

# Prueba ADIP

Marco Méndez Atienza

15/7/2022

## Ejercicio 1

Suponga que se le informa a la Jefa de Gobierno sobre un incremento en el número de robos a pasajeros a bordo de microbús con y sin violencia de mayo de 2021 al mismo mes de 2022. Sin embargo, el solo número de incremento que se le presenta resulta insuficiente para la toma de decisiones. Con base en los datos de carpetas de investigación, diseñar un reporte de máximo 5 láminas en donde se presenten los datos más relevantes que consideres y que le permitan a la Jefa de Gobierno instruir sobre acciones a llevar a cabo para disminuir esta problemática.

Para la resolución de este ejercicio, es necesario utilizar la base de datos de las Carpetas de Investigación de la FGJ, mismas que incluyen datos de julio de 2020 hasta la última información registrada.

Se cargan los paquetes pertinentes:

```
library(pacman)

p_load(tidyverse, viridis, hrbrthemes, ggthemes, RColorBrewer, lubridate,
       DIZtools, janitor, plyr, zoo, reshape2, kableExtra, svy, showtext, glue, pandoc)

options(scipen=999)
```

Se carga la base de datos a utilizar:

```
setwd("C:/Users/marco/OneDrive/Documents/R/Prueba ADIP")

carpetas <- read.csv("~/R/Prueba ADIP/datos/carpetas_completa_mayo_2022.csv",
                      encoding="UTF-8")
```

Se convierten las fechas al formato apropiado, y se filtra un nuevo objeto con las fechas relevantes para este ejercicio.

```
# glimpse(carpetas)

###

carpetas[, c("fecha_hechos", "fecha_inicio")] <- lapply(carpetas[,c("fecha_hechos", "fecha_inicio")],
                                                       as.Date)

# glimpse(carpetas)

###

fecha1 <- as.Date("2021-05-01")
fecha2 <- as.Date("2022-06-01")
```

```

carpetas_21_22 <- carpetas %>%
  filter(fecha_hechos >= fecha1 & fecha_hechos <= fecha2)

###

# unique(carpetas_21_22$delito)

```

Se filtran los dos delitos a estudiar: Robo a pasajero a bordo de pesero colectivo con y sin violencia:

```

carpetas_21_22 <- carpetas_21_22 %>%
  filter(delito == "ROBO A PASAJERO A BORDO DE PESERO COLECTIVO SIN VIOLENCIA" |
         delito == "ROBO A PASAJERO A BORDO DE PESERO COLECTIVO CON VIOLENCIA")

carpetas_21_22$delito <- tolower(carpetas_21_22$delito)
carpetas_21_22$delito <- firstup(carpetas_21_22$delito)

```

Se calcula el número de estos delitos por mes en el periodo señalado:

```

carpetas_21_22_mes <- tabyl(carpetas_21_22, fecha_hechos, delito)

carpetas_21_22_mes <- carpetas_21_22_mes %>%
  mutate(Total = rowSums(.[2:3]),
         Fecha = format(as.Date(fecha_hechos, "%Y-%m-%d"), "%Y-%m")) %>%
  mutate(Fecha = as.yearmon(Fecha), "%Y-%m")

carpetas_21_22_mes <- aggregate(x = carpetas_21_22_mes[c("Robo a pasajero a bordo de pesero colectivo con violencia",
                                                       "Robo a pasajero a bordo de pesero colectivo sin violencia",
                                                       "Total")],
                                    FUN = sum,
                                    by = carpetas_21_22_mes["Fecha"])

carpetas_21_22_mes$Fecha <- as.Date(as.yearmon(carpetas_21_22_mes$Fecha), frac = 1)

```

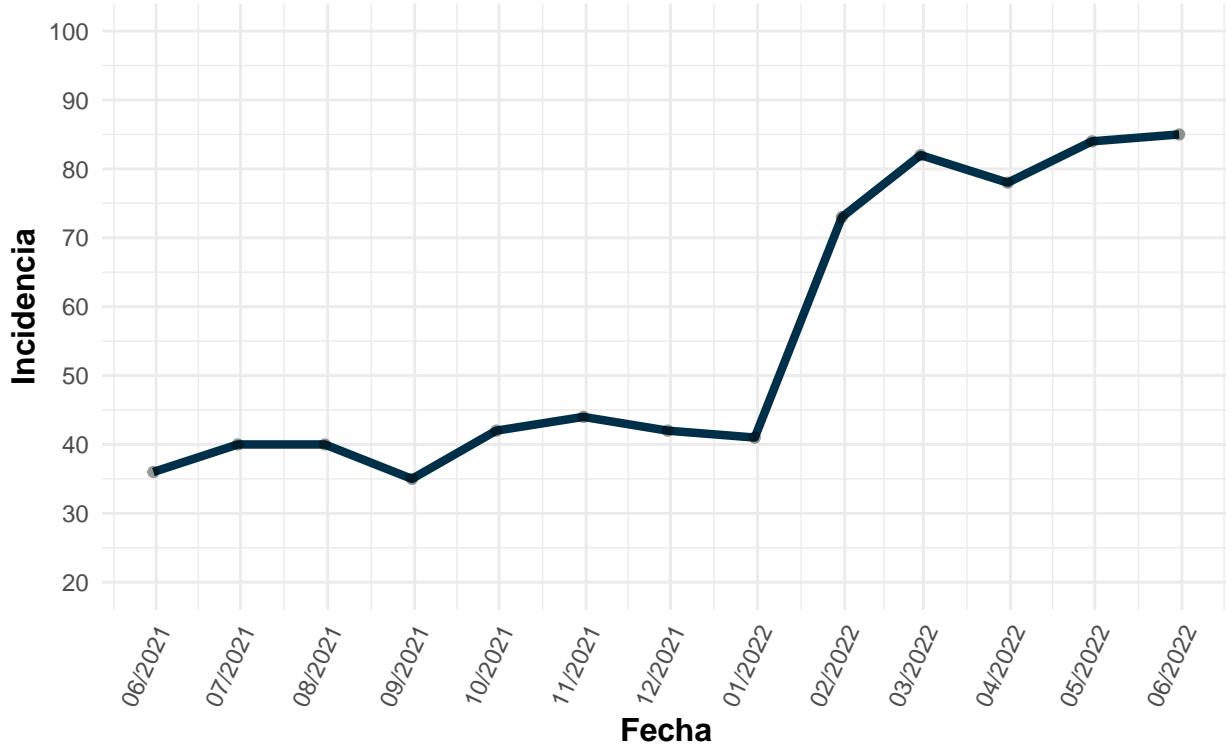
Lo cual puede visualizarse en la siguiente gráfica, en donde es claro que los robos a pasajeros al interior de microbuses han aumentado en la Ciudad de México a partir de enero de 2022. Dicha tendencia continuó en ascenso sustancial durante tres meses y luego se estabilizó.

```

carpetas_21_22_mes %>%
  ggplot() +
  geom_line(aes(Fecha, Total), size = 1.5, color = "#003049") +
  geom_point(aes(Fecha, Total), size = 1.5, color = "black", alpha = 0.4) +
  theme_minimal() +
  labs(
    title="Incidencia mensual de robo a pasajeros en microbús en la CDMX, \n May, 2021 - May, 2022",
    x = "Fecha", y = "Incidencia")+
  scale_x_date(date_breaks = "1 months", date_labels = "%m/%Y") +
  scale_y_continuous(limits = c(20, 100), breaks = seq(20, 100, by = 10)) +
  theme(plot.title = element_text(hjust = 0.5, size = 14, face="bold"),
        axis.title=element_text(size=12, face="bold"),
        axis.text.x = element_text(angle = 65, hjust = 1),
        plot.caption = element_text(hjust = 0))

```

## Incidencia mensual de robo a pasajeros en microbús en la CDMX, May, 2021 – May, 2022



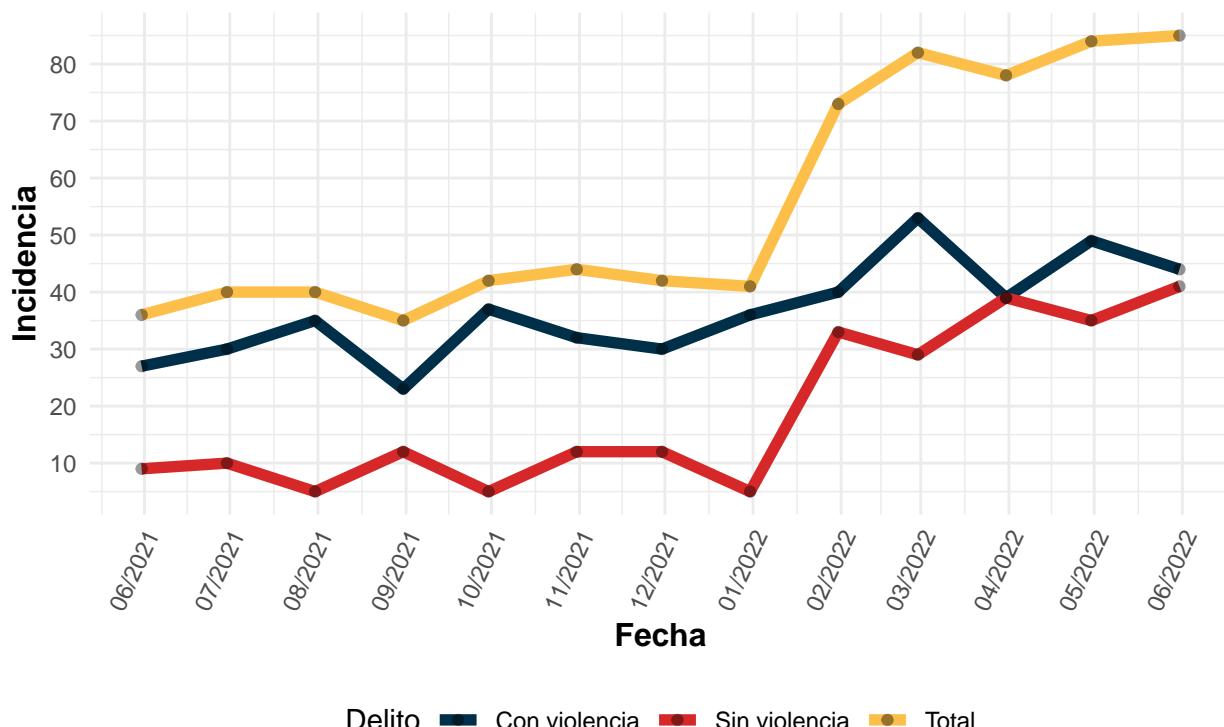
Como primer análisis, es pertinente visualizar la evolución de estos delitos en sus modalidades con y sin violencia, por separado:

```

carpetas_21_22_mes_long <- carpetas_21_22_mes %>%
  gather(Delito, Incidencia, -c(Fecha))

carpetas_21_22_mes_long %>%
  mutate(Delito = recode(Delito,
    "Robo a pasajero a bordo de pesero colectivo con violencia" =
      "Con violencia",
    "Robo a pasajero a bordo de pesero colectivo sin violencia" =
      "Sin violencia",
    "Total" = "Total")) %>%
ggplot(aes(x = Fecha, y = Incidencia, fill = Delito, color = Delito)) +
  geom_line(size = 2) +
  geom_point(size = 1.5, color = "black", alpha = 0.4) +
  labs(title =
    "Robos a pasajeros de microbús, con y sin violencia, \n May, 2021 – May, 2022") +
  scale_x_date(date_breaks = "1 months", date_labels = "%m/%Y") +
  scale_y_continuous(n.breaks = 7) +
  theme_minimal() +
  scale_colour_manual(values=c("#003049", "#d62828", "#fcbf49")) +
  theme(plot.title = element_text(hjust = 0.5, size = 14, face="bold", color = "black"),
    axis.title=element_text(size=12,face="bold"),
    plot.caption = element_text(hjust = 0),
    legend.position = "bottom",
    axis.text.x = element_text(angle = 65, hjust =1))
  
```

## Robos a pasajeros de microbús, con y sin violencia, May, 2021 – May, 2022



La gráfica de arriba sugiere que los robos en microbuses han aumentado significativamente, de manera primordial, por el ascenso de la incidencia de robos sin violencia. Si bien los robos con violencia han experimentado un aumento ligero desde octubre de 2021, aquellos sin violencia han ascendido mucho más dramáticamente.

Por otro lado, conviene visualizar si este incremento ha ocurrido de manera generalizada en toda la capital del país, o bien, si existen alcaldías en donde ha sido más explícito. A continuación, se presentan la incidencia delictiva reportada por alcaldía y mes:

```

carpetas_21_22$alcaldia_hechos <- tolower(carpetas_21_22$alcaldia_hechos)
carpetas_21_22$alcaldia_hechos <- str_to_title(carpetas_21_22$alcaldia_hechos)

carpetas_21_22$fecha_hechos <- as.yearmon(carpetas_21_22$fecha_hechos)

carpetas_21_22_alcaldia <- count(carpetas_21_22, c("fecha_hechos", "alcaldia_hechos"))

carpetas_21_22_alcaldia <- transform(carpetas_21_22_alcaldia,
                                       Fecha = as.Date(fecha_hechos, frac = 1))
carpetas_21_22_alcaldia <- dplyr::rename(carpetas_21_22_alcaldia,
                                           "Alcaldía" = "alcaldia_hechos",
                                           "Incidencia" = "freq")

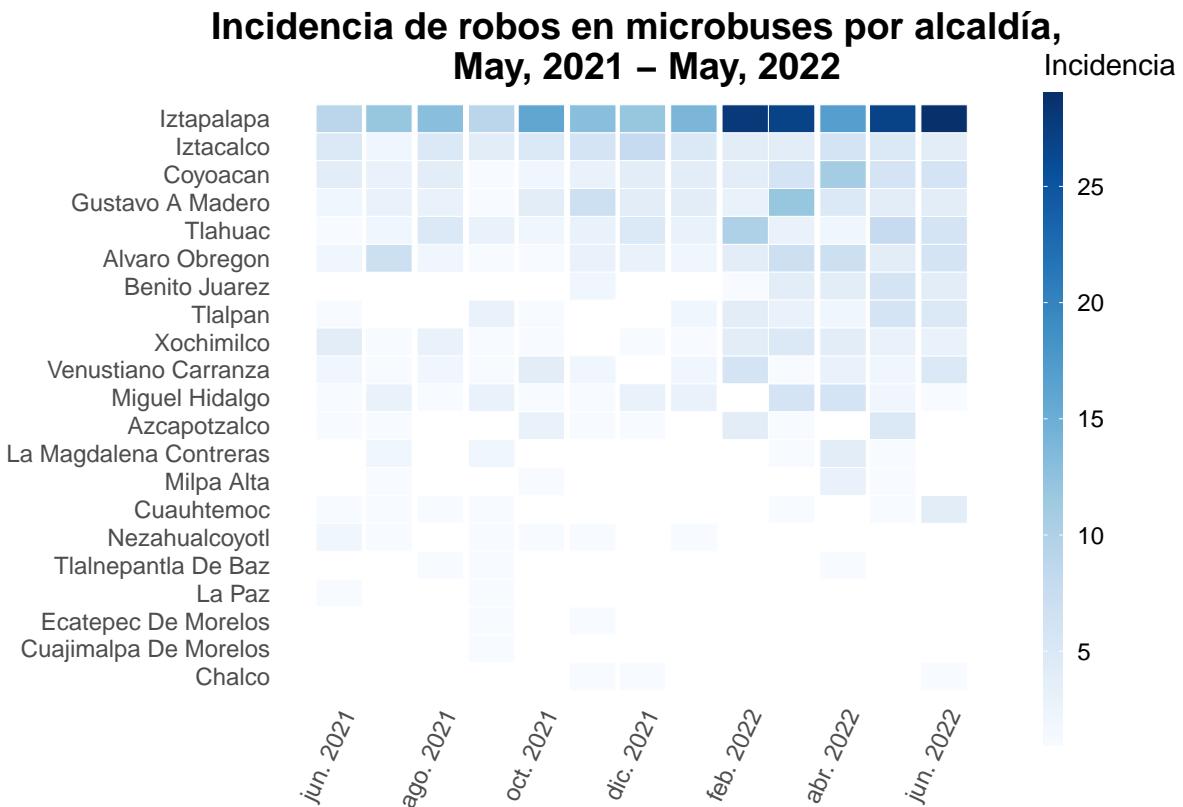
carpetas_21_22_alcaldia %>%
  na.omit() %>%
  ggplot(aes(x = Fecha, y = reorder(Alcaldía, Incidencia), fill = Incidencia)) +
  geom_tile(color = "white", lwd = 0, linetype = 1) +

```

```

scale_fill_gradientn(colors = RColorBrewer::brewer.pal(9, "Blues"),
                     breaks = c(0, 5, 10, 15, 20, 25)) +
  theme_minimal() +
  scale_x_date(date_breaks = "2 months", date_labels = "%b%Y") +
  theme(panel.grid = element_blank()) +
  guides(fill = guide_colourbar(barwidth = 0.5, barheight = 17)) +
  labs(title = "Incidencia de robos en microbuses por alcaldía, \n May, 2021 – May, 2022",
       x = "", y = "") +
  theme(plot.title = element_text(hjust = 0.5, size = 14, face="bold", color = "black"),
        axis.text.x = element_text(angle = 65, hjust =1))

```



Donde es claro que la alcaldía Iztapalapa es la que ha experimentado un alza más pronunciada en los robos en microbuses. De hecho, lo anterior sugiere que el incremento en la Ciudad de México responde mayoritariamente al comportamiento de estos delitos en dicha alcaldía. Aunque de manera más sutil, Iztacalco, Coyoacán y Gustavo A. Madero también exhiben incrementos ligeros desde el inicio de 2022.

La figura de abajo ilustra las colonias con más reportes de robos a microbuses en el periodo estudiado. A partir de esto, se sugiere realizar un análisis más exhaustivo sobre las características sociodemográficas en dichas colonias, las cuales podrían arrojar luz sobre las causas de una incidencia delictiva tan elevada.

```

top_10_colonias <- carpetas_21_22 %>%
  group_by(colonia_hechos) %>%
  dplyr::summarize(Incidencia = n()) %>%
  arrange(desc(Incidencia)) %>%
  top_n(10)

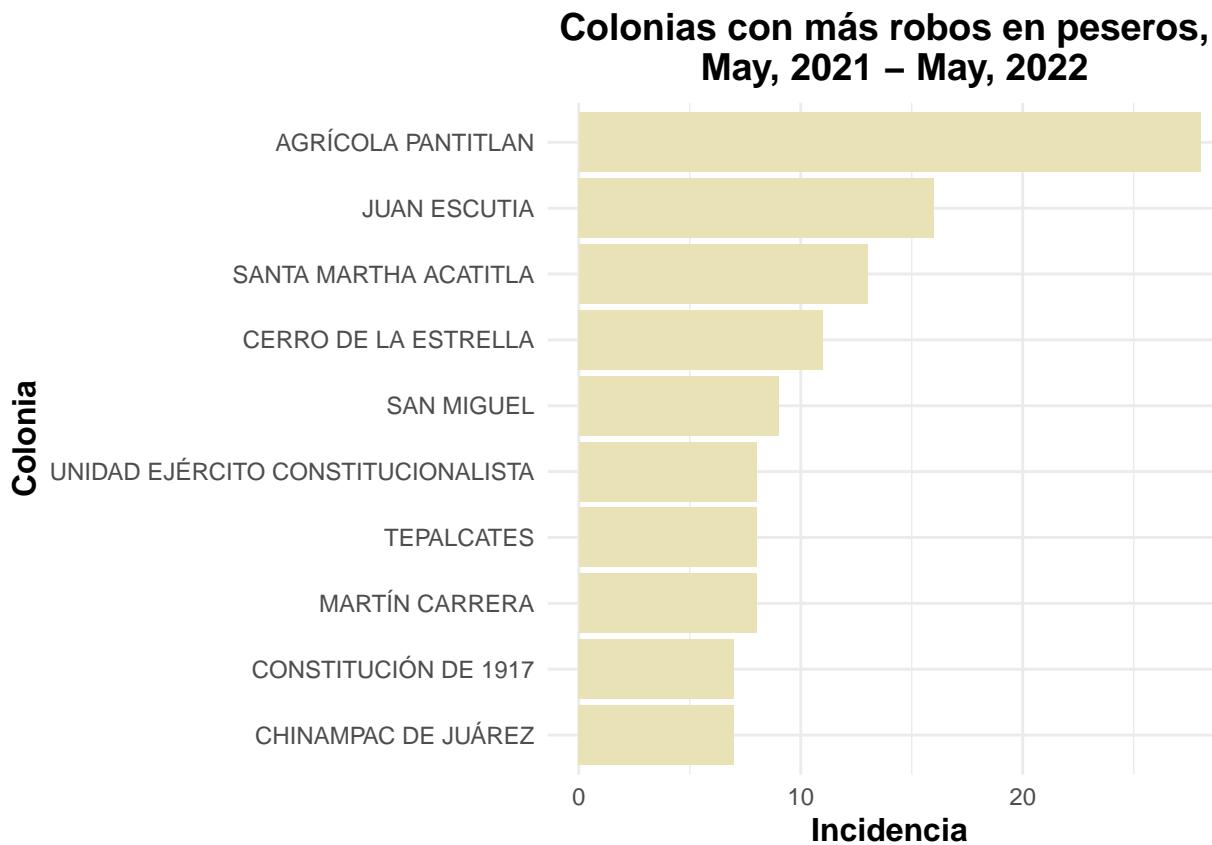
## Selecting by Incidencia

```

```

top_10_colonias %>%
  filter(Incidencia != 67) %>%
  ggplot(aes(x = reorder(colonia_hechos, Incidencia), y = Incidencia)) +
  geom_bar(stat = "identity", position = "dodge", fill = "#eae2b7") +
  labs(title = "Colonias con más robos en peseros, \n May, 2021 – May, 2022",
       y = "Incidencia", x = "Colonia") +
  theme_minimal() +
  coord_flip() +
  theme(plot.title = element_text(hjust = 0.5, size = 14, face = "bold", color = "black"),
        axis.title = element_text(size = 12, face = "bold"),
        plot.caption = element_text(hjust = 0),
        legend.position = "bottom")

```



Finalmente, es pertinente explorar si las denuncias de los robos en microbuses se realizaron en la misma alcaldía en la cual ocurrieron los hechos, toda vez que esto podría sugerir hipótesis favorables o desfavorables sobre la actuación, capacidad y eficiencia de las Fiscalías en las alcaldías.

```

carpetas_21_22$fiscalia <- tolower(carpetas_21_22$fiscalia)
carpetas_21_22$fiscalia <- str_to_title(carpetas_21_22$fiscalia)

carpetas_21_22$alcaldia_hechos[carpetas_21_22$alcaldia_hechos == "Coyoacan"] <- "Coyoacán"
carpetas_21_22$alcaldia_hechos[carpetas_21_22$alcaldia_hechos == "Cuauhtemoc"] <- "Cuauhtémoc"
carpetas_21_22$alcaldia_hechos[carpetas_21_22$alcaldia_hechos == "Gustavo A Madero"] <- "Gustavo A. Madero"
carpetas_21_22$alcaldia_hechos[carpetas_21_22$alcaldia_hechos == "Alvaro Obregon"] <- "Alvaro Obregón"
carpetas_21_22$alcaldia_hechos[carpetas_21_22$alcaldia_hechos == "Tlahuac"] <- "Tláhuac"
carpetas_21_22$alcaldia_hechos[carpetas_21_22$alcaldia_hechos == "Benito Juarez"] <- "Benito Juárez"

```

```

carpetas_21_22$alcaldia_hechos[carpetas_21_22$alcaldia_hechos == "La Magdalena Contreras"] <- "Magdeleno"

i <- sapply(seq_along(carpetas_21_22$alcaldia_hechos),
            function(i) grepl(carpetas_21_22$alcaldia_hechos[i], carpetas_21_22$fiscalia[i]))

carpetas_21_22$Flag <- c(0, 1)[i + 1L]

table(carpetas_21_22$Flag)

##
##      0      1
## 270 437

prop.table(table(carpetas_21_22$Flag))

##
##          0           1
## 0.3818953 0.6181047

```

De las 707 denuncias por robo a pasajeros en microbús en el periodo estudiado, solo 437 fueron registradas en la misma alcaldía en donde ocurrieron los hechos (31.8 %), mientras que 270 en otra fiscalía o instancia de investigación (38.2 %). De manera más específica, Gustavo A. Madero, Azcapotzalco y Venustiano Carranza son las alcaldías donde un mayor porcentaje de los robos a microbuses se reportan ahí mismo (82, 76 y 74 por ciento, respectivamente). Existen alcaldías en donde ningún delito se reportó ahí mismo, tal como muestra la siguiente tabla:

```

class(carpetas_21_22$Flag)

## [1] "numeric"

carpetas_21_22 %>%
  group_by(alcaldia_hechos) %>%
  summarise_at(vars(Flag), list(name = mean)) %>%
  kbl(caption = "Porcentaje de robos reportados en la misma alcaldía que la ocurrencia del hecho",
       col.names = c("Alcaldía", "Porcentaje")) %>%
  kable_classic_2(full_width = F)

```

Cuadro 1: Porcentaje de robos reportados en la misma alcaldía que la ocurrencia del hecho

Alcaldía	Porcentaje
Alvaro Obregón	0.7346939
Azcapotzalco	0.7647059
Benito Juárez	0.3809524
Chalco	0.0000000
Coyoacán	0.6034483
Cuajimalpa De Morelos	0.0000000
Cuauhtémoc	0.3000000
Ecatepec De Morelos	0.0000000
Gustavo A. Madero	0.8214286
Iztacalco	0.6349206
Iztapalapa	0.6858407
La Paz	0.0000000
Magdalena Contreras	0.0000000
Miguel Hidalgo	0.6451613
Milpa Alta	0.5000000
Nezahualcoyotl	0.0000000
Tláhuac	0.5660377
Tlalnepantla De Baz	0.0000000
Tlalpan	0.3333333
Venustiano Carranza	0.7419355
Xochimilco	0.5161290
NA	NA

## Ejercicio 2

De la página de datos abiertos de la Ciudad de México, consultar los datos de víctimas en carpetas de investigación y realizar lo siguiente:

- a. Obtener el total de víctimas por homicidio doloso por alcaldía en 2021. Realizar una gráfica de línea para cada alcaldía para poder observar las tendencias.

Para las siguientes gráficas, se hizo uso de los datos de la Fiscalía General de Justicia de la CDMX; específicamente, la información que contiene datos sobre las víctimas de los delitos registrados por esta instancia. En general, el siguiente análisis exploratorio ilustra que alcaldías como Azcapotzalco, Tláhuac y Xochimilco han experimentado incrementos en el número de homicidios dolosos registrados. El resto de las delimitaciones territoriales han mantenido o, incluso, disminuido la incidencia de este delito.

```
victimas <- read.csv("~/R/Prueba ADIP/datos/victimas_completa_mayo_2022.csv",
  encoding="UTF-8")

victimas$FechaInicio <- as.Date(victimas$FechaInicio, "%d/%m/%Y")
victimas$FechaHecho <- as.Date(victimas$FechaHecho, "%d/%m/%Y")

# glimpse(carpetas)

### 

victimas_21_22 <- victimas %>%
  filter(FechaHecho >= fecha1 & FechaHecho <= fecha2)
```

```

victimas_21_22 <- victimas_21_22 %>%
  filter(Categoría == "HOMICIDIO DOLOSO")

victimas_21_22$Categoría <- tolower(victimas_21_22$Categoría)
victimas_21_22$Categoría <- firstup(victimas_21_22$Categoría)
victimas_21_22$AlcaldiaHechos <- tolower(victimas_21_22$AlcaldiaHechos)
victimas_21_22$AlcaldiaHechos <- str_to_title(victimas_21_22$AlcaldiaHechos)

victimas_21_22_mes <- tabyl(victimas_21_22, FechaHecho, AlcaldiaHechos)

victimas_21_22_mes$mes <- floor_date(victimas_21_22_mes$FechaHecho, "month")

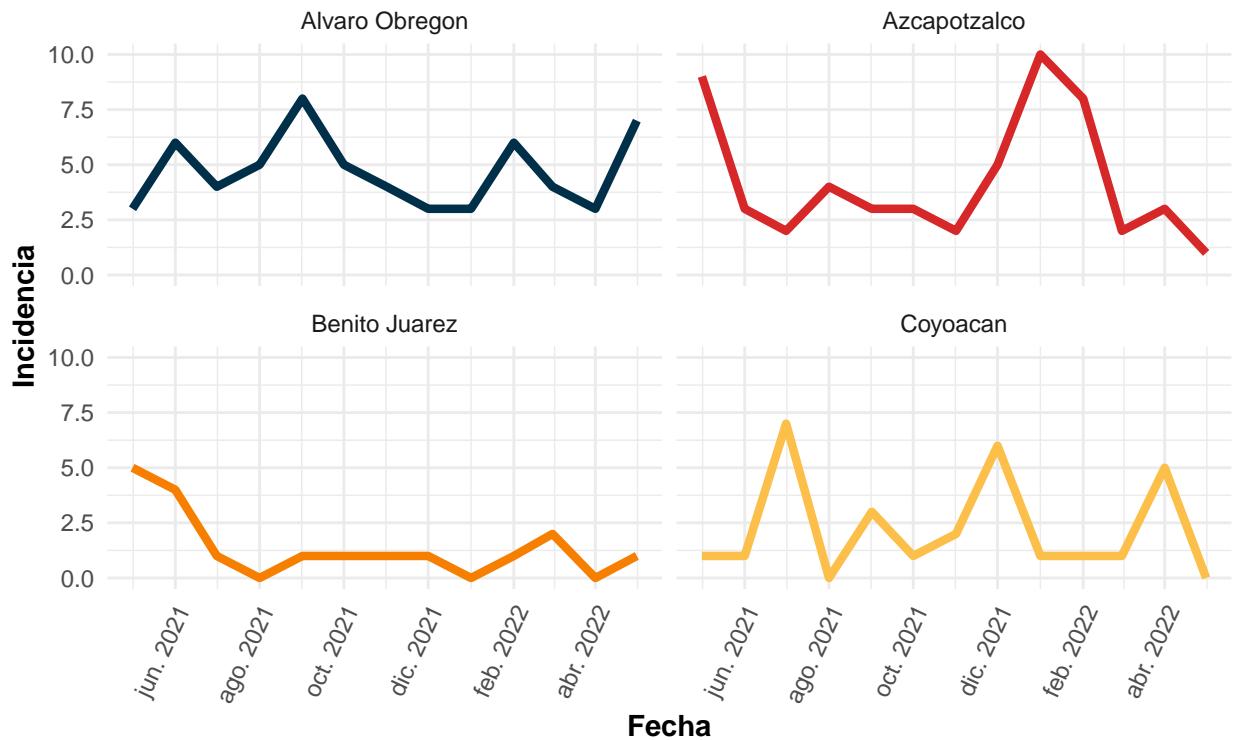
victimas_21_22_mes <- aggregate(x = victimas_21_22_mes[c("Alvaro Obregon", "Azcapotzalco",
  "Benito Juarez", "Coyoacan",
  "Cuajimalpa De Morelos", "Cuauhtemoc",
  "Gustavo A Madero", "Iztacalco",
  "Iztapalapa", "La Magdalena Contreras",
  "Miguel Hidalgo", "Milpa Alta",
  "Tlahuac", "Tlalpan",
  "Venustiano Carranza",
  "Xochimilco")], FUN = sum, by = victimas_21_22)

victimas_21_22_mes_long <- melt(victimas_21_22_mes, id.vars = "mes")

victimas_21_22_mes_long %>%
  filter(variable %in% c("Alvaro Obregon", "Azcapotzalco", "Benito Juarez", "Coyoacan")) %>%
  ggplot(aes(mes, y = value, color = variable)) +
  geom_line(size = 1.5) +
  facet_wrap(~variable) +
  theme_minimal() +
  scale_x_date(date_breaks = "2 months", date_labels = "%b%Y") +
  theme(legend.position = "none", plot.title = element_text(hjust = 0.5,
    size = 14,
    face = "bold"),
    axis.text.x = element_text(angle = 65, hjust = 1),
    axis.title.x = element_text(face = "bold"),
    axis.title.y = element_text(face = "bold")) +
  scale_colour_manual(values=c("#003049", "#d62828", "#f77f00", "#fcbf49")) +
  labs(title = "Homicidios dolosos por alcaldía, \n May, 2021 - May, 2022",
    x = "Fecha",
    y = "Incidencia")

```

## Homicidios dolosos por alcaldía, May, 2021 – May, 2022

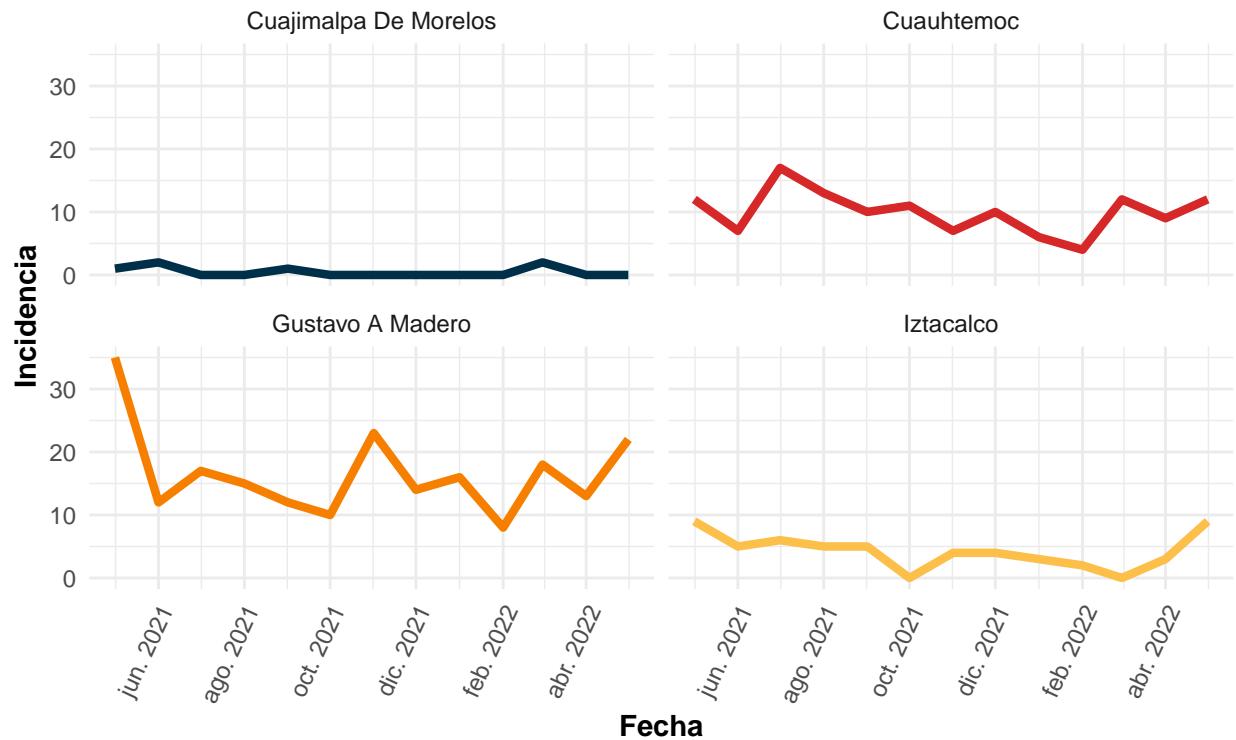


```

victimas_21_22_mes_long %>%
  filter(variable %in% c("Cuajimalpa De Morelos", "Cuauhtemoc", "Gustavo A Madero",
                         "Iztacalco")) %>%
  ggplot(aes(mes, y = value, color = variable)) +
  geom_line(size = 1.5) +
  facet_wrap(~variable) +
  theme_minimal() +
  scale_x_date(date_breaks = "2 months", date_labels = "%b%Y") +
  theme(legend.position = "none", plot.title = element_text(hjust = 0.5,
                                                          size = 14,
                                                          face = "bold"),
        axis.text.x = element_text(angle = 65, hjust = 1),
        axis.title.x = element_text(face = "bold"),
        axis.title.y = element_text(face = "bold")) +
  scale_colour_manual(values=c("#003049", "#d62828", "#f77f00", "#fcbf49")) +
  labs(title = "Homicidios dolosos por alcaldía, \n May, 2021 – May, 2022",
       x = "Fecha", y = "Incidencia")

```

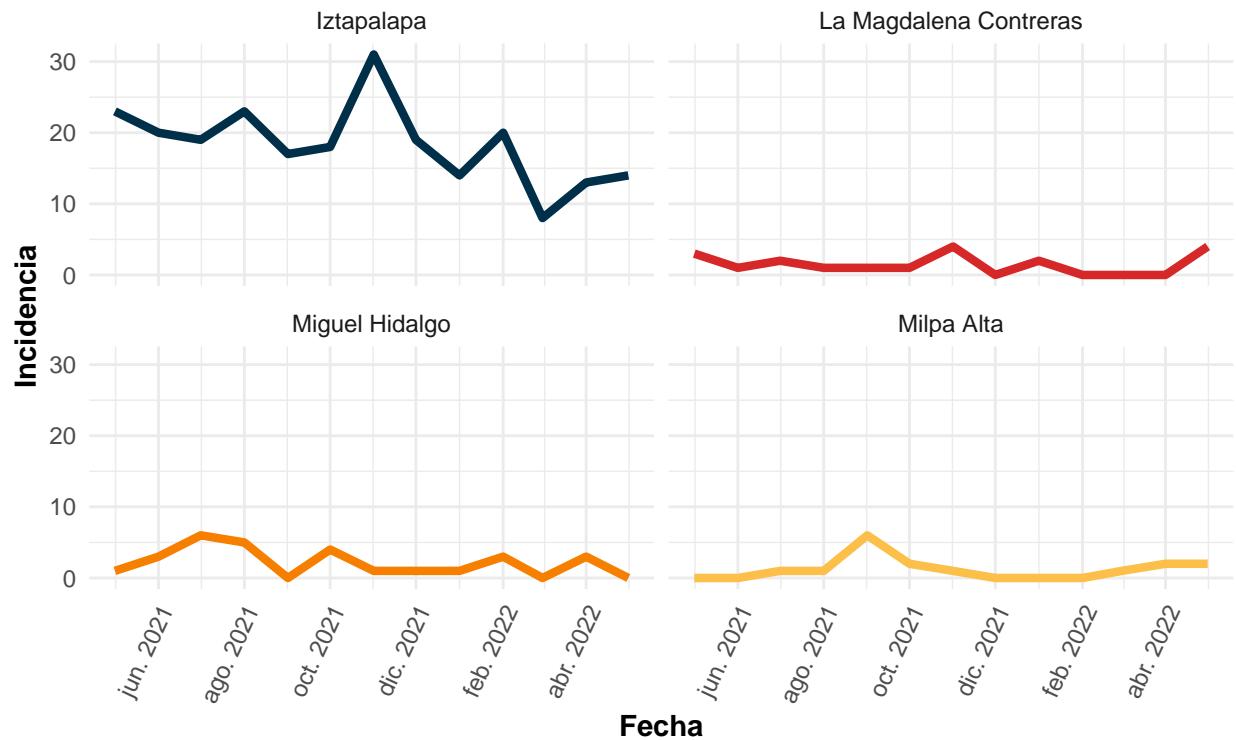
## Homicidios dolosos por alcaldía, May, 2021 – May, 2022



```

victimas_21_22_mes_long %>%
  filter(variable %in% c("Iztapalapa", "La Magdalena Contreras",
    "Miguel Hidalgo", "Milpa Alta")) %>%
  ggplot(aes(mes, y = value, color = variable)) +
  geom_line(size = 1.5) +
  facet_wrap(~variable) +
  theme_minimal() +
  scale_x_date(date_breaks = "2 months", date_labels = "%b%Y") +
  theme(legend.position = "none", plot.title = element_text(hjust = 0.5,
    size = 14,
    face = "bold"),
    axis.text.x = element_text(angle = 65, hjust = 1),
    axis.title.x = element_text(face = "bold"),
    axis.title.y = element_text(face = "bold")) +
  scale_colour_manual(values=c("#003049", "#d62828", "#f77f00", "#fcbf49")) +
  labs(title = "Homicidios dolosos por alcaldía, \n May, 2021 – May, 2022",
    x = "Fecha", y = "Incidencia")
  
```

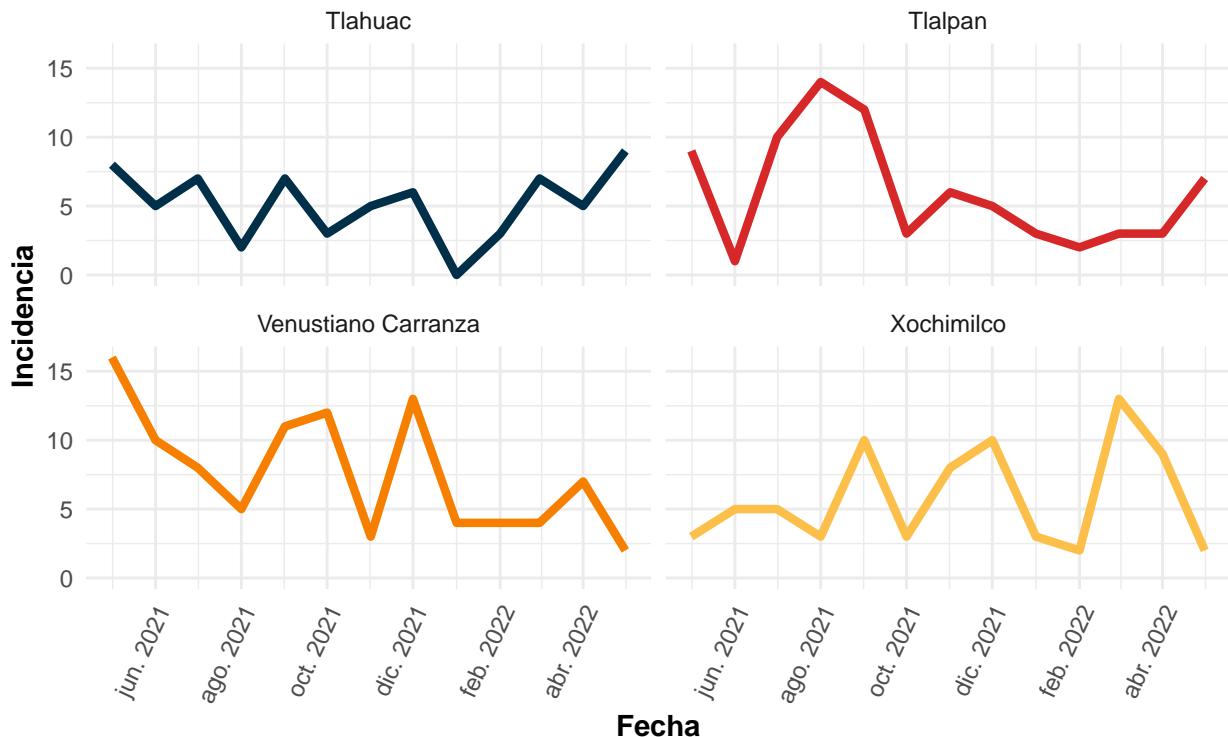
## Homicidios dolosos por alcaldía, May, 2021 – May, 2022



```

victimas_21_22_mes_long %>%
  filter(variable %in% c("Tlalpan", "Tlalpan", "Venustiano Carranza", "Xochimilco")) %>%
  ggplot(aes(mes, y = value, color = variable)) +
  geom_line(size = 1.5) +
  facet_wrap(~variable) +
  theme_minimal() +
  scale_x_date(date_breaks = "2 months", date_labels = "%b%Y") +
  theme(legend.position = "none", plot.title = element_text(hjust = 0.5,
    size = 14,
    face = "bold"),
    axis.text.x = element_text(angle = 65, hjust = 1),
    axis.title.x = element_text(face = "bold"),
    axis.title.y = element_text(face = "bold")) +
  scale_colour_manual(values=c("#003049", "#d62828", "#f77f00", "#fcbf49")) +
  labs(title = "Homicidios dolosos por alcaldía, \n May, 2021 – May, 2022",
    x = "Fecha", y = "Incidencia")
  
```

## Homicidios dolosos por alcaldía, May, 2021 – May, 2022



- b. Con los datos de víctimas por homicidio doloso de 2021, realizar una pirámide poblacional para observar en qué grupos de edad se concentran los casos por sexo.

La figura de abajo exhibe, en primer lugar, que los hombre son mucho más propensos a ser víctimas de un homicidio doloso en el periodo estudiado. Además, en cualquiera de los dos sexos, la mayor concentración de víctimas se encuentra en un rango de entre 20 y 50 años. Nótese que la distribución es ligeramente más alta en edades más bajas dentro de este intervalo, y disminuye conforme las víctimas se acercan a los 50 y 60 años.

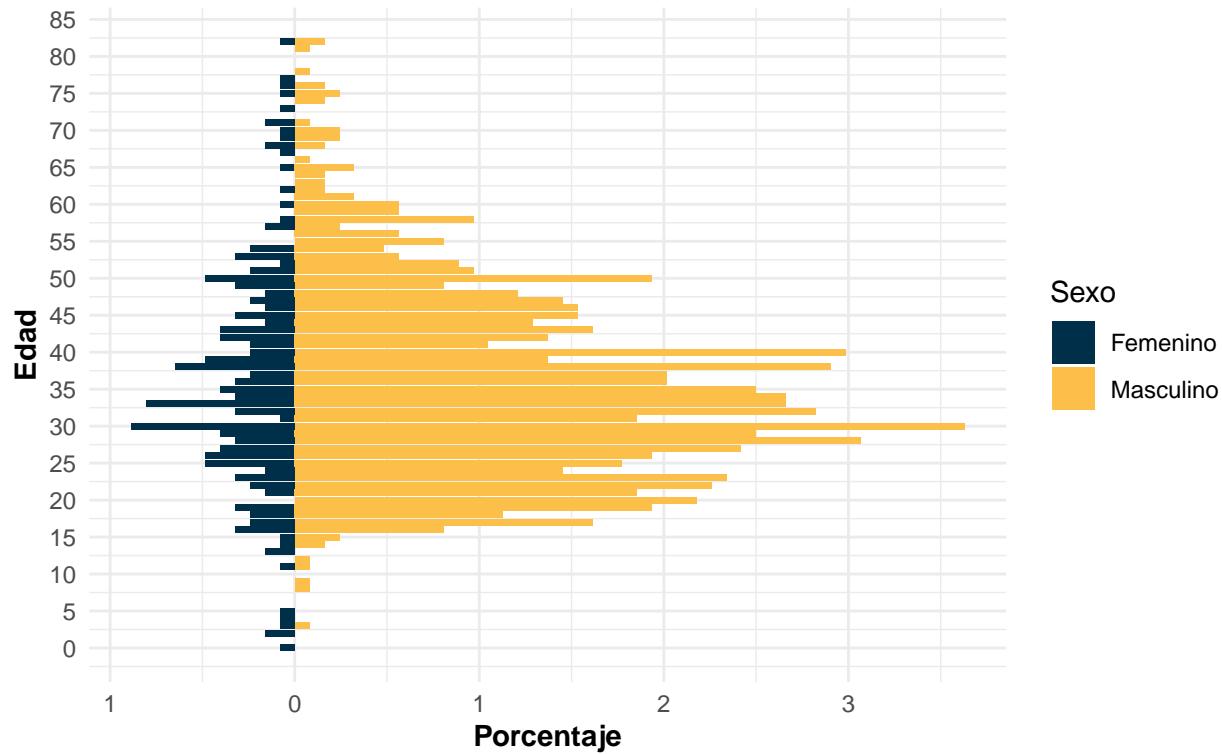
```
piramide <- victimas_21_22 %>%
  count(c("Sexo", "Edad"))

piramide$freq <- piramide$freq / sum(piramide$freq) * 100

ggplot(piramide, aes(x = Edad, fill = Sexo, y = ifelse(test = Sexo == "Femenino",
                                                       yes = -freq, no = freq))) +
  geom_bar(stat = "identity") +
  scale_y_continuous(labels = abs) +
  coord_flip() +
  theme_minimal() +
  scale_x_continuous(breaks = seq(0, 85, by = 5)) +
  theme(plot.title = element_text(hjust = 0.5, size = 14, face = "bold"),
        axis.title.x = element_text(face = "bold"),
        axis.title.y = element_text(face = "bold"),
        legend.position = "right") +
  scale_fill_manual(values=c("#003049", "#fcbf49")) +
  labs(title = "Pirámide poblacional de homicidios dolosos, \n May, 2021 – May, 2022",
       x = "Edad", y = "Porcentaje")
```

```
## Warning: Removed 4 rows containing missing values (position_stack).
```

## Pirámide poblacional de homicidios dolosos, May, 2021 – May, 2022



### Ejercicio 3

De la Encuesta Nacional de Ingresos y Gastos de los Hogares realizar lo siguiente:

```
url <- "https://inegi.org.mx/contenidos/programas/enigh/nc/2020/micrdatos/enigh2020_ns_viviendas_csv.zip"

directorio <- tempdir()

archivo = tempfile(tmpdir = directorio, fileext = ".zip")
download.file(url, archivo)

unzip(archivo, files = "viviendas.csv", exdir = directorio, overwrite = TRUE)

ruta = file.path(directorio, "viviendas.csv")
unlink(directorio)

enigh <- read.csv(ruta) %>%
  rename(folioviv = 1)

enigh <- enigh %>%
  mutate(clave_entidad = case_when(nchar(folioviv) == 9 ~ substr(folioviv, 1, 1),
                                    TRUE ~ substr(folioviv, 1, 2)),
        entidad= case_when(clave_entidad==1 ~ "Aguascalientes",
                           clave_entidad==2 ~ "Baja California",
```

```

clave_entidad==3 ~ "Baja California Sur",
clave_entidad==4 ~ "Campeche",
clave_entidad==5 ~ "Coahuila de Zaragoza",
clave_entidad==6 ~ "Colima",
clave_entidad==7 ~ "Chiapas",
clave_entidad==8 ~ "Chihuahua",
clave_entidad==9 ~ "Ciudad de México",
clave_entidad==10 ~ "Durango",
clave_entidad==11 ~ "Guanajuato",
clave_entidad==12 ~ "Guerrero",
clave_entidad==13 ~ "Hidalgo",
clave_entidad==14 ~ "Jalisco",
clave_entidad==15 ~ "México",
clave_entidad==16 ~ "Michoacán de Ocampo",
clave_entidad==17 ~ "Morelos",
clave_entidad==18 ~ "Nayarit",
clave_entidad==19 ~ "Nuevo León",
clave_entidad==20 ~ "Oaxaca",
clave_entidad==21 ~ "Puebla",
clave_entidad==22 ~ "Querétaro",
clave_entidad==23 ~ "Quintana Roo",
clave_entidad==24 ~ "San Luis Potosí",
clave_entidad==25 ~ "Sinaloa",
clave_entidad==26 ~ "Sonora",
clave_entidad==27 ~ "Tabasco",
clave_entidad==28 ~ "Tamaulipas",
clave_entidad==29 ~ "Tlaxcala",
clave_entidad==30 ~ "Veracruz de Ignacio de la Llave",
clave_entidad==31 ~ "Yucatán",
clave_entidad==32 ~ "Zacatecas"))

```

- Obtener el número de viviendas particulares habitadas por entidad federativa.
- Obtener el coeficiente de variación de las estimaciones del inciso a) Comentar si, de acuerdo con esta medida, las estimaciones son estadísticamente significativas o no.

De acuerdo a los resultados estimados en la siguiente tabla, puede observarse que todas las estimaciones de las viviendas ocupadas en las entidades federativas del país exhiben un coeficiente de variación relativamente bajo entre 1.06 y 2.11. Esto sugiere que el número de viviendas ocupadas estimado es estadísticamente significativo en niveles elevados para todos los casos, lo cual implica una estimación sensata de esta medida.

```

diseño_muestral <- enigh %>%
  as_survey_design(ids = upm, strata = est_dis, weights = factor)

vivnal <- diseño_muestral %>%
  group_by(entidad) %>%
  summarise(viv = survey_total(vartype = c("cv", "ci", "cl")), level = 0.95)) %>%
  mutate(viv_cv = viv_cv *100)

vivnal %>%
  kable(caption = "Viviendas ocupadas estimadas, coeficiente de variación y otras estadísticas",
        col.names = c("Entidad", "Viviendas ocupadas",
                     "Coeficiente de variación",
                     "Límite inferior",

```

Cuadro 2: Viviendas ocupadas estimadas, coeficiente de variación y otras estadísticas

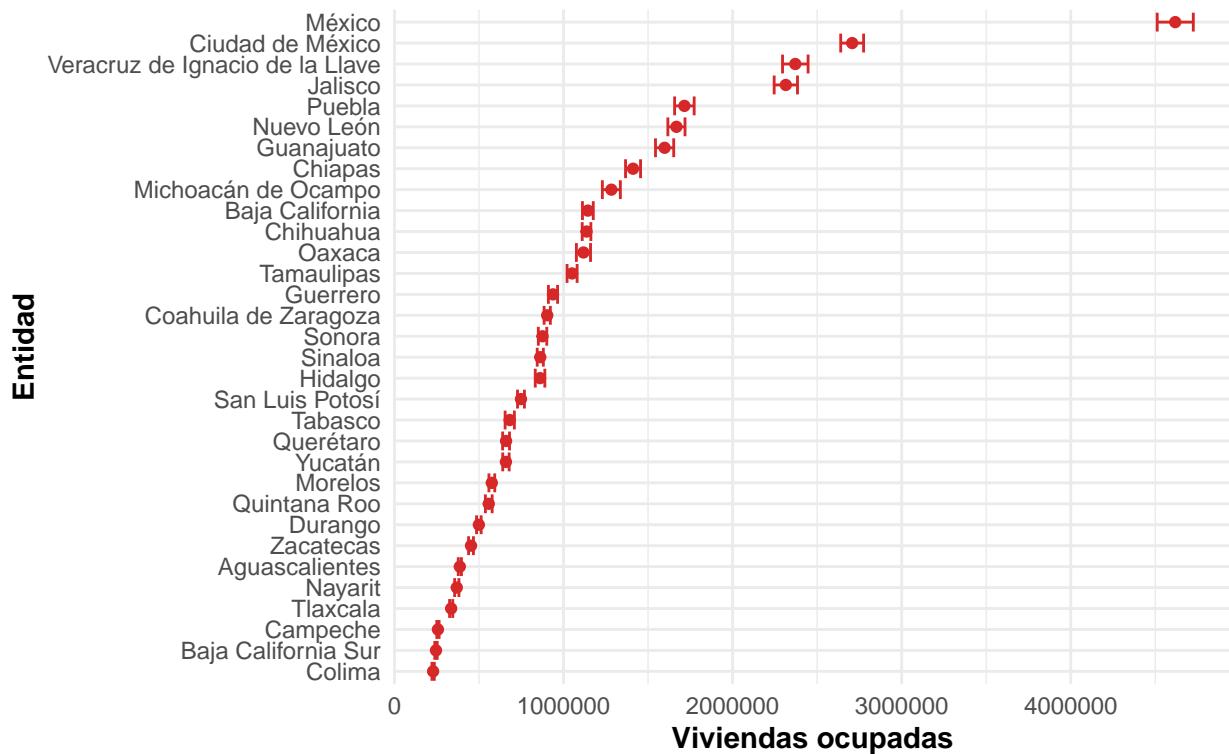
Entidad	Viviendas ocupadas	Coeficiente de variación	Límite inferior	Límite superior
Aguascalientes	386782	1.232079	377440.7	396123.3
Baja California	1144066	1.427566	1112051.2	1176080.8
Baja California Sur	245689	1.485768	238533.5	252844.5
Campeche	257257	1.343160	250483.7	264030.3
Chiapas	1411266	1.606430	1366826.0	1455706.0
Chihuahua	1136150	1.145227	1110644.7	1161655.3
Ciudad de México	2707156	1.273969	2639551.6	2774760.4
Coahuila de Zaragoza	903683	1.040112	885258.3	922107.7
Colima	228765	1.555352	221790.4	235739.6
Durango	499269	1.379436	485768.8	512769.2
Guanajuato	1597797	1.718720	1543966.3	1651627.7
Guerrero	937809	1.483251	910542.3	965075.7
Hidalgo	861222	1.686024	832758.9	889685.1
Jalisco	2314882	1.521004	2245864.0	2383900.0
México	4618111	1.179818	4511308.2	4724913.8
Michoacán de Ocampo	1282994	2.113429	1229842.5	1336145.5
Morelos	576022	1.525767	558794.2	593249.8
Nayarit	368748	1.737586	356188.3	381307.7
Nuevo León	1667880	1.549988	1617204.7	1718555.3
Oaxaca	1118022	1.907791	1076211.6	1159832.4
Puebla	1714877	1.718370	1657113.6	1772640.4
Querétaro	660207	1.571011	639875.8	680538.2
Quintana Roo	557830	1.822686	537899.6	577760.4
San Luis Potosí	748392	1.393653	727947.0	768837.0
Sinaloa	862532	1.069485	844449.7	880614.3
Sonora	875959	1.457195	850938.0	900980.0
Tabasco	681950	2.047947	654573.7	709326.3
Tamaulipas	1050737	1.445884	1020956.6	1080517.4
Tlaxcala	335517	1.408763	326251.8	344782.2
Veracruz de Ignacio de la Llave	2370865	1.625039	2295343.0	2446387.0
Yucatán	659472	1.465583	640526.3	678417.7
Zacatecas	452914	1.611868	438603.7	467224.3

```

    "Límite superior")) %>%
  kable_classic(full_width = F)

vivnal %>%
  ggplot(aes(reorder(entidad, viv), viv)) +
  geom_point(color = "#d62828") +
  geom_errorbar(aes(ymin = viv_low, ymax = viv_upp), color = "#d62828") +
  coord_flip() +
  theme_minimal() +
  theme(plot.title = element_text(hjust = 0.5, size = 14, face = "bold"),
        axis.title.x = element_text(face = "bold"), axis.title.y = element_text(face = "bold")) +
  scale_fill_manual(values=c("#003049", "#fcbf49")) +
  labs(title = "Estimación de viviendas ocupadas \n por entidad federativa", x = "Entidad", y = "Viviendas
  
```

## Estimación de viviendas ocupadas por entidad federativa



### Ejercicio 4

Suponga que se tiene como prioridad identificar aquellos predios con edificaciones cuyo año de construcción es menor o igual a 1975 para poder diseñar una estrategia de renovación del parque inmobiliario. ¿Qué datos ocuparía? ¿En qué alcaldías se concentra el mayor número de edificaciones de este tipo? ¿A qué dependencias del Gobierno de la Ciudad le presentaría este reporte y por qué?

```
require("knitr")
## Loading required package: knitr
opts_knit$set(root.dir = "C:/Users/marco/OneDrive/Documents/R/Prueba ADIP/")
myfiles <- list.files(pattern = "\\.csv$")
CATASTRO <- read.csv(myfiles[1])
for(f in myfiles[-1]){
  catastro <- read.csv(f)
  CATASTRO <- rbind(CATASTRO, catastro)
}
write.csv(CATASTRO, "catastro_cdmx.csv", quote = FALSE, row.names = FALSE)
```

Con los datos del Sistema Abierto de Información Geográfica de la Ciudad de México, se calculó el porcentaje de viviendas construidas antes de 1975 por alcaldía. Puede observarse que Cuauhtémoc, Venustiano Carranza

Cuadro 3: Viviendas construidas antes de 1975 por alcaldía

Alcaldía	Total de viviendas	Construidas antes de 1975	% antes de 1975
ALVARO OBREGON	172875	25245	14.603037
AZCAPOTZALCO	115932	35594	30.702481
BENITO JUAREZ	243019	71987	29.621964
COYOACAN	167629	32270	19.250846
CUAJIMALPA DE MORELOS	65870	2345	3.560043
CUAUHTEMOC	255491	103410	40.475007
GUSTAVO A. MADERO	250430	59007	23.562273
IZTACALCO	101604	24190	23.808118
IZTAPALAPA	370846	48132	12.978972
MAGDALENA CONTRERAS	39399	5082	12.898805
MIGUEL HIDALGO	185549	45698	24.628535
MILPA ALTA	10206	2208	21.634333
TLAHUAC	67725	3239	4.782577
TLALPAN	142903	11977	8.381210
VENUSTIANO CARRANZA	122053	43942	36.002392
XOCHIMILCO	59416	9997	16.825434

y Azcapotzalco son las demarcaciones con mayor porcentaje de viviendas en esta condición. Mientras que Tlalpan, Tláhuac y Cuajimalpa las que menos.

```
CATASTRO <- na.omit(CATASTRO)
```

```
CATASTRO %>%
  group_by(alcaldia_cumplimiento) %>%
  dplyr::summarize(Freq = n(),
    Freq_año = sum(anio_construccion <= 1975)) %>%
  mutate(Pct = Freq_año/Freq*100) %>%
  kable(caption = "Viviendas construidas antes de 1975 por alcaldía",
    col.names = c("Alcaldía", "Total de viviendas", "Construidas antes de 1975", "% antes de 1975"))
  kable_classic(full_width = F)
```

## Ejercicio opcional

Con los datos de víctimas por homicidio doloso de 2021 realizar un mapa de puntos.

Este ejercicio se realizó en QGIS utilizando datos shape sobre la Ciudad de México y las carpetas de investigación de la FGJ debidamente georreferenciados. Además, se incluye una *standard deviational ellipse*, que ilustra la distribución geoespacial de los datos de punto, indicando que la posición, dirección, inclinación y tamaño en donde se concentra un mayor número de homicidios dolosos en la capital del país.

**Homicidios dolosos reportados en la CDMX y su distribución espacial, 2021.**

