

Encuesta Centro Colombo Americano

INGRITH KATHERINE LLANOS DUSSAN

Maestría en Inteligencia de Negocios

Trabajo dirigido: Diseño de un modelo de inteligencia de negocios centralizado basado en el modelo AMAR del Centro Colombo Americano

Table of Contents

Contextualización.....	1
Elección de variables.....	1
Variables.....	2
Resumen estadístico.....	3

Contextualización.

El objetivo de la aplicación del instrumento de medición interna es evaluar la percepción de los colaboradores de la Gerencia de Planeación y Desarrollo y Gerencia de Servicio Compartidos del Centro Colombo Americano sobre la efectividad de los procesos de gestión de datos dentro de la institución, de acuerdo a las variables: generación y recolección de datos, interoperabilidad de herramientas, agregación, explotación, y gobierno de datos. Mediante un análisis de medidas de tendencia central y medidas de variabilidad a todas las respuestas, para luego identificar fortalezas y áreas de oportunidad en la gestión de la información, permitiendo así fundamentar el diseño de un modelo de inteligencia de negocios centralizado. Este modelo no solo optimizará la gestión de datos y mejorará la toma de decisiones, sino que también alineará las prácticas institucionales con los objetivos estratégicos del modelo AMAR, promoviendo una cultura de datos sostenible en el tiempo.

Elección de variables.

Variable	Dimensión	Código	Precedencia	Es	Naturaleza
Generación y recolección de datos	Percepción sobre la efectividad de datos integrales	V1	P1 - P4	Ordinal	Cualitativa

Variable	Dimensión	Código	Preferencias	Es cala	Naturaleza
Interoperabilidad de las herramientas	Percepción de la integración y compatibilidad de los sistemas	V2	P5 - P7	Ordinal	Cualitativa
Agregación de datos	Consolidación y calidad de los datos	V3	P8 - P11	Ordinal	Cualitativa
Explotación de datos	Percepción sobre la capacidad analítica y utilización efectiva de datos	V4	P12 - P15	Ordinal	Cualitativa
Gobierno de datos	Percepción sobre la efectividad de la gestión de los datos	V5	P16 - P18	Ordinal	Cualitativa

Variables

Con La función `srt()` se puede visualizar La estructura de La base de datos

```
str(Data_CCA)
```

```
## tibble [29 × 19] (S3: tbl_df/tbl/data.frame)
## $ Id : num [1:29] 1 2 3 4 5 6 7 8 9 10 ...
## $ P1 : num [1:29] 4 5 4 4 5 4 4 4 1 4 ...
## $ P2 : num [1:29] 3 3 4 5 5 5 3 4 1 2 ...
## $ P3 : num [1:29] 3 2 4 4 4 4 3 3 1 2 ...
## $ P4 : num [1:29] 4 3 5 4 4 5 4 3 1 3 ...
## $ P5 : num [1:29] 2 4 4 5 5 4 2 2 2 2 ...
## $ P6 : num [1:29] 2 3 4 5 5 4 2 3 2 1 ...
## $ P7 : num [1:29] 1 3 4 5 4 4 1 3 1 2 ...
## $ P8 : num [1:29] 3 3 4 4 4 4 3 3 2 2 ...
## $ P9 : num [1:29] 4 3 4 5 4 5 4 4 2 3 ...
## $ P10: num [1:29] 4 4 4 5 5 4 4 3 2 2 ...
## $ P11: num [1:29] 3 3 4 4 4 4 3 3 2 2 ...
## $ P12: num [1:29] 3 5 4 4 4 4 3 3 2 1 ...
## $ P13: num [1:29] 4 4 4 5 4 4 4 4 2 2 ...
## $ P14: num [1:29] 4 3 4 4 5 4 4 4 2 2 ...
## $ P15: num [1:29] 4 4 4 4 5 4 4 5 2 2 ...
## $ P16: num [1:29] 2 3 4 5 4 4 2 2 2 2 ...
```

```
## $ P17: num [1:29] 2 5 4 5 5 5 2 2 3 3 ...
## $ P18: num [1:29] 3 4 4 4 5 4 3 2 2 2 ...
```

Con La función names() Los nombres de Las variables.

```
names(Data_CCA)
```

```
## [1] "Id" "P1" "P2" "P3" "P4" "P5" "P6" "P7" "P8" "P9" "P10"
"P11"
## [13] "P12" "P13" "P14" "P15" "P16" "P17" "P18"
```

Una vez hayan explorado estas dos funciones, oculten este bloque de código.

Resumen estadístico

```
summary(Data_CCA)
```

```
##      Id          P1          P2          P3          P4
## Min.   : 1   Min.   :1.000   Min.   :1.000   Min.   :1.000   Min.
:1.000
## 1st Qu.: 8   1st Qu.:4.000   1st Qu.:3.000   1st Qu.:3.000   1st
Qu.:3.000
## Median :15   Median :4.000   Median :4.000   Median :4.000   Median
:4.000
## Mean   :15   Mean    :4.034   Mean    :3.586   Mean    :3.552   Mean
:3.759
## 3rd Qu.:22   3rd Qu.:4.000   3rd Qu.:4.000   3rd Qu.:4.000   3rd
Qu.:4.000
## Max.   :29   Max.    :5.000   Max.    :5.000   Max.    :5.000   Max.
:5.000
##      P5          P6          P7          P8
P9
## Min.   :2.000   Min.   :1.000   Min.   :1.000   Min.   :2.000   Min.
:2
## 1st Qu.:2.000   1st Qu.:2.000   1st Qu.:2.000   1st Qu.:3.000   1st
Qu.:4
## Median :3.000   Median :3.000   Median :3.000   Median :3.000
Median :4
## Mean   :3.138   Mean    :2.759   Mean    :2.862   Mean    :3.241   Mean
:4
## 3rd Qu.:4.000   3rd Qu.:4.000   3rd Qu.:4.000   3rd Qu.:4.000   3rd
Qu.:5
## Max.   :5.000   Max.    :5.000   Max.    :5.000   Max.    :5.000   Max.
:5
##      P10         P11         P12         P13
## Min.   :2.000   Min.   :2.000   Min.   :1.000   Min.   :1.000
## 1st Qu.:3.000   1st Qu.:3.000   1st Qu.:3.000   1st Qu.:3.000
## Median :4.000   Median :4.000   Median :4.000   Median :4.000
## Mean   :3.931   Mean    :3.483   Mean    :3.621   Mean    :3.414
```

```

## 3rd Qu.:5.000 3rd Qu.:4.000 3rd Qu.:4.000 3rd Qu.:4.000
## Max. :5.000 Max. :5.000 Max. :5.000 Max. :5.000
## P14 P15 P16 P17
P18
## Min. :2.000 Min. :1.000 Min. :2.000 Min. :1.000 Min.
:1
## 1st Qu.:3.000 1st Qu.:3.000 1st Qu.:2.000 1st Qu.:3.000 1st
Qu.:2
## Median :4.000 Median :4.000 Median :3.000 Median :3.000
Median :3
## Mean :3.448 Mean :3.552 Mean :3.172 Mean :3.241 Mean
:3
## 3rd Qu.:4.000 3rd Qu.:4.000 3rd Qu.:4.000 3rd Qu.:4.000 3rd
Qu.:4
## Max. :5.000 Max. :5.000 Max. :5.000 Max. :5.000 Max.
:5

```

Calcular el coeficiente alfa de Cronbach para Las variables P1 a P18

```

#install.packages("psych")
#library(psych)

```

```

cronbach_result <- psych::alpha(Data_CCA[, -1])

```

```

# Mostrar el resultado
print(cronbach_result)

```

```

##
## Reliability analysis
## Call: psych::alpha(x = Data_CCA[, -1])
##
## raw_alpha std.alpha G6(smc) average_r S/N ase mean sd median_r
## 0.93 0.93 0.97 0.42 13 0.019 3.4 0.65 0.43
##
## 95% confidence boundaries
## lower alpha upper
## Feldt 0.88 0.93 0.96
## Duhachek 0.89 0.93 0.97
##
## Reliability if an item is dropped:
## raw_alpha std.alpha G6(smc) average_r S/N alpha se var.r med.r
## P1 0.93 0.93 0.97 0.44 13 0.019 0.018 0.44
## P2 0.92 0.92 0.97 0.42 12 0.021 0.021 0.43
## P3 0.92 0.93 0.97 0.43 13 0.020 0.020 0.44
## P4 0.92 0.93 0.97 0.43 13 0.020 0.020 0.44
## P5 0.92 0.92 0.97 0.42 12 0.021 0.020 0.43
## P6 0.92 0.93 0.97 0.42 12 0.021 0.021 0.44
## P7 0.92 0.92 0.97 0.42 12 0.021 0.021 0.41
## P8 0.92 0.92 0.96 0.41 12 0.021 0.020 0.41
## P9 0.92 0.93 0.97 0.42 12 0.021 0.021 0.43

```

```

## P10      0.92      0.93      0.97      0.42  13      0.020 0.021  0.44
## P11      0.92      0.92      0.97      0.42  12      0.021 0.020  0.42
## P12      0.92      0.93      0.97      0.43  13      0.020 0.022  0.44
## P13      0.93      0.93      0.97      0.43  13      0.020 0.018  0.43
## P14      0.92      0.93      0.97      0.42  12      0.020 0.021  0.44
## P15      0.93      0.93      0.97      0.43  13      0.020 0.020  0.44
## P16      0.92      0.93      0.97      0.42  12      0.021 0.021  0.42
## P17      0.92      0.93      0.97      0.43  13      0.020 0.019  0.44
## P18      0.92      0.93      0.97      0.42  12      0.021 0.021  0.43

```

```
##
```

```
## Item statistics
```

```

##      n raw.r std.r r.cor r.drop mean  sd
## P1  29  0.46  0.48  0.46  0.41  4.0 0.82
## P2  29  0.74  0.74  0.73  0.70  3.6 1.02
## P3  29  0.64  0.65  0.64  0.59  3.6 0.91
## P4  29  0.63  0.65  0.64  0.59  3.8 0.87
## P5  29  0.73  0.72  0.72  0.68  3.1 1.03
## P6  29  0.72  0.69  0.68  0.66  2.8 1.30
## P7  29  0.77  0.75  0.74  0.72  2.9 1.25
## P8  29  0.82  0.82  0.82  0.79  3.2 0.79
## P9  29  0.70  0.72  0.71  0.66  4.0 0.85
## P10 29  0.66  0.67  0.67  0.61  3.9 0.96
## P11 29  0.74  0.76  0.75  0.71  3.5 0.74
## P12 29  0.62  0.63  0.60  0.57  3.6 0.86
## P13 29  0.58  0.59  0.58  0.52  3.4 0.95
## P14 29  0.67  0.69  0.68  0.63  3.4 0.78
## P15 29  0.57  0.56  0.54  0.51  3.6 1.02
## P16 29  0.72  0.71  0.71  0.67  3.2 1.10
## P17 29  0.65  0.62  0.61  0.59  3.2 1.06
## P18 29  0.72  0.71  0.70  0.67  3.0 1.10

```

```
##
```

```
## Non missing response frequency for each item
```

```

##      1  2  3  4  5 miss
## P1  0.03 0.00 0.10 0.62 0.24  0
## P2  0.03 0.07 0.38 0.31 0.21  0
## P3  0.03 0.07 0.31 0.48 0.10  0
## P4  0.03 0.00 0.31 0.48 0.17  0
## P5  0.00 0.34 0.28 0.28 0.10  0
## P6  0.21 0.24 0.24 0.21 0.10  0
## P7  0.21 0.14 0.31 0.28 0.07  0
## P8  0.00 0.17 0.45 0.34 0.03  0
## P9  0.00 0.07 0.14 0.52 0.28  0
## P10 0.00 0.10 0.17 0.41 0.31  0
## P11 0.00 0.10 0.34 0.52 0.03  0
## P12 0.03 0.03 0.31 0.52 0.10  0
## P13 0.03 0.17 0.17 0.59 0.03  0
## P14 0.00 0.14 0.31 0.52 0.03  0
## P15 0.03 0.14 0.21 0.48 0.14  0
## P16 0.00 0.38 0.21 0.28 0.14  0

```

```

## P17 0.03 0.17 0.48 0.14 0.17 0
## P18 0.10 0.21 0.34 0.28 0.07 0

library(dplyr)

# Subconjunto de datos para el grupo Generacion_Recoleccion y Las
preguntas P1 a P4
data_subsetv1 <- Data_CCA %>%
  select(starts_with("P")) %>%
  mutate_if(is.character, as.factor)

# Crear una función para calcular Las métricas
calcular_metricasv1 <- function(preguntav1) {
  data.frame(
    Preguntav1 = preguntav1,
    Validov1 = sum(!is.na(data_subsetv1[[preguntav1]])),
    Perdidov1 = sum(is.na(data_subsetv1[[preguntav1]])),
    Mediav1 = mean(data_subsetv1[[preguntav1]], na.rm = TRUE),
    Medianav1 = median(data_subsetv1[[preguntav1]], na.rm = TRUE),
    Modav1 = DescTools::Mode(data_subsetv1[[preguntav1]]),
    Desviacion_Estandarv1 = sd(data_subsetv1[[preguntav1]], na.rm =
TRUE),
    Varianzav1 = var(data_subsetv1[[preguntav1]], na.rm = TRUE)
  )
}

# Calcular Las métricas para Las preguntas P1 a P4
resultadosv1 <- lapply(c("P1", "P2", "P3", "P4"), calcular_metricasv1)
%>%
  bind_rows()

# Mostrar Los resultados
print(resultadosv1)

## Preguntav1 Validov1 Perdidov1 Mediav1 Medianav1 Modav1
Desviacion_Estandarv1
## 1 P1 29 0 4.034483 4 4
0.8230066
## 2 P2 29 0 3.586207 4 3
1.0183054
## 3 P3 29 0 3.551724 4 4
0.9097177
## 4 P4 29 0 3.758621 4 4
0.8724011
## Varianzav1
## 1 0.6773399
## 2 1.0369458
## 3 0.8275862
## 4 0.7610837

```

```

# Crear Gráfica de barras para cada pregunta
for (preguntav1 in c("P1", "P2", "P3", "P4")) {
  # Calcular La media para La pregunta actual
  media_preguntav1 <- mean(data_subsetv1[[preguntav1]], na.rm = TRUE)

  # Crear La Gráfica de barras
  pv1 <- ggplot(data_subsetv1, aes(x = !!sym(preguntav1))) +
    geom_histogram(aes(y = ..density..), binwidth = 1, fill = "#2F5597",
color = "#2F5597", alpha = 0.7) +
    geom_density(color = "#446C88", size = 1) +
    geom_vline(aes(xintercept = media_preguntav1), color = "#B4C7E7",
linetype = "dashed", size = 1) +
    labs(
      title = paste("Gráfica de barras de generacion y recoleccion de
datos Pregunta", preguntav1),
      x = preguntav1,
      y = "Densidad"
    ) +
    theme_minimal()

  # Mostrar La gráfica de barras en el dispositivo gráfico
  print(pv1)

  # Guardar La gráfica de barras como imagen
  ggsave(filename = paste0("histograma_", preguntav1, ".png"), plot =
pv1)
}

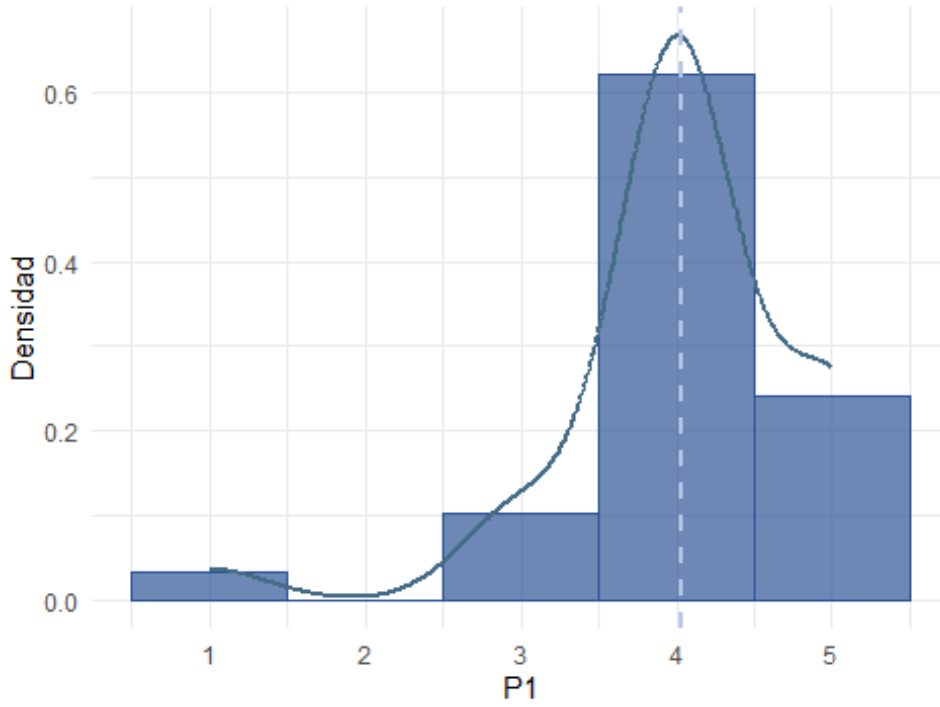
## Warning: Using `size` aesthetic for lines was deprecated in ggplot2
3.4.0.
## i Please use `linewidth` instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning
was
## generated.

## Warning: The dot-dot notation (`..density..`) was deprecated in
ggplot2 3.4.0.
## i Please use `after_stat(density)` instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning
was
## generated.

## Saving 5 x 4 in image

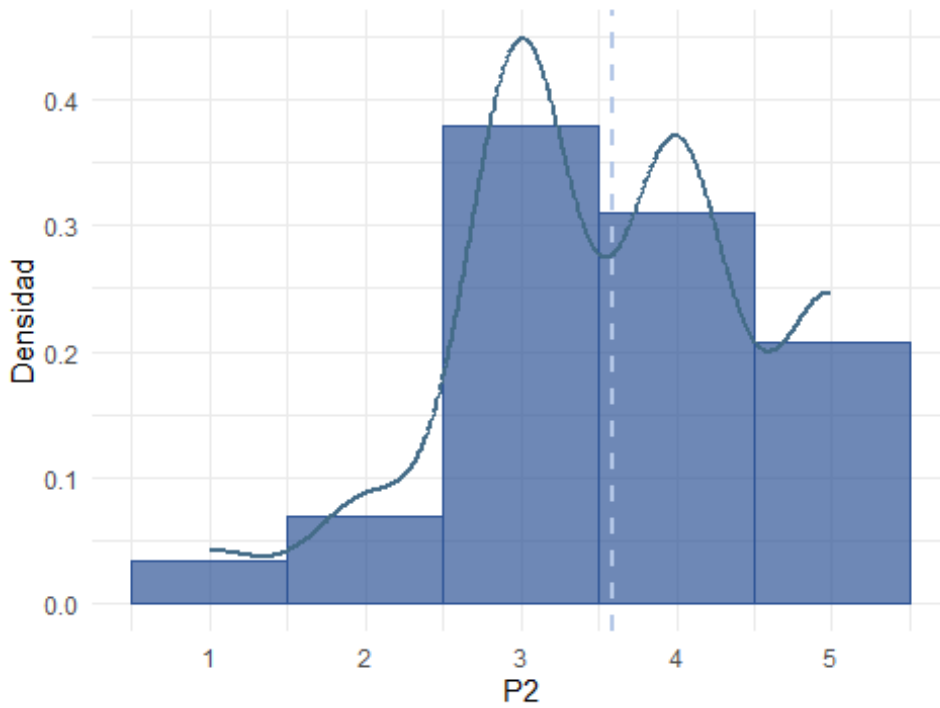
```

Gráfica de barras de generación y recolección de dato



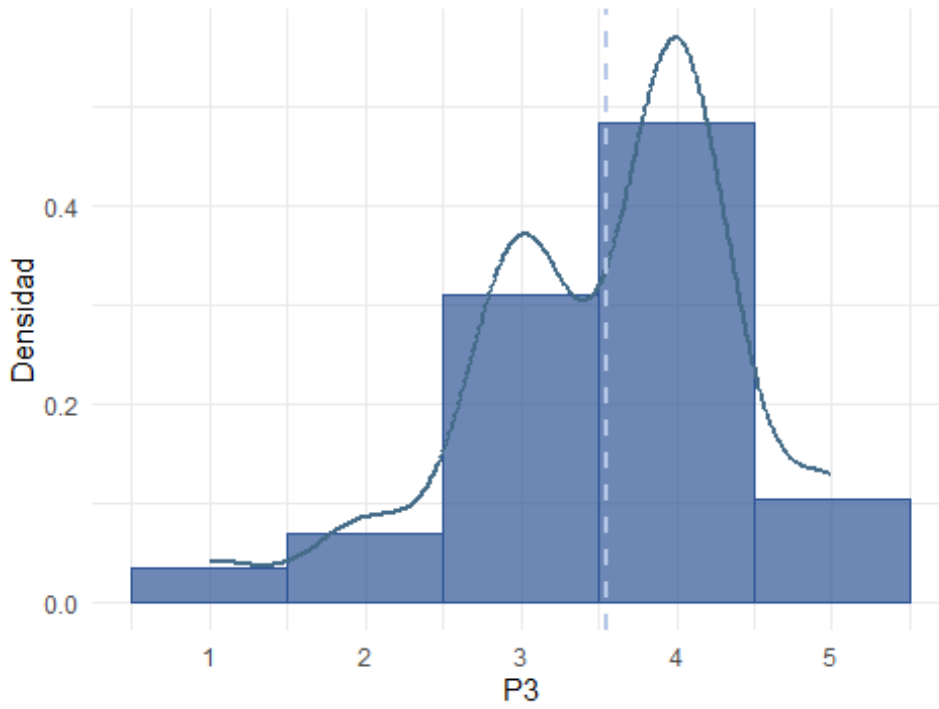
Saving 5 x 4 in image

Gráfica de barras de generación y recolección de dato



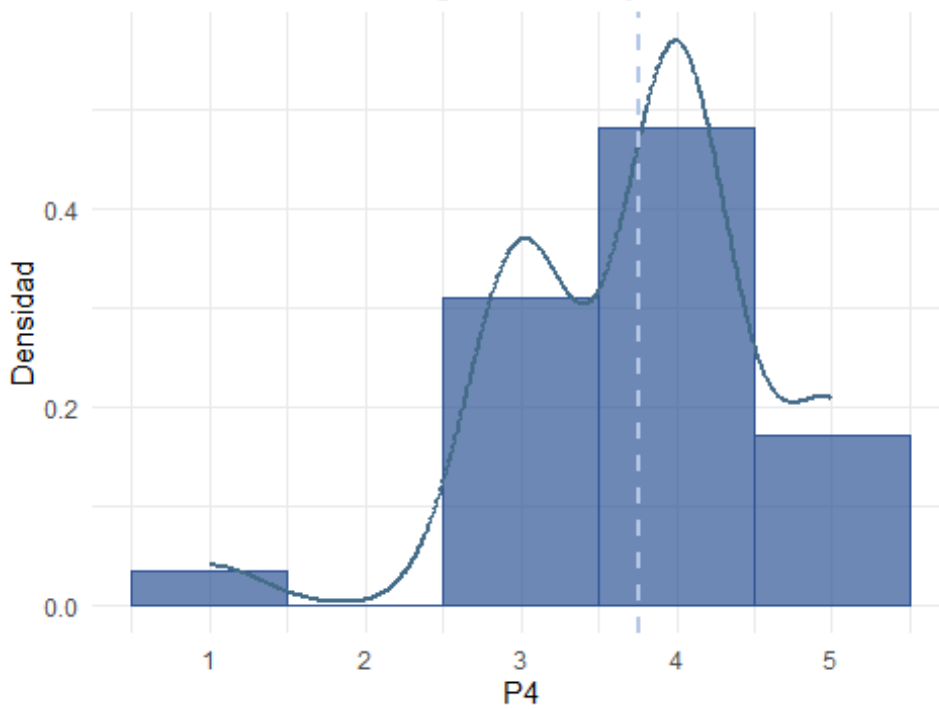
Saving 5 x 4 in image

Gráfica de barras de generacion y recoleccion de dato



Saving 5 x 4 in image

Gráfica de barras de generacion y recoleccion de dato



Subconjunto de datos para el grupo Interoperabilidad de Las herramientas y Las preguntas P5 a P7

```

data_subsetv2 <- Data_CCA %>%
  select(starts_with("P")) %>%
  mutate_if(is.character, as.factor)

# Crear una función para calcular Las métricas
calcular_metricasv2 <- function(preguntav2) {
  modav2 <- DescTools::Mode(data_subsetv2[[preguntav2]])
  if (length(modav2) > 1) {
    modav2 <- paste(modav2, collapse = ", ")
  } else {
    modav2 <- as.character(modav2) # Asegurar que modav2 sea de tipo
character
  }
  data.frame(
    Preguntav2 = preguntav2,
    Validov2 = sum(!is.na(data_subsetv2[[preguntav2]])),
    Perdido2 = sum(is.na(data_subsetv2[[preguntav2]])),
    Mediav2 = mean(data_subsetv2[[preguntav2]], na.rm = TRUE),
    Medianav2 = median(data_subsetv2[[preguntav2]], na.rm = TRUE),
    Modav2 = modav2,
    Desviacion_Estandarv2 = sd(data_subsetv2[[preguntav2]], na.rm =
TRUE),
    Varianzav2 = var(data_subsetv2[[preguntav2]], na.rm = TRUE)
  )
}

# Calcular Las métricas para Las preguntas P5 a P7
resultadosv2 <- lapply(c("P5", "P6", "P7"), calcular_metricasv2)

# Convertir La Lista a un data frame, manejar casos donde no hay moda
encontrada
resultadosv2 <- bind_rows(resultadosv2)

# Mostrar Los resultados
print(resultadosv2)

##   Preguntav2 Validov2 Perdido2   Mediav2 Medianav2 Modav2
Desviacion_Estandarv2
## 1          P5       29         0 3.137931         3      2
1.025536
## 2          P6       29         0 2.758621         3     2, 3
1.299867
## 3          P7       29         0 2.862069         3      3
1.245682
##   Varianzav2
## 1  1.051724
## 2  1.689655
## 3  1.551724

```

```

# Crear gráfica de barras para cada pregunta con línea de media y
densidad de probabilidad
for (preguntav2 in c("P5", "P6", "P7")) {
  # Calcular La media para La pregunta actual
  media_preguntav2 <- mean(data_subsetv2[[preguntav2]], na.rm = TRUE)

  # Crear La gráfica de barras
  pv2 <- ggplot(data_subsetv2, aes(x = !!sym(preguntav2))) +
    geom_histogram(aes(y = ..density..), binwidth = 1, fill = "#2F5597",
color = "#2F5597", alpha = 0.7) +
    geom_density(color = "#446C88", size = 1) +
    geom_vline(aes(xintercept = media_preguntav2), color = "#B4C7E7",
linetype = "dashed", size = 1) +
    labs(
      title = paste("Gráfica de barras de interoperabilidad de las
herramientas Pregunta", preguntav2),
      x = preguntav2,
      y = "Densidad"
    ) +
    theme_minimal()

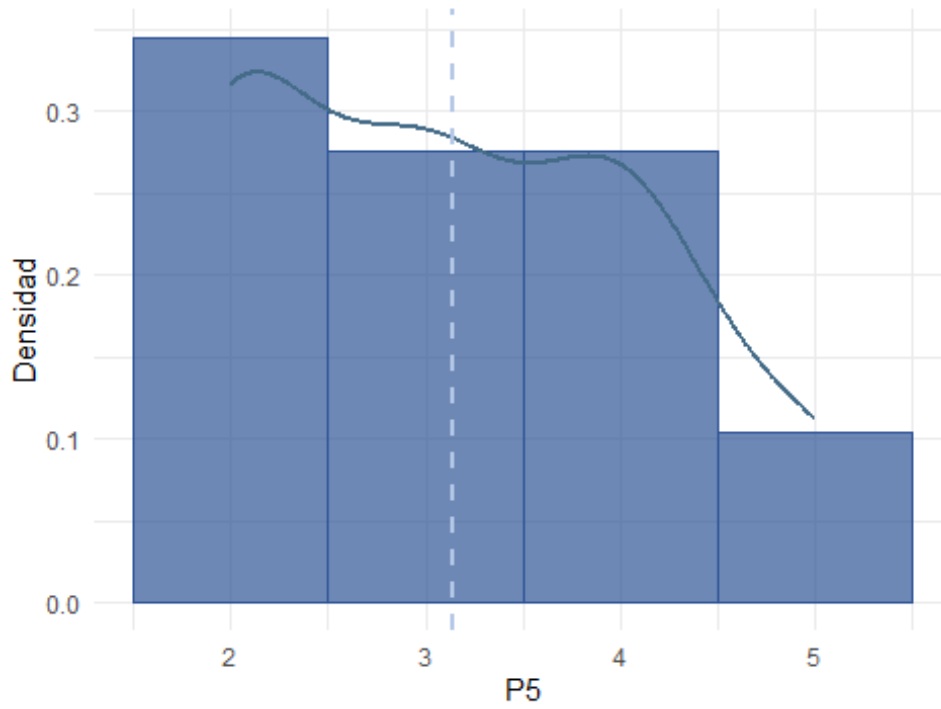
  # Mostrar La gráfica de barras en el dispositivo gráfico
  print(pv2)

  # Guardar La gráfica de barras como imagen
  ggsave(filename = paste0("histograma_", preguntav2, ".png"), plot =
pv2)
}

## Saving 5 x 4 in image

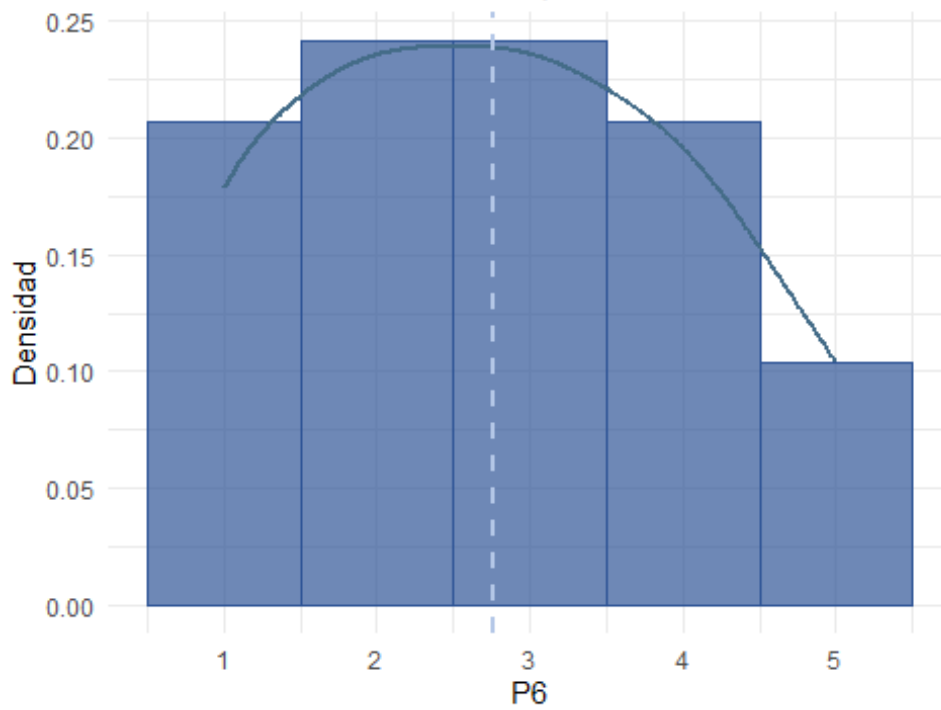
```

Gráfica de barras de interoperabilidad de las herramie



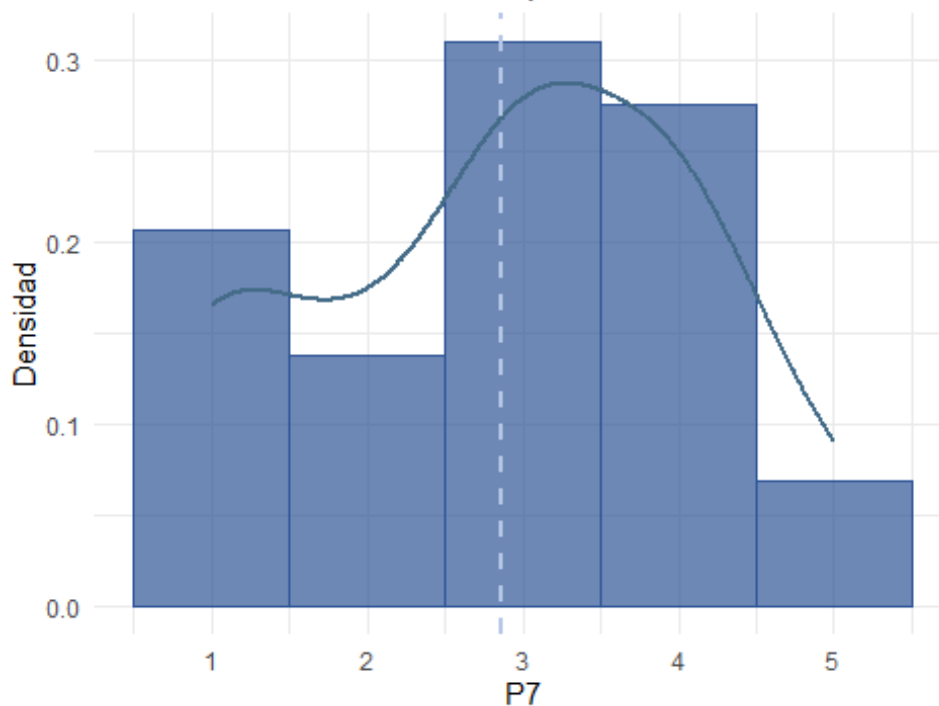
Saving 5 x 4 in image

Gráfica de barras de interoperabilidad de las herramie



Saving 5 x 4 in image

Gráfica de barras de interoperabilidad de las herramie



```
# Subconjunto de datos para el grupo Agregación de datos y Las preguntas
P8 a P11
data_subsetv3 <- Data_CCA %>%
  select(starts_with("P")) %>%
  mutate_if(is.character, as.factor)

# Crear una función para calcular Las métricas
calcular_metricasv3 <- function(preguntav3) {
  modav3 <- DescTools::Mode(data_subsetv3[[preguntav3]])
  if (length(modav3) > 1) {
    modav3 <- paste(modav3, collapse = ", ")
  } else {
    modav3 <- as.character(modav3) # Asegurar que modav3 sea de tipo
character
  }
  data.frame(
    Preguntav3 = preguntav3,
    Validov3 = sum(!is.na(data_subsetv3[[preguntav3]])),
    Perdidov3 = sum(is.na(data_subsetv3[[preguntav3]])),
    Mediov3 = mean(data_subsetv3[[preguntav3]], na.rm = TRUE),
    Medianav3 = median(data_subsetv3[[preguntav3]], na.rm = TRUE),
    Modav3 = modav3,
    Desviacion_Estandarv3 = sd(data_subsetv3[[preguntav3]], na.rm =
TRUE),
    Varianzav3 = var(data_subsetv3[[preguntav3]], na.rm = TRUE)
  )
}
```

```

# Calcular Las métricas para Las preguntas P8 a P11
resultadosv3 <- lapply(c("P8", "P9", "P10", "P11"), calcular_metricasv3)

# Convertir La Lista a un data frame, manejar casos donde no hay moda
encontrada
resultadosv3 <- bind_rows(resultadosv3)

# Mostrar Los resultados
print(resultadosv3)

## Preguntav3 Validov3 Perdidov3 Mediav3 Medianav3 Modav3
Desviacion_Estandarv3
## 1 P8 29 0 3.241379 3 3
0.7862739
## 2 P9 29 0 4.000000 4 4
0.8451543
## 3 P10 29 0 3.931034 4 4
0.9610647
## 4 P11 29 0 3.482759 4 4
0.7377906
## Varianzav3
## 1 0.6182266
## 2 0.7142857
## 3 0.9236453
## 4 0.5443350

# Crear La gráfica de barras para cada pregunta con Línea de media y
densidad de probabilidad
for (preguntav3 in c("P8", "P9", "P10", "P11")) {
  # Calcular La media para La pregunta actual
  media_preguntav3 <- mean(data_subsetv3[[preguntav3]], na.rm = TRUE)

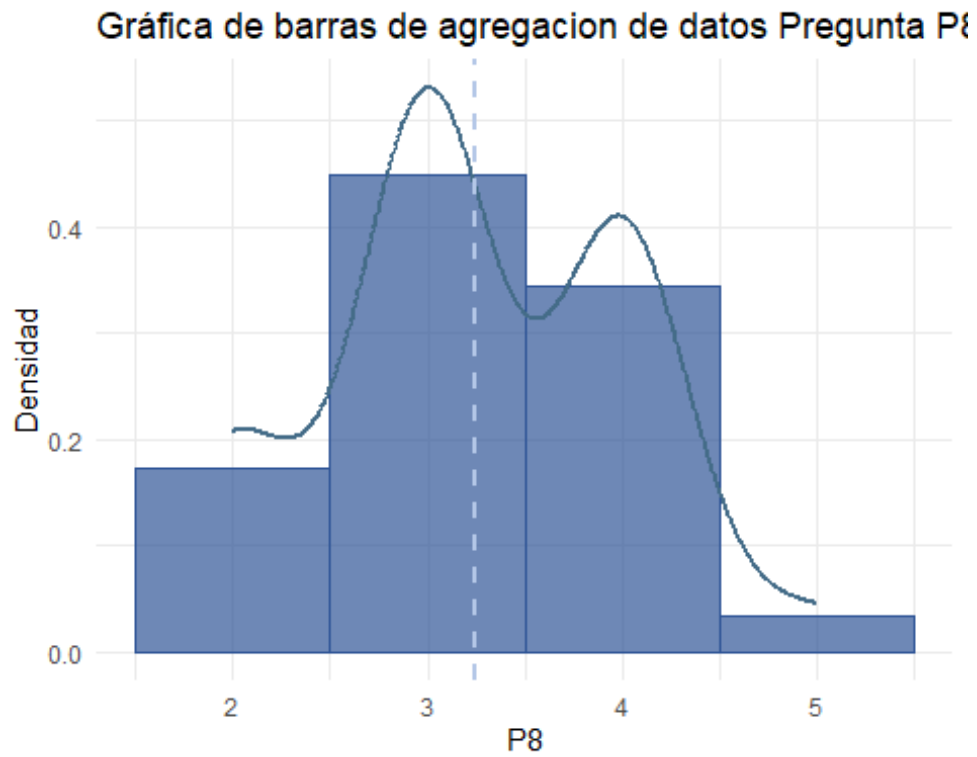
  # Crear La gráfica de barras
  pv3 <- ggplot(data_subsetv3, aes(x = !!sym(preguntav3))) +
    geom_histogram(aes(y = ..density..), binwidth = 1, fill = "#2F5597",
color = "#2F5597", alpha = 0.7) +
    geom_density(color = "#446C88", size = 1) +
    geom_vline(aes(xintercept = media_preguntav3), color = "#B4C7E7",
linetype = "dashed", size = 1) +
    labs(
      title = paste("Gráfica de barras de agregacion de datos Pregunta",
preguntav3),
      x = preguntav3,
      y = "Densidad"
    ) +
    theme_minimal()

  # Mostrar La gráfica de barras en el dispositivo gráfico
  print(pv3)
}

```

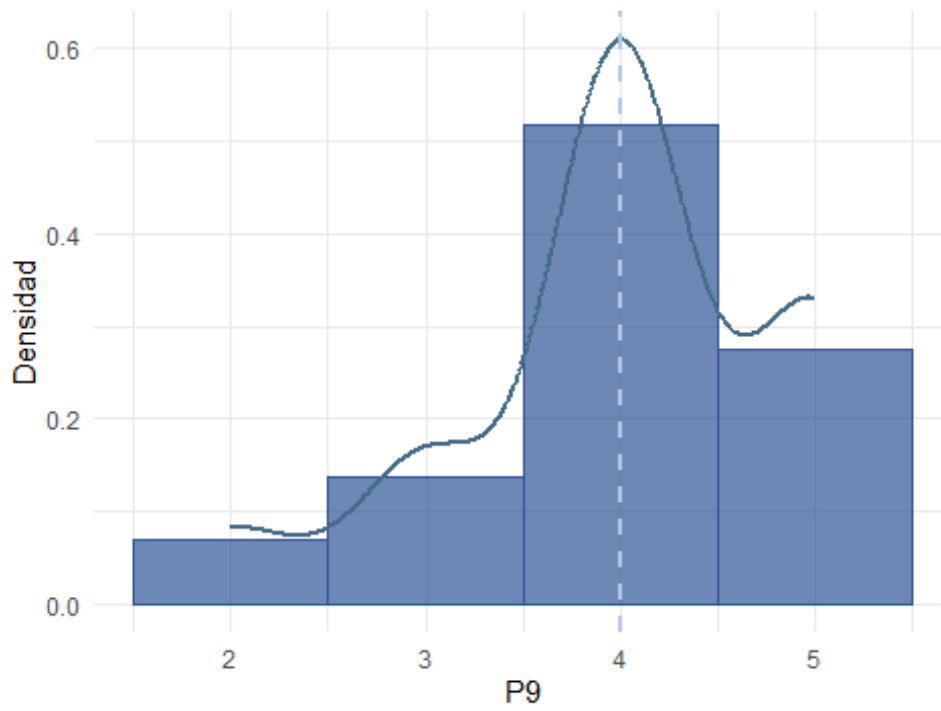
```
# Guardar La gráfica de barras como imagen
ggsave(filename = paste0("histograma_", preguntav3, ".png"), plot =
pv3)
}
```

```
## Saving 5 x 4 in image
```



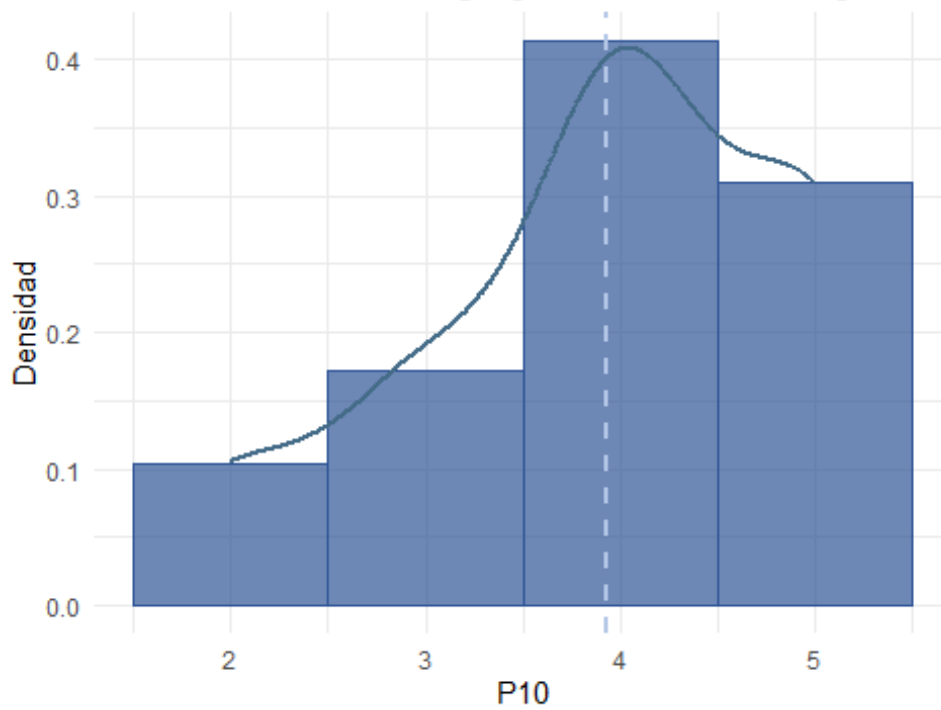
```
## Saving 5 x 4 in image
```

Gráfica de barras de agregacion de datos Pregunta P9



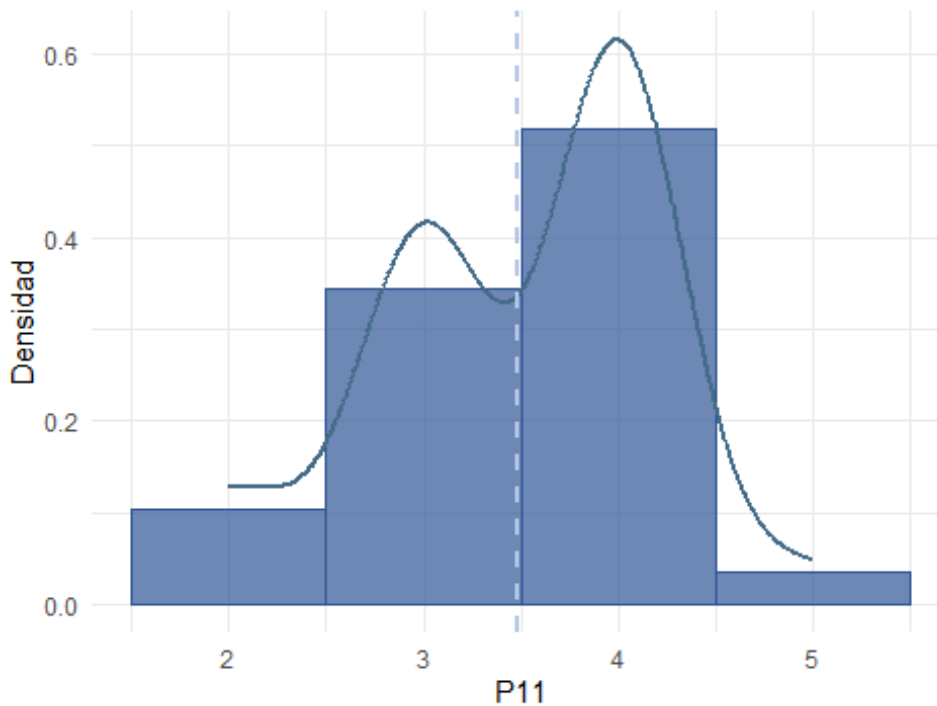
Saving 5 x 4 in image

Gráfica de barras de agregacion de datos Pregunta P10



Saving 5 x 4 in image

Gráfica de barras de agregacion de datos Pregunta P11



```
# Subconjunto de datos para el grupo Explotación de datos y Las preguntas
P12 a P15
data_subsetv4 <- Data_CCA %>%
  select(starts_with("P")) %>%
  mutate_if(is.character, as.factor)

# Crear una función para calcular Las métricas
calcular_metricasv4 <- function(preguntav4) {
  modav4 <- DescTools::Mode(data_subsetv4[[preguntav4]])
  if (length(modav4) > 1) {
    modav4 <- paste(modav4, collapse = ", ")
  } else {
    modav4 <- as.character(modav4) # Asegurar que modav4 sea de tipo
character
  }
  data.frame(
    Preguntav4 = preguntav4,
    Validov4 = sum(!is.na(data_subsetv4[[preguntav4]])),
    Perdidov4 = sum(is.na(data_subsetv4[[preguntav4]])),
    Mediov4 = mean(data_subsetv4[[preguntav4]], na.rm = TRUE),
    Medianav4 = median(data_subsetv4[[preguntav4]], na.rm = TRUE),
    Modav4 = modav4,
    Desviacion_Estandarv4 = sd(data_subsetv4[[preguntav4]], na.rm =
TRUE),
    Varianzav4 = var(data_subsetv4[[preguntav4]], na.rm = TRUE)
  )
}
```

```

# Calcular Las métricas para Las preguntas P12 a P15
resultadosv4 <- lapply(c("P12", "P13", "P14", "P15"),
  calcular_metricasv4)

# Convertir La Lista a un data frame, manejar casos donde no hay moda
encontrada
resultadosv4 <- bind_rows(resultadosv4)

# Mostrar Los resultados
print(resultadosv4)

## Preguntav4 Validov4 Perdidov4 Mediav4 Medianav4 Modav4
Desviacion_Estandarv4
## 1 P12 29 0 3.620690 4 4
0.8624630
## 2 P13 29 0 3.413793 4 4
0.9455626
## 3 P14 29 0 3.448276 4 4
0.7831350
## 4 P15 29 0 3.551724 4 4
1.0207213
## Varianzav4
## 1 0.7438424
## 2 0.8940887
## 3 0.6133005
## 4 1.0418719

# Crear La gráfica de barras para cada pregunta con Línea de media y
densidad de probabilidad
for (preguntav4 in c("P12", "P13", "P14", "P15")) {
  # Calcular La media para La pregunta actual
  media_preguntav4 <- mean(data_subsetv4[[preguntav4]], na.rm = TRUE)

  # Crear La gráfica de barras
  pv4 <- ggplot(data_subsetv4, aes(x = !!sym(preguntav4))) +
    geom_histogram(aes(y = ..density..), binwidth = 1, fill = "#2F5597",
  color = "#2F5597", alpha = 0.7) +
    geom_density(color = "#446C88", size = 1) +
    geom_vline(aes(xintercept = media_preguntav4), color = "#B4C7E7",
  linetype = "dashed", size = 1) +
    labs(
      title = paste("Gráfica de barras de explotacion de datos Pregunta",
  preguntav4),
      x = preguntav4,
      y = "Densidad"
    ) +
    theme_minimal()

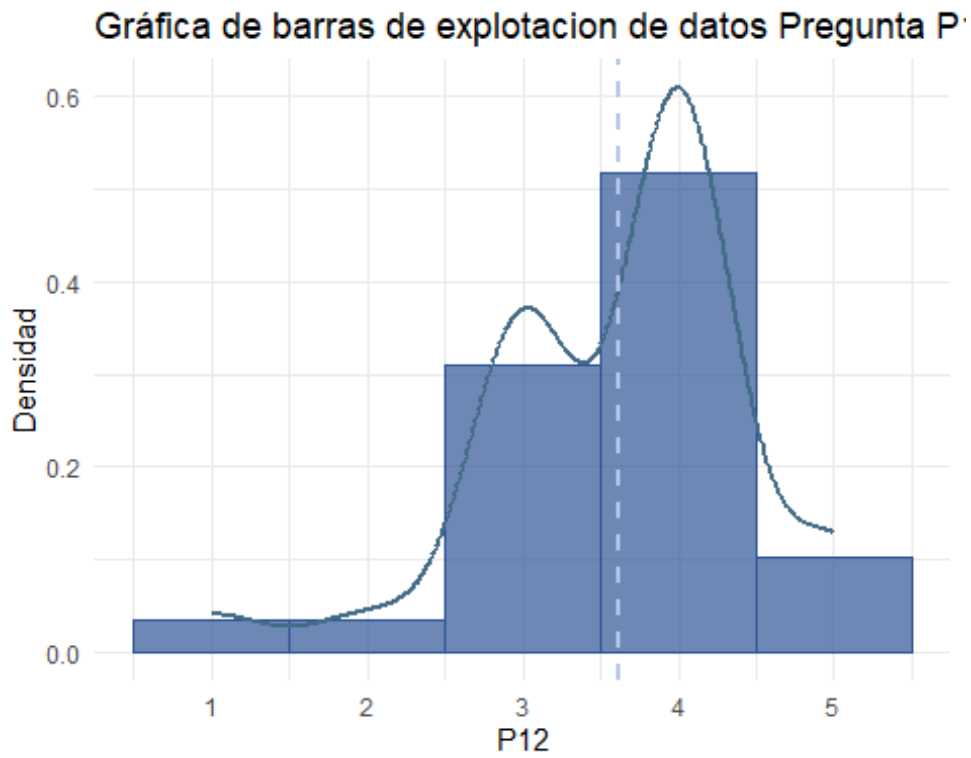
  # Mostrar La gráfica de barras en el dispositivo gráfico

```

```
print(pv4)

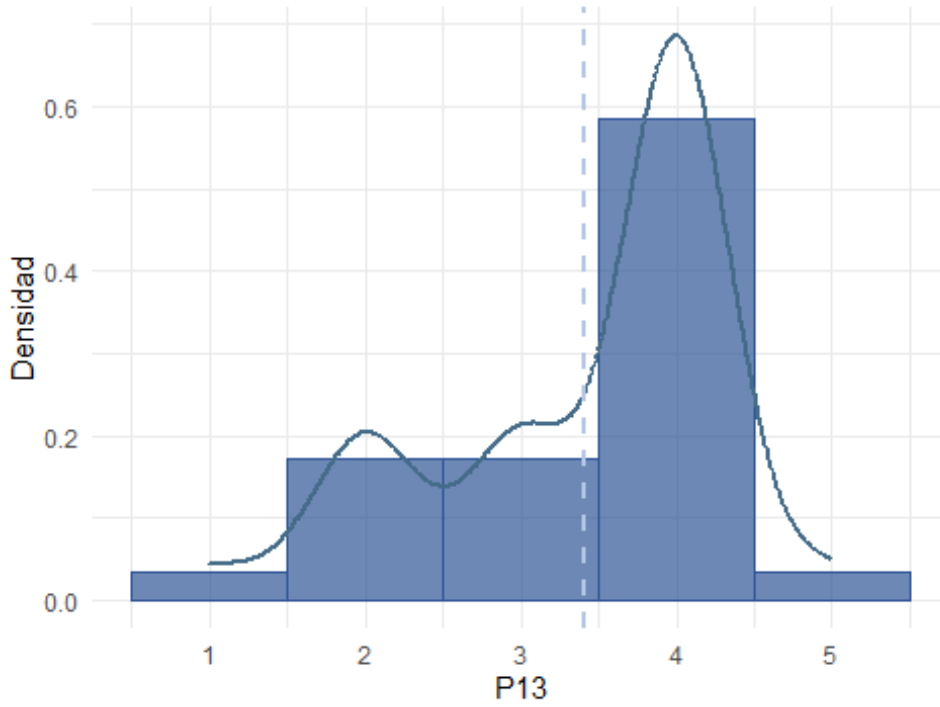
# Guardar La gráfica de barras como imagen
ggsave(filename = paste0("histograma_", preguntav4, ".png"), plot =
pv4)
}

## Saving 5 x 4 in image
```



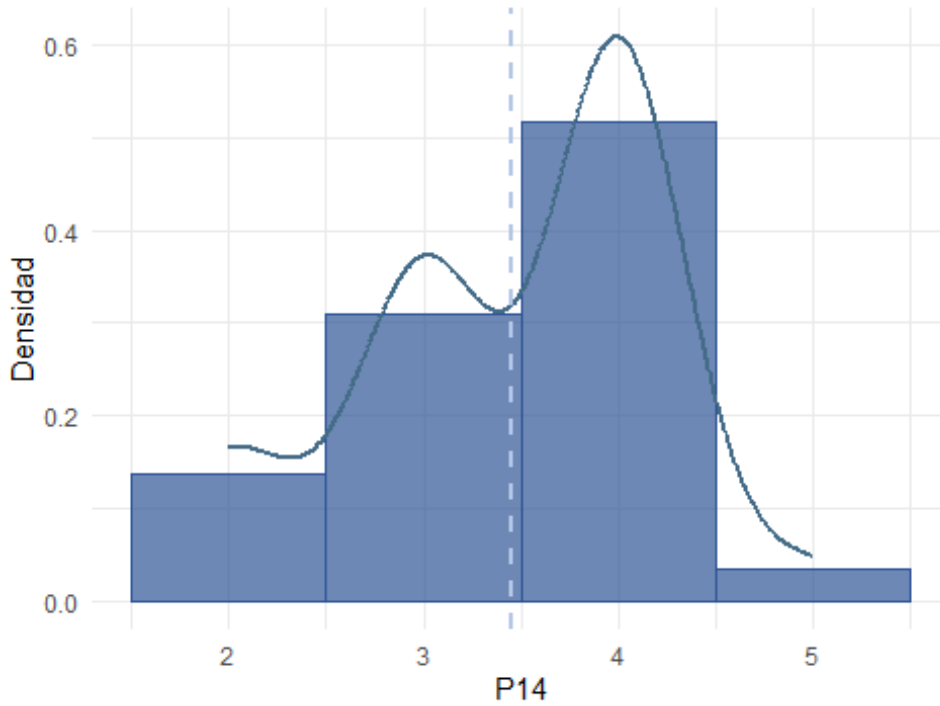
```
## Saving 5 x 4 in image
```

Gráfica de barras de explotación de datos Pregunta P



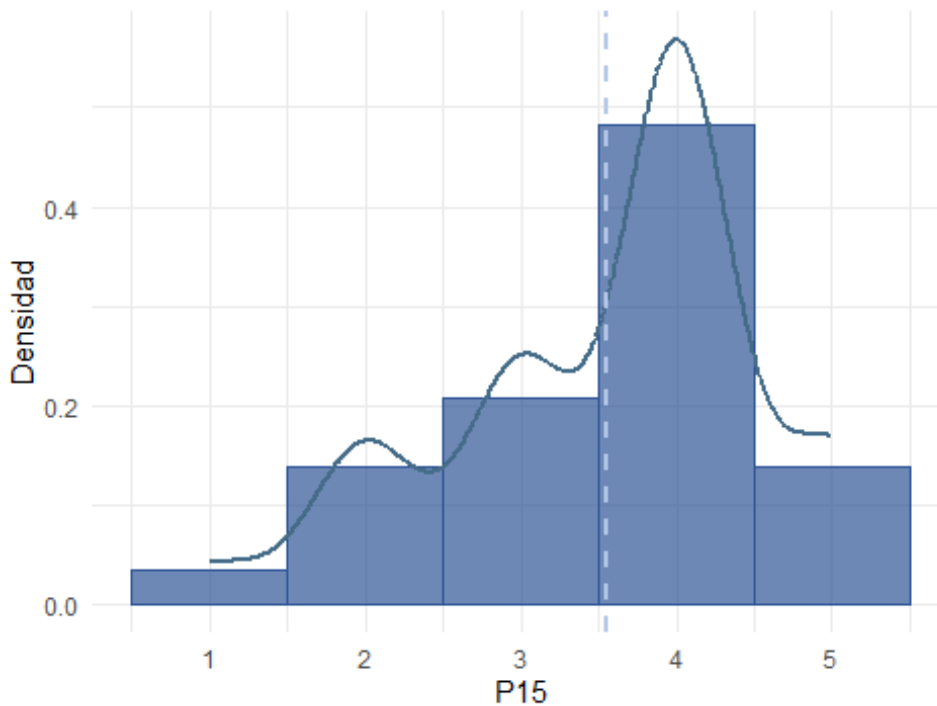
Saving 5 x 4 in image

Gráfica de barras de explotación de datos Pregunta P



Saving 5 x 4 in image

Gráfica de barras de explotación de datos Pregunta P



```
# Subconjunto de datos para el grupo Explotación de datos y Las preguntas
P16 a P18
data_subsetv5 <- Data_CCA %>%
  select(starts_with("P")) %>%
  mutate_if(is.character, as.factor)

# Crear una función para calcular Las métricas
calcular_metricasv5 <- function(preguntav5) {
  modav5 <- DescTools::Mode(data_subsetv5[[preguntav5]])
  if (length(modav5) > 1) {
    modav5 <- paste(modav5, collapse = ", ")
  } else {
    modav5 <- as.character(modav5) # Asegurar que modav5 sea de tipo
character
  }
  data.frame(
    Preguntav5 = preguntav5,
    Validov5 = sum(!is.na(data_subsetv5[[preguntav5]])),
    Perdidov5 = sum(is.na(data_subsetv5[[preguntav5]])),
    Mediov5 = mean(data_subsetv5[[preguntav5]], na.rm = TRUE),
    Medianav5 = median(data_subsetv5[[preguntav5]], na.rm = TRUE),
    Modav5 = modav5,
    Desviacion_Estandarv5 = sd(data_subsetv5[[preguntav5]], na.rm =
TRUE),
    Varianzav5 = var(data_subsetv5[[preguntav5]], na.rm = TRUE)
  )
}
```

```

# Calcular Las métricas para Las preguntas P16 a P18
resultadosv5 <- lapply(c("P16", "P17", "P18"), calcular_metricasv5)

# Convertir La Lista a un data frame, manejar casos donde no hay moda
encontrada
resultadosv5 <- bind_rows(resultadosv5)

# Mostrar Los resultados
print(resultadosv5)

## Preguntav5 Validov5 Perdidov5 Mediav5 Medianav5 Modav5
Desviacion_Estandarv5
## 1 P16 29 0 3.172414 3 2
1.104179
## 2 P17 29 0 3.241379 3 3
1.057462
## 3 P18 29 0 3.000000 3 3
1.101946
## Varianzav5
## 1 1.219212
## 2 1.118227
## 3 1.214286

# Crear La gráfica de barras para cada pregunta con Línea de media y
densidad de probabilidad
for (preguntav5 in c("P16", "P17", "P18")) {
  # Calcular La media para La pregunta actual
  media_preguntav5 <- mean(data_subsetv5[[preguntav5]], na.rm = TRUE)

  # Crear La gráfica de barras
  pv5 <- ggplot(data_subsetv5, aes(x = !!sym(preguntav5))) +
    geom_histogram(aes(y = ..density..), binwidth = 1, fill = "#2F5597",
color = "#2F5597", alpha = 0.7) +
    geom_density(color = "#446C88", size = 1) +
    geom_vline(aes(xintercept = media_preguntav5), color = "#B4C7E7",
linetype = "dashed", size = 1) +
    labs(
      title = paste("Histograma de ExplotaciOn de datos Pregunta",
preguntav5),
      x = preguntav5,
      y = "Densidad"
    ) +
    theme_minimal()

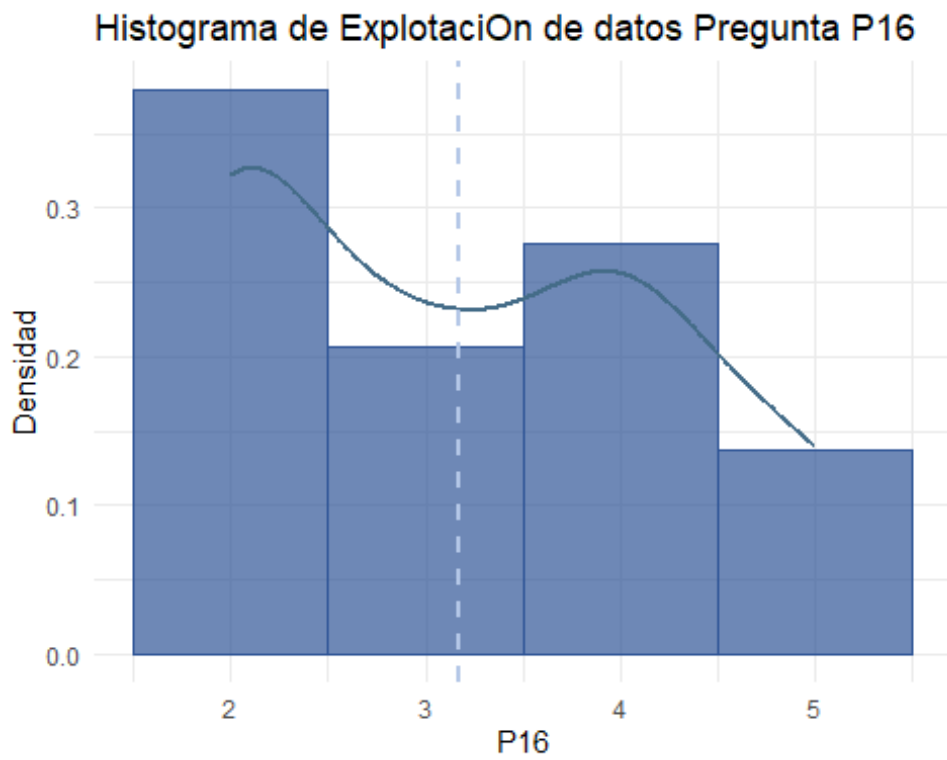
  # Mostrar La gráfica de barras en el dispositivo gráfico
  print(pv5)

  # Guardar La gráfica de barras como imagen
  ggsave(filename = paste0("histograma_", preguntav5, ".png"), plot =

```

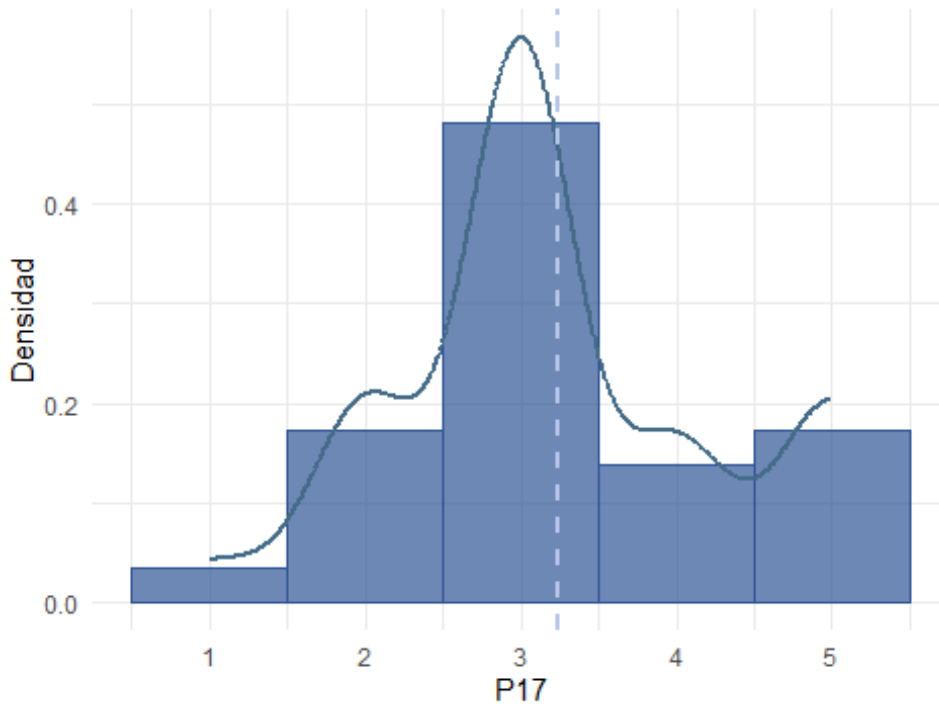
```
pv5)  
}
```

```
## Saving 5 x 4 in image
```



```
## Saving 5 x 4 in image
```

Histograma de ExplotaciOn de datos Pregunta P17



```
## Saving 5 x 4 in image
```

Histograma de ExplotaciOn de datos Pregunta P18

