

Laborato 3 Correspondencias Simples y Multiples

Grupo 6

Integrantes

- Edwin Sanchez
- Stephanie Tamayo
- Andres Felipe Torres
- Fredy Urrea
- Sergio Velasquez
- Manuel Espitia

Introduccion

Carga de archivos y organizaicon de los datos

```
library("FactoMineR")
library("dplyr")
library("kableExtra")
library(readxl)
library(FactoMineR)
library(factoextra)
library(kableExtra)
library(readr)
library(tidyr)
library(dplyr)
library(tibble)

encuesta <- read_csv2("ECC_completa_19426.csv")
ecc<- read_csv2("ECC_completa_19426.csv")
```

Capitulo 5 Análisis de correspondencias simples (ACS)

Punto 1

Con la metodologia del ejemplo 5.4 hacer un ACS para la tabla de contingencias p17b x p21 con las preguntas “p17_b” (en las filas) sobre la facilidad para cumplir la ley y la preferencia por hacer acuerdos p21. Identificar patrones o tendencias si los hat comentar los resultados.

Punto 2

Construir la tabla de contingencias p17b x ciudad que le correspondio al frupo y las ciudades de Asuncion y Montevideo. Yuxtaponerla a la tabla p17b x p21 del ejercicio 1 y utilizarla como variables suplementarias para averiguar si se puede identificar algun patron o tendecia en la facilidad para cumplir la ley en las ciudades suplementarias con respecto a las tendencias de la ciudad que le correspondio al grupo

Punto 3

Realizar un ACS a la tabla de contingencias “p17_b” (filas) vs preferencia para hacer acuerdo p21 en un ejercicio similar al 1.

```
tabla <- table(encuesta$p17_b, encuesta$p21)

tabla_df <- as.data.frame(tabla)

tabla_bonita <- tabla_df %>%
  pivot_wider(names_from = Var2, values_from = Freq, values_fill = 0) %>%
  rename(`Facilidad para cumplir la ley (p17_b)` = Var1)

kable(tabla_bonita,
      caption = "Tabla de contingencia entre p17b y p21",
      format = "latex",
      booktabs = TRUE,
      escape = TRUE) %>%
  kable_styling(latex_options = c("striped", "hold_position"), font_size = 8) %>%
  column_spec(1, latex_column_spec = "p{3cm}") %>%
  column_spec(2, latex_column_spec = "p{2.5cm}") %>%
  column_spec(3, latex_column_spec = "p{2.5cm}") %>%
  column_spec(4, latex_column_spec = "p{2.5cm}") %>%
  column_spec(5, latex_column_spec = "p{2.5cm}")
```

Table 1: Tabla de contingencia entre p17b y p21

Facilidad para cumplir la ley (p17_b)	p21_HACACU=1_n	p21_HACACU=2_cnp21_HACACU=3_cs	p21_HACACU=4_s	
p17b_FACL=1_n	88	86	171	129
p17b_FACL=2_cn	315	262	746	428
p17b_FACL=3_cs	1088	1022	3705	2087
p17b_FACL=4_s	1420	1077	3262	3144

Ejecucion de acs

```
require(FactoMineR)
acsp17p21= CA(tabla, graph = T)
```

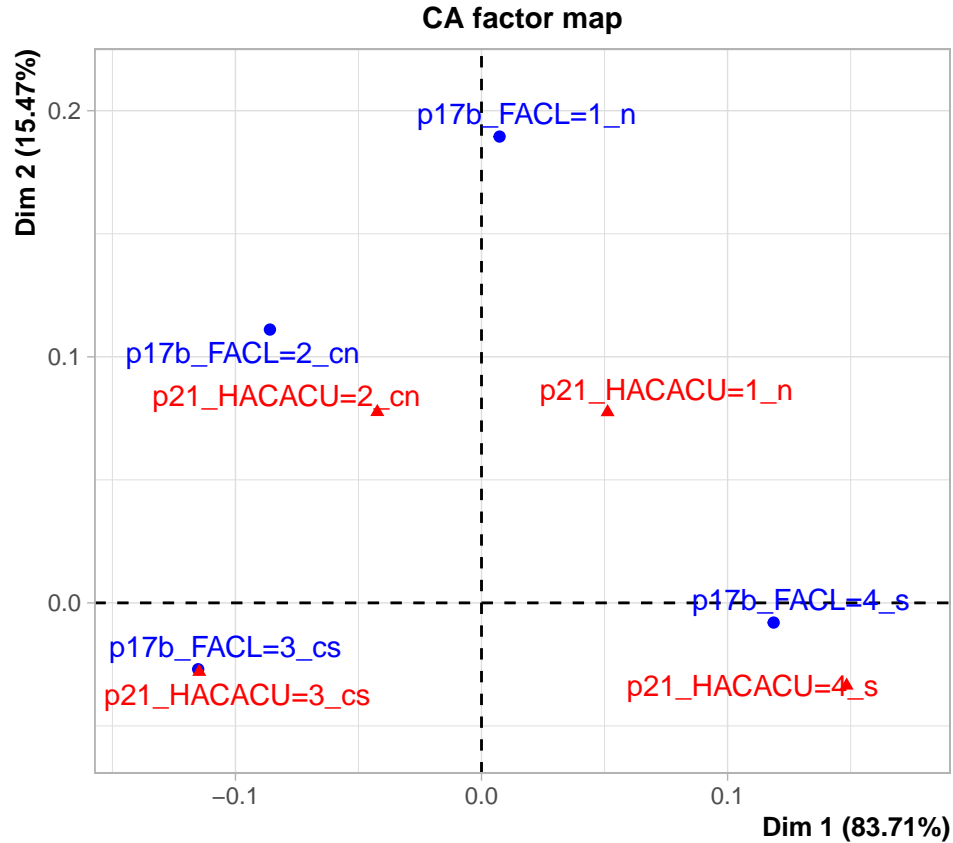


Tabla de valores propios y varianza acumulada

```
valores_propios <- acsp17p21$eig
kable(valores_propios,
      col.names = c("Valor propio", "% de varianza", "% varianza acumulada"),
      caption = "Tabla de valores propios y varianza explicada del ACS") %>%
kable_styling(latex_options = c("hold_position", "striped", "scale_down"))
```

```
## Warning in styling_latex_scale(out, table_info, "down"): Longtable cannot be
## resized.
```

Table 2: Tabla de valores propios y varianza explicada del ACS

	Valor propio	% de varianza	% varianza acumulada
dim 1	0.0127797	83.7061683	83.70617
dim 2	0.0023611	15.4651828	99.17135
dim 3	0.0001265	0.8286489	100.00000

Tabla contribuciones columnas

```
# Extraer solo primeras 2 dimensiones
coord_col <- acsp17p21$col$coord[, 1:2]
contrib_col <- acsp17p21$col$contrib[, 1:2]
cos2_col <- acsp17p21$col$cos2[, 1:2]
```

```

# Combinar en una sola tabla
tabla_columnas <- cbind(coord_col, contrib_col, cos2_col)

# Asignar nombres adecuados
colnames(tabla_columnas) <- c(
  "Coord_Dim1", "Coord_Dim2",
  "Contrib_Dim1", "Contrib_Dim2",
  "Cos2_Dim1", "Cos2_Dim2"
)

library(knitr)

## Warning: package 'knitr' was built under R version 4.3.3

library(kableExtra)

kable(tabla_columnas,
      digits = 3,
      caption = "Coordenadas, contribuciones y cosenos cuadrados de las columnas (p21) - primeras 2 dim",
      format = "latex",
      booktabs = TRUE) %>%
  kable_styling(latex_options = c("striped", "hold_position"), font_size = 8)

```

Table 3: Coordenadas, contribuciones y cosenos cuadrados de las columnas (p21) - primeras 2 dimensiones

	Coord_Dim1	Coord_Dim2	Contrib_Dim1	Contrib_Dim2	Cos2_Dim1	Cos2_Dim2
p21_HACACU=1_n	0.051	0.078	3.143	38.914	0.292	0.668
p21_HACACU=2_cn	-0.042	0.078	1.804	32.764	0.215	0.723
p21_HACACU=3_cs	-0.115	-0.028	42.656	13.774	0.943	0.056
p21_HACACU=4_s	0.148	-0.034	52.397	14.549	0.951	0.049

Grafico

```

library(factoextra)
library(ggplot2)
library(grid)

# Crear gráfico base
f_col <- fviz_ca_col(acsp17p21, labelsize = 5,
  xlim = c(-0.15, 0.15), ylim = c(-0.1, 0.1),
  title = "Perfiles columna: p21")

# Extraer coordenadas
coords <- f_col$data

# Orden correcto de las modalidades (¡ya corregido!)
orden_modalidades <- c("p21_HACACU=1_n",
  "p21_HACACU=2_cn",
  "p21_HACACU=3_cs",
  "p21_HACACU=4_s")

# Filtrar y ordenar según ese orden
coords_linea <- coords[match(orden_modalidades, coords$name), ]

```

```

# Crear segmentos entre puntos consecutivos
segmentos <- data.frame(
  x = coords_linea$x[-nrow(coords_linea)],
  y = coords_linea$y[-nrow(coords_linea)],
  xend = coords_linea$x[-1],
  yend = coords_linea$y[-1]
)

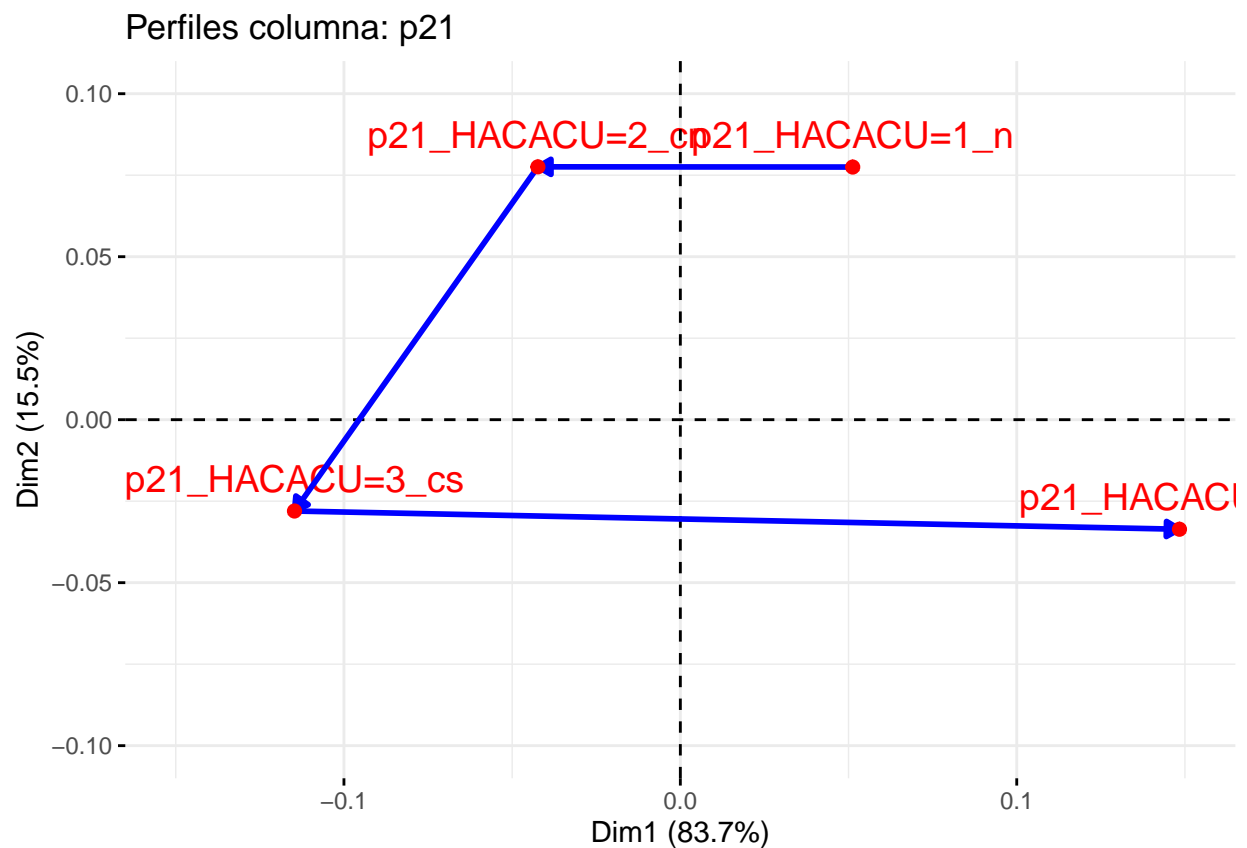
# Agregar flechas y puntos al gráfico base
f_col +
  geom_segment(data = segmentos,
    aes(x = x, y = y, xend = xend, yend = yend),
    arrow = arrow(length = unit(0.2, "cm"), type = "closed"),
    color = "blue", size = 1) +
  geom_point(data = coords_linea, aes(x = x, y = y),
    color = "red", size = 2)

```

```

## Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use `linewidth` instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.

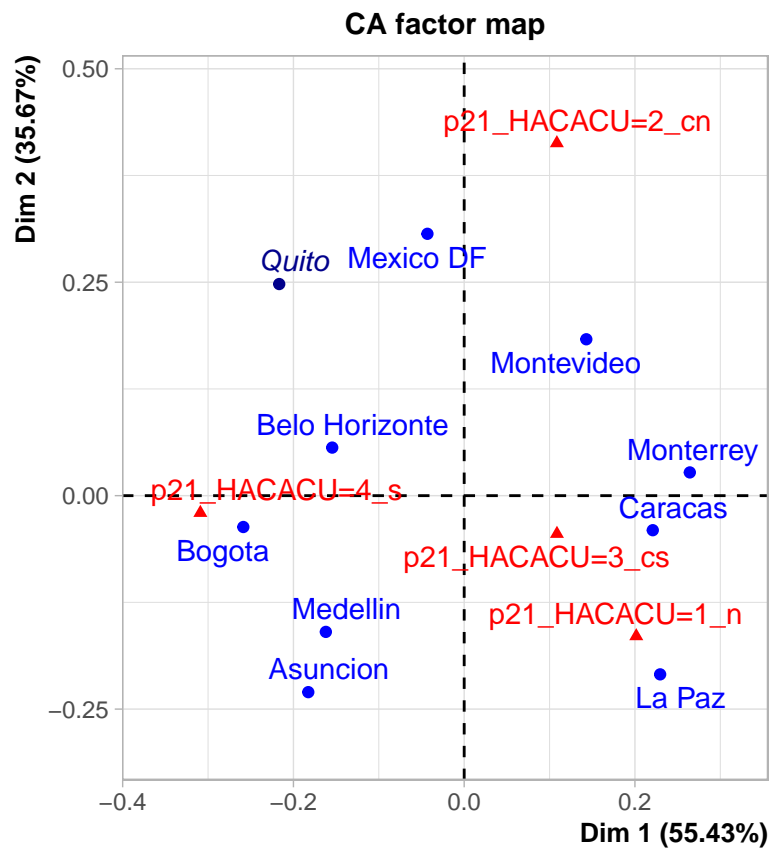
```



Punto 4

Apilar como ilustrativa la tabla ciudades (filas) vs preferencias para hacer acuerdos p21 para investigar si hay algun patron o tendencia en las ciudades respecto a las tendencias de la ciudad que le correspondio al grupo

```
#GRAFICO
tab_ciud_p21 <- with(ecc, table(ciudad2, p21))
ciudad_grupo <- "Quito"
fila_sup <- which(rownames(tab_ciud_p21) == ciudad_grupo)
res.ca <- CA(tab_ciud_p21,
             graph = TRUE,
             row.sup = fila_sup)
```



El mapa de factores del Análisis de Correspondencia (CA) nos permite visualizar las relaciones y patrones de asociación entre las diferentes ciudades (puntos azules) y las categorías de “preferencia para hacer acuerdos p21” (triángulos rojos). La ciudad de Quito ha sido tratada como un punto suplementario o ilustrativo, lo que significa que su posición se proyecta en el mapa sin influir en la construcción de los ejes principales, permitiéndonos observar su tendencia en relación con las variables activas y las demás ciudades.

El análisis de correspondencias muestra que las dos primeras dimensiones capturan un 91,10 % de la inercia total: la Dimensión 1 (eje horizontal) aporta un 55,43 % y la Dimensión 2 (eje vertical) un 35,67 %.

En la Dimensión 1, el extremo derecho se asocia con la categoría **p21_HACACU = 1** (“nunca”), donde se agrupan Monterrey, Caracas y La Paz; el izquierdo, con **p21_HACACU = 3** (“casi siempre”), muy cerca de Bogotá, Medellín, Asunción y la ciudad suplementaria, Quito. Así, esta dimensión distingue claramente el polo “nunca” del polo “casi siempre”.

La Dimensión 2, por su parte, separa en su parte superior la preferencia **p21_HACACU = 2** (“casi

nunca”), representada especialmente por Ciudad de México y Montevideo, mientras que en la zona inferior no emerge un vínculo fuerte con ninguna categoría, aunque Medellín, Asunción y La Paz se extienden ligeramente hacia ese lado.

Si observamos los grupos de ciudades, se distinguen tres patrones:

- **“Siempre”**: Quito, Bogotá y Belo Horizonte se ubican en el cuadrante superior-izquierdo, muy próximos a la categoría **4 (“siempre”)**, lo que indica que en estas urbes predomina la preferencia por formalizar acuerdos de forma invariable.
- **“Casi nunca / Nunca”**: Monterrey, Caracas y La Paz aparecen juntos en el cuadrante inferior-derecho, alineados con “nunca” y “casi nunca”, lo que revela su reticencia a hacer acuerdos.
- **“Intermedio”**: Ciudad de México y Montevideo se sitúan en una posición media, entre “casi nunca” y “casi siempre”, mostrando un patrón de respuestas más equilibrado.

En definitiva, existe un claro agrupamiento según la propensión a pactar acuerdos: Quito forma parte del grupo “siempre”, mientras que La Paz, Caracas y Monterrey constituyen el polo opuesto, y México DF y Montevideo ocupan un punto intermedio.

```
#TABLA
nuevos_nombres_p21 <- c("SIEMPRE", "CASI SIEMPRE", "NUNCA", "CASI NUNCA")
colnames(tab_ciud_p21) <- nuevos_nombres_p21
# Generar la tabla en formato kable para LaTeX
kable(
  tab_ciud_p21, # Tu tabla de contingencia de ciudades vs. p21
  "latex",
  booktabs = TRUE,
  caption = "Tabla de Contingencia: Ciudades vs. Preferencia para hacer acuerdos (p21)",
  label = "tab_ciud_p21_contingency" # Un label descriptivo para la tabla
) %>%
kable_styling(latex_options = c("striped", "hold_position"))
```

Table 4: Tabla de Contingencia: Ciudades vs. Preferencia para hacer acuerdos (p21)

	SIEMPRE	CASI SIEMPRE	NUNCA	CASI NUNCA
Asuncion	55	12	299	215
Belo Horizonte	202	188	481	507
Bogota	502	344	1141	1431
Caracas	465	274	937	424
La Paz	519	161	957	421
Medellin	311	128	956	820
Mexico DF	230	534	970	749
Monterrey	363	300	1067	355
Montevideo	216	248	516	290
Quito	87	287	618	602

Para ilustrar cómo se agrupan las ciudades según su propensión a pactar acuerdos, proyectamos la tabla de frecuencias de cada ciudad frente a las cuatro categorías de p21 (“siempre”, “casi siempre”, “nunca” y “casi nunca”) en el plano factorial definido por las dos primeras dimensiones, que explican en conjunto el 91,10 % de la inercia (55,43 % en el eje 1 y 35,67 % en el eje 2).

El panorama tras proyectar las ciudades revela tres zonas bien diferenciadas:

1. **Grupo “Siempre”** Asunción, Bogotá y Belo Horizonte se sitúan en el extremo izquierdo del eje 1, muy cerca de la modalidad **“siempre”**. Esto indica que en estas urbes una proporción notable de encuestados prefiere hacer acuerdos de forma invariable.

2. **Grupo “Nunca / Casi nunca”** En el cuadrante inferior-derecho aparecen con claridad La Paz, Caracas y Monterrey, muy alineadas con “nunca” y “casi nunca”. Aquí predomina la reticencia a formalizar acuerdos o se tiende a evitarlos casi siempre.
3. **Posición intermedia** Ciudad de México y Montevideo ocupan una posición central, equidistante de “casi nunca” y “casi siempre”, lo que sugiere un patrón de respuestas más equilibrado o mixto.

La ciudad suplementaria **Quito**, proyectada sobre el mismo plano, queda próxima al clúster “casi siempre” y “siempre”, reforzando su afinidad con Asunción, Bogotá y Belo Horizonte.

En definitiva, el análisis confirma un claro agrupamiento: un polo “siempre” (Asunción, Bogotá, Belo Horizonte y Quito), otro polo “nunca/casi nunca” (La Paz, Caracas, Monterrey) y un segmento intermedio (Ciudad de México y Montevideo), lo que evidencia la variedad de tendencias al hacer acuerdos en las diferentes ciudades estudiadas.

Punto 5

Seleccionar un pregunta del “Formulario Generico ECC” que el rupo considere de interes para realizar un ACS de la tabla de contingencia de esa pregunta con las ciudades e identificar tendencias o patrones por ciudades con respecto a esa pregunta

```
# TABLA DE CONTINGENCIA -----

tab <- ecc %>%
  select(ciudad = ciudad2, transp = p9) %>%
  filter(!is.na(ciudad) & !is.na(transp)) %>%
  table()

# Distancia 2 y clustering -----

perf <- sweep(tab, 1, rowSums(tab), "/")
w_col <- colSums(tab) / sum(tab)
chi_d <- dist( sweep(perf, 2, sqrt(w_col), "/" ) )

hc <- hclust(chi_d, method = "ward.D2")
k <- 3

pal_hex <- c("#00AFBB", "#E7B800", "#FC4E07") # grupo 1-2-3
names(pal_hex) <- pal_hex # ← clave = HEX

labs_hex <- c(
  `#00AFBB` = "Transporte masivo + taxi",
  `#E7B800` = "Bus dominante",
  `#FC4E07` = "Bicicleta dominante"
)

# TABLA DE RESUMEN -----

df_resumen <- data.frame(
  Ciudad = rownames(tab),
  Grupo = cutree(hc, k),
  Descripcion = labs_hex[ pal_hex[ cutree(hc, k) ] ],
  row.names = NULL
)
```



```
print(df_resumen)
```

```
##          Ciudad Grupo      Descripcion
## 1      Asuncion      1 Transporte masivo + taxi
## 2 Belo Horizonte      2      Bus dominante
## 3      Bogota        2      Bus dominante
## 4      Caracas        2      Bus dominante
## 5      La Paz         2      Bus dominante
## 6      Medellin       2      Bus dominante
## 7      Mexico DF      2      Bus dominante
## 8      Monterrey      2      Bus dominante
## 9      Montevideo     1 Transporte masivo + taxi
## 10     Quito          3      Bicicleta dominante
```

```
# DENDOGRAMA -----

p <- fviz_dend(
  hc,
  k          = k,
  k_colors   = pal_hex,
  rect       = TRUE,
  rect_fill  = TRUE,
  cex        = 0.85,
  main       = "Agrupaciones de ciudades latinoamericanas según patrón de transporte (p9)",
  xlab       = "Ciudades latinoamericanas",
  ylab       = "Altura"
) +

  #COLOR TIPO HEX?(?)
  scale_color_identity(
    name     = "Dominant pattern",
    breaks   = names(labs_hex),
    labels   = labs_hex
  ) +

  guides(colour = guide_legend(override.aes = list(size = 2))) +

  theme(
    legend.position = "top",
    legend.title    = element_text(face = "bold")
  )
```

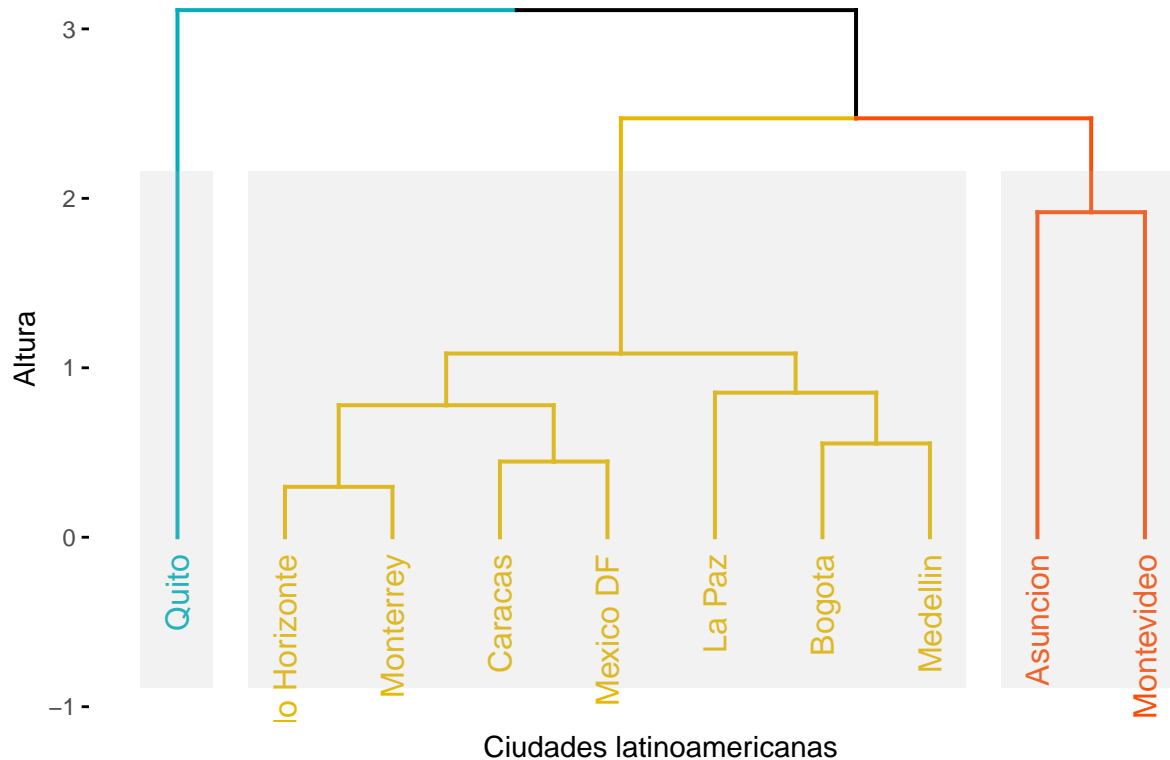
```
## Warning: The `scale` argument of `guides()` cannot be `FALSE`. Use "none" instead as
## of ggplot2 3.3.4.
```

```
## i The deprecated feature was likely used in the factoextra package.
## Please report the issue at <https://github.com/kassambara/factoextra/issues>.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
```

```
## Scale for colour is already present.
## Adding another scale for colour, which will replace the existing scale.
```

```
print(p)
```

Agrupaciones de ciudades latinoamericanas según patrón de transporte (p9)



Punto 6

Explorar las posibles asociaciones de la pregunta seleccionada en el punto 5 con la pregunta sobre la facilidad para cumplir la ley “p17_b” adicionandola como varibale ilustrativa.

Este análisis busca explorar la asociación entre los patrones de transporte (p9) y la percepción de facilidad para cumplir la ley (p17_b). Se aplica un Análisis de Correspondencias Simples (ACS) a la tabla de contingencia Ciudad × Transporte, y se proyecta p17_b como variable ilustrativa para identificar posibles asociaciones.

TABLA DE CONTINGENCIA Ciudad × Transporte (p9) -----

```
tab_p9 <- ecc %>%
  select(ciudad = ciudad2, transporte = p9) %>%
  filter(!is.na(ciudad), !is.na(transporte)) %>%
  count(ciudad, transporte) %>%
  pivot_wider(names_from = transporte, values_from = n, values_fill = 0) %>%
  column_to_rownames("ciudad")
```

```
## Warning in grep("^.[.](?:.[.][1-9][0-9]*)$", names): unable to translate
## 'Autom<a2>vil particular' to a wide string
## Warning in grep("^.[.](?:.[.][1-9][0-9]*)$", names): input string 3 is
## invalid
```

```
print(tab_p9)
```

```
##           A pie Autom\xa2vil particular Bicicleta Motocicleta Otros
## Asuncion      45                158          1          38      4
## Belo Horizonte 129              345          6          34     14
## Bogota        485              344          94          60      5
## Caracas       121              635          0          56     107
## La Paz        309               63           2           4     122
## Medellin      296              215          22          196      5
## Mexico DF     60               621          13           21     32
## Monterrey     47               635          11           37      0
## Montevideo    290              155          15           0       2
## Quito         86               118          755          43      10
##           Sistema de transport Taxi Autobus, colectivo
## Asuncion                325      9              0
## Belo Horizonte          16     16              824
## Bogota                   530  129             1765
## Caracas                   361   55              765
## La Paz                    0     62             1538
## Medellin                  132  147             1205
## Mexico DF                 355   83             1309
## Monterrey                  29  102             1357
## Montevideo                 274  500              51
## Quito                      9   249             282
```

La tabla de contingencia Ciudad \times Transporte (p9) evidencia diferencias claras en los patrones de movilidad urbana: Bogotá, Medellín y Ciudad de México muestran una fuerte dependencia del autobús o colectivo, mientras que Montevideo destaca por un uso inusualmente alto del taxi y Quito por un notable predominio de la bicicleta, lo que sugiere la influencia de factores estructurales, culturales y de política pública en la elección del transporte. Estas disparidades justifican el uso del Análisis de Correspondencias Simples (ACS) para identificar agrupamientos y explorar cómo estos patrones se relacionan con variables como la percepción sobre el cumplimiento de la ley (p17_b), permitiendo así una comprensión más profunda y crítica de la movilidad urbana en cada contexto.

```
# ANÁLISIS DE CORRESPONDENCIAS (ACS) -----
```

```
res.ca <- CA(tab_p9, graph = FALSE)
```

```
# ASOCIAR p17_b A CADA CIUDAD -----
```

```
ley_ciudad <- ecc %>%
  filter(!is.na(ciudad2), !is.na(p17_b)) %>%
  group_by(ciudad2, p17_b) %>%
  summarise(n = n(), .groups = "drop") %>%
  group_by(ciudad2) %>%
  slice_max(n, n = 1, with_ties = FALSE) %>% # moda
  rename(ciudad = ciudad2, ley = p17_b)
```

```
# RECODIFICAR p17_b -----
```

```
ley_ciudad <- ley_ciudad %>%
  mutate(ley = case_when(
```

```

ley == "p17b_FACL=1_n" ~ "Nada",
ley == "p17b_FACL=2_cn" ~ "Casi nada",
ley == "p17b_FACL=3_cs" ~ "Casi sí",
ley == "p17b_FACL=4_s" ~ "Sí",
TRUE ~ as.character(ley)
))

# EXTRAER COORDENADAS DEL CA Y UNIR CON p17_b -----

coord_ciudades <- as.data.frame(res.ca$row$coord)
colnames(coord_ciudades) <- gsub(" ", ".", colnames(coord_ciudades)) # Convertir "Dim 1" a "Dim.1"
coord_ciudades$ciudad <- rownames(coord_ciudades)

# Unimos coordenadas con la ley por ciudad
coords_df <- left_join(coord_ciudades, ley_ciudad, by = "ciudad")

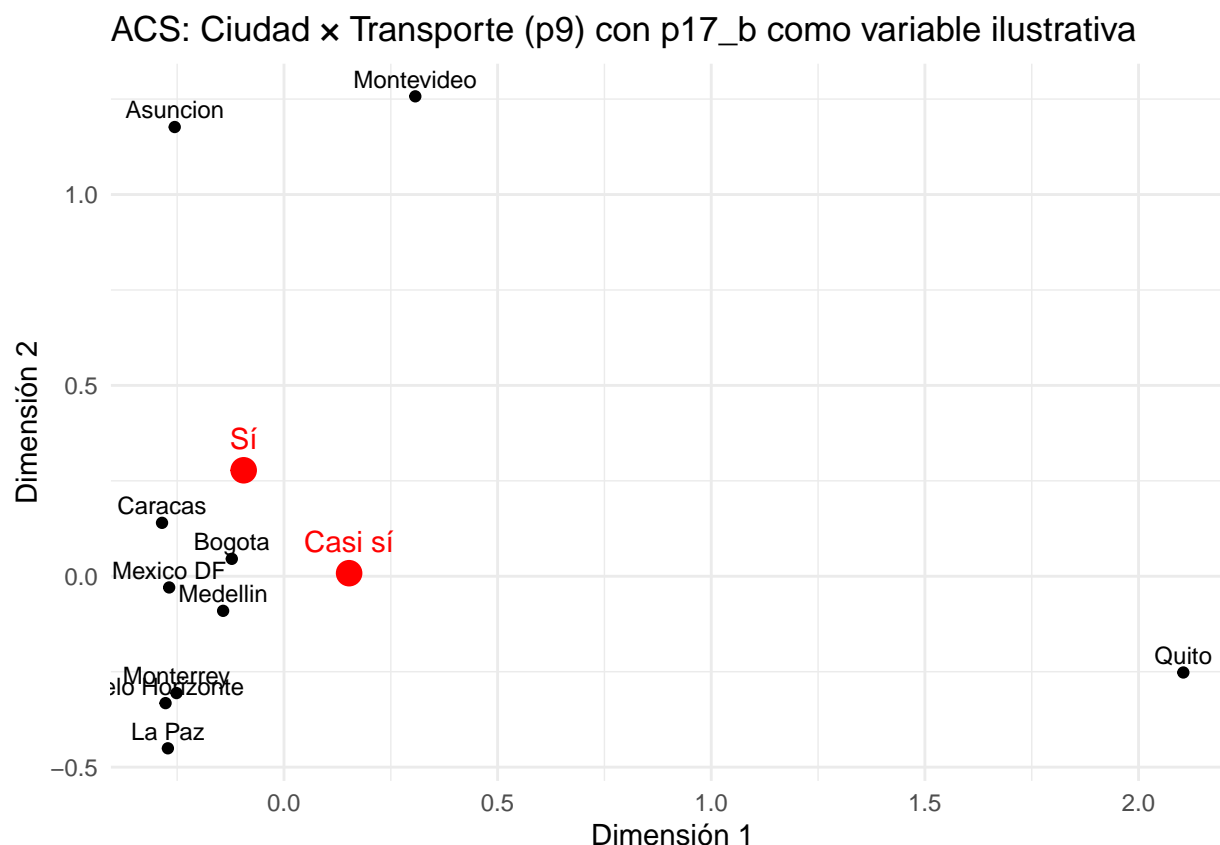
# CALCULAR CENTROIDES DE CADA NIVEL DE p17_b -----

centroides <- coords_df %>%
  group_by(ley) %>%
  summarise(across(starts_with("Dim"), \ (x) mean(x, na.rm = TRUE)))

# GRAFICAR RESULTADOS -----

ggplot(coords_df, aes(x = Dim.1, y = Dim.2)) +
  geom_point(color = "black") +
  geom_text(aes(label = ciudad), vjust = -0.5, size = 3) +
  geom_point(data = centroides, aes(x = Dim.1, y = Dim.2), color = "red", size = 4) +
  geom_text(data = centroides, aes(x = Dim.1, y = Dim.2, label = ley),
            color = "red", vjust = -1, size = 4) +
  labs(
    title = "ACS: Ciudad x Transporte (p9) con p17_b como variable ilustrativa",
    x = "Dimensión 1",
    y = "Dimensión 2"
  ) +
  theme_minimal()

```



El análisis revela una asociación diferenciada entre la percepción sobre la facilidad para cumplir la ley (variable p17_b) y los patrones de movilidad urbana en distintas ciudades. La categoría “Sí”, que indica una percepción favorable, se relaciona principalmente con ciudades como Asunción y Montevideo, sugiriendo que en estos contextos la movilidad y el cumplimiento normativo podrían estar más alineados. Por otro lado, la categoría “Casi sí” agrupa a ciudades como Bogotá, Medellín y Ciudad de México, ubicadas en el centro del gráfico, lo que indica una percepción intermedia sobre el cumplimiento de la ley en entornos urbanos con patrones de transporte más convencionales. Quito se destaca por su clara separación del resto, reflejando un patrón de transporte distinto, probablemente vinculado al uso masivo de la bicicleta, y sin una asociación directa con las categorías de percepción sobre la ley mencionadas. Esta desconexión sugiere que factores específicos de movilidad alternativa pueden influir en la percepción ciudadana de manera diferente. En conjunto, estos hallazgos apuntan a una posible relación entre ciertos patrones de movilidad urbana y percepciones más favorables sobre el cumplimiento de la ley, lo que puede ser relevante para el diseño de políticas públicas que integren movilidad y gobernanza urbana de manera más efectiva y contextualizada.

Capítulo 6 Análisis de correspondencias Múltiples (ACM)

Punto 1

Utilizar el archivo ECC_completa_19426.csv y los datos de la ciudad que le correspondió al grupo para el laboratorio de ACS para realizar un ACM con las siguientes preguntas como variables activas : p_20_a a p20_k, p21, p27 y p33_a a p33_p.

Punto 2

Utilizar como variables ilustrativas el nivel socioeconómico (NSE), el sexo (p5) y el nivel educativo (p7_NEd) e identificar si hay alguna tendencias o patrón de asociación con las variables activas.