

# Proyecto Final Módulo R

Viteri Ayala Flavia Kamila

2023-08-06

## Tabla de Contenidos

<b>Parte 1. Setup</b>	<b>1</b>
1.1. Librerías . . . . .	1
1.2. Establecer formato de números . . . . .	2
1.3. Bases de datos . . . . .	2
<b>Parte 2. Manipulación de datos</b>	<b>3</b>
2.1. Base data_codebook . . . . .	3
2.2. Base data_balances . . . . .	4
2.3. Base data_ciiu . . . . .	4
2.4. Base empresas (creación) . . . . .	5
<b>Parte 3. Análisis</b>	<b>6</b>
3.1. Número de empresas por actividad económica por cantón . . . . .	6
3.2. Liquidez y solvencia patrimonial por status y provincia . . . . .	8
3.4. Endeudamiento del activo por tamaño de empresa . . . . .	12
3.5. Liquidez por número de trabajadores . . . . .	13
3.6. Apalancamiento por empresas . . . . .	15

## Parte 1. Setup

### 1.1. Librerías

Para este proyecto se utilizan las siguientes librerías:

```
library(openxlsx)
library(magrittr)
library(dplyr)
library(tidyverse)
library(readxl)
library(ggplot2)
library(patchwork)
library(janitor)
library(ggrepel)
```

## 1.2. Establecer formato de números

Este parámetro se utiliza para controlar la forma en que los números se imprimen en la consola cuando se presenta el resultado de un cálculo o una operación.

```
options(scipen = 999)
```

## 1.3. Bases de datos

Para el desarrollo de este proyecto se utilizan tres bases principales las cuales contienen información oficial emitida por la Superintendencia de Compañías, Valores y Seguros (SUPERCÍAS).

```
data_balances <- read.xlsx("Data/balances_2014.xlsx", sheet = 1, detectDates = TRUE)

data_ciiu <- read.xlsx("Data/ciiu.xlsx", sheet = 1, detectDates = TRUE)

data_codebook <- read.xlsx("Data/cias_codebook.xlsx", sheet = 1, detectDates = TRUE)
```

La base de datos `data_balances` contiene información de diversas empresas ecuatorianas. Tiene **47033** filas, donde cada una representa un registro en el catastro empresarial y cuenta con **347** columnas que incluyen el número de expediente, el Registro Único de Contribuyentes (RUC), nombre de la compañía, situación legal (activa/inactiva/en liquidación, etc.), tipo de compañía, fecha de constitución, ubicación geográfica, código ciiu, características de empleo, tamaño de la empresa, año de constitución e indicadores financieros. Esta base de datos es una fuente rica y diversa de información sobre empresas en Ecuador, útil para análisis económicos y empresariales.

A continuación, se presenta la estructura de la base `data_balances` con una muestra de sus primeras 2 columnas más 3 columnas aleatorias.

```
data_balances |>
select(1:2, sample(3:ncol(data_balances), 3)) |>
glimpse()
```

```
## Rows: 47,033
## Columns: 5
## $ expediente <dbl> 1, 2, 3, 11, 21, 22, 41, 49, 58, 63, 69, 76, 78, 79, 83, 85~
## $ ruc         <chr> "1790013731001", "1790004724001", "1790008959001", "1790044~
## $ v7292       <dbl> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, NA, NA, 0, 0, 0, NA, NA, 0, 0, 0~
## $ v7341       <dbl> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, NA, NA, 0, 0, 0, NA, NA, 0, 0, 0~
## $ v7473       <dbl> 0.00, 23905.86, 0.00, 0.00, 0.00, 0.00, 0.00, 0.00, 6.97, 0.00, N~
```

Como segundo punto, la base de datos `data_ciiu` constituye una clasificación jerárquica que abarca diversas actividades vinculadas con los sectores económicos del Ecuador. Específicamente, se trata de una adaptación de la Clasificación Internacional Uniforme (CIIU) para este país. En esta base de datos, las actividades se encuentran estructuradas en diferentes niveles jerárquicos, donde el primer nivel representa categorías generales, y los niveles subsiguientes detallan subcategorías más específicas. Estos niveles desembocan en un código único para cada actividad, proporcionando así una identificación única para cada una de ellas.

A continuación, se presenta la estructura de la base `data_ciiu`:

```
data_ciiu |>
glimpse()
```

```
## Rows: 3,044
## Columns: 3
## $ CODIGO      <chr> "A", "A01", "A011", "A0111", "A0111.1", "A0111.11", "A0111~
## $ DESCRIPCION <chr> "AGRICULTURA, GANADERÍA, SILVICULTURA Y PESCA.", "AGRICUL~
## $ NIVEL       <dbl> 1, 2, 3, 4, 5, 6, 6, 6, 5, 6, 6, 6, 6, 5, 6, 6, 6, 4, 5~
```

Por último, la base de datos `data_codebook` se encuentra estructurada de manera vertical, donde cada columna representa un registro y las filas son las variables asociadas a los mismos. Los campos incluidos en esta base de datos corresponden al diccionario de variables de la base de datos “data\_balance”, es decir, para poder entender la base de datos en mención, es necesario recurrir a la base de datos “data\_codebook”.

A continuación, se presenta la estructura de la base `data_codebook` :

```
data_codebook |>
glimpse()

## Rows: 348
## Columns: 2
## $ expediente <chr> "ruc", "nombre_cia", "situacion", "tipo", "fecha_const", "p~
## $ EXPEDIENTE <chr> "RUC", "NOMBRE DE LA COMPAÑÍA", "SITUACIÓN LEGAL", "TIPO", ~
```

## Parte 2. Manipulación de datos

### 2.1. Base `data_codebook`

Se manipula esta base en primera instancia porque la misma nos permite identificar las variables necesarias para el desarrollo y cálculo de los indicadores financieros establecidos para la ejecución del proyecto.

```
## Identificar variables de interés ----
data_codebook <- data_codebook |>
  filter(
    str_detect(EXPEDIENTE, "NOMBRE") |
    str_detect(EXPEDIENTE, "SITUACIÓN") |
    str_detect(EXPEDIENTE, "TIPO") |
    str_detect(EXPEDIENTE, "PAÍS") |
    str_detect(EXPEDIENTE, "PROVINCIA") |
    str_detect(EXPEDIENTE, "CANTÓN") |
    str_detect(EXPEDIENTE, "CIUDAD") |
    str_detect(EXPEDIENTE, "NÚMERO DE TRABAJADORES DIRECTIVOS") |
    str_detect(EXPEDIENTE, "NÚMERO DE TRABAJADORES ADMINISTRATIVOS") |
    str_detect(EXPEDIENTE, "NÚMERO DE TRABAJADORES PRODUCCIÓN") |
    str_detect(EXPEDIENTE, "NÚMERO DE TRABAJADORES OTROS") |
    str_detect(EXPEDIENTE, "TAMAÑO") |
    str_detect(EXPEDIENTE, "^CIIU4 NIVEL 1") |
    str_detect(EXPEDIENTE, "CIIU4 NIVEL 6") |
    str_detect(EXPEDIENTE, "TOTAL ACTIVOS CORRIENTE") |
    str_detect(EXPEDIENTE, "TOTAL PASIVOS CORRIENTES") |
    str_detect(EXPEDIENTE, "TOTAL DEL ACTIVO") |
    str_detect(EXPEDIENTE, "TOTAL DEL PASIVO") |
    str_detect(EXPEDIENTE, "TOTAL PATRIMONIO NETO") |
    str_detect(EXPEDIENTE, "TOTAL ACTIVOS NO CORRIENTES")
  )
```

En este sentido, se filtra `data_codebook` sobrescribiendo el mismo dataframe. Es menester mencionar que se decide usar este tipo de filtro porque facilita la identificación de las variables dentro de la base y además el código se muestra más amigable para su lectura y comprensión.

## 2.2. Base data\_balances

Se manipula esta base tomando en cuenta la base `data_codebook` con el objetivo de identificar el nombre de cada indicador para su análisis.

```
## 1. Seleccionar variables ----
data_balances <- data_balances |>
  select(all_of(data_codebook$expediente))

## 2. Cambiar nombres ----
names(data_balances) <- data_codebook$EXPEDIENTE

## 3. Limpiar nombres de variables----
data_balances <- data_balances |>
  clean_names()
```

En este sentido, se filtra `data_balances` sobrescribiendo el mismo dataframe. A continuación se presenta la estructura de la base en mención:

```
data_balances |>
  glimpse()

## Rows: 47,033
## Columns: 20
## $ nombre_de_la_compania      <chr> "ACEITES TROPICALES SOCIEDAD AN~
## $ situacion_legal            <chr> "ACTIVA", "ACTIVA", "ACTIVA", "~
## $ tipo                       <chr> "ANÓNIMA", "ANÓNIMA", "ANÓNIMA"~
## $ pais                      <chr> "ECUADOR", "ECUADOR", "ECUADOR"~
## $ provincia                 <chr> "SANTO DOMINGO DE LOS TSACHILAS~
## $ canton                    <chr> "SANTO DOMINGO", "QUITO", "QUIT~
## $ ciudad                    <chr> "SANTO DOMINGO", "QUITO", "QUIT~
## $ ciu4_nivel_1               <chr> "A", "C", "G", "H", "G", "G", "~
## $ ciu4_nivel_6              <chr> "A0126.01", "C2410.25", "G4663.~
## $ numero_de_trabajadores_directivos <dbl> 8, 9, 3, 1, 1, 1, 1, 2, 1, 9, 2~
## $ numero_de_trabajadores_administrativos <dbl> 4, 384, 70, NA, 6, 3, NA, 4, 5,~
## $ numero_de_trabajadores_produccion <dbl> 5, 980, 77, NA, NA, NA, NA, 10,~
## $ numero_de_trabajadores_otros <dbl> NA, NA, NA, NA, 19, NA, NA, NA,~
## $ tamano                    <chr> "PEQUEÑA", "GRANDE", "GRANDE", ~
## $ total_activos_corrientes    <dbl> 38558.00, 211859589.00, 1120702~
## $ total_activos_no_corrientes <dbl> 881354.91, 130000116.00, 117938~
## $ total_del_activo            <dbl> 919912.9, 341859705.0, 12386414~
## $ total_pasivos_corrientes    <dbl> 12204.78, 115561889.00, 3629656~
## $ total_del_pasivo            <dbl> 41875.19, 196826509.00, 8383147~
## $ total_patrimonio_neto       <dbl> 878037.72, 145033196.00, 400326~
```

## 2.3. Base data\_ciuu

El objetivo de manipular esta base se enmarca en dejarla lista para posteriormente cargar en la nueva base que nos permitirá hacer los análisis de investigación.

```
## Seleccionar variables de interés ----
data_ciiu <- data_ciiu |>
  select(-NIVEL) |>

## Limpiar nombres de variables ----
clean_names()
```

A continuación se presenta la estructura de la base data\_ciiu:

```
data_ciiu |>
glimpse()

## Rows: 3,044
## Columns: 2
## $ codigo      <chr> "A", "A01", "A011", "A0111", "A0111.1", "A0111.11", "A0111~
## $ descripcion <chr> "AGRICULTURA, GANADERÍA, SILVICULTURA Y PESCA.", "AGRICUL~
```

## 2.4. Base empresas (creación)

La base empresas nos permitirá abordar el y la resolución de las preguntas de investigación.

```
## Cálculo de indicadores ----
empresas <- data_balances |>
  mutate(
    ind_liquidez_corriente = total_activos_corrientes / total_pasivos_corrientes,
    ind_endeudamiento_del_activo = total_del_pasivo / total_del_activo,
    ind_endeudamiento_patrimonial = total_del_pasivo / total_patrimonio_netto,
    ind_endeudamiento_del_activo_fijo = total_patrimonio_netto / total_activos_no_corrientes,
    ind_apalancamiento = total_del_activo / total_patrimonio_netto
  )

## Quitar variables ya utilizadas ----
empresas <- empresas |>
  select(-starts_with("total"))

## Agregar descripción de los CIIU nivel 1 ----
empresas <- empresas |>
  left_join(data_ciiu, by = c("ciiu4_nivel_1" = "codigo")) |>
  rename(desc_ciiu_nivel_1 = descripcion)

## Agregar descripción de los CIIU nivel 1 ----
empresas <- empresas |>
  left_join(data_ciiu, by = c("ciiu4_nivel_6" = "codigo")) |>
  rename(desc_ciiu_nivel_6 = descripcion)

## Colocar columnas en orden ----
empresas <- empresas |>
  relocate(desc_ciiu_nivel_1, .after = ciiu4_nivel_1) |>
  relocate(desc_ciiu_nivel_6, .after = ciiu4_nivel_6)

## Cambiar datos infinitos a NA ----
empresas <- empresas |>
```

```
mutate(across(starts_with("ind"), function(x) {
  case_when(
    is.infinite(x) ~ NA_real_,
    TRUE ~ x
  )
}))
)
```

A continuación se presenta la estructura de la base de datos **empresas**:

```
empresas |>
glimpse()
```

```
## Rows: 47,033
## Columns: 21
## $ nombre_de_la_compania      <chr> "ACEITES TROPICALES SOCIEDAD AN~
## $ situacion_legal            <chr> "ACTIVA", "ACTIVA", "ACTIVA", "~
## $ tipo                      <chr> "ANÓNIMA", "ANÓNIMA", "ANÓNIMA"~
## $ pais                      <chr> "ECUADOR", "ECUADOR", "ECUADOR"~
## $ provincia                 <chr> "SANTO DOMINGO DE LOS TSACHILAS~
## $ canton                    <chr> "SANTO DOMINGO", "QUITO", "QUIT~
## $ ciudad                    <chr> "SANTO DOMINGO", "QUITO", "QUIT~
## $ ciuu4_nivel_1             <chr> "A", "C", "G", "H", "G", "G", "~
## $ desc_ciiu_nivel_1         <chr> "AGRICULTURA, GANADERÍA, SILVI~
## $ ciuu4_nivel_6             <chr> "A0126.01", "C2410.25", "G4663.~
## $ desc_ciiu_nivel_6        <chr> "Cultivo de palmas de aceite (p~
## $ numero_de_trabajadores_directivos <dbl> 8, 9, 3, 1, 1, 1, 1, 2, 1, 9, 2~
## $ numero_de_trabajadores_administrativos <dbl> 4, 384, 70, NA, 6, 3, NA, 4, 5,~
## $ numero_de_trabajadores_produccion <dbl> 5, 980, 77, NA, NA, NA, NA, 10,~
## $ numero_de_trabajadores_otros <dbl> NA, NA, NA, NA, 19, NA, NA, NA,~
## $ tamano                    <chr> "PEQUEÑA", "GRANDE", "GRANDE", ~
## $ ind_liquidez_corriente    <dbl> 3.15925400, 1.83329981, 3.08762~
## $ ind_endeudamiento_del_activo <dbl> 0.04552082, 0.57575229, 0.67680~
## $ ind_endeudamiento_patrimonial <dbl> 0.04769179, 1.35711351, 2.09407~
## $ ind_endeudamiento_del_activo_fijo <dbl> 0.9962363, 1.1156390, 3.3943681~
## $ ind_apalancamiento        <dbl> 1.047692, 2.357114, 3.094077, --
```

## Parte 3. Análisis

### 3.1. Número de empresas por actividad económica por cantón

```
## Número total de empresas ----
empresas |>
  group_by(canton, ciuu4_nivel_1, desc_ciiu_nivel_1) |>
  count() |>
  rename("Cantón" = canton,
         "Ciiu Nivel 1" = ciuu4_nivel_1,
         "Descripción Ciiu" = desc_ciiu_nivel_1,
         "Número de empresas" = n)
```

```
## # A tibble: 1,309 x 4
## # Groups:   Cantón, Ciiu Nivel 1, Descripción Ciiu [1,309]
##   Cantón      `Ciiu Nivel 1` `Descripción Ciiu` `Número de empresas`
##   <chr>      <chr>      <chr>      <int>
## 1 24 DE MAYO      F      CONSTRUCCIÓN.      1
## 2 24 DE MAYO      G      COMERCIO AL POR M~    1
## 3 24 DE MAYO      H      TRANSPORTE Y ALMA~    4
## 4 AGUARICO        F      CONSTRUCCIÓN.      2
## 5 ALAUSÍ          F      CONSTRUCCIÓN.      1
## 6 ALAUSÍ          H      TRANSPORTE Y ALMA~    3
## 7 ALAUSÍ          I      ACTIVIDADES DE AL~    1
## 8 ALFREDO BAQUERIZO MOR~ A      AGRICULTURA, GANA~    1
## 9 ALFREDO BAQUERIZO MOR~ F      CONSTRUCCIÓN.      1
## 10 ALFREDO BAQUERIZO MOR~ G      COMERCIO AL POR M~    2
## # i 1,299 more rows
```

```
## Número total de cantones de participación ----
```

```
cantones <- empresas |>
  distinct(canton, .keep_all = TRUE)
```

```
## Número total de actividades realizadas ----
```

```
actividades <- empresas |>
  distinct(desc_ciiu_nivel_1, .keep_all = TRUE)
```

**RESULTADO:** Existe un total de **47033** empresas que realizan **22** actividades distintas distribuidas en **217** cantones ecuatorianos.

En el siguiente gráfico se presenta el top 10 de actividades con mayor número de empresas a nivel nacional y su participación en los 4 principales cantones con mayor número de empresas.

```
## GRÁFICO: Top 10 actividades con mayor número de empresas a nivel nacional
## y su participación en los 4 principales cantones.
```

```
top_5_actividades <- empresas |>
  count(ciiu4_nivel_1, desc_ciiu_nivel_1) |>
  slice_max(n, n = 5)

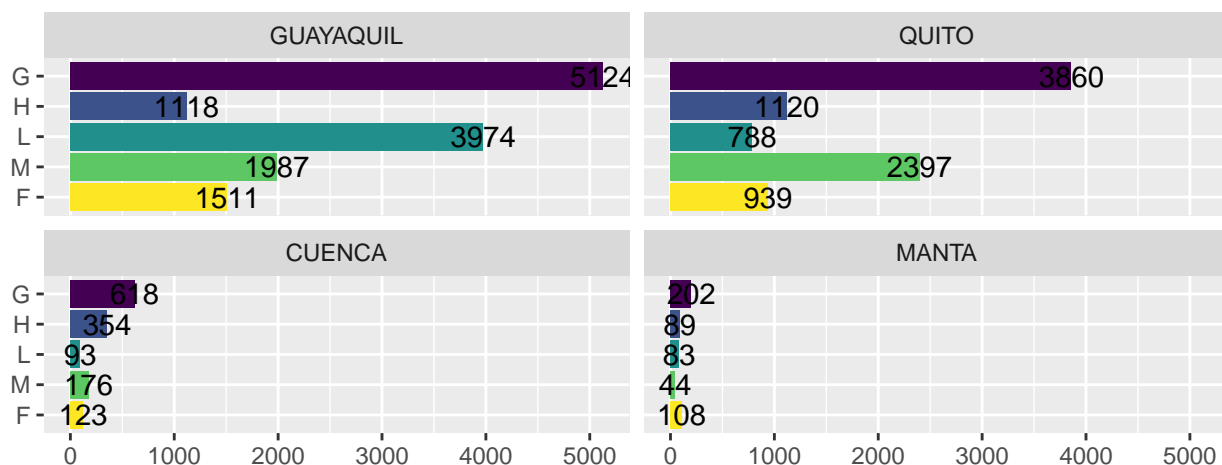
top_4_cantones <- empresas |>
  semi_join(top_5_actividades, by = "ciiu4_nivel_1") |>
  count(canton) |>
  slice_max(n, n = 4)

conteo_empresas <- empresas |>
  semi_join(top_5_actividades, by = "ciiu4_nivel_1") |>
  semi_join(top_4_cantones, by = "canton") |>
  count(canton, ciiu4_nivel_1, desc_ciiu_nivel_1) |>
  mutate(ciiu4_nivel_1 = factor(ciiu4_nivel_1,
                                levels = top_5_actividades$ciiu4_nivel_1,
                                ordered = TRUE),
         canton = factor(canton,
                         levels = top_4_cantones$canton,
                         ordered = TRUE),
         desc_ciiu_nivel_1 = factor(desc_ciiu_nivel_1,
                                     levels = top_5_actividades$desc_ciiu_nivel_1,
```

```
ordered = TRUE))

conteo_empresas |>
  ggplot(aes(n, fct_rev(ciiu4_nivel_1), fill = desc_ciiu_nivel_1)) +
  geom_col() +
  geom_text(aes(label = n)) +
  facet_wrap(~canton) +
  labs(title = "Top 10 actividades con mayor número de empresas a nivel nacional",
        fill = "Actividades") +
  theme(
    legend.position = "bottom",
    legend.direction = "vertical",
    axis.title.x = element_blank(),
    axis.title.y = element_blank())
```

Top 10 actividades con mayor número de empresas a nivel nacional



tividades

- COMERCIO AL POR MAYOR Y AL POR MENOR; REPARACIÓN DE VEHÍCULOS AUTOMOTORES Y MOTOCICL
- TRANSPORTE Y ALMACENAMIENTO.
- ACTIVIDADES INMOBILIARIAS.
- ACTIVIDADES PROFESIONALES, CIENTÍFICAS Y TÉCNICAS.
- CONSTRUCCIÓN.

Guayaquil, Quito, Cuenca y Manta son los principales cantones a nivel nacional que poseen el mayor número de empresas. En términos de actividades, Comercio, Transporte, Actividades Inmobiliarias, Actividades Profesionales y Construcción son las actividades que cuentan con el mayor número de empresas.

### 3.2. Liquidez y solvencia patrimonial por status y provincia

En este gráfico se pretende realizar un resumen en términos de media, máximo y mínimo de los indicadores de liquidez corriente y endeudamiento patrimonial para describir la solvencia, por situación legal y provincia.



```

liquidez_solvencia_provincia_status <- empresas |>
  group_by(provincia, situacion_legal) |>
  summarise(liq_mean = mean(ind_liquidez_corriente, na.rm = TRUE),
            liq_max = max(ind_liquidez_corriente, na.rm = TRUE),
            liq_min = min(ind_liquidez_corriente, na.rm = TRUE),
            sol_mean = mean(ind_endeudamiento_patrimonial, na.rm = TRUE),
            sol_max = max(ind_endeudamiento_patrimonial, na.rm = TRUE),
            sol_min = min(ind_endeudamiento_patrimonial, na.rm = TRUE)) |>
  mutate(across(starts_with(c("liq", "sol")), function(x) {
    case_when(
      is.infinite(x) ~ NA_real_,
      is.nan(x) ~ NA_real_,
      TRUE ~ x
    )
  }))

liquidez_solvencia_provincia_status

```

```

## # A tibble: 117 x 8
## # Groups:   provincia [24]
##   provincia situacion_legal liq_mean liq_max liq_min sol_mean sol_max sol_min
##   <chr>      <chr>          <dbl>  <dbl>  <dbl>   <dbl>   <dbl>  <dbl>
## 1 AZUAY      ACTIVA             110.   9.21e+4 0        25.6    1.46e+4 -1.29e+3
## 2 AZUAY      DISOLUC. LIQUI~     3.98   1.46e+2 0         7.46    5.01e+2 -2.28e+2
## 3 AZUAY      DISOLUC. LIQUI~     6.41   1.15e+1 1.27e+0 -0.368   9.48e-2 -1.20e+0
## 4 AZUAY      DISOLUC. Y LIQ~    79.2   1.16e+3 0         0.937   1.94e+1 -1.82e+0
## 5 AZUAY      DISOLUC. Y LIQ~     0.999  2.49e+0 7.22e-3 -0.873   6.73e-1 -2.01e+0
## 6 AZUAY      INACTIVA           12.1   4.42e+1 1.76e-1 17.4     1.92e+2 0
## 7 AZUAY      LIQUIDAC. DE P~    0.386  3.86e-1 3.86e-1 -2.04    -2.04e+0 -2.04e+0
## 8 BOLIVAR     ACTIVA             21.9   2.83e+2 0         1.03    8.38e+0 -1.55e+0
## 9 BOLIVAR     DISOLUC. LIQUI~     1.23   1.23e+0 1.23e+0 1.45     4.35e+0 0
## 10 BOLIVAR    DISOLUC. LIQUI~    14.6   1.46e+1 1.46e+1 0.0576   5.76e-2 5.76e-2
## # i 107 more rows

```

```

tipo_principal <- empresas |>
  count(situacion_legal) |>
  slice_max(n, n = 1)

top_3_provincias <- empresas |>
  semi_join(tipo_principal, by = "situacion_legal") |>
  count(provincia) |>
  slice_max(n, n = 4)

indicadores_activas <- tipo_principal |>
  left_join(liquidez_solvencia_provincia_status, by = "situacion_legal")

provincias_mayor_participacion <- top_3_provincias |>
  left_join(indicadores_activas, by = "provincia")

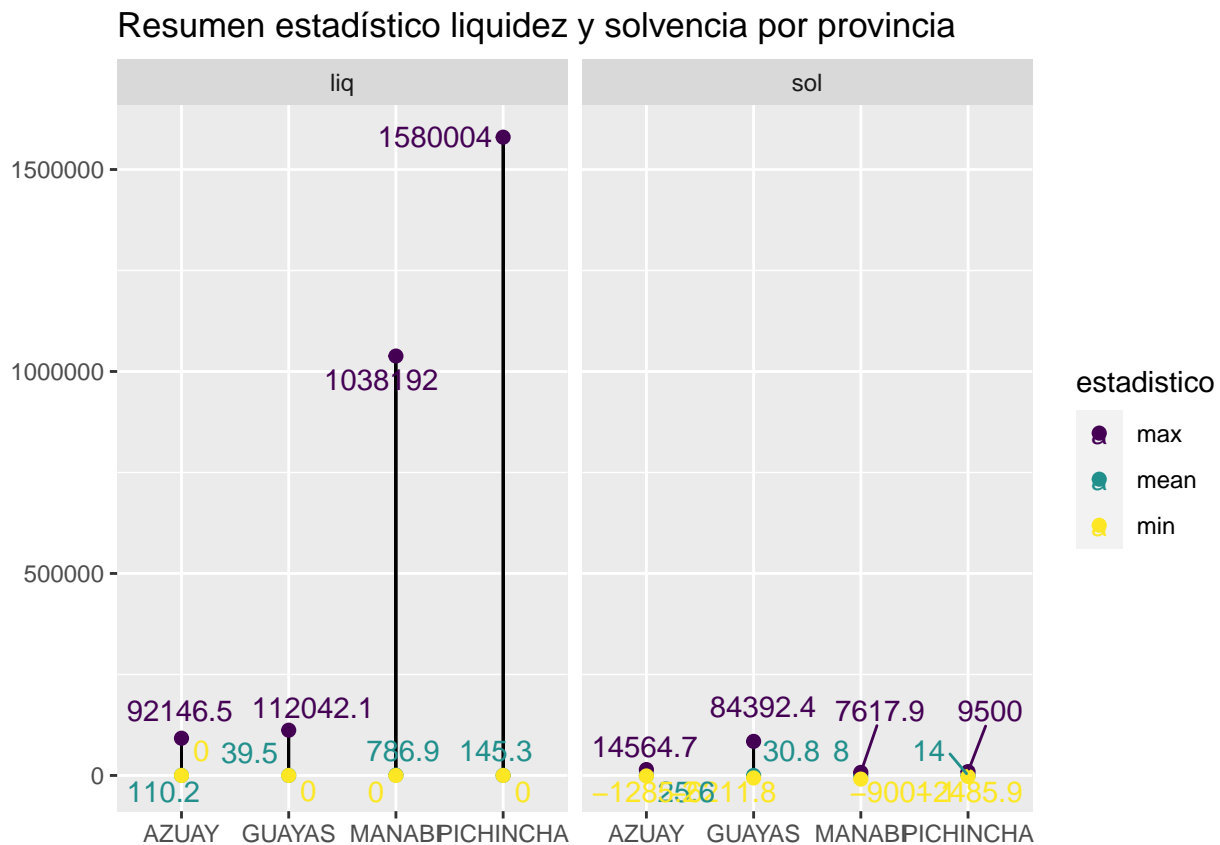
provincias_mayor_participacion |>
  pivot_longer(
    cols = starts_with("liq") | starts_with("sol"),
    names_to = c("indicador", "estadistico"),

```

```

names_sep = "_",
values_to = "valor") |>
ggplot(aes(provincia, valor, label = round(valor, 1))) +
  geom_line(aes(group = provincia)) +
  geom_point(aes(color = estadistico), size = 2) +
  geom_text_repel(aes(color = estadistico)) +
  scale_color_viridis_d() +
  facet_wrap(~indicador) +
  labs(title = "Resumen estadístico liquidez y solvencia por provincia",
       fill = "Estadísticos") +
  theme(
    legend.position = "right",
    legend.direction = "vertical",
    axis.title.x = element_blank(),
    axis.title.y = element_blank())

```



### ## 3.3. Liquidez y solvencia patrimonial (endeudamiento patrimonial) por tipo de empresa

En este gráfico se pretende mostrar la liquidez y solvencia (tomando en cuenta el endeudamiento patrimonial) por tipo de empresa.

```

liquidez_solvencia_tipo_empresa <- empresas |>
  group_by(tipo) |>
  summarise(
    liq_mean = min(ind_liquidez_corriente, na.rm = TRUE),
    liq_max = max(ind_liquidez_corriente, na.rm = TRUE),
    liq_min = min(ind_liquidez_corriente, na.rm = TRUE),
    sol_mean = mean(ind_endeudamiento_patrimonial, na.rm = TRUE),

```

```

      sol_max = max(ind_endeudamiento_patrimonial, na.rm = TRUE),
      sol_min = min(ind_endeudamiento_patrimonial, na.rm = TRUE)) |>
mutate(across(starts_with("ind"), function(x) {
  case_when(
    is.infinite(x) ~ NA_real_,
    TRUE ~ x
  )
}))
)

liquidez_solvencia_tipo_empresa

```

```

## # A tibble: 7 x 7
##   tipo                                liq_mean liq_max liq_min sol_mean sol_max sol_min
##   <chr>                                <dbl>   <dbl>   <dbl>   <dbl>   <dbl>   <dbl>
## 1 ANÓNIMA                             0       1.04e6  0        23.0   84392. -9.00e+3
## 2 ANÓNIMA EN PREDIOS RÚSTIC~ 0.0227  2.03e1  0.0227   28.7    413.   5.65e-4
## 3 ANÓNIMA MULTINACIONAL ANDI~ 0.556   3.93e1  0.556    4.45    14.8   2.61e-2
## 4 ASOCIACIÓN O CONSORCIO         0       5.45e0  0       -0.313   68.8 -7.01e+1
## 5 ECONOMÍA MIXTA                 0.501   1.03e2  0.501    1.79    13.7   0
## 6 RESPONSABILIDAD LIMITADA       0       1.58e6  0        13.5   18945. -1.80e+3
## 7 SUCURSAL EXTRANJERA            0       7.99e3  0        56.0   14565. -6.63e+1

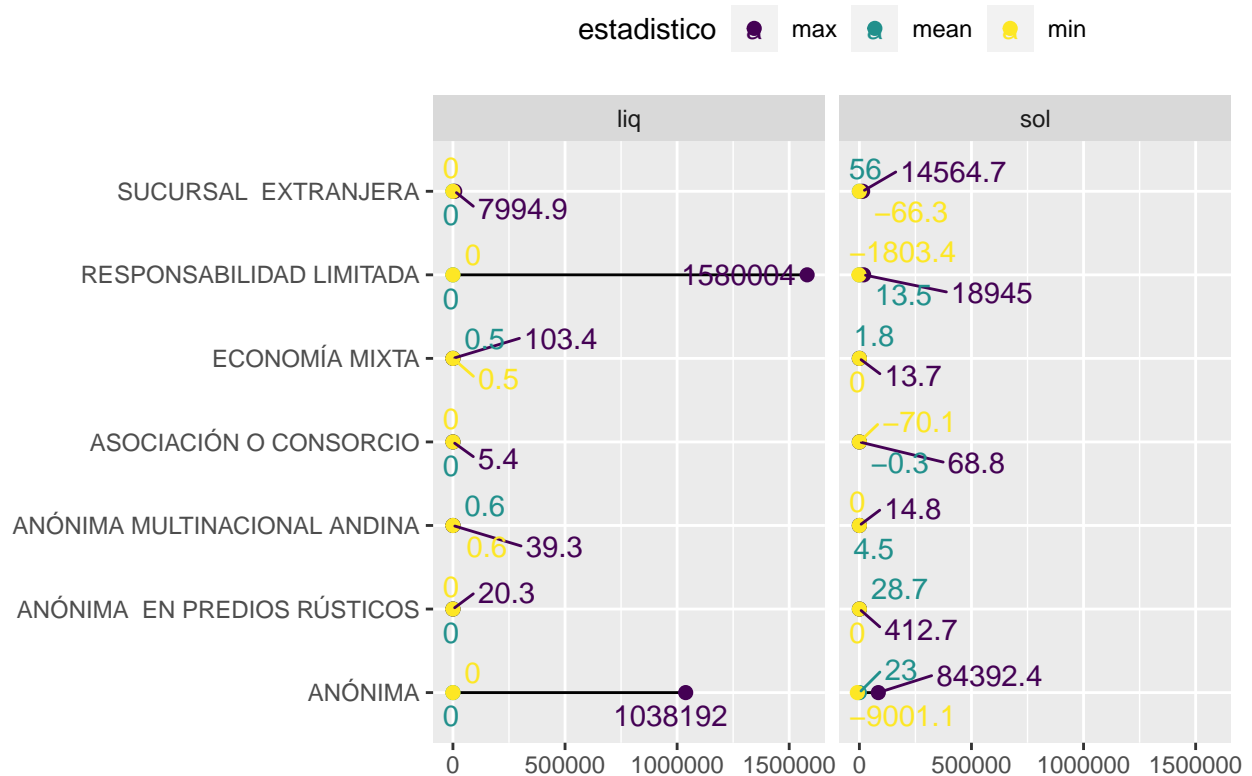
```

```

## Liquidez y solvencia por tipo de empresa
liquidez_solvencia_tipo_empresa |>
  pivot_longer(
    cols = starts_with("liq") | starts_with("sol"),
    names_to = c("indicador", "estadistico"),
    names_sep = "_",
    values_to = "valor") |>
  ggplot(aes(valor, tipo, label = round(valor, 1))) +
  geom_line(aes(group = tipo)) +
  geom_point(aes(color = estadistico), size = 2) +
  geom_text_repel(aes(color = estadistico)) +
  scale_color_viridis_d() +
  facet_wrap(~indicador) +
  labs(title = "Resumen estadístico liquidez y solvencia por tipo de empresa") +
  theme(
    legend.position = "top",
    legend.direction = "horizontal",
    axis.title.x = element_blank(),
    axis.title.y = element_blank())

```

## Resumen estadístico liquidez y solvencia por tipo de



### 3.4. Endeudamiento del activo por tamaño de empresa

En este gráfico se pretende mostrar el endeudamiento del activo por tamaño de empresa (grande, mediana, micro, no definido y pequeña).

```
endeudamiento_tamano_empresa <- empresas |>
  group_by(tamano) |>
  summarise(end_mean = mean(ind_endeudamiento_del_activo, na.rm = TRUE),
            end_max = max(ind_endeudamiento_del_activo, na.rm = TRUE),
            end_min = min(ind_endeudamiento_del_activo, na.rm = TRUE))
```

```
endeudamiento_tamano_empresa
```

```
## # A tibble: 5 x 4
##   tamano      end_mean end_max end_min
##   <chr>      <dbl>   <dbl>   <dbl>
## 1 GRANDE      1.23    687.      0
## 2 MEDIANA     0.721    85.0      0
## 3 MICRO       4.97   33061.     0
## 4 NO DEFINIDO  1.03     170.      0
## 5 PEQUEÑA     1.23    3437.      0
```

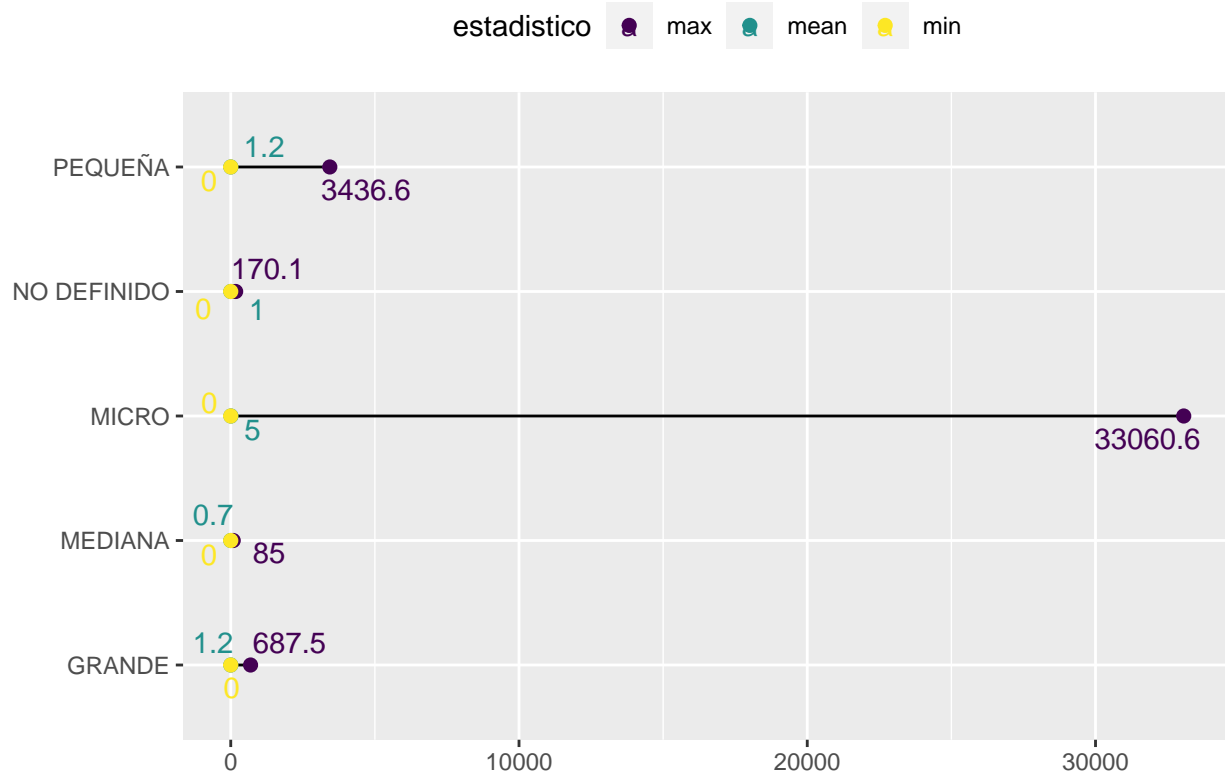
```
## Endeudamiento del activo por tamaño de empresa
endeudamiento_tamano_empresa |>
```

```

pivot_longer(
  cols = starts_with("end"),
  names_to = c("indicador", "estadistico"),
  names_sep = "_",
  values_to = "valor") |>
ggplot(aes(valor, tamano, label = round(valor, 1))) +
  geom_line(aes(group = tamano)) +
  geom_point(aes(color = estadistico), size = 2) +
  geom_text_repel(aes(color = estadistico)) +
  scale_color_viridis_d() +
  labs(title = "Resumen estadístico endeudamiento por tamaño de empresa",
       fill = "Actividades") +
  theme(
    legend.position = "top",
    legend.direction = "horizontal",
    axis.title.x = element_blank(),
    axis.title.y = element_blank())

```

## Resumen estadístico endeudamiento por tamaño de empresa



**RESULTADO:** Con el resultado expuesto en el gráfico anterior, se puede observar que el promedio de endeudamiento del activo de las empresas micro y pequeñas es mayor al endeudamiento del activo de las empresas grandes.

### 3.5. Liquidez por número de trabajadores

En este gráfico se pretende observar la liquidez corriente por número de empleados que trabaja en cada empresa. Para este análisis se toma en cuenta a las empresas que cumplen los dos criterios de tener más de 60 trabajadores

directos y que contar con 100 a 800 trabajadores administrativos y se las cataloga como “Cumple ambos”; por otro lado se encuentran las empresas que cumplen con el criterio de tener más de 60 trabajadores directos y se las cataloga como “Cumple criterio 1”; y por último se toman en cuenta a las empresas que cumplen con el criterio de contar con 100 a 800 trabajadores administrativos a la cual se le cataloga como “Cumple criterio 2”; a todas las demás que no se encuentran dentro de este grupo se les cataloga como “No cumple”.

```
liquidez_trabajadores_empresa <- empresas |>
  mutate(tipo_por_trabajadores = case_when(
    is.na(numero_de_trabajadores_directivos) &
    is.na(numero_de_trabajadores_administrativos) ~ NA,
    numero_de_trabajadores_produccion > 60 &
    between(numero_de_trabajadores_administrativos, 100, 800) ~ "Cumple ambos",
    numero_de_trabajadores_produccion > 60 ~ "Cumple criterio 1",
    between(numero_de_trabajadores_administrativos, 100, 800) ~ "Cumple criterio 2",
    TRUE ~ "No cumple"
  )) |>
  group_by(tipo_por_trabajadores) |>
  summarise(liq_mean = round(mean(ind_liquidez_corriente, na.rm = TRUE), 1),
            n = n())

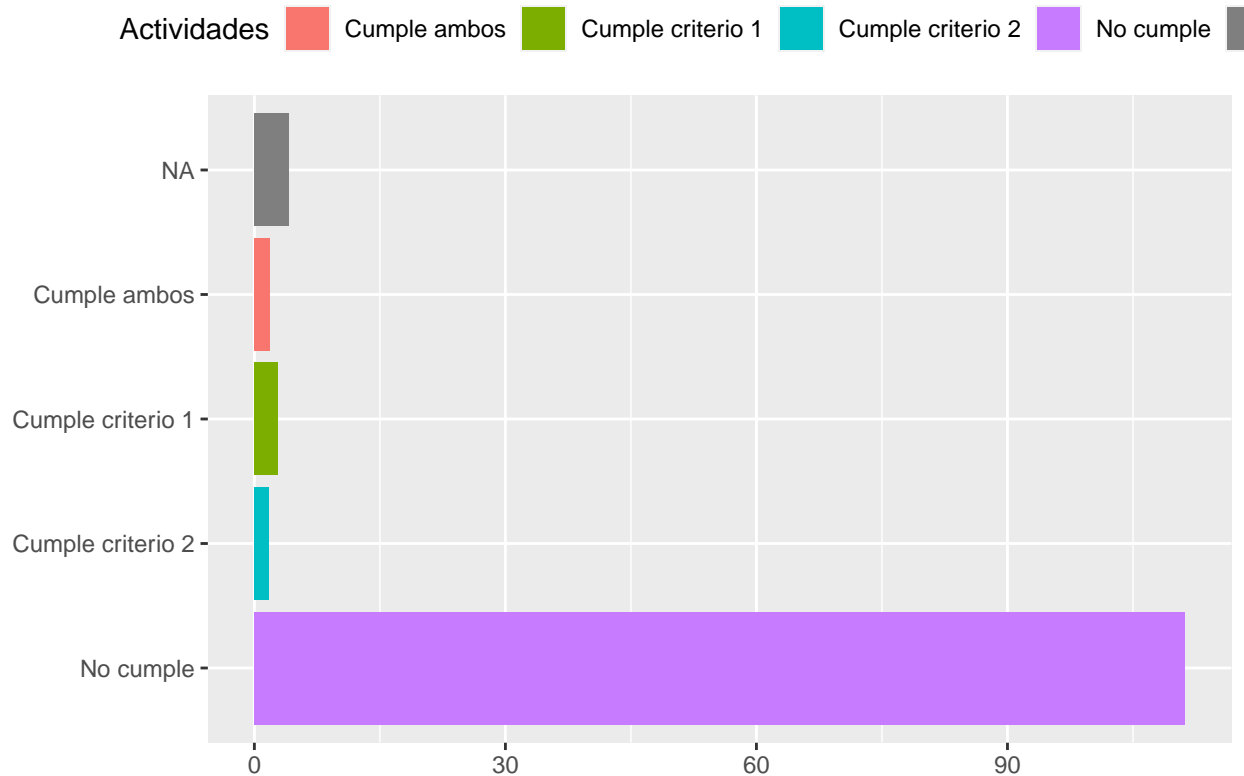
liquidez_trabajadores_empresa
```

```
## # A tibble: 5 x 3
##   tipo_por_trabajadores liq_mean    n
##   <chr>                <dbl> <int>
## 1 Cumple ambos         1.8    147
## 2 Cumple criterio 1    2.8   1317
## 3 Cumple criterio 2    1.7    79
## 4 No cumple           111.  45328
## 5 <NA>                 4.1    162
```

*## Liquidez por número de trabajadores*

```
liquidez_trabajadores_empresa |>
  ggplot(aes(liq_mean,
             fct_rev(tipo_por_trabajadores),
             fill = tipo_por_trabajadores)) +
  geom_col() +
  scale_color_viridis_d() +
  labs(title = "Promedio liquidez corriente por número de trabajadores",
       fill = "Actividades") +
  theme(
    legend.position = "top",
    legend.direction = "horizontal",
    axis.title.x = element_blank(),
    axis.title.y = element_blank())
```

## Promedio liquidez corriente por número de trabajadores



**RESULTADO:** La liquidez corriente promedio es mayor en las empresas que no cumplen con ningún criterio, sin embargo, las empresas que cumplen con tener más de 60 trabajadores directos tienen la liquidez corriente promedio más alta en términos de los demás criterios descritos en el primer párrafo de este análisis.

### 3.6. Apalancamiento por empresas

En este análisis se pretende observar el top 10 de empresas que cuentan con el volumen más alto de apalancamiento.

```
apalancamiento_empresa <- empresas |>
  select(nombre_de_la_compania, ind_apalancamiento) |>
  slice_max(ind_apalancamiento, n = 10) |>
  mutate(
    ind_apalancamiento = round(ind_apalancamiento, 1),
    ranking = min_rank(-ind_apalancamiento))

apalancamiento_empresa
```

##	nombre_de_la_compania	ind_apalancamiento	ranking
## 1	ADELCA DEL LITORAL S.A.	84393.4	1
## 2	FARIBALL HOLDING CORP. C.A. FARIBALLCORP	28877.5	2
## 3	HIROAKY S.A.	26048.3	3
## 4	MEGATROPIC S.A.	23802.6	4
## 5	MOCANATI CIA.LTDA.	18946.0	5
## 6	MINERA EL PARAISO S.A. MINELPARSA	17063.5	6
## 7	VOITH HYDRO LTDA.	14565.7	7

## 8	HOLDING IN FERLE HOLFERLE S.A.	11042.6	8
## 9	TERRARIUM S.A.	10272.5	9
## 10	ECUADESK S.A.	9601.9	10

**RESULTADO:** El resultado de este análisis lo demuestro en el siguiente gráfico que contiene las 10 empresas con el volumen más alto de apalancamiento.

```

apalancamiento_empresa |>
  ggplot(aes(ind_apalancamiento, fct_inorder(nombre_de_la_compania) |>
    fct_rev(), fill = nombre_de_la_compania)) +
  geom_col() +
  scale_fill_viridis_d() +
  labs(title = "Top 10 empresas con mayor volumen de apalancamiento") +
  theme(legend.position = "none",
    axis.title.x = element_blank(),
    axis.title.y = element_blank())

```

