservadora se estimará que no existe una asociación significativa.

### Vulnerabilidad y vulnerabilidad energética en función de la autopercepción de clase social

La encuesta EINSFOESSA incluye una pregunta en su cuestionario a partir de la cual podemos analizar la auto-percepción de los hogares en relación a su clase social:

Pregunta E.70 ¿Cómo calificaría a su hogar teniendo en cuenta la situación económica del mismo durante los últimos 12 meses (o, en su caso, en los que lleva constituido el hogar)?

Las respuestas posibles a esta pregunta, y que configuran la variable clase\_subjetiva son:

levels(foessa2$clase\_subjetiva)

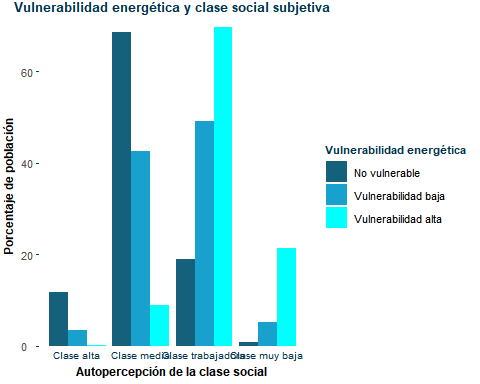
## [1] "Rico" "Por encima de la media" "En la media"   
## [4] "Por debajo de la media" "Casi pobre" "Pobre"

Con esta variable, podemos ver como se distribuye la población afectada por alguno de los indicadores de pobreza energética en función de su autopercepción de clase:

foessa2$clase\_subjetiva <- fct\_recode(foessa2$clase\_subjetiva,   
 "Clase alta" = "Rico",   
 "Clase alta" = "Por encima de la media",   
 "Clase media" = "En la media",   
 "Clase trabajadora" = "Por debajo de la media",   
 "Clase trabajadora" = "Casi pobre",   
 "Clase muy baja" = "Pobre")  
  
crear\_tabla("Vulnerabilidad\_Energetica", "clase\_subjetiva", "Vulnerabilidad energética y clase social subjetiva", "Autopercepción de la clase social", 4)

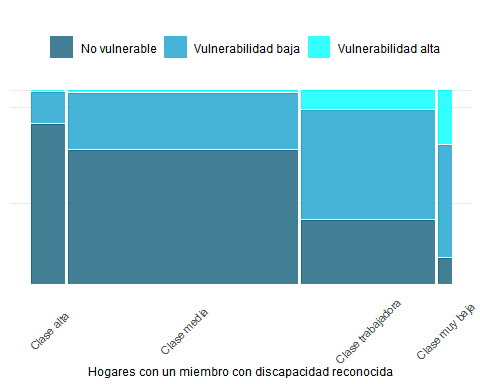
| **Vulnerabilidad energética y clase social subjetiva** | | | | | |
| --- | --- | --- | --- | --- | --- |
|  | **Cálculo** | **Autopercepción de la clase social** | | | |
| Vulnerabilidad\_Energetica |  | Clase alta | Clase media | Clase trabajadora | Clase muy baja |
| **No vulnerable** | *Frequency* | 788 | 4588 | 1264 | 51 |
| *Row Pct* | 11.78% | 68.57% | 18.88% | 0.77% |
| *Col Pct* | 84.18% | 70.48% | 33.43% | 13.36% |
| **Vulnerabilidad baja** | *Frequency* | 148 | 1877 | 2161 | 225 |
| *Row Pct* | 3.35% | 42.55% | 49% | 5.1% |
| *Col Pct* | 15.78% | 28.83% | 57.19% | 58.43% |
| **Vulnerabilidad alta** | *Frequency* | 0 | 45 | 355 | 109 |
| *Row Pct* | 0.08% | 8.88% | 69.68% | 21.35% |
| *Col Pct* | 0.05% | 0.69% | 9.38% | 28.21% |

crear\_grafico(clase\_subjetiva,   
 Vulnerabilidad\_Energetica,   
 "Vulnerabilidad energética y clase social subjetiva",   
 "Autopercepción de la clase social")



#### Test de independencia y medidas de asociación entre variables

# Creamos una table de contingencia ponderada  
t1 <- wtable(foessa2$Vulnerabilidad\_Energetica,  
 foessa2$clase\_subjetiva,   
 weights=foessa2$peso, mar = FALSE)  
  
#Establecemos la paleta de colores para el gráfico   
colors <- c("#15607a","#18a1cd","cyan")  
  
# Visualizamos la distribución de datos  
ggplot(data = subset(foessa2, !is.na(clase\_subjetiva))) + # Subset si queremos eliminar NA de facet\_grid  
 geom\_mosaic(aes(weight = peso,   
 x = ggmosaic::product(Vulnerabilidad\_Energetica,clase\_subjetiva),   
 fill= Vulnerabilidad\_Energetica),  
 na.rm = TRUE) +  
 labs(x="Hogares con un miembro con discapacidad reconocida",   
 y= "",   
 title='') +   
 theme(axis.text.y = element\_blank(),  
 axis.ticks.y = element\_blank(),   
 axis.text.x = element\_text(angle = 45, size = 8),   
 plot.title = element\_text(size = 10.5, hjust = 0.5, face = "bold"),   
 axis.title.x = element\_text(size = 9)) +   
 scale\_fill\_manual(name="Vulnerabilidad\nEnergética",   
 values=colors)



# + facet\_grid(~entorno\_degradado) # Tercera variable

##### Chi Cuadrado (x2), V de Cramer y Coeficiente de contingencia

A continuación ejecutaremos la prueba de independencia en relación a la hipótesis nula *“No existe relación entre el nivel de vulnerabilidad energética y el tamaño del hogar”* y también analizaremos la medida de asociación a través del test V de Cramer:

chisq1 <- chisq.test(t1) #Relación significativa  
chisq1 # 2.2e-16

##   
## Pearson's Chi-squared test  
##   
## data: t1  
## X-squared = 2447, df = 6, p-value < 2.2e-16

assocstats(t1)

## X^2 df P(> X^2)  
## Likelihood Ratio 2358.2 6 0  
## Pearson 2447.0 6 0  
##   
## Phi-Coefficient : NA   
## Contingency Coeff.: 0.417   
## Cramer's V : 0.325

Vemos que la relación entre variables es significativa, dado que el valor p-value es menor de 0.05. En cuanto al V de Cramer este es de 0.325, que nos indica una relación de asociación moderada. ##### Test de Goodman and Kruskal

install.packages("MESS")

## Warning: package 'MESS' is in use and will not be installed

library(MESS)  
  
gkgamma(t1)

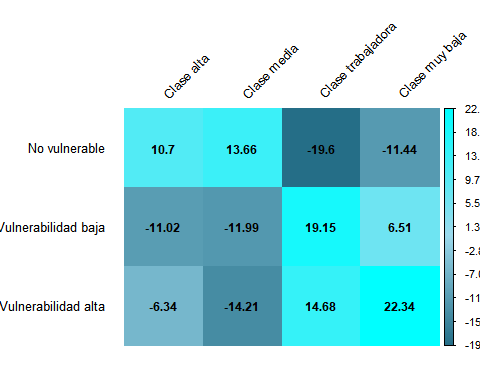
##   
## Goodman-Kruskal's gamma for ordinal categorical data  
##   
## data: t1  
## Z = 49.421, p-value < 2.2e-16  
## 95 percent confidence interval:  
## 0.6439354 0.6843637  
## sample estimates:  
## Goodman-Kruskal's gamma   
## 0.6641496

Dado que ambas variables son de tipo ordinal, podemos aplicar esta medida. El valor de Gamma de Goodman-Kruskal que obtenemos es de 0.66 indicando una relación de asociación positiva fuerte.

#### Análisis de residuos de Pearson

A continuación, podemos visualizar los residuos de Pearson que hemos extraído a través de la prueba de Chi Cuadrado, para ver qué interacciones contribuyen en mayor medida al cálculo del estadístico.

#install.packages("corrplot")  
library(corrplot)  
corrplot(chisq1$residuals,   
 type = "full",   
 tl.col= "black",   
 tl.srt = 45,   
 is.cor = FALSE,  
 tl.offset = 1,  
 method = "color",   
 addCoef.col = "black",  
 number.cex=0.75,  
 tl.cex = 0.8,   
 cl.cex = 0.7,   
 cl.align.text = "l",  
 col=colorRampPalette(c("#15607a","#9ddaed","cyan"))(200))



c("turquoise","white","blue")

## [1] "turquoise" "white" "blue"

A partir del análisis de residuos vemos claramente como identificarse como parte de la clase alta o media se correlaciona con no ser vulnerable energéticamente, mientras que identificarse como clase trabajadora o baja se correlaciona con estar en situación de vulnerabilidad energética.

## Dimensión 3: Características de la vivienda

### Vulnerabilidad y precariedad energética en función de si el hogar dispone o no de calefacción

Este indicador corresponde con un indicador de caracterización incluído en la Estrategia Nacional contra la Pobreza Energética.

crear\_tabla ("Vulnerabilidad\_Energetica", "dispone\_calef", "Vulnerabilidad energética y disponibilidad de calefacción", "Disponibilidad de calefacción", 2)

| **Vulnerabilidad energética y disponibilidad de calefacción** | | | |
| --- | --- | --- | --- |
|  | **Cálculo** | **Disponibilidad de calefacción** | |
| Vulnerabilidad\_Energetica |  | Si | No |
| **No vulnerable** | *Frequency* | 5840 | 867 |
| *Row Pct* | 87.08% | 12.92% |
| *Col Pct* | 63.86% | 35.09% |
| **Vulnerabilidad baja** | *Frequency* | 3102 | 1300 |
| *Row Pct* | 70.47% | 29.53% |
| *Col Pct* | 33.93% | 52.62% |
| **Vulnerabilidad alta** | *Frequency* | 202 | 304 |
| *Row Pct* | 39.98% | 60.02% |
| *Col Pct* | 2.21% | 12.29% |

#### Indicadores individuales

##### Indicador de temperatura inadecuada en hogares en función de la disponibilidad de calefacción

crear\_tabla ("temp\_adecuada", "dispone\_calef", "Temperatura adecuada del hogar y disponibilidad de calefacción", "Disponibilidad de calefacción", 2)

| **Temperatura adecuada del hogar y disponibilidad de calefacción** | | | |
| --- | --- | --- | --- |
|  | **Cálculo** | **Disponibilidad de calefacción** | |
| temp\_adecuada |  | Si | No |
| **Si** | *Frequency* | 8172 | 1515 |
| *Row Pct* | 84.36% | 15.64% |
| *Col Pct* | 89.36% | 61.35% |
| **No** | *Frequency* | 973 | 955 |
| *Row Pct* | 50.48% | 49.52% |
| *Col Pct* | 10.64% | 38.65% |

##### Indicador de retrasos en el pago de suministros en hogares en función de la disponibilidad de calefacción

crear\_tabla ("retrasos1", "dispone\_calef", "Retrasos en el pago de suministros y disponibilidad de calefacción", "Disponibilidad de calefacción", 2)

| **Retrasos en el pago de suministros y disponibilidad de calefacción** | | | |
| --- | --- | --- | --- |
|  | **Cálculo** | **Disponibilidad de calefacción** | |
| retrasos1 |  | Si | No |
| **No** | *Frequency* | 8287 | 1798 |
| *Row Pct* | 82.17% | 17.83% |
| *Col Pct* | 90.73% | 73.05% |
| **Sí** | *Frequency* | 847 | 664 |
| *Row Pct* | 56.07% | 43.93% |
| *Col Pct* | 9.27% | 26.95% |

##### Indicador de gasto desproporcionado energético (2M) en hogares en función de la disponibilidad de calefacción

crear\_tabla ("TWO\_M", "dispone\_calef", "Gasto energético desproporcionado y disponibilidad de calefacción", "Disponibilidad de calefacción", 2)

| **Gasto energético desproporcionado y disponibilidad de calefacción** | | | |
| --- | --- | --- | --- |
|  | **Cálculo** | **Disponibilidad de calefacción** | |
| TWO\_M |  | Si | No |
| **No** | *Frequency* | 7835 | 1906 |
| *Row Pct* | 80.43% | 19.57% |
| *Col Pct* | 85.67% | 77.17% |
| **Sí** | *Frequency* | 1310 | 564 |
| *Row Pct* | 69.91% | 30.09% |
| *Col Pct* | 14.33% | 22.83% |

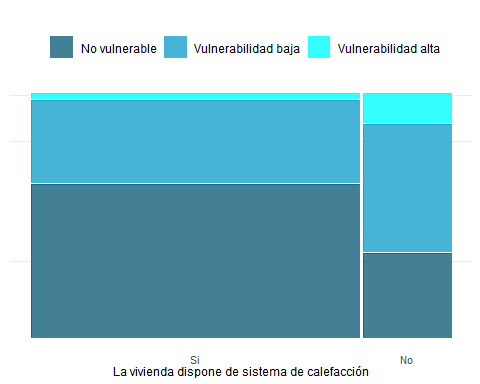
##### Indicador de pobreza energética escondida (M/2) en hogares en función de la disponibilidad de calefacción

crear\_tabla ("HEP", "dispone\_calef", "Pobreza energética escondida y disponibilidad de calefacción", "Disponibilidad de calefacción", 2)

| **Pobreza energética escondida y disponibilidad de calefacción** | | | |
| --- | --- | --- | --- |
|  | **Cálculo** | **Disponibilidad de calefacción** | |
| HEP |  | Si | No |
| **No** | *Frequency* | 8057 | 1920 |
| *Row Pct* | 80.75% | 19.25% |
| *Col Pct* | 88.1% | 77.74% |
| **Sí** | *Frequency* | 1088 | 550 |
| *Row Pct* | 66.44% | 33.56% |
| *Col Pct* | 11.9% | 22.26% |

#### Test de independencia y medidas de asociación entre variables

# Creamos una table de contingencia ponderada  
t1 <- wtable(foessa2$Vulnerabilidad\_Energetica,  
 foessa2$dispone\_calef,   
 weights=foessa2$peso, mar = FALSE)  
  
#Establecemos la paleta de colores para el gráfico   
colors <- c("#15607a","#18a1cd","cyan")  
  
# Visualizamos la distribución de datos  
ggplot(data = subset(foessa2, !is.na(dispone\_calef))) + # Subset si queremos eliminar NA de facet\_grid  
 geom\_mosaic(aes(weight = peso,   
 x = ggmosaic::product(Vulnerabilidad\_Energetica, dispone\_calef),   
 fill= Vulnerabilidad\_Energetica),  
 na.rm = TRUE) +  
 labs(x="La vivienda dispone de sistema de calefacción",   
 y= "",   
 title='') +   
 theme(axis.text.y = element\_blank(),  
 axis.ticks.y = element\_blank(),   
 axis.text.x = element\_text(size = 8),   
 plot.title = element\_text(size = 10.5, hjust = 0.5, face = "bold"),   
 axis.title.x = element\_text(size = 9)) +   
 scale\_fill\_manual(name="Vulnerabilidad\nEnergética",   
 values=colors)



# + facet\_grid(~entorno\_degradado) # Tercera variable

##### Chi Cuadrado (x2), V de Cramer y Coeficiente de contingencia

A continuación ejecutaremos la prueba de independencia en relación a la hipótesis nula *“No existe relación entre el nivel de vulnerabilidad energética y el tamaño del hogar”* y también analizaremos la medida de asociación a través del test V de Cramer:

chisq1 <- chisq.test(t1) #Relación significativa  
chisq1 # 2.2e-16

##   
## Pearson's Chi-squared test  
##   
## data: t1  
## X-squared = 911.87, df = 2, p-value < 2.2e-16

assocstats(t1)

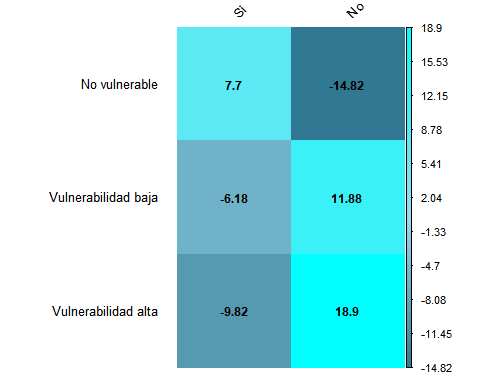
## X^2 df P(> X^2)  
## Likelihood Ratio 834.30 2 0  
## Pearson 911.87 2 0  
##   
## Phi-Coefficient : NA   
## Contingency Coeff.: 0.27   
## Cramer's V : 0.28

Vemos que la relación entre ambas variables es significativa, por lo que rechazamos la hipótesis nula. Por otra parte, vemos que la prueba de V de Cramer nos indica una medida de asociación de 0.28, tratándose de una medida muy baja, ligeramente por encima del límite de 0.2. El coeficiente de contingencia nos indica igualmente un valor muy bajo.

##### Análisis de residuos de Pearson

A continuación, podemos visualizar los residuos de Pearson que hemos extraído a través de la prueba de Chi Cuadrado, para ver qué interacciones contribuyen en mayor medida al cálculo del estadístico.

corrplot(chisq1$residuals,   
 type = "full",   
 tl.col= "black",   
 tl.srt = 45,   
 is.cor = FALSE,  
 tl.offset = 1,  
 method = "color",   
 addCoef.col = "black",  
 number.cex=0.75,  
 tl.cex = 0.8,   
 cl.cex = 0.7,   
 cl.align.text = "l",  
 col=colorRampPalette(c("#15607a","#9ddaed","cyan"))(200))

 A partir del análisis de residuos, vemos que los únicos residuos que destacan en mayor medida es la relación de asociación positiva entre vulnerabilidad energética alta y no disponer de sistema de calefacción, así como la relación negativa entre no disponer de calefación y no ser vulnerable.

### Vulnerabilidad energética y disponibilidad de servicios y equipamiento en la vivienda

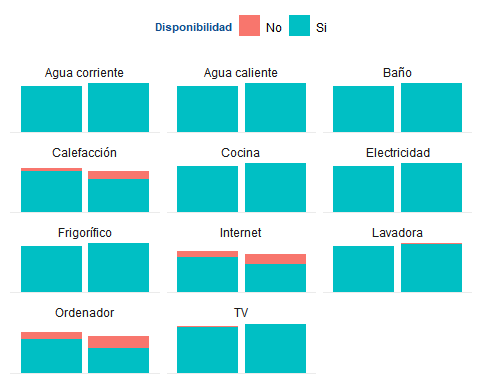
En este apartado, analizaremos el impacto de PE en relación con la disponibilidad de diferentes servicios y dispositivos en la vivienda. Los servicios seleccionados son: \* Suministros básicos: agua corriente, agua caliente y suministro de electricidad \* Instalaciones básicas: cocina y baño completo en la vivienda.  
\* Electrodomésticos básicos como lavadora y frigorífico \* Servicio de comunicaciones a través de connexión a internet, tv y ordenador

Para mostrar los datos disponibles, los resumiremos en el siguiente gráfico:

df <- foessa2 %>% select (starts\_with("dispone"), PE)  
names(df)

## [1] "dispone\_agua" "dispone\_agua\_cal" "dispone\_elect" "dispone\_calef"   
## [5] "dispone\_baño" "dispone\_cocina" "dispone\_frigo" "dispone\_lava"   
## [9] "dispone\_pc" "dispone\_internet" "dispone\_tv" "PE"

# Modificar los datos para que sean largos   
  
df\_long <- gather(df, key="measure", value="value", c("dispone\_agua","dispone\_agua\_cal", "dispone\_elect", "dispone\_calef", "dispone\_baño", "dispone\_cocina", "dispone\_frigo", "dispone\_lava", "dispone\_pc", "dispone\_internet", "dispone\_tv"))  
  
  
# Creamos el nombre de variables   
variable\_names <- list(  
 "dispone\_agua" = "Agua corriente" ,  
 "dispone\_agua\_cal" = "Agua caliente",   
 "dispone\_elect" = "Electricidad",   
 "dispone\_calef" = "Calefacción",   
 "dispone\_baño" = "Baño",   
 "dispone\_cocina" = "Cocina",   
 "dispone\_frigo" = "Frigorífico",   
 "dispone\_lava" = "Lavadora",   
 "dispone\_pc" = "Ordenador",   
 "dispone\_internet" = "Internet",   
 "dispone\_tv" = "TV"  
)  
  
variable\_labeller <- function(variable,value){  
 return(variable\_names[value])  
}  
  
  
  
# Creamos el gráfico   
df\_long %>% filter(!(value %in% NA)) %>%  
 ggplot(aes(x=PE, y=value, fill = value))+  
 geom\_bar(stat='identity', position = "stack")+  
 geom\_col (position = position\_dodge2(preserve = "single"))+  
 facet\_wrap(~ measure, ncol=3, labeller=variable\_labeller) +   
 labs(x = "Vulnerabilidad energética", y = "Porcentaje de población", fill = "Disponibilidad") +  
 theme(axis.ticks.y = element\_blank(),  
 axis.text.y = element\_blank(),  
 legend.title = element\_text(colour="dodgerblue4", size=8, face="bold"))



### Vulnerabilidad energética y régimen de tenencia

En primer lugar, recodificaremos la variable tenencia, que en su forma inicial está formada por 11 categorías, para simplificar su interpretación. Para hacerlo, reducimos y agrupamos las 11 categorías anteriores en 4 categorías.

foessa2 <- foessa2 %>% mutate("tenencia\_rec" = case\_when(  
 tenencia == "Por compra, totalmente pagada" ~ "Propiedad",   
 tenencia == "Por compra, por pagos pendientes" ~ "Propiedad",   
 tenencia == "Por herencia o donación" ~ "Propiedad",   
 tenencia == "Por otras personas hogares o instituciones" ~ "(Semi)gratuita",   
 tenencia == "Por patrón/empresa por razón de trabajo" ~ "(Semi)gratuita",  
 tenencia == "Particular, con muebles" ~ "Alquiler",  
 tenencia == "Particular, sin muebles" ~ "Alquiler",  
 tenencia == "Realquilada" ~ "Alquiler",  
 tenencia == "Alquiler social" ~ "Alquiler",  
 tenencia == "Otras" ~ "Otras",  
 tenencia == "Ocupada ilegalmente" ~ "Otras"  
))  
  
foessa2$tenencia\_rec <- as.factor(foessa2$tenencia\_rec)  
summary(foessa2$tenencia\_rec)

## (Semi)gratuita Alquiler Otras Propiedad NA's   
## 223 3112 129 8166 25

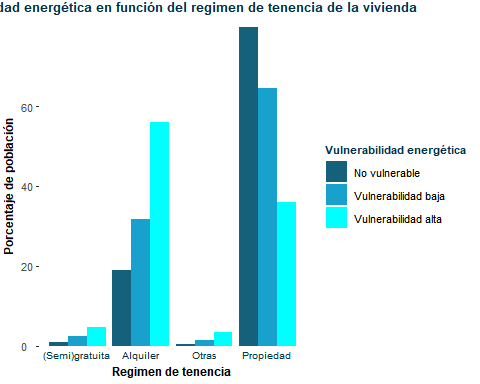
Una vez realizada la recodificación, vemos que obtenemos cuatro categorías (Alquiler, Gratuita o semigratuita, Propiedad y Otras) y, residualmente, vemos que hay 25 casos perdidos. A partir de aquí, con nuestra nueva variable recodificada, podremos analizar los datos y crear las tablas de contingencia para su posterior interpretación.

crear\_tabla("Vulnerabilidad\_Energetica", "tenencia\_rec", "Vulnerabilidad energética y avisos de cortes de suministros", "Avisos por cortes de suministros", 4)

| **Vulnerabilidad energética y avisos de cortes de suministros** | | | | | |
| --- | --- | --- | --- | --- | --- |
|  | **Cálculo** | **Avisos por cortes de suministros** | | | |
| Vulnerabilidad\_Energetica |  | (Semi)gratuita | Alquiler | Otras | Propiedad |
| **No vulnerable** | *Frequency* | 59 | 1269 | 22 | 5362 |
| *Row Pct* | 0.89% | 18.91% | 0.32% | 79.88% |
| *Col Pct* | 32.38% | 42.92% | 22.73% | 63.88% |
| **Vulnerabilidad baja** | *Frequency* | 100 | 1403 | 56 | 2850 |
| *Row Pct* | 2.27% | 31.83% | 1.27% | 64.63% |
| *Col Pct* | 54.62% | 47.45% | 59.08% | 33.95% |
| **Vulnerabilidad alta** | *Frequency* | 24 | 285 | 17 | 182 |
| *Row Pct* | 4.69% | 56.05% | 3.39% | 35.86% |
| *Col Pct* | 13% | 9.64% | 18.2% | 2.17% |

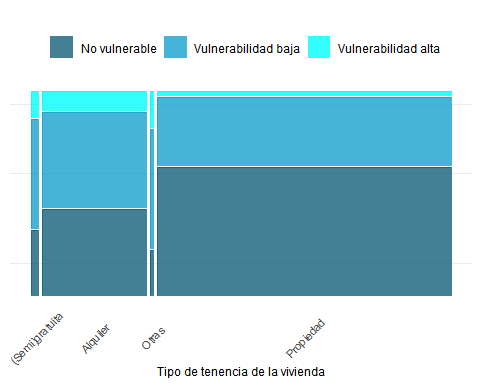
Los datos presentados en la tabla indican que la población en regimen de alquiler tiene una mayor probabilidad de estar en situación de vulnerabilidad. Del total de población en situación de alta vulnerabilidad energética, el 56.05% viven en regimen de alquiler.

crear\_grafico(tenencia\_rec,   
 Vulnerabilidad\_Energetica,   
 "Vulnerabilidad energética en función del regimen de tenencia de la vivienda",   
 "Regimen de tenencia")



#### Test de independencia y medidas de asociación entre variables

# Creamos una table de contingencia ponderada  
t1 <- wtable(foessa2$Vulnerabilidad\_Energetica,  
 foessa2$tenencia\_rec,   
 weights=foessa2$peso, mar = FALSE)  
  
#Establecemos la paleta de colores para el gráfico   
colors <- c("#15607a","#18a1cd","cyan")  
  
# Visualizamos la distribución de datos  
ggplot(data = subset(foessa2, !is.na(tenencia\_rec))) + # Subset si queremos eliminar NA de facet\_grid  
 geom\_mosaic(aes(weight = peso,   
 x = ggmosaic::product(Vulnerabilidad\_Energetica, tenencia\_rec),   
 fill= Vulnerabilidad\_Energetica),  
 na.rm = TRUE) +  
 labs(x="Tipo de tenencia de la vivienda",   
 y= "",   
 title='') +   
 theme(axis.text.y = element\_blank(),  
 axis.ticks.y = element\_blank(),   
 axis.text.x = element\_text(angle = 45, size = 8),   
 plot.title = element\_text(size = 10.5, hjust = 0.5, face = "bold"),   
 axis.title.x = element\_text(size = 9)) +   
 scale\_fill\_manual(name="Vulnerabilidad\nEnergética",   
 values=colors)



# + facet\_grid(~entorno\_degradado) # Tercera variable

##### Chi Cuadrado (x2), V de Cramer y Coeficiente de contingencia

A continuación ejecutaremos la prueba de independencia en relación a la hipótesis nula *“No existe relación entre el nivel de vulnerabilidad energética y el régimen de tenencia de la vivienda”* y también analizaremos la medida de asociación a través del test V de Cramer:

chisq1 <- chisq.test(t1, simulate.p.value = TRUE) #Relación significativa  
chisq1 # 2.2e-16

##   
## Pearson's Chi-squared test with simulated p-value (based on 2000  
## replicates)  
##   
## data: t1  
## X-squared = 691.51, df = NA, p-value = 0.0004998

assocstats(t1)

## X^2 df P(> X^2)  
## Likelihood Ratio 646.95 6 0  
## Pearson 691.51 6 0  
##   
## Phi-Coefficient : NA   
## Contingency Coeff.: 0.237   
## Cramer's V : 0.172

En este caso, hemos aplicado el test de independencia Chi Cuadrado con un p-value simulado, dado que utilizando los valores reales, las frecuencias en algunas celdas eran demasiado bajas. Aún así, se debe ser cauto ya que la simulación de los valores se hace a partir de la distribución normal de los valores existentes, por lo que las estimaciones no serán totalmente fidedignas.

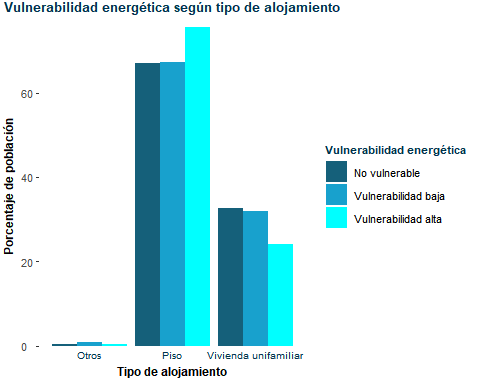
A partir del análisis de Chi Cuadrado, obtenemos un valor signficativo por lo que deberíamos descartar la hipótesis nula, pero la medida de asociación de V de Cramer es baja, con un 0.172, que no supera el umbral para indicar asociación.

### Vulnerabilidad energética y tipo de alojamiento

foessa2$alojamiento\_rec <- fct\_recode(foessa2$alojamiento,  
 Otros = "Chabola",   
 Otros = "Cueva",   
 Otros = "Barracón, prefabricado o similar",   
 Otros = "Otras",   
 Otros = "Bajera, garaje",   
 "Vivienda unifamiliar" = "En vivienda unifamiliar",   
 "Piso" = "En piso"  
 )  
  
crear\_tabla("Vulnerabilidad\_Energetica", "alojamiento\_rec", "Vulnerabilidad energética y tipo de alojamiento", "Tipo de alojamiento", 3)

| **Vulnerabilidad energética y tipo de alojamiento** | | | | |
| --- | --- | --- | --- | --- |
|  | **Cálculo** | **Tipo de alojamiento** | | |
| Vulnerabilidad\_Energetica |  | Otros | Piso | Vivienda unifamiliar |
| **No vulnerable** | *Frequency* | 24 | 4509 | 2184 |
| *Row Pct* | 0.36% | 67.12% | 32.52% |
| *Col Pct* | 41.09% | 57.24% | 58.75% |
| **Vulnerabilidad baja** | *Frequency* | 33 | 2984 | 1412 |
| *Row Pct* | 0.74% | 67.38% | 31.88% |
| *Col Pct* | 55.55% | 37.88% | 37.97% |
| **Vulnerabilidad alta** | *Frequency* | 2 | 385 | 122 |
| *Row Pct* | 0.39% | 75.62% | 23.99% |
| *Col Pct* | 3.36% | 4.88% | 3.28% |

crear\_grafico(x = alojamiento\_rec,   
 y = Vulnerabilidad\_Energetica,   
 title = "Vulnerabilidad energética según tipo de alojamiento",   
 xtitle = "Tipo de alojamiento")

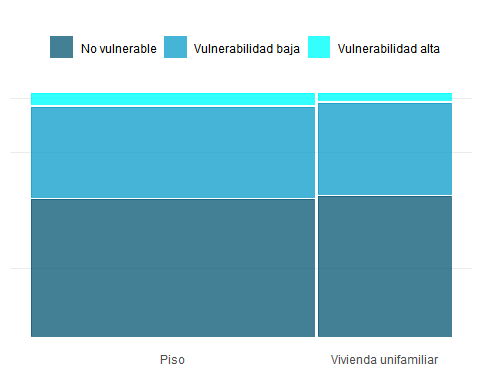


#### Test de independencia y medidas de asociación entre variables

Para la aplicación del test de independencia de Chi Cuadrado, ya vemos en la tabla de contingencia que no se cumplen los requisitos para su aplicación dado las bajas frecuencias de la categoría otros. Por este motivo, se decide excluir esta categoría y considerarla valores perdidos para este análisis.

# Recodificamos la variable de alojamiento en una nueva variable de tipo de vivienda  
  
foessa2$tipo\_vivienda <- recode\_factor(foessa2$alojamiento\_rec,   
 "Otros" = NA\_character\_)

# Creamos una table de contingencia ponderada  
t1 <- wtable(foessa2$Vulnerabilidad\_Energetica,  
 foessa2$tipo\_vivienda,   
 weights=foessa2$peso, mar = FALSE)  
  
#Establecemos la paleta de colores para el gráfico   
colors <- c("#15607a","#18a1cd","cyan")  
  
# Visualizamos la distribución de datos  
ggplot(data = subset(foessa2, !is.na(tipo\_vivienda))) + # Subset si queremos eliminar NA de facet\_grid  
 geom\_mosaic(aes(weight = peso,   
 x = ggmosaic::product(Vulnerabilidad\_Energetica, tipo\_vivienda),   
 fill= Vulnerabilidad\_Energetica),  
 na.rm = TRUE) +  
 labs(x="",   
 y= "",   
 title='') +   
 theme(axis.text.y = element\_blank(),  
 axis.ticks.y = element\_blank(),   
 axis.text.x = element\_text(size = 9),   
 plot.title = element\_text(size = 10.5, hjust = 0.5, face = "bold"),   
 axis.title.x = element\_text(size = 9)) +   
 scale\_fill\_manual(name="Vulnerabilidad\nEnergética",   
 values=colors)



# + facet\_grid(~entorno\_degradado) # Tercera variable

##### Chi Cuadrado (x2), V de Cramer y Coeficiente de contingencia

A continuación ejecutaremos la prueba de independencia en relación a la hipótesis nula *“No existe relación entre el nivel de vulnerabilidad energética y el tamaño del hogar”* y también analizaremos la medida de asociación a través del test V de Cramer:

chisq1 <- chisq.test(t1) #Relación significativa  
chisq1 # 2.2e-16

##   
## Pearson's Chi-squared test  
##   
## data: t1  
## X-squared = 15.842, df = 2, p-value = 0.000363

assocstats(t1)

## X^2 df P(> X^2)  
## Likelihood Ratio 16.623 2 0.00024571  
## Pearson 15.842 2 0.00036304  
##   
## Phi-Coefficient : NA   
## Contingency Coeff.: 0.037   
## Cramer's V : 0.037

Por una parte vemos que la relación entre ambas variables es significativa, por lo que rechazamos la hipótesis nula. Por otra parte, vemos que la prueba de V de Cramer nos indica una medida de asociación de 0.037, tratándose de una medida muy baja, por debajo de 0.2, por lo que no se estimará asociación. El coeficiente de contingencia nos indica igualmente un valor muy bajo.

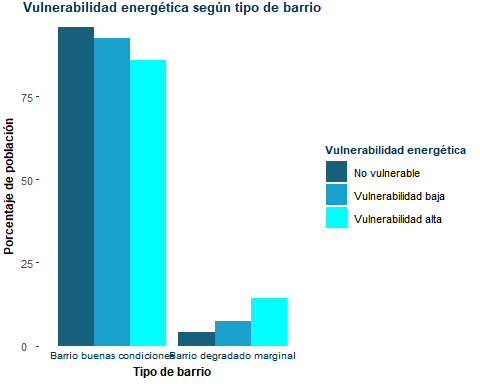
## Dimensión 4: Características del entorno urbano

### Vulnerabilidad energética y tipo de barrio

crear\_tabla("Vulnerabilidad\_Energetica", "barrio", "Vulnerabilidad energética y avisos de cortes de suministros", "Avisos por cortes de suministros", 4)

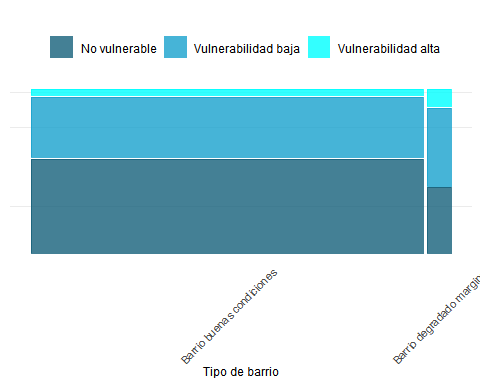
| **Vulnerabilidad energética y avisos de cortes de suministros** | | | | | |
| --- | --- | --- | --- | --- | --- |
|  | **Cálculo** | **Avisos por cortes de suministros** | | | |
| Vulnerabilidad\_Energetica |  | Zona marginal | Barrio o zona deteriorado | Barrio o zona en buenas condiciones | Zona de clase de media-alta |
| **No vulnerable** | *Frequency* | 13 | 261 | 5719 | 724 |
| *Row Pct* | 0.19% | 3.88% | 85.14% | 10.78% |
| *Col Pct* | 26.74% | 41.89% | 57.07% | 75.36% |
| **Vulnerabilidad baja** | *Frequency* | 25 | 301 | 3872 | 231 |
| *Row Pct* | 0.55% | 6.8% | 87.43% | 5.22% |
| *Col Pct* | 50.15% | 48.34% | 38.63% | 24.04% |
| **Vulnerabilidad alta** | *Frequency* | 11 | 61 | 431 | 6 |
| *Row Pct* | 2.22% | 11.97% | 84.68% | 1.13% |
| *Col Pct* | 23.11% | 9.78% | 4.3% | 0.6% |

crear\_grafico(x = barrio\_dummy,   
 y = Vulnerabilidad\_Energetica,   
 title = "Vulnerabilidad energética según tipo de barrio",   
 xtitle = "Tipo de barrio")



#### Test de independencia y medidas de asociación entre variables

# Creamos una table de contingencia ponderada  
t1 <- wtable(foessa2$Vulnerabilidad\_Energetica,  
 foessa2$barrio\_dummy,   
 weights=foessa2$peso, mar = FALSE)  
  
#Establecemos la paleta de colores para el gráfico   
colors <- c("#15607a","#18a1cd","cyan")  
  
# Visualizamos la distribución de datos  
ggplot(data = subset(foessa2, !is.na(barrio\_dummy))) + # Subset si queremos eliminar NA de facet\_grid  
 geom\_mosaic(aes(weight = peso,   
 x = ggmosaic::product(Vulnerabilidad\_Energetica, barrio\_dummy),   
 fill= Vulnerabilidad\_Energetica),  
 na.rm = TRUE) +  
 labs(x="Tipo de barrio",   
 y= "",   
 title='') +   
 theme(axis.text.y = element\_blank(),  
 axis.ticks.y = element\_blank(),   
 axis.text.x = element\_text(angle = 45, size = 8),   
 plot.title = element\_text(size = 10.5, hjust = 0.5, face = "bold"),   
 axis.title.x = element\_text(size = 9)) +   
 scale\_fill\_manual(name="Vulnerabilidad\nEnergética",   
 values=colors)



# + facet\_grid(~entorno\_degradado) # Tercera variable

##### Chi Cuadrado (x2), V de Cramer y Coeficiente de contingencia

A continuación ejecutaremos la prueba de independencia en relación a la hipótesis nula *“No existe relación entre el nivel de vulnerabilidad energética y el tipo de barrio”* y también analizaremos la medida de asociación a través del test V de Cramer:

chisq1 <- chisq.test(t1) #Relación significativa  
chisq1 # 2.2e-16

##   
## Pearson's Chi-squared test  
##   
## data: t1  
## X-squared = 122.22, df = 2, p-value < 2.2e-16

assocstats(t1)

## X^2 df P(> X^2)  
## Likelihood Ratio 106.23 2 0  
## Pearson 122.22 2 0  
##   
## Phi-Coefficient : NA   
## Contingency Coeff.: 0.102   
## Cramer's V : 0.102

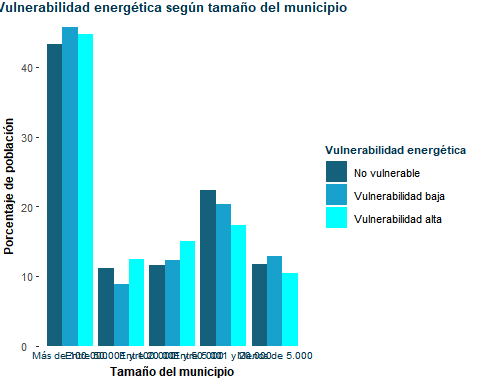
Por una parte vemos que la relación entre ambas variables es significativa, por lo que rechazamos la hipótesis nula. Por otra parte, vemos que la prueba de V de Cramer nos indica una medida de asociación de 0.102, tratándose de una medida muy baja, por debajo de 0.2, por lo que no se estimará asociación. El coeficiente de contingencia nos indica igualmente un valor muy bajo.

### Vulnerabilidad energética y tamaño del municipio

crear\_tabla("Vulnerabilidad\_Energetica", "tamano\_municipio", "Vulnerabilidad energética y tamaño del municipio", "Tamaño del municipio", 5)

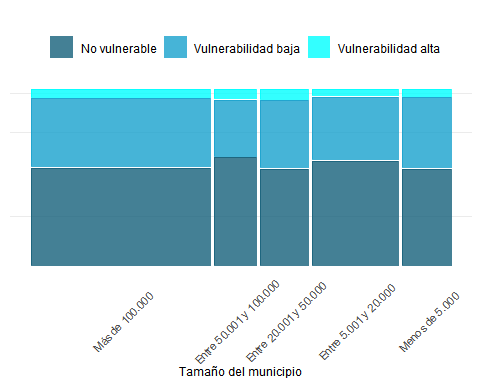
| **Vulnerabilidad energética y tamaño del municipio** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Cálculo** | **Tamaño del municipio** | | | | |
| Vulnerabilidad\_Energetica |  | Más de 100.000 | Entre 50.001 y 100.000 | Entre 20.001 y 50.000 | Entre 5.001 y 20.000 | Menos de 5.000 |
| **No vulnerable** | *Frequency* | 2908 | 750 | 775 | 1499 | 785 |
| *Row Pct* | 43.29% | 11.17% | 11.53% | 22.32% | 11.68% |
| *Col Pct* | 56.33% | 62.35% | 55.65% | 60.3% | 55.66% |
| **Vulnerabilidad baja** | *Frequency* | 2027 | 390 | 541 | 899 | 573 |
| *Row Pct* | 45.76% | 8.8% | 12.21% | 20.3% | 12.93% |
| *Col Pct* | 39.26% | 32.37% | 38.85% | 36.16% | 40.61% |
| **Vulnerabilidad alta** | *Frequency* | 228 | 64 | 77 | 88 | 53 |
| *Row Pct* | 44.76% | 12.5% | 15.06% | 17.34% | 10.35% |
| *Col Pct* | 4.41% | 5.28% | 5.5% | 3.55% | 3.73% |

crear\_grafico(x = tamano\_municipio,   
 y = Vulnerabilidad\_Energetica,   
 title = "Vulnerabilidad energética según tamaño del municipio",   
 xtitle = "Tamaño del municipio")



#### Test de independencia y medidas de asociación entre variables

# Creamos una table de contingencia ponderada  
t1 <- wtable(foessa2$Vulnerabilidad\_Energetica,  
 foessa2$tamano\_municipio,   
 weights=foessa2$peso, mar = FALSE)  
  
#Establecemos la paleta de colores para el gráfico   
colors <- c("#15607a","#18a1cd","cyan")  
  
# Visualizamos la distribución de datos  
ggplot(data = subset(foessa2, !is.na(tamano\_municipio))) + # Subset si queremos eliminar NA de facet\_grid  
 geom\_mosaic(aes(weight = peso,   
 x = ggmosaic::product(Vulnerabilidad\_Energetica, tamano\_municipio),   
 fill= Vulnerabilidad\_Energetica),  
 na.rm = TRUE) +  
 labs(x="Tamaño del municipio",   
 y= "",   
 title='') +   
 theme(axis.text.y = element\_blank(),  
 axis.ticks.y = element\_blank(),   
 axis.text.x = element\_text(angle = 45, size = 8),   
 plot.title = element\_text(size = 10.5, hjust = 0.5, face = "bold"),   
 axis.title.x = element\_text(size = 9)) +   
 scale\_fill\_manual(name="Vulnerabilidad\nEnergética",   
 values=colors)



# + facet\_grid(~entorno\_degradado) # Tercera variable

##### Chi Cuadrado (x2), V de Cramer y Coeficiente de contingencia

A continuación ejecutaremos la prueba de independencia en relación a la hipótesis nula *“No existe relación entre el nivel de vulnerabilidad energética y el tipo de barrio”* y también analizaremos la medida de asociación a través del test V de Cramer:

chisq1 <- chisq.test(t1) #Relación significativa  
chisq1 # 2.2e-16

##   
## Pearson's Chi-squared test  
##   
## data: t1  
## X-squared = 39.896, df = 8, p-value = 3.35e-06

assocstats(t1)

## X^2 df P(> X^2)  
## Likelihood Ratio 40.118 8 3.0457e-06  
## Pearson 39.896 8 3.3500e-06  
##   
## Phi-Coefficient : NA   
## Contingency Coeff.: 0.058   
## Cramer's V : 0.041

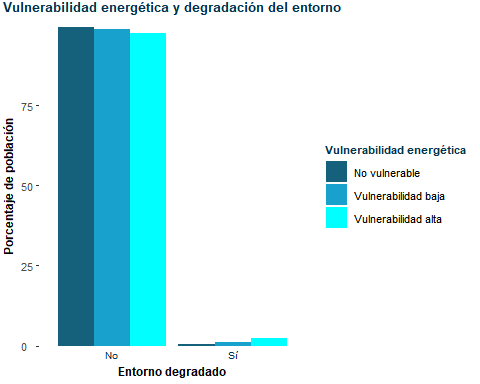
Por una parte vemos que la relación entre ambas variables es significativa, por lo que rechazamos la hipótesis nula. Por otra parte, vemos que la prueba de V de Cramer nos indica una medida de asociación de 0.041, tratándose de una medida muy baja, por debajo de 0.2, por lo que no se estimará asociación. El coeficiente de contingencia nos indica igualmente un valor muy bajo.

### Vulnerabilidad energética y entorno degradado

crear\_tabla("Vulnerabilidad\_Energetica", "entorno\_degradado", "Vulnerabilidad energética y degradación del entorno", "Entorno degradado", 2)

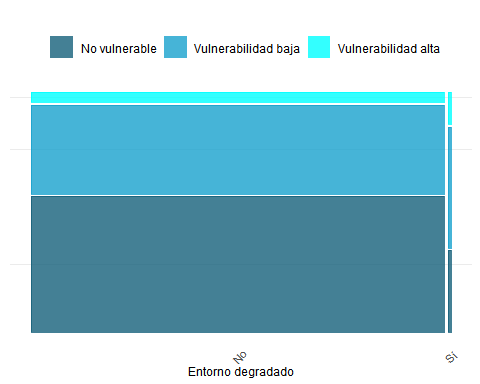
| **Vulnerabilidad energética y degradación del entorno** | | | |
| --- | --- | --- | --- |
|  | **Cálculo** | **Entorno degradado** | |
| Vulnerabilidad\_Energetica |  | No | Sí |
| **No vulnerable** | *Frequency* | 6686 | 31 |
| *Row Pct* | 99.54% | 0.46% |
| *Col Pct* | 57.82% | 34.65% |
| **Vulnerabilidad baja** | *Frequency* | 4382 | 47 |
| *Row Pct* | 98.95% | 1.05% |
| *Col Pct* | 37.89% | 51.75% |
| **Vulnerabilidad alta** | *Frequency* | 496 | 12 |
| *Row Pct* | 97.59% | 2.41% |
| *Col Pct* | 4.29% | 13.6% |

crear\_grafico(x = entorno\_degradado,   
 y = Vulnerabilidad\_Energetica,   
 title = "Vulnerabilidad energética y degradación del entorno",   
 xtitle = "Entorno degradado")



#### Test de independencia y medidas de asociación entre variables

# Creamos una table de contingencia ponderada  
t1 <- wtable(foessa2$Vulnerabilidad\_Energetica,  
 foessa2$entorno\_degradado,   
 weights=foessa2$peso, mar = FALSE)  
  
#Establecemos la paleta de colores para el gráfico   
colors <- c("#15607a","#18a1cd","cyan")  
  
# Visualizamos la distribución de datos  
ggplot(data = subset(foessa2, !is.na(barrio\_dummy))) + # Subset si queremos eliminar NA de facet\_grid  
 geom\_mosaic(aes(weight = peso,   
 x = ggmosaic::product(Vulnerabilidad\_Energetica, entorno\_degradado),   
 fill= Vulnerabilidad\_Energetica),  
 na.rm = TRUE) +  
 labs(x="Entorno degradado",   
 y= "",   
 title='') +   
 theme(axis.text.y = element\_blank(),  
 axis.ticks.y = element\_blank(),   
 axis.text.x = element\_text(angle = 45, size = 8),   
 plot.title = element\_text(size = 10.5, hjust = 0.5, face = "bold"),   
 axis.title.x = element\_text(size = 9)) +   
 scale\_fill\_manual(name="Vulnerabilidad\nEnergética",   
 values=colors)



# + facet\_grid(~entorno\_degradado) # Tercera variable

##### Chi Cuadrado (x2), V de Cramer y Coeficiente de contingencia

A continuación ejecutaremos la prueba de independencia en relación a la hipótesis nula *“No existe relación entre el nivel de vulnerabilidad energética y la degradación del entorno”* y también analizaremos la medida de asociación a través del test V de Cramer:

chisq1 <- chisq.test(t1, simulate.p.value = TRUE) #Relación significativa  
chisq1

##   
## Pearson's Chi-squared test with simulated p-value (based on 2000  
## replicates)  
##   
## data: t1  
## X-squared = 30.776, df = NA, p-value = 0.0004998

assocstats(t1)

## X^2 df P(> X^2)  
## Likelihood Ratio 25.229 2 3.3229e-06  
## Pearson 30.776 2 2.0751e-07  
##   
## Phi-Coefficient : NA   
## Contingency Coeff.: 0.051   
## Cramer's V : 0.051

En este caso, hemos aplicado el test de independencia Chi Cuadrado con un p-value simulado, dado que utilizando los valores reales, las frecuencias en algunas celdas eran demasiado bajas. Aún así, se debe ser cauto ya que la simulación de los valores se hace a partir de la distribución normal de los valores existentes, por lo que las estimaciones no serán totalmente fidedignas.

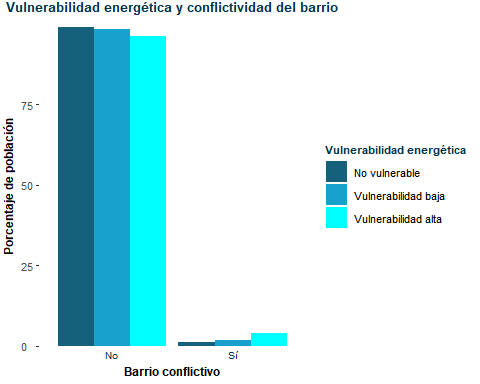
A partir del análisis de Chi Cuadrado, obtenemos un valor signficativo por lo que deberíamos descartar la hipótesis nula, pero la medida de asociación de V de Cramer es baja, con un 0.051, que no supera el umbral para indicar asociación.

### Vulnerabilidad energética y conflictividad

crear\_tabla("Vulnerabilidad\_Energetica", "barrio\_conflictivo", "Vulnerabilidad energética y conflictividad del barrio", "Barrio conflictivo", 2)

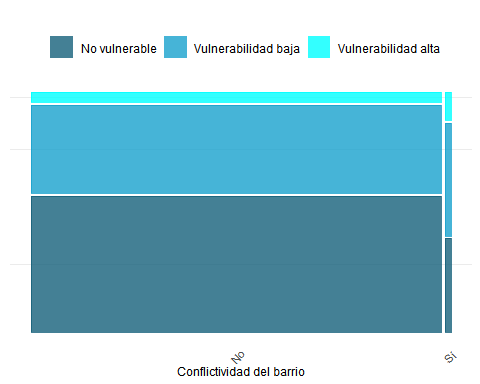
| **Vulnerabilidad energética y conflictividad del barrio** | | | |
| --- | --- | --- | --- |
|  | **Cálculo** | **Barrio conflictivo** | |
| Vulnerabilidad\_Energetica |  | No | Sí |
| **No vulnerable** | *Frequency* | 6654 | 64 |
| *Row Pct* | 99.05% | 0.95% |
| *Col Pct* | 57.89% | 39.67% |
| **Vulnerabilidad baja** | *Frequency* | 4350 | 78 |
| *Row Pct* | 98.24% | 1.76% |
| *Col Pct* | 37.85% | 48.37% |
| **Vulnerabilidad alta** | *Frequency* | 489 | 19 |
| *Row Pct* | 96.2% | 3.8% |
| *Col Pct* | 4.26% | 11.96% |

crear\_grafico(x = barrio\_conflictivo,   
 y = Vulnerabilidad\_Energetica,   
 title = "Vulnerabilidad energética y conflictividad del barrio",   
 xtitle = "Barrio conflictivo")



#### Test de independencia y medidas de asociación entre variables

# Creamos una table de contingencia ponderada  
t1 <- wtable(foessa2$Vulnerabilidad\_Energetica,  
 foessa2$barrio\_conflictivo,   
 weights=foessa2$peso, mar = FALSE)  
  
#Establecemos la paleta de colores para el gráfico   
colors <- c("#15607a","#18a1cd","cyan")  
  
# Visualizamos la distribución de datos  
ggplot(data = subset(foessa2, !is.na(barrio\_conflictivo))) + # Subset si queremos eliminar NA de facet\_grid  
 geom\_mosaic(aes(weight = peso,   
 x = ggmosaic::product(Vulnerabilidad\_Energetica, barrio\_conflictivo),   
 fill= Vulnerabilidad\_Energetica),  
 na.rm = TRUE) +  
 labs(x="Conflictividad del barrio",   
 y= "",   
 title='') +   
 theme(axis.text.y = element\_blank(),  
 axis.ticks.y = element\_blank(),   
 axis.text.x = element\_text(angle = 45, size = 8),   
 plot.title = element\_text(size = 10.5, hjust = 0.5, face = "bold"),   
 axis.title.x = element\_text(size = 9)) +   
 scale\_fill\_manual(name="Vulnerabilidad\nEnergética",   
 values=colors)



# + facet\_grid(~entorno\_degradado) # Tercera variable

##### Chi Cuadrado (x2), V de Cramer y Coeficiente de contingencia

A continuación ejecutaremos la prueba de independencia en relación a la hipótesis nula *“No existe relación entre el nivel de vulnerabilidad energética y la degradación del entorno”* y también analizaremos la medida de asociación a través del test V de Cramer:

chisq1 <- chisq.test(t1) #Relación significativa  
chisq1

##   
## Pearson's Chi-squared test  
##   
## data: t1  
## X-squared = 35.404, df = 2, p-value = 2.052e-08

assocstats(t1)

## X^2 df P(> X^2)  
## Likelihood Ratio 29.257 2 4.4356e-07  
## Pearson 35.404 2 2.0520e-08  
##   
## Phi-Coefficient : NA   
## Contingency Coeff.: 0.055   
## Cramer's V : 0.055

A partir del análisis de Chi Cuadrado, obtenemos un valor significativo por lo que deberíamos descartar la hipótesis nula, pero la medida de asociación de V de Cramer es baja, con un 0.055, que no supera el umbral para indicar asociación, por lo que no se estima asociación entre las variables.

## Dimensión 5: Impactos sobre la salud y relaciones sociales

### Salud

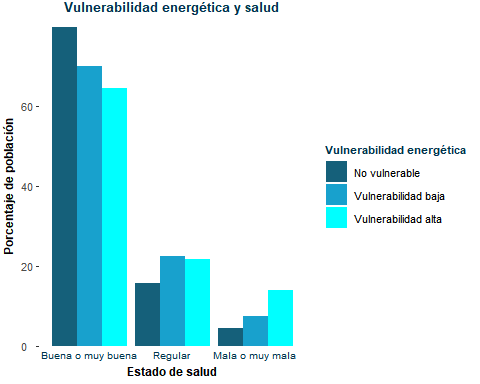
#Comprobación de niveles de la variable salud   
levels(foessa2$salud)

## [1] "Muy buena" "Bastante buena" "Regular" "Más bien mala"   
## [5] "Francamente mala"

# Simplificación de categorias de la variable salud  
foessa2$salud\_rec <- fct\_recode(foessa2$salud,   
 "Buena o muy buena" = "Muy buena",   
 "Buena o muy buena" = "Bastante buena",   
 Regular = "Regular",   
 "Mala o muy mala" = "Más bien mala",   
 "Mala o muy mala" = "Francamente mala")  
  
crear\_tabla("Vulnerabilidad\_Energetica", "salud\_rec", "Vulnerabilidad energética y salud", "Estado de salud", 3)

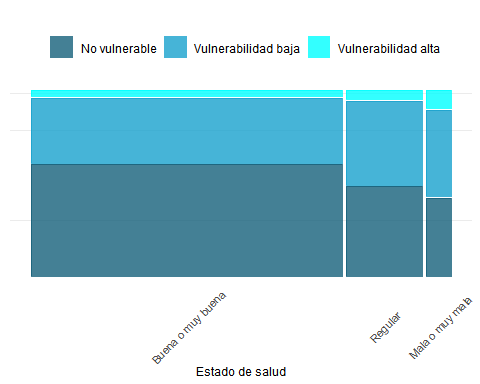
| **Vulnerabilidad energética y salud** | | | | |
| --- | --- | --- | --- | --- |
|  | **Cálculo** | **Estado de salud** | | |
| Vulnerabilidad\_Energetica |  | Buena o muy buena | Regular | Mala o muy mala |
| **No vulnerable** | *Frequency* | 5362 | 1053 | 301 |
| *Row Pct* | 79.84% | 15.68% | 4.48% |
| *Col Pct* | 61% | 48.85% | 42.69% |
| **Vulnerabilidad baja** | *Frequency* | 3101 | 993 | 333 |
| *Row Pct* | 70.04% | 22.43% | 7.52% |
| *Col Pct* | 35.27% | 46.06% | 47.25% |
| **Vulnerabilidad alta** | *Frequency* | 328 | 110 | 71 |
| *Row Pct* | 64.48% | 21.58% | 13.94% |
| *Col Pct* | 3.73% | 5.09% | 10.06% |

crear\_grafico(x = salud\_rec,   
 y = Vulnerabilidad\_Energetica,   
 title = "Vulnerabilidad energética y salud",   
 xtitle = "Estado de salud")



#### Test de independencia y medidas de asociación entre variables

# Creamos una table de contingencia ponderada  
t1 <- wtable(foessa2$Vulnerabilidad\_Energetica,  
 foessa2$salud\_rec,   
 weights=foessa2$peso, mar = FALSE)  
  
#Establecemos la paleta de colores para el gráfico   
colors <- c("#15607a","#18a1cd","cyan")  
  
# Visualizamos la distribución de datos  
ggplot(data = subset(foessa2, !is.na(salud\_rec))) + # Subset si queremos eliminar NA de facet\_grid  
 geom\_mosaic(aes(weight = peso,   
 x = ggmosaic::product(Vulnerabilidad\_Energetica, salud\_rec),   
 fill= Vulnerabilidad\_Energetica),  
 na.rm = TRUE) +  
 labs(x="Estado de salud",   
 y= "",   
 title='') +   
 theme(axis.text.y = element\_blank(),  
 axis.ticks.y = element\_blank(),   
 axis.text.x = element\_text(angle = 45, size = 8),   
 plot.title = element\_text(size = 10.5, hjust = 0.5, face = "bold"),   
 axis.title.x = element\_text(size = 9)) +   
 scale\_fill\_manual(name="Vulnerabilidad\nEnergética",   
 values=colors)



# + facet\_grid(~entorno\_degradado) # Tercera variable

##### Chi Cuadrado (x2), V de Cramer y Coeficiente de contingencia

A continuación ejecutaremos la prueba de independencia en relación a la hipótesis nula *“No existe relación entre el nivel de vulnerabilidad energética y el tamaño del hogar”* y también analizaremos la medida de asociación a través del test V de Cramer:

chisq1 <- chisq.test(t1) #Relación significativa  
chisq1 # 2.2e-16

##   
## Pearson's Chi-squared test  
##   
## data: t1  
## X-squared = 206.42, df = 4, p-value < 2.2e-16

assocstats(t1)

## X^2 df P(> X^2)  
## Likelihood Ratio 193.37 4 0  
## Pearson 206.42 4 0  
##   
## Phi-Coefficient : NA   
## Contingency Coeff.: 0.132   
## Cramer's V : 0.094

Por una parte vemos que la relación entre ambas variables es significativa, por lo que rechazamos la hipótesis nula. Por otra parte, vemos que la prueba de V de Cramer nos indica una medida de asociación de 0.09, tratándose de una medida muy baja, por debajo de 0.2, por lo que no se estimará asociación. El coeficiente de contingencia nos indica igualmente un valor muy bajo.

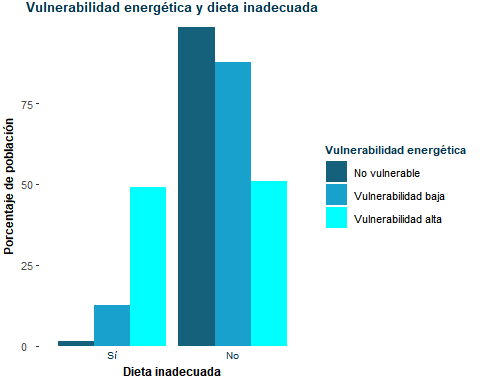
### Impactos sobre la alimentación

#### Dieta inadecuada

crear\_tabla("Vulnerabilidad\_Energetica", "dieta\_inadec", "Vulnerabilidad energética y dieta inadecuada", "Dieta inadecuada", 2)

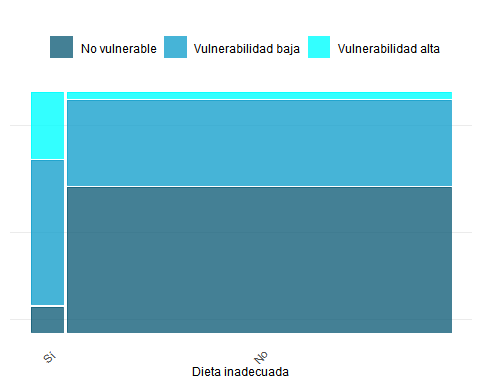
| **Vulnerabilidad energética y dieta inadecuada** | | | |
| --- | --- | --- | --- |
|  | **Cálculo** | **Dieta inadecuada** | |
| Vulnerabilidad\_Energetica |  | Sí | No |
| **No vulnerable** | *Frequency* | 94 | 6623 |
| *Row Pct* | 1.4% | 98.6% |
| *Col Pct* | 10.57% | 61.58% |
| **Vulnerabilidad baja** | *Frequency* | 550 | 3873 |
| *Row Pct* | 12.43% | 87.57% |
| *Col Pct* | 61.63% | 36.01% |
| **Vulnerabilidad alta** | *Frequency* | 248 | 259 |
| *Row Pct* | 48.95% | 51.05% |
| *Col Pct* | 27.79% | 2.4% |

crear\_grafico(x = dieta\_inadec,   
 y = Vulnerabilidad\_Energetica,   
 title = "Vulnerabilidad energética y dieta inadecuada",   
 xtitle = "Dieta inadecuada")



#### Test de independencia y medidas de asociación entre variables

# Creamos una table de contingencia ponderada  
t1 <- wtable(foessa2$Vulnerabilidad\_Energetica,  
 foessa2$dieta\_inadec,   
 weights=foessa2$peso, mar = FALSE)  
  
#Establecemos la paleta de colores para el gráfico   
colors <- c("#15607a","#18a1cd","cyan")  
  
# Visualizamos la distribución de datos  
ggplot(data = subset(foessa2, !is.na(salud\_rec))) + # Subset si queremos eliminar NA de facet\_grid  
 geom\_mosaic(aes(weight = peso,   
 x = ggmosaic::product(Vulnerabilidad\_Energetica, dieta\_inadec),   
 fill= Vulnerabilidad\_Energetica),  
 na.rm = TRUE) +  
 labs(x="Dieta inadecuada",   
 y= "",   
 title='') +   
 theme(axis.text.y = element\_blank(),  
 axis.ticks.y = element\_blank(),   
 axis.text.x = element\_text(angle = 45, size = 8),   
 plot.title = element\_text(size = 10.5, hjust = 0.5, face = "bold"),   
 axis.title.x = element\_text(size = 9)) +   
 scale\_fill\_manual(name="Vulnerabilidad\nEnergética",   
 values=colors)



# + facet\_grid(~entorno\_degradado) # Tercera variable

##### Chi Cuadrado (x2), V de Cramer y Coeficiente de contingencia

A continuación ejecutaremos la prueba de independencia en relación a la hipótesis nula *“No existe relación entre el nivel de vulnerabilidad energética y el tamaño del hogar”* y también analizaremos la medida de asociación a través del test V de Cramer:

chisq1 <- chisq.test(t1) #Relación significativa  
chisq1 # 2.2e-16

##   
## Pearson's Chi-squared test  
##   
## data: t1  
## X-squared = 1734.9, df = 2, p-value < 2.2e-16

assocstats(t1)

## X^2 df P(> X^2)  
## Likelihood Ratio 1282.9 2 0  
## Pearson 1734.9 2 0  
##   
## Phi-Coefficient : NA   
## Contingency Coeff.: 0.36   
## Cramer's V : 0.386

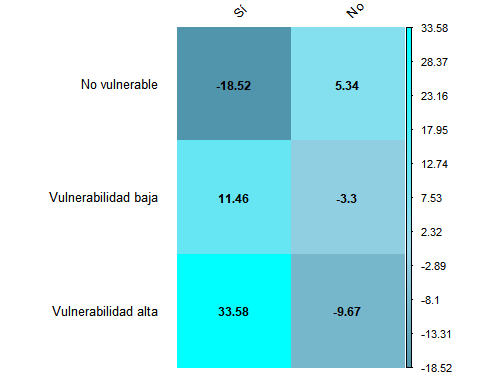
Por una parte vemos que la relación entre ambas variables es significativa, por lo que rechazamos la hipótesis nula. Por otra parte, vemos que la prueba de V de Cramer nos indica una medida de asociación de 0.386, tratándose de una medida que indica asociación moderada entre variables.

Dado que la variable de dieta inadecuada es dicotómica, no aplicaremos la medida de Gamma de Goodman-Kruskal, más adecuada cuando ambas variables son ordinales. Para ahondar en esta relación, procederemos con el análisis de residuos.

##### Análisis de residuos de Pearson

A continuación, podemos visualizar los residuos de Pearson que hemos extraído a través de la prueba de Chi Cuadrado, para ver qué interacciones contribuyen en mayor medida al cálculo del estadístico.

corrplot(chisq1$residuals,   
 type = "full",   
 tl.col= "black",   
 tl.srt = 45,   
 is.cor = FALSE,  
 tl.offset = 1,  
 method = "color",   
 addCoef.col = "black",  
 number.cex=0.75,  
 tl.cex = 0.8,   
 cl.cex = 0.7,   
 cl.align.text = "l",  
 col=colorRampPalette(c("#15607a","#9ddaed","cyan"))(200))

 A partir de los resultados, podemos ver que los residuos más elevados se encuentran en la celda que relaciona positivamente estar en una situación de alta vulnerabilidad y no tener una dieta adecuada, mientra que se relaciona negativamente no ser vulnerable y tener una dieta inadecuada.

#### Reducción del gasto en alimentación

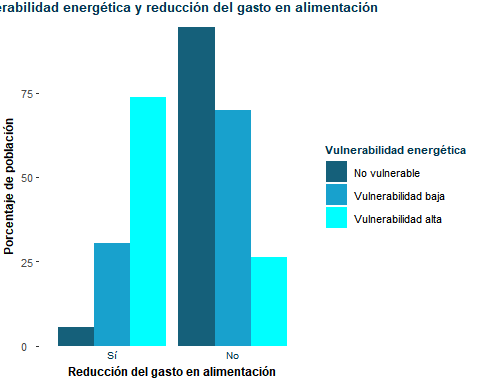
levels(foessa2$reducir\_galim)

## [1] "Sí" "No"

crear\_tabla("Vulnerabilidad\_Energetica", "reducir\_galim", "Vulnerabilidad energética y reducción del gasto en alimentación", "Reducción del gasto en alimentación", 2)

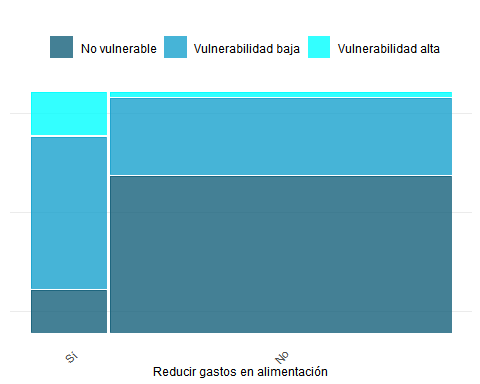
| **Vulnerabilidad energética y reducción del gasto en alimentación** | | | |
| --- | --- | --- | --- |
|  | **Cálculo** | **Reducción del gasto en alimentación** | |
| Vulnerabilidad\_Energetica |  | Sí | No |
| **No vulnerable** | *Frequency* | 369 | 6348 |
| *Row Pct* | 5.49% | 94.51% |
| *Col Pct* | 17.7% | 66.34% |
| **Vulnerabilidad baja** | *Frequency* | 1342 | 3086 |
| *Row Pct* | 30.31% | 69.69% |
| *Col Pct* | 64.35% | 32.26% |
| **Vulnerabilidad alta** | *Frequency* | 374 | 134 |
| *Row Pct* | 73.64% | 26.36% |
| *Col Pct* | 17.95% | 1.4% |

crear\_grafico(x = reducir\_galim,   
 y = Vulnerabilidad\_Energetica,   
 title = "Vulnerabilidad energética y reducción del gasto en alimentación",   
 xtitle = "Reducción del gasto en alimentación")



#### Test de independencia y medidas de asociación entre variables

# Creamos una table de contingencia ponderada  
t1 <- wtable(foessa2$Vulnerabilidad\_Energetica,  
 foessa2$reducir\_galim,   
 weights=foessa2$peso, mar = FALSE)  
  
#Establecemos la paleta de colores para el gráfico   
colors <- c("#15607a","#18a1cd","cyan")  
  
# Visualizamos la distribución de datos  
ggplot(data = subset(foessa2, !is.na(reducir\_galim))) + # Subset si queremos eliminar NA de facet\_grid  
 geom\_mosaic(aes(weight = peso,   
 x = ggmosaic::product(Vulnerabilidad\_Energetica, reducir\_galim),   
 fill= Vulnerabilidad\_Energetica),  
 na.rm = TRUE) +  
 labs(x="Reducir gastos en alimentación",   
 y= "",   
 title='') +   
 theme(axis.text.y = element\_blank(),  
 axis.ticks.y = element\_blank(),   
 axis.text.x = element\_text(angle = 45, size = 8),   
 plot.title = element\_text(size = 10.5, hjust = 0.5, face = "bold"),   
 axis.title.x = element\_text(size = 9)) +   
 scale\_fill\_manual(name="Vulnerabilidad\nEnergética",   
 values=colors)



# + facet\_grid(~entorno\_degradado) # Tercera variable

##### Chi Cuadrado (x2), V de Cramer y Coeficiente de contingencia

A continuación ejecutaremos la prueba de independencia en relación a la hipótesis nula *“No existe relación entre el nivel de vulnerabilidad energética y la reducción de gastos en la alimentación”* y también analizaremos la medida de asociación a través del test V de Cramer:

chisq1 <- chisq.test(t1) #Relación significativa  
chisq1 # 2.2e-16

##   
## Pearson's Chi-squared test  
##   
## data: t1  
## X-squared = 2242.5, df = 2, p-value < 2.2e-16

assocstats(t1)

## X^2 df P(> X^2)  
## Likelihood Ratio 2071.4 2 0  
## Pearson 2242.5 2 0  
##   
## Phi-Coefficient : NA   
## Contingency Coeff.: 0.402   
## Cramer's V : 0.439

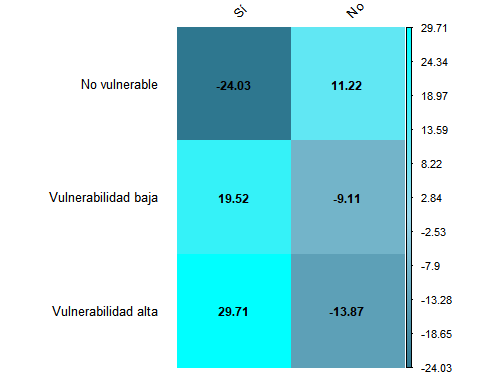
La relación entre ambas variables es significativa, por lo que rechazamos la hipótesis nula. Por otra parte, vemos que la prueba de V de Cramer nos indica una medida de asociación de 0.439, tratándose de una medida que indica asociación moderada entre variables.

Dado que la variable de dieta inadecuada es dicotómica, no aplicaremos la medida de Gamma de Goodman-Kruskal, más adecuada cuando ambas variables son ordinales. Para ahondar en esta relación, procederemos con el análisis de residuos.

##### Análisis de residuos de Pearson

A continuación, podemos visualizar los residuos de Pearson que hemos extraído a través de la prueba de Chi Cuadrado, para ver qué interacciones contribuyen en mayor medida al cálculo del estadístico.

corrplot(chisq1$residuals,   
 type = "full",   
 tl.col= "black",   
 tl.srt = 45,   
 is.cor = FALSE,  
 tl.offset = 1,  
 method = "color",   
 addCoef.col = "black",  
 number.cex=0.75,  
 tl.cex = 0.8,   
 cl.cex = 0.7,   
 cl.align.text = "l",  
 col=colorRampPalette(c("#15607a","#9ddaed","cyan"))(200))

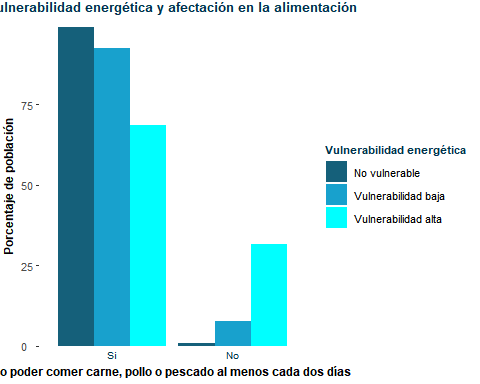
 A partir de los residuos, podemos ver que existe una clara asociación positiva entre encontrarse en situación de vulnerabilidad (tanto alta como moderada o baja) con la reducción de gastos en alimentación. También encontramos una asociación negativa, inversa, entre no ser vulnerable y la misma reducción de gastos.

#### No poder comer carne, pollo o pescado al menos cada dos días

crear\_tabla("Vulnerabilidad\_Energetica", "proteina", "Vulnerabilidad energética y afectación en la alimentación", "No poder comer carne, pollo o pescado al menos cada dos días", 2)

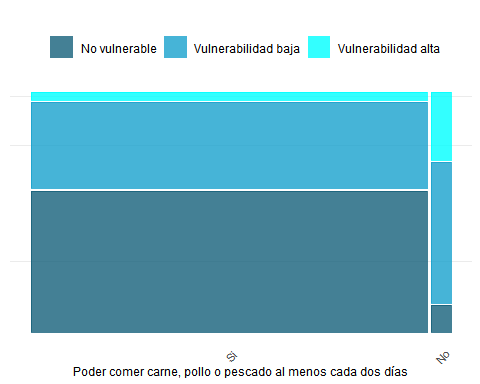
| **Vulnerabilidad energética y afectación en la alimentación** | | | |
| --- | --- | --- | --- |
|  | **Cálculo** | **No poder comer carne, pollo o pescado al menos cada dos días** | |
| Vulnerabilidad\_Energetica |  | Si | No |
| **No vulnerable** | *Frequency* | 6644 | 63 |
| *Row Pct* | 99.07% | 0.93% |
| *Col Pct* | 60.07% | 11.29% |
| **Vulnerabilidad baja** | *Frequency* | 4070 | 332 |
| *Row Pct* | 92.45% | 7.55% |
| *Col Pct* | 36.8% | 59.97% |
| **Vulnerabilidad alta** | *Frequency* | 347 | 159 |
| *Row Pct* | 68.51% | 31.49% |
| *Col Pct* | 3.13% | 28.74% |

crear\_grafico(x = proteina,   
 y = Vulnerabilidad\_Energetica,   
 title = "Vulnerabilidad energética y afectación en la alimentación",   
 xtitle = "No poder comer carne, pollo o pescado al menos cada dos días")



#### Test de independencia y medidas de asociación entre variables

# Creamos una table de contingencia ponderada  
t1 <- wtable(foessa2$Vulnerabilidad\_Energetica,  
 foessa2$proteina,   
 weights=foessa2$peso, mar = FALSE)  
  
#Establecemos la paleta de colores para el gráfico   
colors <- c("#15607a","#18a1cd","cyan")  
  
# Visualizamos la distribución de datos  
ggplot(data = subset(foessa2, !is.na(proteina))) + # Subset si queremos eliminar NA de facet\_grid  
 geom\_mosaic(aes(weight = peso,   
 x = ggmosaic::product(Vulnerabilidad\_Energetica, proteina),   
 fill= Vulnerabilidad\_Energetica),  
 na.rm = TRUE) +  
 labs(x="Poder comer carne, pollo o pescado al menos cada dos días",   
 y= "",   
 title='') +   
 theme(axis.text.y = element\_blank(),  
 axis.ticks.y = element\_blank(),   
 axis.text.x = element\_text(angle = 45, size = 8),   
 plot.title = element\_text(size = 10.5, hjust = 0.5, face = "bold"),   
 axis.title.x = element\_text(size = 9)) +   
 scale\_fill\_manual(name="Vulnerabilidad\nEnergética",   
 values=colors)



# + facet\_grid(~entorno\_degradado) # Tercera variable

##### Chi Cuadrado (x2), V de Cramer y Coeficiente de contingencia

A continuación ejecutaremos la prueba de independencia en relación a la hipótesis nula *“No existe relación entre el nivel de vulnerabilidad energética y el tamaño del hogar”* y también analizaremos la medida de asociación a través del test V de Cramer:

chisq1 <- chisq.test(t1) #Relación significativa  
chisq1 # 2.2e-16

##   
## Pearson's Chi-squared test  
##   
## data: t1  
## X-squared = 1087.1, df = 2, p-value < 2.2e-16

assocstats(t1)

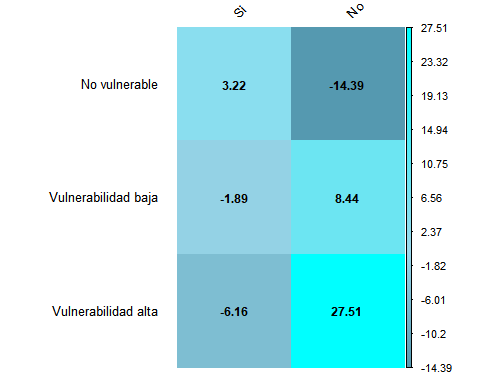
## X^2 df P(> X^2)  
## Likelihood Ratio 757.84 2 0  
## Pearson 1087.05 2 0  
##   
## Phi-Coefficient : NA   
## Contingency Coeff.: 0.293   
## Cramer's V : 0.306

En primer lugar, vemos que el Chi Cuadrado con un p-valor inferior a 0.05 nos indica que existe una relación significativa entre variables. Por otro lado, la medida de V de Cramer de 0.306 nos indica una intensidad de la asociación moderada, por lo que ahondaremos en el análisis a partir de los residuos de Pearson.

##### Análisis de residuos de Pearson

A continuación, podemos visualizar los residuos de Pearson que hemos extraído a través de la prueba de Chi Cuadrado, para ver qué interacciones contribuyen en mayor medida al cálculo del estadístico.

corrplot(chisq1$residuals,   
 type = "full",   
 tl.col= "black",   
 tl.srt = 45,   
 is.cor = FALSE,  
 tl.offset = 1,  
 method = "color",   
 addCoef.col = "black",  
 number.cex=0.75,  
 tl.cex = 0.8,   
 cl.cex = 0.7,   
 cl.align.text = "l",  
 col=colorRampPalette(c("#15607a","#9ddaed","cyan"))(200))

 En este caso, vemos que la combinación que contribuye claramente a esta asociación es la relación positiva entre encontrarse en situación de vulnerabilidad y no poder comer carne, pollo o pescado al menos cada dos días.

### Reducción del gasto en ocio

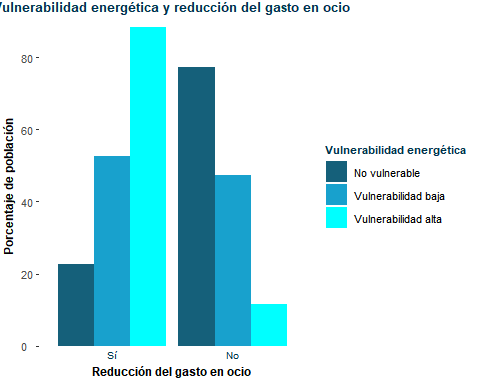
levels(foessa2$reducir\_galim)

## [1] "Sí" "No"

crear\_tabla("Vulnerabilidad\_Energetica", "reducir\_ocio", "Vulnerabilidad energética y reducción del gasto en ocio", "Reducción del gasto en ocio", 2)

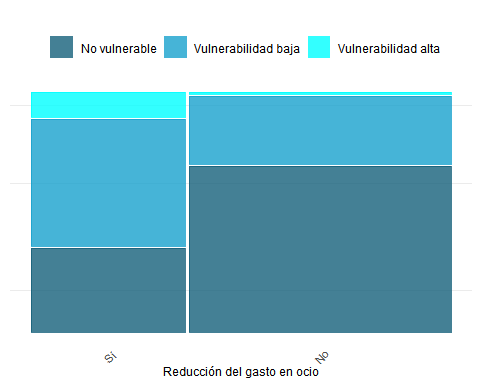
| **Vulnerabilidad energética y reducción del gasto en ocio** | | | |
| --- | --- | --- | --- |
|  | **Cálculo** | **Reducción del gasto en ocio** | |
| Vulnerabilidad\_Energetica |  | Sí | No |
| **No vulnerable** | *Frequency* | 1528 | 5187 |
| *Row Pct* | 22.76% | 77.24% |
| *Col Pct* | 35.55% | 70.65% |
| **Vulnerabilidad baja** | *Frequency* | 2320 | 2096 |
| *Row Pct* | 52.54% | 47.46% |
| *Col Pct* | 53.98% | 28.55% |
| **Vulnerabilidad alta** | *Frequency* | 450 | 59 |
| *Row Pct* | 88.49% | 11.51% |
| *Col Pct* | 10.47% | 0.8% |

crear\_grafico(x = reducir\_ocio,   
 y = Vulnerabilidad\_Energetica,   
 title = "Vulnerabilidad energética y reducción del gasto en ocio",   
 xtitle = "Reducción del gasto en ocio")



#### Test de independencia y medidas de asociación entre variables

# Creamos una table de contingencia ponderada  
t1 <- wtable(foessa2$Vulnerabilidad\_Energetica,  
 foessa2$reducir\_ocio,   
 weights=foessa2$peso, mar = FALSE)  
  
#Establecemos la paleta de colores para el gráfico   
colors <- c("#15607a","#18a1cd","cyan")  
  
# Visualizamos la distribución de datos  
ggplot(data = subset(foessa2, !is.na(reducir\_ocio))) + # Subset si queremos eliminar NA de facet\_grid  
 geom\_mosaic(aes(weight = peso,   
 x = ggmosaic::product(Vulnerabilidad\_Energetica, reducir\_ocio),   
 fill= Vulnerabilidad\_Energetica),  
 na.rm = TRUE) +  
 labs(x="Reducción del gasto en ocio",   
 y= "",   
 title='') +   
 theme(axis.text.y = element\_blank(),  
 axis.ticks.y = element\_blank(),   
 axis.text.x = element\_text(angle = 45, size = 8),   
 plot.title = element\_text(size = 10.5, hjust = 0.5, face = "bold"),   
 axis.title.x = element\_text(size = 9)) +   
 scale\_fill\_manual(name="Vulnerabilidad\nEnergética",   
 values=colors)



# + facet\_grid(~entorno\_degradado) # Tercera variable

##### Chi Cuadrado (x2), V de Cramer y Coeficiente de contingencia

A continuación ejecutaremos la prueba de independencia en relación a la hipótesis nula *“No existe relación entre el nivel de vulnerabilidad energética y reducción del gasto en ocio”* y también analizaremos la medida de asociación a través del test V de Cramer:

chisq1 <- chisq.test(t1) #Relación significativa  
chisq1 # 2.2e-16

##   
## Pearson's Chi-squared test  
##   
## data: t1  
## X-squared = 1621.4, df = 2, p-value < 2.2e-16

assocstats(t1)

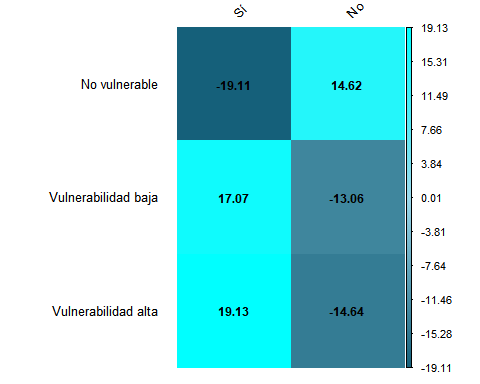
## X^2 df P(> X^2)  
## Likelihood Ratio 1654.3 2 0  
## Pearson 1621.4 2 0  
##   
## Phi-Coefficient : NA   
## Contingency Coeff.: 0.35   
## Cramer's V : 0.373

En este caso, vemos que el Chi Cuadrado con un p-valor inferior a 0.05 nos indica que existe una relación significativa entre variables. Por otro lado, la medida de V de Cramer de 0.373 nos indica una intensidad de la asociación moderada, por lo que ahondaremos en el análisis a partir de los residuos de Pearson.

##### Análisis de residuos de Pearson

A continuación, podemos visualizar los residuos de Pearson que hemos extraído a través de la prueba de Chi Cuadrado, para ver qué interacciones contribuyen en mayor medida al cálculo del estadístico.

corrplot(chisq1$residuals,   
 type = "full",   
 tl.col= "black",   
 tl.srt = 45,   
 is.cor = FALSE,  
 tl.offset = 1,  
 method = "color",   
 addCoef.col = "black",  
 number.cex=0.75,  
 tl.cex = 0.8,   
 cl.cex = 0.7,   
 cl.align.text = "l",  
 col=colorRampPalette(c("#15607a","#9ddaed","cyan"))(200))

 El análisis de residuos nos confirma que existe una relación de asociación positiva entre encontrarse en situación de vulnerabilidad energética y la existencia de reducción de gastos en ocio, mientras que esta misma asociación se da inversamente en las categorías contrarias.

### Pérdida de relaciones sociales

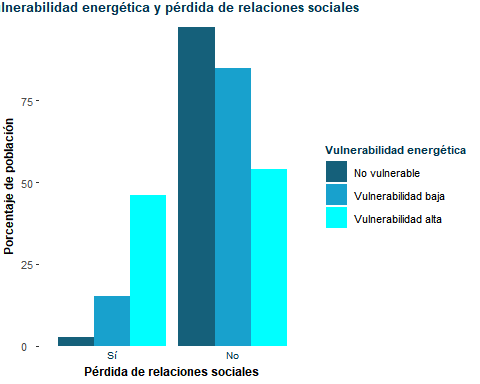
levels(foessa2$barrio\_conflictivo)

## [1] "No" "Sí"

crear\_tabla("Vulnerabilidad\_Energetica", "perdida\_relaciones", "Vulnerabilidad energética y pérdida de relaciones sociales", "Pérdida de relaciones sociales ", 2)

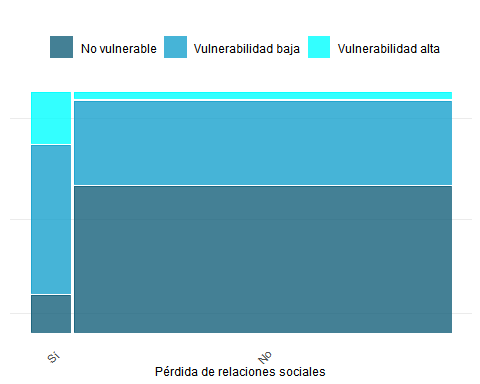
| **Vulnerabilidad energética y pérdida de relaciones sociales** | | | |
| --- | --- | --- | --- |
|  | **Cálculo** | **Pérdida de relaciones sociales** | |
| Vulnerabilidad\_Energetica |  | Sí | No |
| **No vulnerable** | *Frequency* | 169 | 6547 |
| *Row Pct* | 2.51% | 97.49% |
| *Col Pct* | 15.69% | 61.93% |
| **Vulnerabilidad baja** | *Frequency* | 674 | 3750 |
| *Row Pct* | 15.24% | 84.76% |
| *Col Pct* | 62.64% | 35.47% |
| **Vulnerabilidad alta** | *Frequency* | 233 | 274 |
| *Row Pct* | 45.99% | 54.01% |
| *Col Pct* | 21.67% | 2.59% |

crear\_grafico(x = perdida\_relaciones,   
 y = Vulnerabilidad\_Energetica,   
 title = "Vulnerabilidad energética y pérdida de relaciones sociales",   
 xtitle = "Pérdida de relaciones sociales")



#### Test de independencia y medidas de asociación entre variables

# Creamos una table de contingencia ponderada  
t1 <- wtable(foessa2$Vulnerabilidad\_Energetica,  
 foessa2$perdida\_relaciones,   
 weights=foessa2$peso, mar = FALSE)  
  
#Establecemos la paleta de colores para el gráfico   
colors <- c("#15607a","#18a1cd","cyan")  
  
# Visualizamos la distribución de datos  
ggplot(data = subset(foessa2, !is.na(perdida\_relaciones))) + # Subset si queremos eliminar NA de facet\_grid  
 geom\_mosaic(aes(weight = peso,   
 x = ggmosaic::product(Vulnerabilidad\_Energetica, perdida\_relaciones),   
 fill= Vulnerabilidad\_Energetica),  
 na.rm = TRUE) +  
 labs(x="Pérdida de relaciones sociales",   
 y= "",   
 title='') +   
 theme(axis.text.y = element\_blank(),  
 axis.ticks.y = element\_blank(),   
 axis.text.x = element\_text(angle = 45, size = 8),   
 plot.title = element\_text(size = 10.5, hjust = 0.5, face = "bold"),   
 axis.title.x = element\_text(size = 9)) +   
 scale\_fill\_manual(name="Vulnerabilidad\nEnergética",   
 values=colors)



# + facet\_grid(~entorno\_degradado) # Tercera variable

##### Chi Cuadrado (x2), V de Cramer y Coeficiente de contingencia

A continuación ejecutaremos la prueba de independencia en relación a la hipótesis nula *“No existe relación entre el nivel de vulnerabilidad energética y pérdida de relaciones sociales”* y también analizaremos la medida de asociación a través del test V de Cramer:

chisq1 <- chisq.test(t1) #Relación significativa  
chisq1 # 2.2e-16

##   
## Pearson's Chi-squared test  
##   
## data: t1  
## X-squared = 1369.4, df = 2, p-value < 2.2e-16

assocstats(t1)

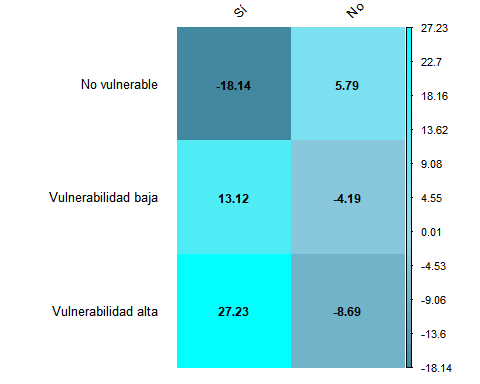
## X^2 df P(> X^2)  
## Likelihood Ratio 1123.1 2 0  
## Pearson 1369.4 2 0  
##   
## Phi-Coefficient : NA   
## Contingency Coeff.: 0.324   
## Cramer's V : 0.343

En este caso, vemos que el Chi Cuadrado con un p-valor inferior a 0.05 nos indica que existe una relación significativa entre variables. Por otro lado, la medida de V de Cramer de 0.343 nos indica una intensidad de la asociación moderada, por lo que ahondaremos en el análisis a partir de los residuos de Pearson.

##### Análisis de residuos de Pearson

A continuación, podemos visualizar los residuos de Pearson que hemos extraído a través de la prueba de Chi Cuadrado, para ver qué interacciones contribuyen en mayor medida al cálculo del estadístico.

corrplot(chisq1$residuals,   
 type = "full",   
 tl.col= "black",   
 tl.srt = 45,   
 is.cor = FALSE,  
 tl.offset = 1,  
 method = "color",   
 addCoef.col = "black",  
 number.cex=0.75,  
 tl.cex = 0.8,   
 cl.cex = 0.7,   
 cl.align.text = "l",  
 col=colorRampPalette(c("#15607a","#9ddaed","cyan"))(200))

 Al relacionar ambas variables y analizar los residuos obtenidos entre las frecuencias esperadas y las obtenidas, vemos que existe una clara asociación positiva entre vulnerabilidad energética alta y pérdida de relaciones sociales, mientras que la relación se da inversamente para las personas que no se encuentran en situación de vulnerabilidad.

### Discriminación

En primer lugar creamos una variable general sobre discriminación que agrupe las personas que han expresado haberse sentido discriminadas por alguna de las siguientes causas: etnia o raza, aspecto físico, por ser mujer o por su orientación sexual. Todas las personas que han expresado no haberse sentido nunca discriminadas o bien que no han contestado a la pregunta, se considera que no han sido discriminadas.

class(foessa2$discrim\_etnia)

## [1] "factor"

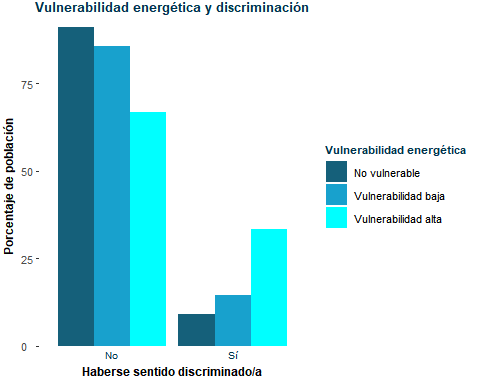
foessa2 <- foessa2 %>% mutate(discriminacion = case\_when(  
 discrim\_etnia == "Sí" ~ "Sí",   
 discrim\_fisico == "Sí" ~ "Sí",   
 discrim\_mujer == "Sí" ~ "Sí",   
 discrim\_sexual == "Sí" ~ "Sí",   
 discrim\_nunca == "No" ~ "No",   
 TRUE ~ "No"  
))  
  
foessa2$discriminacion <- as.factor(foessa2$discriminacion)

Una vez disponemos de la variable, crearemos las tablas básicas de análisis:

crear\_tabla("Vulnerabilidad\_Energetica", "discriminacion", "Vulnerabilidad energética y discriminación", "Haberse sentido discriminado/a", 2)

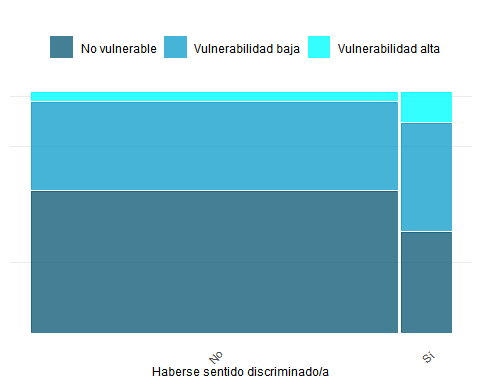
| **Vulnerabilidad energética y discriminación** | | | |
| --- | --- | --- | --- |
|  | **Cálculo** | **Haberse sentido discriminado/a** | |
| Vulnerabilidad\_Energetica |  | No | Sí |
| **No vulnerable** | *Frequency* | 6120 | 598 |
| *Row Pct* | 91.1% | 8.9% |
| *Col Pct* | 59.72% | 42.49% |
| **Vulnerabilidad baja** | *Frequency* | 3789 | 639 |
| *Row Pct* | 85.57% | 14.43% |
| *Col Pct* | 36.98% | 45.43% |
| **Vulnerabilidad alta** | *Frequency* | 339 | 170 |
| *Row Pct* | 66.61% | 33.39% |
| *Col Pct* | 3.31% | 12.07% |

crear\_grafico(x = discriminacion,   
 y = Vulnerabilidad\_Energetica,   
 title = "Vulnerabilidad energética y discriminación",   
 xtitle = "Haberse sentido discriminado/a")



#### Test de independencia y medidas de asociación entre variables

# Creamos una table de contingencia ponderada  
t1 <- wtable(foessa2$Vulnerabilidad\_Energetica,  
 foessa2$discriminacion,   
 weights=foessa2$peso, mar = FALSE)  
  
#Establecemos la paleta de colores para el gráfico   
colors <- c("#15607a","#18a1cd","cyan")  
  
# Visualizamos la distribución de datos  
ggplot(data = subset(foessa2, !is.na(discriminacion))) + # Subset si queremos eliminar NA de facet\_grid  
 geom\_mosaic(aes(weight = peso,   
 x = ggmosaic::product(Vulnerabilidad\_Energetica, discriminacion),   
 fill= Vulnerabilidad\_Energetica),  
 na.rm = TRUE) +  
 labs(x="Haberse sentido discriminado/a",   
 y= "",   
 title='') +   
 theme(axis.text.y = element\_blank(),  
 axis.ticks.y = element\_blank(),   
 axis.text.x = element\_text(angle = 45, size = 8),   
 plot.title = element\_text(size = 10.5, hjust = 0.5, face = "bold"),   
 axis.title.x = element\_text(size = 9)) +   
 scale\_fill\_manual(name="Vulnerabilidad\nEnergética",   
 values=colors)



# + facet\_grid(~entorno\_degradado) # Tercera variable

##### Chi Cuadrado (x2), V de Cramer y Coeficiente de contingencia

A continuación ejecutaremos la prueba de independencia en relación a la hipótesis nula *“No existe relación entre el nivel de vulnerabilidad energética y el tamaño del hogar”* y también analizaremos la medida de asociación a través del test V de Cramer:

chisq1 <- chisq.test(t1) #Relación significativa  
chisq1 # 2.2e-16

##   
## Pearson's Chi-squared test  
##   
## data: t1  
## X-squared = 304.9, df = 2, p-value < 2.2e-16

assocstats(t1)

## X^2 df P(> X^2)  
## Likelihood Ratio 248.87 2 0  
## Pearson 304.90 2 0  
##   
## Phi-Coefficient : NA   
## Contingency Coeff.: 0.16   
## Cramer's V : 0.162

En este último análisis vemos que la relación entre ambas variables es significativa, por lo que rechazamos la hipótesis nula. Por otra parte, vemos que la prueba de V de Cramer nos indica una medida de asociación de 0.162, tratándose de una medida muy baja, por debajo de 0.2, por lo que no se estimará asociación. El coeficiente de contingencia nos indica igualmente un valor muy bajo.