

La evolución del número delitos en Chile y su relación con la percepción subjetiva del aumento de delincuencia en Chile.

Rafael Bergoeing

2024-12-19

Creación y limpieza de las bases de datos

A continuación documentaremos la creación de las base de datos que utilizaremos para estudiar la relación entre el número de delitos mensuales en Chile y la percepción de los habitantes respecto a esta como ha evolucionado esta variable.

El objetivo de este código es construir una base de datos a nivel comunal que nos entregue, para cada comuna:

- número de delitos mensuales entre 2021 y 2023
- población entre 2021 y 2023
- indicadores socioeconómicos comunales
- indicadores de percepción de inseguridad (de forma preliminar, utilizaremos como indicador la percepción del encuestado respecto a si la el número de delitos en su comuna pareciera haber aumentado/mantenido/disminuido)

Por lo tanto, para poder desarrollar esta base de datos deberemos limpiar y unir las 3 fuentes de información siguientes:

Las fuentes de datos utilizadas son las siguientes:

“Estadísticas Delictuales” (<https://cead.spd.gov.cl/estadisticas-delictuales/>) - número de delitos mensuales en las comunas chilenas entre 2021 y 2023 “Proyecciones de Población” (<https://regiones.ine.cl/biobio/estadisticas-regionales/sociales/demografia-y-vitales/proyecciones-de-poblacion>) - proyecciones de población total a nivel comuna en Chile “Encuesta Nacional Urbana de Seguridad Ciudadana (ENUSC)” (<https://www.ine.gob.cl/estadisticas/sociales/seguridad-publica-y-justicia/seguridad-ciudadana>) - Encuesta de hogares con información de características socioeconómicas y percepciones o vivencias ligadas al crimen y la seguridad

Hipótesis

I) Base de datos de delitos mensuales por comuna

Importamos archivo excel con delitos mensuales entre 2021 y 2023:

```
file_path <- "data/totaldelitos_2021_2023_porcomuna_pormes.xlsx"
totaldelitos_original <- read_excel(file_path, skip = 22)
```

Problema con formato de variables de texto

PROBLEMA! Al intentar exportar el archivo markdown como pdf (knit) Hay caracteres Unicode, lo q impide procesar con latex, por lo tanto deberemos remplazar los unicodes por caracteres ASCII.

Específicamente, vemos que los unicode characters encontrados son vocales con tildes, por lo que remplazaremos las vocales con tildes por vocales sin tildes.

```
find_unicode <- function(column) {  
  if (is.character(column)) {  
    unique(unlist(regmatches(column, gregexpr("[^\\x01-\\x7F]", column))))  
  } else {  
    NULL  
  }  
}  
  
unicode_chars <- lapply(totaldelitos_original, find_unicode)  
unicode_chars <- unicode_chars[lengths(unicode_chars) > 0]  
print(unicode_chars)
```

```
## $`UNIDAD TERRITORIAL`  
## [1] "Í" " " "ó" "á" "ñ" "ú" "í" "é" "Ñ" "Á" "Ú"
```

Finalmente confirmamos que vocales con tildes han sido rempladas por vocales sin tilde.

```
head(totaldelitos_original[, 1:5])
```

```
##          UNIDAD.TERRITORIAL          X1...2          X2...3  
## 1          TOTAL PAIS 129.636,0000000000 115.961,0000000000  
## 2 Region de Arica y Parinacota 2.100,0000000000 2.232,0000000000  
## 3          Provincia de Arica 2.069,0000000000 2.193,0000000000  
## 4              Arica 2.052,0000000000 2.179,0000000000  
## 5              Camarones 170000000000 140000000000  
## 6  Provincia de Parinacota 310000000000 390000000000  
##          X3...4          X4...5  
## 1 138.950,0000000000 140.564,0000000000  
## 2 2.440,0000000000 2.172,0000000000  
## 3 2.404,0000000000 2.149,0000000000  
## 4 2.378,0000000000 2.121,0000000000  
## 5 260000000000 280000000000  
## 6 360000000000 230000000000
```

Limpieza de la base de datos:

Exploramos las columnas para limpiar la base de datos

```
str(totaldelitos_original)  
colnames(totaldelitos_original)  
ncol(totaldelitos_original)
```

Vemos que hay varios problemas: 1) Hay q cambiar nombre de 36 columnas a 202101,202102,202103...202312
2) Arreglar formato de valores: a) borrar 10 ceros al final de cada valor b) eliminar caracteres “.” y “,” c)
finalmente convertir a numerico 3) Eliminar filas que corresponden a agrupamientos de comunas

1) Cambiar nombres a variables

```
year_month <- rep(2021:2023, each = 12)
months <- sprintf("%02d", 1:12)
year_month_names <- c(paste0(2021, months), paste0(2022, months), paste0(2023, months[1:12]))
colnames(totaldelitos_original) <- c("UNIDAD_TERRITORIAL", year_month_names)
```

2) arreglar formato números

hay 10 ceros al final de cada celda, eliminar “.” y “,” , convertir a numerico

```
#Hacemos copia para guardar el dataframe original
totaldelitos <- totaldelitos_original
# Editamos solamente las 36 columnas correspondientes a los 36 meses entre 2021 y 2023
totaldelitos[2:37] <- lapply(totaldelitos[2:37], function(x) {
  x <- substr(x, 1, nchar(x) - 10) # a) eliminar ultimos 10 caracteres
  gsub("[,\\.]", "", x) # b) eliminar ',' y '.'
})
totaldelitos[2:37] <- lapply(totaldelitos[2:37], as.numeric) # c) convertir a numérico
```

3) Eliminar filas que corresponden a agrupamientos de comunas

```
totaldelitos <- totaldelitos %>%
  filter(!grepl("(Region|Provincia)", `UNIDAD_TERRITORIAL`, ignore.case = TRUE))
#head(totaldelitos)
```

4) Chequeo de la base de datos limpiada

Para ver que la limpieza que tuvo que hacerse genero datos consistentes, calcularemos el total de delitos nivel país y lo compararemos con el total país entregado originalmente por la base de datos. También graficaremos los datos y veremos estadística descriptiva

```
totaldelitos_pais_sum <- totaldelitos %>%
  filter(`UNIDAD_TERRITORIAL` != "TOTAL PAÍS") %>% # no incluimos la fila TOTAL PAÍS
  summarise(across(-`UNIDAD_TERRITORIAL`, \ (x) sum(x, na.rm = TRUE))) # sumamos a través de todas las

totaldelitos_pais_sum
totaldelitos_pais_original <- totaldelitos[totaldelitos$UNIDAD_TERRITORIAL == "TOTAL PAÍS", !(colnames(
totaldelitos_pais_original
```

4.1) Calculamos total de delitos mensuales a nivel país con la base de datos obtenida luego de la limpieza de la base de datos original. Ahora crearemos un dataframe que una la variable TOTAL PAÍS original y la variable creada por nosotros para así ver sus diferencias porcentuales

```

t_totaldelitos_pais_sum <- pivot_longer(
  data = totaldelitos_pais_sum,
  cols = everything(),           # Incluimos todas las columnas
  names_to = "Meses",           # Columna con nombre de las variables antiguas
  values_to = "Delitos"         # Nombre de la columna con los valores de las variables antiguas
)

t_totaldelitos_pais_original <- pivot_longer(
  data = totaldelitos_pais_original,
  cols = everything(),           # Incluimos todas las columnas
  names_to = "Meses",           # Columna con nombre de las variables antiguas
  values_to = "Delitos"         # Nombre de la columna con los valores de las variables antiguas
)

#t_totaldelitos_pais_sum
#t_totaldelitos_pais_original

total_delitos_comparacion <- t_totaldelitos_pais_original %>%
  full_join(t_totaldelitos_pais_sum, by = "Meses") %>%
  mutate(Percentage_Difference = ((Delitos.x - Delitos.y) / Delitos.y) * 100)
head(total_delitos_comparacion)

```

```

## # A tibble: 6 x 4
##   Meses Delitos.x Delitos.y Percentage_Difference
##   <chr>      <dbl>      <dbl>          <dbl>
## 1 202101    129636    129333          0.234
## 2 202102    115961    115689          0.235
## 3 202103    138950    138534          0.300
## 4 202104    140564    140085          0.342
## 5 202105    119644    119420          0.188
## 6 202106    112613    112426          0.166

```

```
summary(total_delitos_comparacion$Percentage_Difference)
```

```

##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.04223 0.05202 0.06598 0.11187 0.15559 0.34194

```

Vemos que hay en promedio un 0.11% de diferencia entre los delitos TOTAL PAÍS Entregados por la base de datos y el mismo valor pero calculado sumando los delitos por comuna posteriormente a haber limpiado la base. Si bien desconocemos el por qué existe una discrepancia, podemos de todas formas estar tranquilos que la diferencia es mínima por lo que la limpieza de datos no debe haber alterado los valores de la base de datos.

4.2) Graficamos la variable TOTAL PAÍS: grafico de numero de delitos mensuales a nivel país entre 2021 y 2023 Convertimos base de Wide a Long para poder graficar más fácilmente la serie de tiempo

```

totaldelitos_long <- totaldelitos %>%
  pivot_longer(
    cols = starts_with("202"),
    names_to = "Mes",

```

```

    values_to = "Delitos"
  )

# Creamos una nueva columna para identificar los meses de enero y julio, esto nos permitirá identificar
totaldelitos_long <- totaldelitos_long %>%
  mutate(
    Mes_2 = case_when(
      grepl("01$", Mes) ~ "Enero",
      grepl("07$", Mes) ~ "Julio",
      TRUE ~ "."
    )
  )

```

```

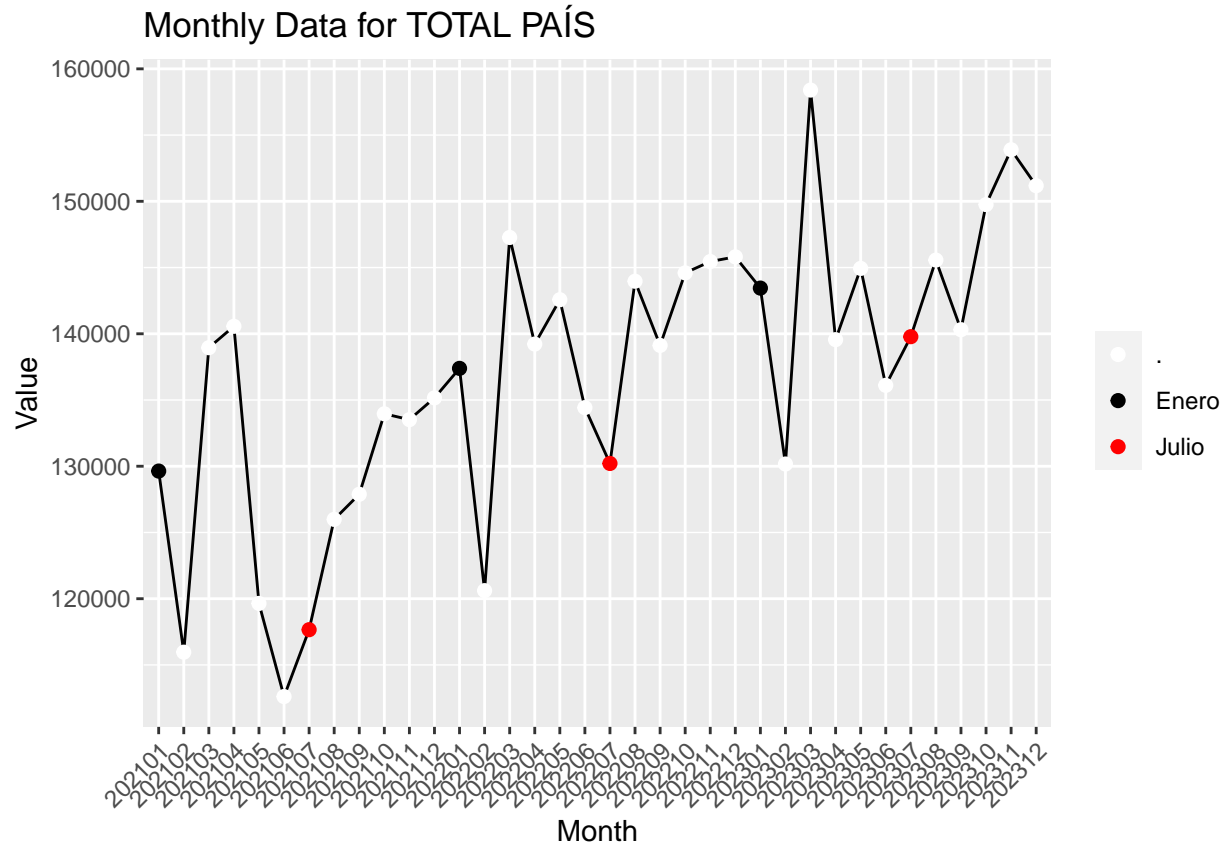
ggplot(
  totaldelitos_long %>% filter(UNIDAD_TERRITORIAL == "TOTAL PAIS"),
  aes(x = Mes, y = Delitos, group = 1)) +
  geom_line() +

  geom_point(aes(color = Mes_2), size = 2) +

  scale_color_manual(
    values = c("Enero" = "black", "Julio" = "red", "." = "white")
  ) +

  labs(
    title = "Monthly Data for TOTAL PAÍS",
    x = "Month",
    y = "Value"
  ) +
  theme(
    axis.text.x = element_text(angle = 45, hjust = 1),
    legend.title = element_blank()
  )

```



5) Conclusión

En conclusión, hemos construido una base de datos que contiene el número de delitos mensuales para cada comuna entre 2021 y 2023.

II) Limpieza de base de datos con la población por comuna entre 2021 y 2023 y posterior union entre base poblacion y base delitos

1) Base de datos con códigos y nombres de comunas

Para unir la base de delitos y la base con la población a nivel comunal, es necesario obtener los códigos asociados a cada columna para así poder unir las bases basandonos en la variable con los códigos “cod.comuna”

```
file_path <- "C:/Users/rafae/Documents/github/proyectos/R/delitos_percepcion/data/df_comunas_dict.csv"
df_comunas_dict <- read.csv(file_path, header = TRUE, fileEncoding = "Latin1", encoding = "Latin1")
```

2) Preparamos ambas bases para poder hacer match

Primero debemos asegurarnos de quitar las tildes de la variable con los nombres de las comunas

```
df_comunas_dict$UNIDAD.TERRITORIAL <- replace_unicode(df_comunas_dict$UNIDAD.TERRITORIAL)
head(df_comunas_dict$UNIDAD.TERRITORIAL)
```

Luego nos aseguramos de que variables que utilizaremos para hacer el match sigan el mismo formato

```
library(dplyr)

totaldelitos <- totaldelitos %>%
  rename(UNIDAD.TERRITORIAL = `UNIDAD_TERRITORIAL`)

str(df_comunas_dict$UNIDAD.TERRITORIAL)
str(totaldelitos$UNIDAD.TERRITORIAL)

df_comunas_dict$UNIDAD.TERRITORIAL <- trimws(df_comunas_dict$UNIDAD.TERRITORIAL)
totaldelitos$UNIDAD.TERRITORIAL <- trimws(totaldelitos$UNIDAD.TERRITORIAL)

str(df_comunas_dict$UNIDAD.TERRITORIAL)
str(totaldelitos$UNIDAD.TERRITORIAL)

#probamos otra forma:
df_comunas_dict$UNIDAD.TERRITORIAL <- gsub("^\\s+|\\s+$", "", df_comunas_dict$UNIDAD.TERRITORIAL)
totaldelitos$UNIDAD.TERRITORIAL <- gsub("^\\s+|\\s+$", "", totaldelitos$UNIDAD.TERRITORIAL)

#Vemos q hay espacios al principio!
str(df_comunas_dict$UNIDAD.TERRITORIAL)
str(totaldelitos$UNIDAD.TERRITORIAL)

head(df_comunas_dict)
head(totaldelitos)
```

3) Finalmente agregamos a la base de delitos los codigos de las comunas. Para eso hacemos match entre las variables UNIDAD.TERRITORIAL, las cuales tienen los nombres de las comunas en el mismo formato

```
library(dplyr)

df_comunas_dict$UNIDAD.TERRITORIAL <- trimws(df_comunas_dict$UNIDAD.TERRITORIAL)
totaldelitos$UNIDAD.TERRITORIAL <- trimws(totaldelitos$UNIDAD.TERRITORIAL)

merged_df <- full_join(df_comunas_dict, totaldelitos, by = "UNIDAD.TERRITORIAL")

head(merged_df)

delitos_2023_comunas <- merged_df
```

4) Unimos base de delitos y poblacion

```
file_path <- "data/ine_estimaciones-y-proyecciones-2002-2035_base-2017_comunas0381d25bc2224f51b9770a705"

# Read the CSV file (assuming the first row contains variable names)
```

```
data <- read.csv(file_path, header = TRUE)
#data <- read.csv(file_path, header = TRUE, fileEncoding = "UTF-8", encoding = "UTF-8")
data <- read.csv(file_path, header = TRUE, fileEncoding = "Latin1", encoding = "Latin1")
head(data)
```

crear base de poblacion

```
data_filtered <- data %>%
  select(
    Region,
    `Nombre.Region`,
    Provincia,
    `Nombre.Provincia`,
    Comuna,
    `Nombre.Comuna`,
    `Poblacion.2023`,
    `Poblacion.2022`,
    `Poblacion.2021`
  )

head(data_filtered)
```

elegir columnas de interés

```
data_grouped <- data_filtered %>%
  group_by(Comuna) %>%
  summarise(
    Poblacion.2023 = sum(Poblacion.2023, na.rm = TRUE),
    Poblacion.2022 = sum(Poblacion.2022, na.rm = TRUE),
    Poblacion.2021 = sum(Poblacion.2021, na.rm = TRUE),
    #Region = first(Region),
    Nombre.Region = first(Nombre.Region),
    #Provincia = first(Provincia),
    #Nombre.Provincia = first(Nombre.Provincia),
    Nombre.Comuna = first(Nombre.Comuna)
  )

head(data_grouped)

poblacion_2023_comunas <- data_grouped
```

Agrupamos la población por comuna


```

delitos_2023_comunas

poblacion_2023_comunas

str(delitos_2023_comunas$cod.comuna)
str(poblacion_2023_comunas$Comuna)

str(colnames(df_comunas_dict))

poblacion_2023_comunas <- poblacion_2023_comunas %>%
  rename(cod.comuna = Comuna)

colnames(poblacion_2023_comunas)

```

Confirmamos que variable de matching (codigo comuna) esté en mismo formato en ambas bases

```

merged_df <- full_join(poblacion_2023_comunas, delitos_2023_comunas, by = "cod.comuna")
merged_df

```

Finalmente unimos delitos y poblacion

5) Creamos variable de analisis: crecimiento del delito a nivel comunal como porcentaje de la poblacion

```

library(dplyr)

months <- paste0("2022", sprintf("%02d", 1:12))

merged_df <- merged_df %>%
  mutate(across(all_of(months), ~ . / Poblacion.2022,
    .names = "{.col}_porc"))

head(merged_df)

months <- paste0("2023", sprintf("%02d", 1:12))

merged_df <- merged_df %>%
  mutate(across(all_of(months), ~ . / Poblacion.2023,
    .names = "{.col}_porc"))

head(merged_df)

```

primero creamos variable con ratio cantidad de delitos sobre poblacion a nivel comuna

en segundo lugar calculamos la media de los ratios de delitos mensuales entre Julio y diciembre para 2022 y 2023. Esto nos permitirá estimar cuanto aumentó el número de delitos en 1 año. La razón de por qué utilizar solo Julio a diciembre y no enero a diciembre es que las cuarentenas de la pandemia provocaron una caída de delitos, por lo que utilizamos solamente datos desde Julio 2022 para evitar capturar ese efecto

```
library(dplyr)

months_2022 <- paste0("2022", sprintf("%02d", 7:12), "_perc")
months_2023 <- paste0("2023", sprintf("%02d", 7:12), "_perc")

merged_df <- merged_df %>%
  mutate(
    `202207_12` = rowMeans(select(., all_of(months_2022)), na.rm = TRUE),
    `202307_12` = rowMeans(select(., all_of(months_2023)), na.rm = TRUE),

    `perc_increase_202307_12_from_202207_12` =
      ((`202307_12` - `202207_12`) / `202207_12`) * 100
  )

head(merged_df)
```

```
## # A tibble: 6 x 70
##   cod.comuna Poblacion.2023 Poblacion.2022 Poblacion.2021 Nombre.Region
##   <int>          <int>          <int>          <int> <chr>
## 1      1101      230595      229072      227127 Tarapacá
## 2      1107      140282      137263      134085 Tarapacá
## 3      1401      18452       18172       17853 Tarapacá
## 4      1402       1379       1383       1384 Tarapacá
## 5      1403       1569       1579       1588 Tarapacá
## 6      1404       3080       3063       3043 Tarapacá
## # i 65 more variables: Nombre.Comuna <chr>, UNIDAD.TERRITORIAL <chr>,
## #   `202101` <dbl>, `202102` <dbl>, `202103` <dbl>, `202104` <dbl>,
## #   `202105` <dbl>, `202106` <dbl>, `202107` <dbl>, `202108` <dbl>,
## #   `202109` <dbl>, `202110` <dbl>, `202111` <dbl>, `202112` <dbl>,
## #   `202201` <dbl>, `202202` <dbl>, `202203` <dbl>, `202204` <dbl>,
## #   `202205` <dbl>, `202206` <dbl>, `202207` <dbl>, `202208` <dbl>,
## #   `202209` <dbl>, `202210` <dbl>, `202211` <dbl>, `202212` <dbl>, ...
```

Conclusión

En conclusión, hemos construido la base de datos que contiene nuestra variable de interés: El crecimiento porcentual del ratio de número de delitos entre 2022 y 2023. A continuación crearemos una base de datos con información de la percepción de los habitantes de las comunas respecto a cuanto aumentaron los delitos en ese periodo, de forma a estudiar como se relaciona el aumento de delitos

III) Base de datos con percepción subjetiva de inseguridad (ENUSC)

Utilizaremos la base de datos enusc para contruir la variable de percepcion de inseguridad

1) Explorar base

```
file_path <- "C:/Users/rafae/Documents/github/proyectos/R/delitos_percepcion/data/base-usuario-20-enusc"

data_enusc <- read.csv(file_path, encoding = "Latin1")

head(data_enusc)

summary(data_enusc$P_AUMENTO_COM)
table(data_enusc$P_AUMENTO_COM)

data_enusc %>%
  select(P_AUMENTO_COM) %>%
  mutate(P_AUMENTO_COM = ifelse(is.na(P_AUMENTO_COM), "Missing", as.character(P_AUMENTO_COM))) %>%
  count(P_AUMENTO_COM)
```

2) Construcción de indicador de percepción de inseguridad en la comuna y de % de clase alta en la comuna

Ahora calcularemos características de las comunas: - utilizaremos el promedio comunal de P_AUMENTO_COM para ver el promedio del índice de percepción de inseguridad - utilizaremos el promedio comunal de rph_nse para ver el % de individuos clase alta en la comuna

```
library(dplyr)

data_enusc %>%
  select(enc_rpc, P_AUMENTO_COM)

table(data_enusc$P_AUMENTO_COM)

data_enusc <- data_enusc %>%
  mutate(P_AUMENTO_COM = case_when(
    P_AUMENTO_COM == 3 ~ -1,
    P_AUMENTO_COM == 2 ~ 0,
    P_AUMENTO_COM == 1 ~ 1,
    P_AUMENTO_COM > 3 ~ NA,
    TRUE ~ P_AUMENTO_COM
  ))

summary(data_enusc$P_AUMENTO_COM)

data_enusc %>%
  select(enc_rpc, P_AUMENTO_COM)

data_enusc_feq <- data_enusc %>%
  group_by(enc_rpc) %>%
  summarise(
    Frequency = n(), # Total count of rows per group
    Proportion_rph_nse_1 = mean(rph_nse %in% c(1), na.rm = TRUE), # Proportion where rph_nse == 1 or 2
```

```

    Mean_P_AUMENTO_COM = mean(P_AUMENTO_COM[P_AUMENTO_COM != 88 & P_AUMENTO_COM != 99], na.rm = TRUE) #
  ) %>%
  arrange(desc(Frequency))

summary(data_enusc$Mean_P_AUMENTO_COM)

print(data_enusc_feq)

```

IV) Base de datos final: union entre crecimiento de delitos a nivel comunal y percepcion de inseguridad a nivel comunal

1) Unimos base delitos y base percepciion inseguridad

Ahora uniremos ambas bases para así poder estudiar la relación entre: - el aumento real (objetivo) del número de delitos a nivel comunal entre 2022 y 2023 (base delitos) - el aumento/disminucion de la percepcion de inseguridad (base enusc)

```

### Ahora, tenemos una base de datos con delitos como % de habitantes de cada comuna y otra base con da
#### Join delitos y enusc

merged_df
data_enusc_feq

library(dplyr)

merged_df <- merged_df %>%
  mutate(cod.comuna = as.character(cod.comuna))

data_enusc_feq <- data_enusc_feq %>%
  mutate(enc_rpc = as.character(enc_rpc))

final_data <- full_join(merged_df, data_enusc_feq,
                        by = c("cod.comuna" = "enc_rpc"))

head(final_data)

```

2) Análisis exploratorio a Nivel Comunal

A continuacion estudiaremos la relación entre aumento objetivo de delitos y percepcion subjetiva de aumento de delitos a nivel comunal.

```

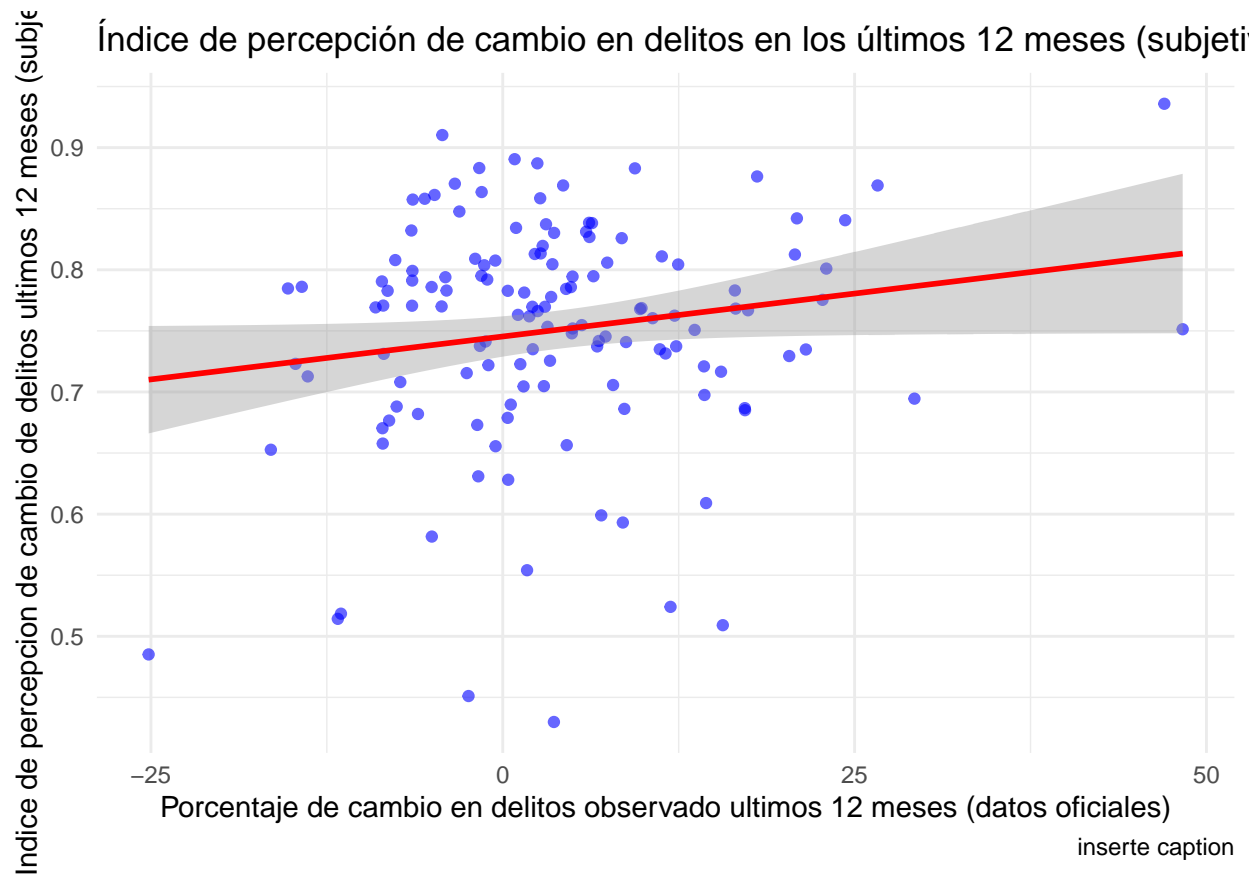
library(ggplot2)
library(dplyr)

filtered_data <- final_data %>%
  filter(!is.na(Mean_P_AUMENTO_COM))

ggplot(filtered_data, aes(x = perc_increase_202307_12_from_202207_12, y = Mean_P_AUMENTO_COM)) +
  geom_point(color = "blue", alpha = 0.6) +
  geom_smooth(method = "lm", color = "red", se = TRUE) +

```

```
theme_minimal() +
labs(
  title = "Índice de percepción de cambio en delitos en los últimos 12 meses (subjetivo, indice<0 => j",
  x = "Porcentaje de cambio en delitos observado ultimos 12 meses (datos oficiales)",
  y = "Indice de percepcion de cambio de delitos ultimos 12 meses (subjetivo)",
  caption = "inserte caption"
)
```



```
model <- lm(perc_increase_202307_12_from_202207_12 ~ Mean_P_AUMENTO_COM, data = filtered_data)
summary(model)
```

Base de datos final 2 : Análisis a nivel individual

Ahora veremos la relación entre la percepción de aumento de delitos por el individuo y el real aumento de delitos en la comuna

1) Agregamos a encuesta ENUC (nivel individuo) los datos a nivel comunal

```
library(dplyr)

head(data_enusc %>%
```

```

select(enc_rpc, P_AUMENTO_COM))

table(data_enusc$P_AUMENTO_COM)

data_enusc <- data_enusc %>%
  mutate(P_AUMENTO_COM = case_when(
    P_AUMENTO_COM == 3 ~ -1,
    P_AUMENTO_COM == 2 ~ 0,
    P_AUMENTO_COM == 1 ~ 1,
    P_AUMENTO_COM > 3 ~ NA,
    TRUE ~ P_AUMENTO_COM
  ))

summary(data_enusc$P_AUMENTO_COM)

head(data_enusc %>%
  select(enc_rpc, P_AUMENTO_COM))

head(merged_df)
head(data_enusc)

merged_result <- merge(merged_df, data_enusc, by.x = "cod.comuna", by.y = "enc_rpc", all.x = TRUE)
head(merged_result)

average_data <- merged_result %>%
  group_by(P_AUMENTO_COM) %>%
  summarise(
    avg_perc_increase = mean(perc_increase_202307_12_from_202207_12, na.rm = TRUE),
    count = n()
  )

ggplot(average_data, aes(x = as.factor(P_AUMENTO_COM), y = avg_perc_increase, fill = as.factor(P_AUMENTO_COM))) +
  geom_col(width = 0.6, color = "black", size = 0.5) +
  geom_text(aes(label = paste0("N=", count)), vjust = -0.5, size = 4, fontface = "bold") +
  ggtitle("Aumento Observado del Delito por Comuna \n según Percepción Subjetiva") +
  xlab("Percepción de Evolución de Delitos Comunales\n(-1: Disminuyó, 0: Se mantuvo, 1: Aumentó)") +
  ylab("Aumento Porcentual Promedio del Delito") +
  theme_minimal(base_size = 14) +
  theme(
    plot.title = element_text(hjust = 0.5, face = "bold", size = 16),
    axis.title.x = element_text(face = "bold"),
    axis.title.y = element_text(face = "bold"),
    panel.grid.major = element_line(color = "gray80"),
    panel.grid.minor = element_blank()
  ) +
  scale_fill_manual(values = c("-1" = "red", "0" = "gray", "1" = "green", "NA" = "darkgray")) +
  guides(fill = "none")

```

