

Resultados de Percepción sobre Organismos Genéticamente Modificados

Daniel Quiroz

Contents

Cargar Librerías	1
Importación de datos	1
Limpieza de datos	2
Carreras repetidas	2
Rangos de edad	2
Semestres	2
Descripción del espacio muestral	3
Medios de Comunicación	6
Conocimiento sobre OGM	7
Nociones básicas sobre OGM	7
Con qué se los relaciona	9
Consumo de OGMs	10
OGM comercializados en EC	12
Percepción de nuevas tecnologías	13

Este documento tiene como finalidad analizar los resultados obtenidos a partir de las encuestas realizadas en la Universidad Regional Amazónica Ikiam.

Cargar Librerías

Como primer paso, se cargarán las librerías necesarias para realizar los análisis.

```
#### Load Packages ####  
library(readxl) # Read Excel Files  
library(tidyverse) # Easily Install and Load the 'Tidyverse'  
library(stringr) # Simple, Consistent Wrappers for Common String Operations  
library(ggalluvial)  
library(magrittr)  
library(ggsci)  
library(scales)  
library(ggpubr)
```

Importación de datos

Como primer paso se importarán los datos crudos, para esto se usará una función iterativa, map.

```
raw_data <- map(2:12, .f = function(x){
  read_xlsx(path = 'data/Encuestas.xlsx', sheet = x) %>%
    mutate(Sheet = x)
}) %>% bind_rows()
```

Limpieza de datos

Como todos los datos del mundo real, tienen errores de tabulación los cuales deben ser corregidos. En esta sección se irán recopilando los pasos para limpiar los datos.

Carreras repetidas

En los datos crudos se encontraron errores de tipeo, los cuales llevan a condierar como diferentes carreras.

```
clean_data <- raw_data %>% mutate(P4a = tolower(P4a) ) %>%
  # Clean Data
  # Repeated careers
  mutate(P4a = ifelse(P4a %in% 'ingenieria en ecosistema',
    "ingeniería en ecosistemas", P4a)) %>%
  mutate(P4a = ifelse(P4a %in% c("ingeniería en ciencias del agua",
    "ingeniería en ciencias del agua"),
    "ingeniería en ciencias del agua", P4a)) %>%
  mutate(P4a = ifelse(grepl("ingeniería", P4a), "Ingeniería", P4a) ) %>%
  mutate(P4a = ifelse(P4a %in% c('agroecología', 'arquitectura',
    'ciencias experimentales',
    'licenciatura en biocomercio'),
    "Licenciaturas", P4a) ) %>%
  mutate(P4a = ifelse(P4a %in% c('administrativo'), 'otros', P4a) ) %>%
  mutate(P4a = ifelse(P4a %in% c('docente'), 'Docentes', P4a) )
```

Ahora, asignaremos las etiquetas de *masculino* y *femenino* para los valores 1 y 2 tabulados

```
clean_data %<>% mutate(P2 = factor(P2, levels = c(1,2),
  labels = c("Masculino", "Femenino")))
```

Rangos de edad

En este paso se asignarán rangos de edad a los valores numéricos

```
clean_data %<>%
  mutate(P3 = ifelse(P3 == 1, "17-20", P3)) %>%
  mutate(P3 = ifelse(P3 == 2, "21-25", P3)) %>%
  mutate(P3 = ifelse(P3 == 3, "26-30", P3)) %>%
  mutate(P3 = ifelse(P3 == 4, "30-40", P3)) %>%
  mutate(P3 = ifelse(P3 == 5, "> 40", P3)) %>%
  mutate(P3 = ifelse(P3 == 6, "> 40", P3))
  #mutate(P3 = factor(P3, levels = c("17-20", "21-25", "26-30", "30-40", "> 40")))
```

Semestres

Se diferenciará mayormente entre tronco común y semestres de carrera

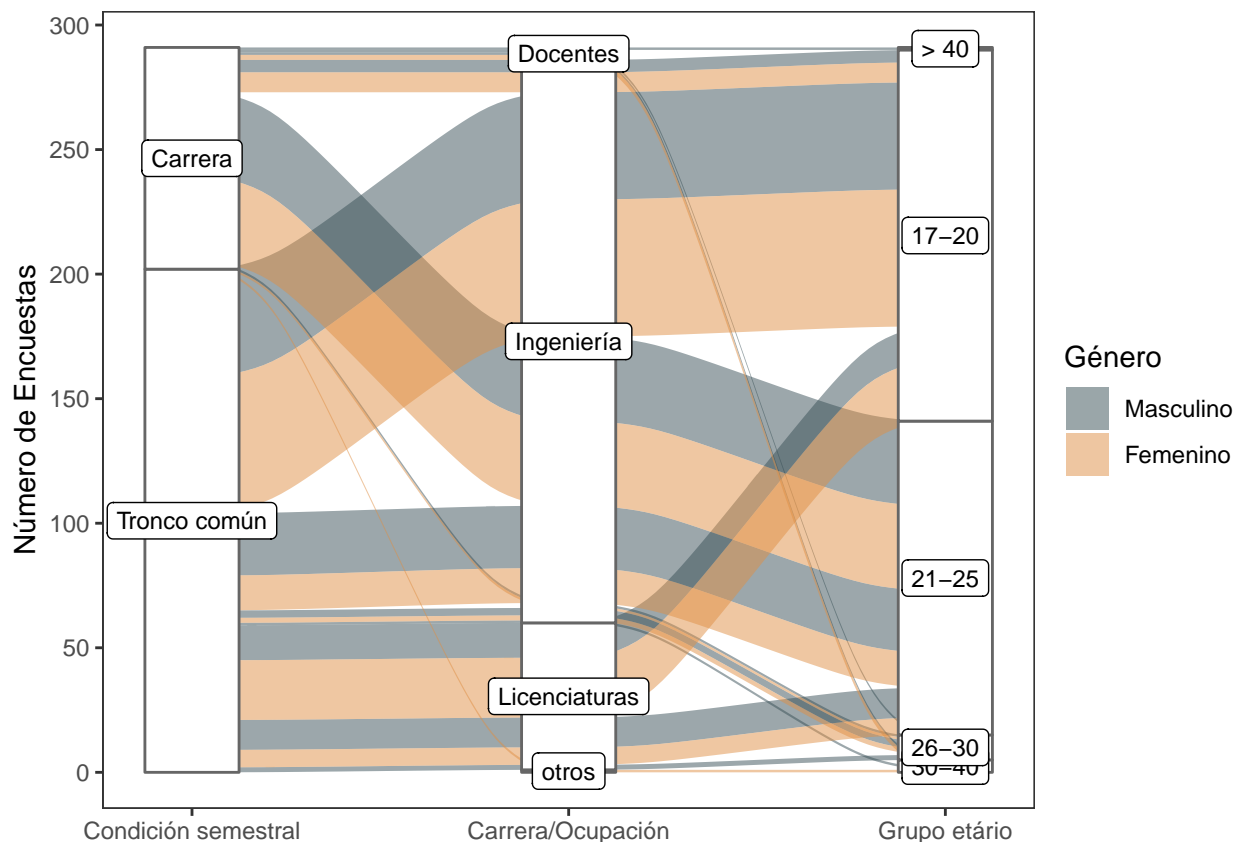
```
clean_data %<>%
  mutate(`5` = ifelse(`5` > 4, "Carrera", "Tronco común"))
```

Descripción del espacio muestral

Como primera instancia, debemos la composición del espacio muestral. Para esto, emplearemos gráficos descriptivos.

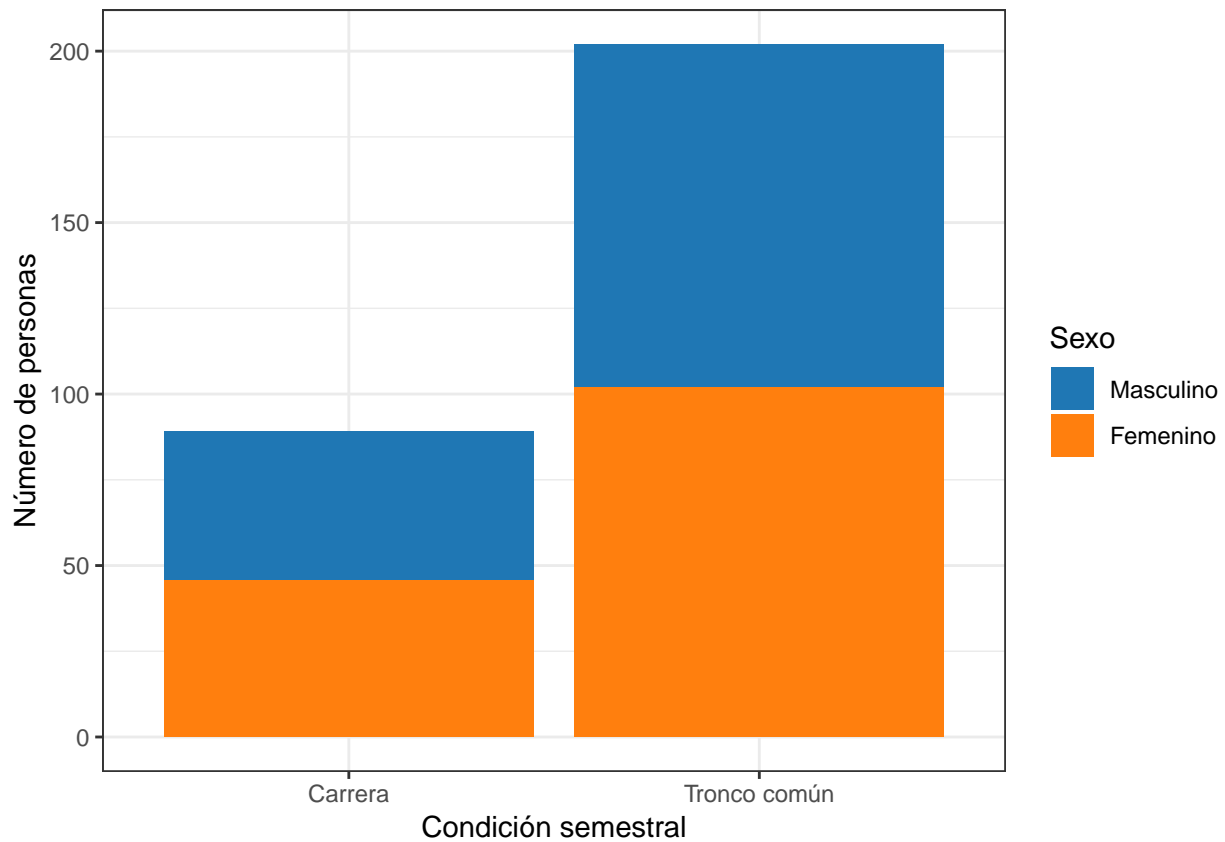
```
# Summarising data
sample_space <- clean_data %>% group_by(P2, P3, P4a, `5`) %>%
  summarise(N = n()) %>% arrange(P4a) %>% filter(!is.na(`5`))
sspace_plot <- ggplot(sample_space, aes(axis1 = `5`, axis2 = P4a, axis3 = P3, y = N)) +
  geom_alluvium(aes(fill = P2), width = 1/8) +
  geom_stratum(width = 1/4, fill = "white", color = "grey40") +
  geom_label(stat = "stratum", label.strata = T, size = 3) +
  scale_x_continuous(breaks = 1:3, labels = c("Condición semestral",
                                              "Carrera/Ocupación",
                                              "Grupo etário")) +

  scale_fill_jama() + theme_test() +
  scale_y_continuous(breaks = seq(0, 300, by = 50),
                    labels = seq(0, 300, by = 50)) +
  labs(fill = "Género", y = "Número de Encuestas")
sspace_plot
```

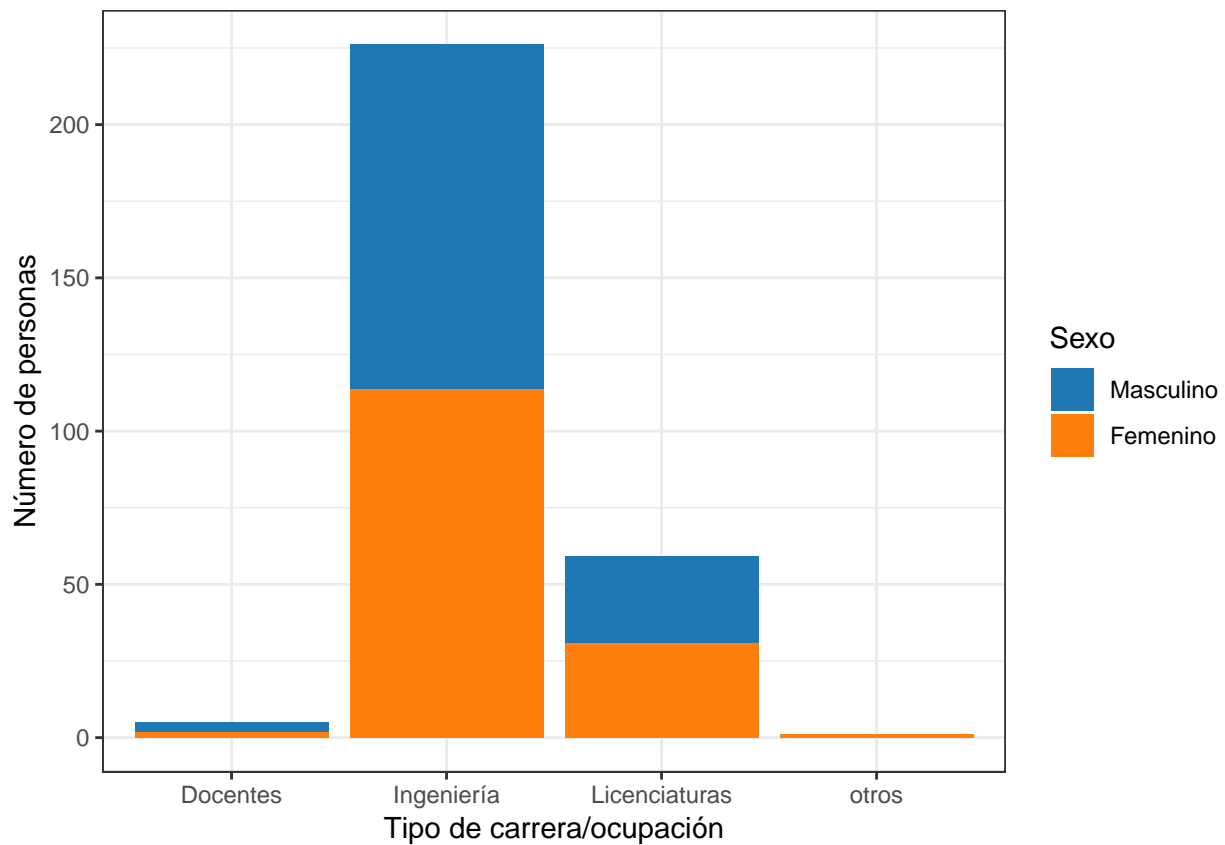


```
semestral <- ggplot(sample_space, aes(x = `5`, y = N, fill = P2)) +
  geom_bar(stat = "identity") + theme_bw() +
```

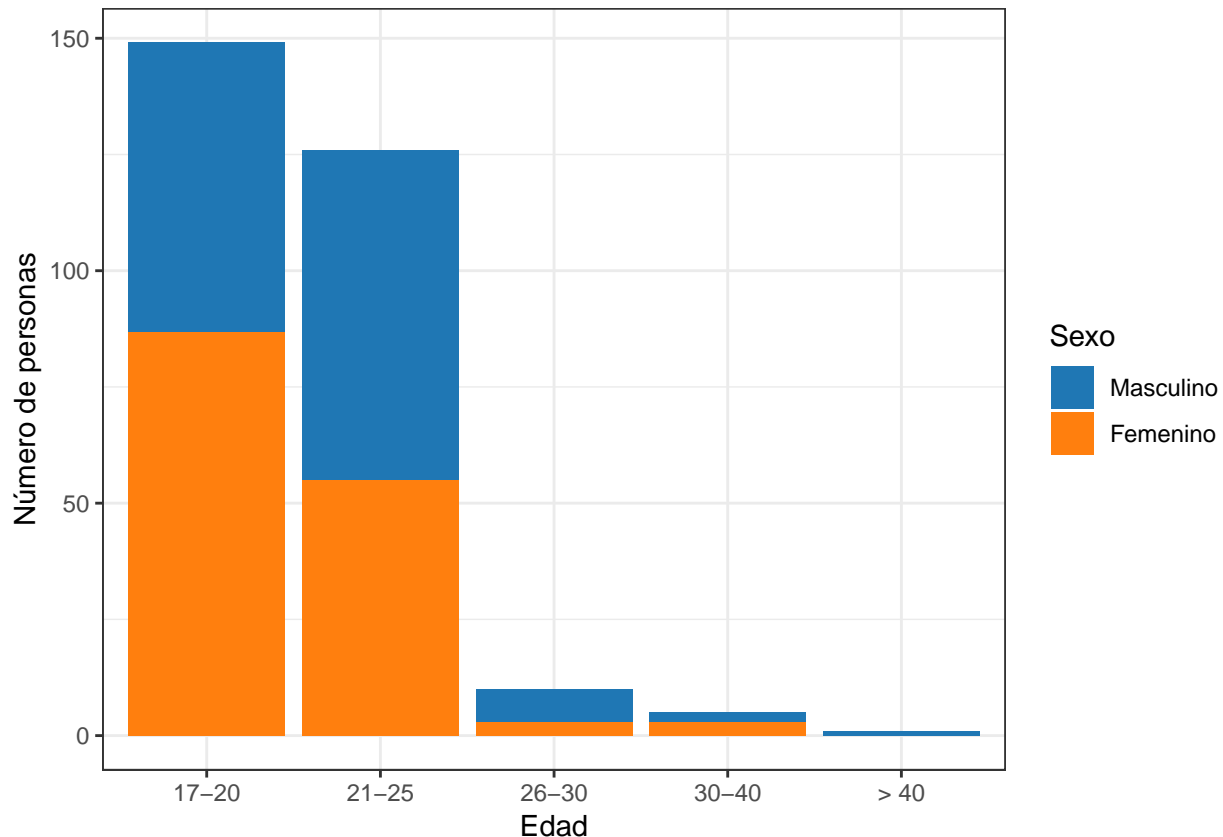
```
scale_fill_d3() +
labs(x = "Condición semestral", y = "Número de personas", fill = "Sexo")
semestral
```



```
ocupacion <- ggplot(sample_space, aes(x= P4a, y = N, fill = P2) ) +
  geom_bar(stat = "identity") + theme_bw() +
  scale_fill_d3() +
  labs(x= "Tipo de carrera/ocupación", y = "Número de personas", fill = "Sexo")
ocupacion
```



```
edad <- sample_space %>% ungroup() %>%
  mutate(P3 = factor(P3,
                     levels = c("17-20", "21-25", "26-30", "30-40", "> 40"))) %>%
  ggplot(aes(x= P3, y = N, fill = P2) ) +
  geom_bar(stat = "identity") + theme_bw() +
  scale_fill_d3() +
  labs(x = "Edad", y = "Número de personas", fill = "Sexo")
edad
```



Medios de Comunicación

En este apartado, se analizará el número de medios de comunicación que emplea cada encuestado para informarse.

```
medios <- clean_data %>% select(`6a`:`6g`) %>% mutate(Index = 1:n()) %>%
  gather(key = "Medio", value = "Valor", `6a`:`6g`) %>%
  mutate(Medio = ifelse(Medio %in% "6a", "Radio AM/FM", Medio)) %>%
  mutate(Medio = ifelse(Medio %in% "6b", "Televisión abierta", Medio)) %>%
  mutate(Medio = ifelse(Medio %in% "6c", "Televisión por cable", Medio)) %>%
  mutate(Medio = ifelse(Medio %in% "6d", "Periódicos/Revistas", Medio)) %>%
  mutate(Medio = ifelse(Medio %in% "6e", "Internet/Redes Sociales", Medio)) %>%
  mutate(Medio = ifelse(Medio %in% "6f", "Revistas científicas", Medio)) %>%
  mutate(Medio = ifelse(Medio %in% "6g", "Otros", Medio)) %>%
  filter(Valor <= 1)
```

Una vez tenemos asignado las etiquetas asignadas a cada código, procedemos a obtener la probabilidad de que determinado medio sea empleado, cuando son empleados n medios de comunicación.

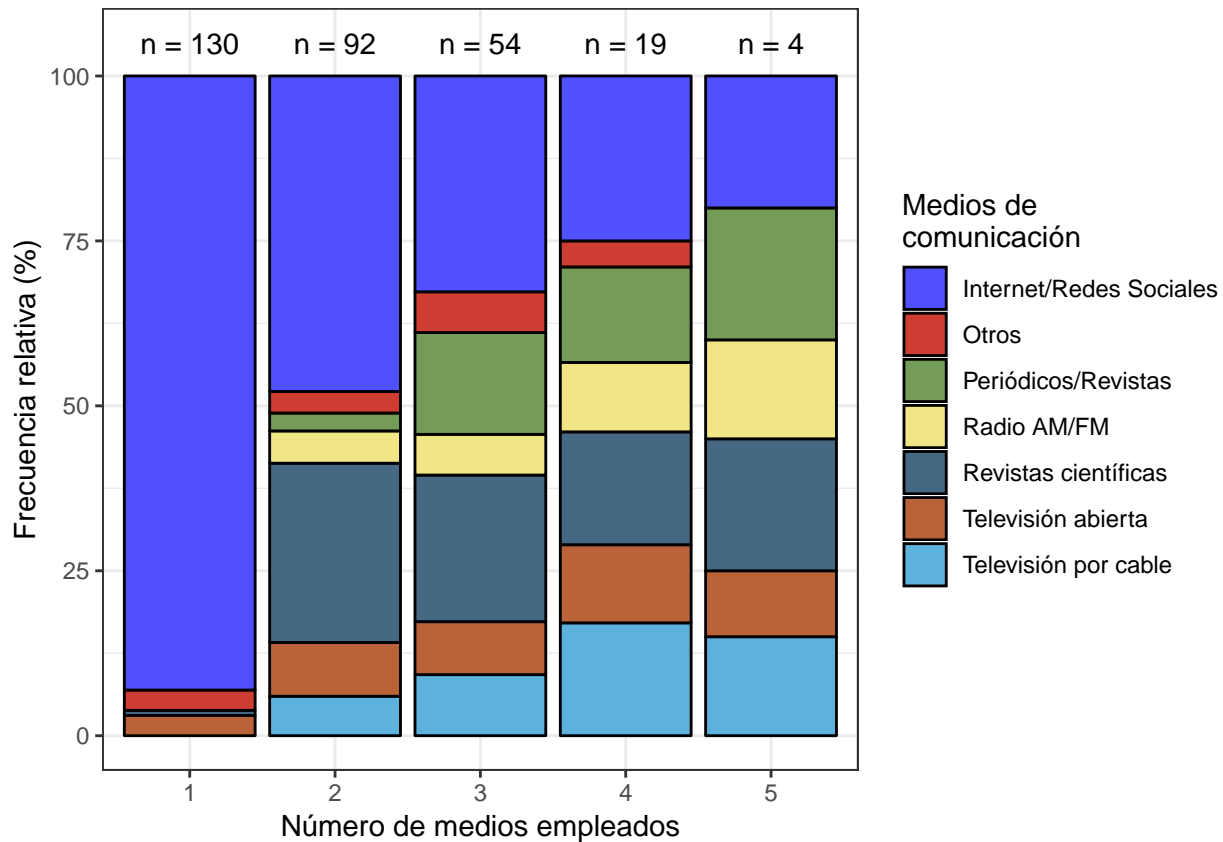
```
# Obtener el n por medios
labels <- medios %>% filter(Valor != 0) %>%
  arrange(Index) %>%
  group_by(Index) %>% mutate(total = sum(Valor)) %>%
  count(total) %>% ungroup() %>% count(n) %>%
  mutate(nn = paste0("n = ", nn))

medios_f <- medios %>% filter(Valor != 0) %>% arrange(Index) %>%
```

```

group_by(Index) %>% mutate(total = sum(Valor)) %>%
select(-Valor)
medios_plot <- ggplot(medios_f) +
  geom_bar(aes(factor(total), fill = Medio),
    position = "fill", color = "black") +
  geom_text(data = labels, aes(x= factor(n), y = 1.05, label = nn) ) +
  ggsci::scale_fill_igv() + theme_bw() +
  scale_y_continuous(breaks = seq(0, 1, by = 0.25),
    labels = seq(0, 100, by = 25)) +
  scale_x_discrete(limits = 1:5) +
  labs(x = "Número de medios empleados", y = "Frecuencia relativa (%)",
    fill = "Medios de\ncomunicación")
medios_plot

```



Por ejemplo, de los 126 encuestados que solo emplean un medio de comunicación, al rededor del 90% emplea solo Internet y/o redes sociales. Por otro lado, cuando emplean dos medios de comunicación, el uso de internet decrece a un 50% y las revistas científicas tienen un 20% de probabilidad de ser empleadas como medio de comunicación.

Conocimiento sobre OGM

Nociones básicas sobre OGM

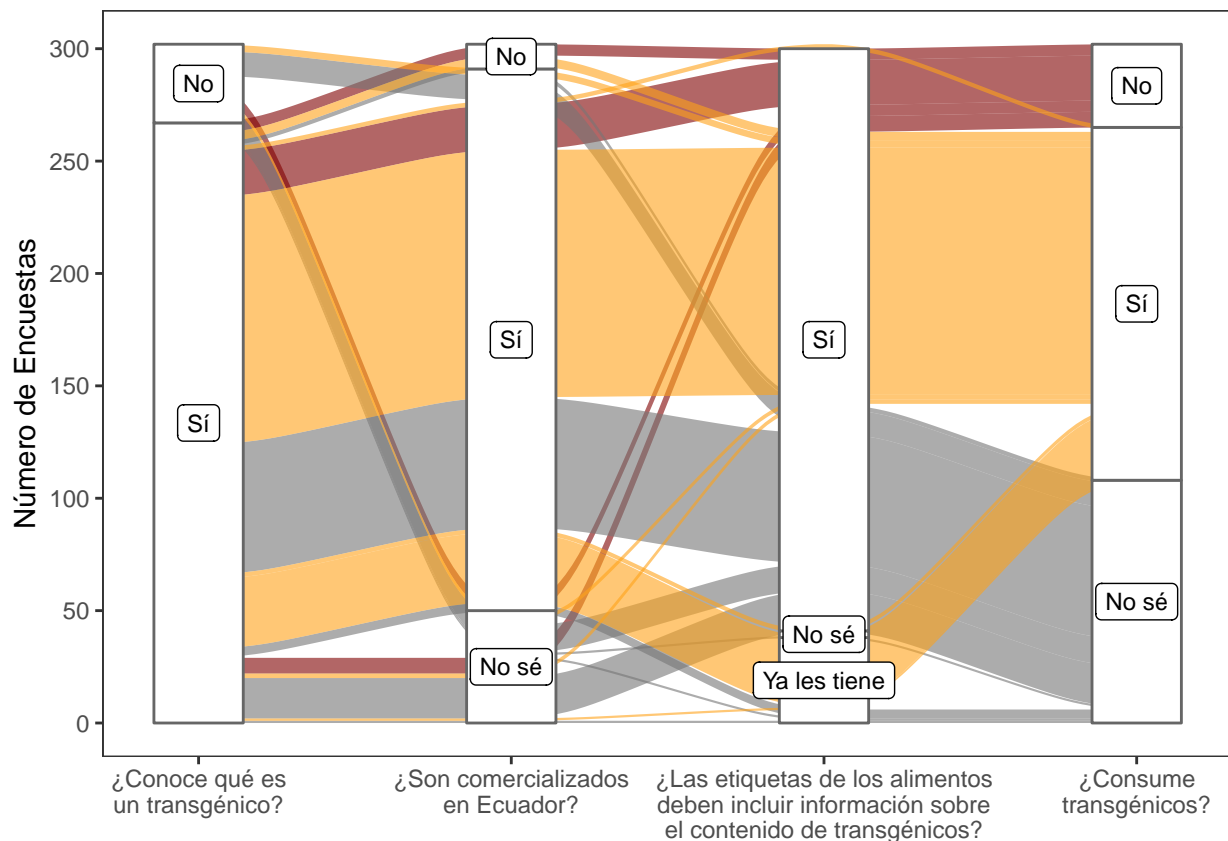
Empezamos por analizar el conocimiento o nociones básicas que tienen las personas encuestadas. Para esto, obtendremos el número de personas encuestadas, conforme a todas las combinaciones para

```
transform <- function(x) {
  ifelse(x == 0, "No", ifelse(x == 1, "Sí",
                             ifelse(x == 2, "No sé", "Ya les tiene")))
}
knowledge <- clean_data %>% select(`7`, `8`, `10`, `17`) %>%
  mutate_all(transform) %>% mutate_all(as.factor) %>%
  group_by(`7`, `8`, `10`, `17`) %>% summarise(N = n())
```

```
## Warning: Factor `17` contains implicit NA, consider using
## `forcats::fct_explicit_na`
```

Con las frecuencias obtenidas, realizared el gráfico de tipo *alluvial*.

```
labs_know <- c("¿Conoce qué es\nun transgénico?",
              "¿Son comercializados\nen Ecuador?",
              "¿Las etiquetas de los alimentos
deben incluir información sobre
el contenido de transgénicos?",
              "¿Consume\ntransgénicos?")
main_plot <- ggplot(knowledge, aes(axis1 = `7`, axis2 = `8`, axis4 = `10`,
                                   axis3 = `17`, y = N)) +
  geom_alluvium(aes(fill = `10`), width = 1/20, alpha = 0.6) +
  geom_stratum(width = 1/3.5, fill = "white", color = "grey40") +
  geom_label(stat = "stratum", label.strata = TRUE, size = 3) +
  scale_x_continuous(breaks = 1:4, labels = labs_know) +
  scale_fill_uchicago() + theme_test() +
  labs(y = "Número de Encuestas") +
  scale_y_continuous(breaks = seq(0, 300, by = 50),
                    labels = seq(0, 300, by = 50)) +
  guides(fill = F)
main_plot
```

Con qué se los relaciona

En este paso obtendremos a que riesgos o beneficios se relaciona los transgénicos. Por lo que la frecuencia de las respuestas

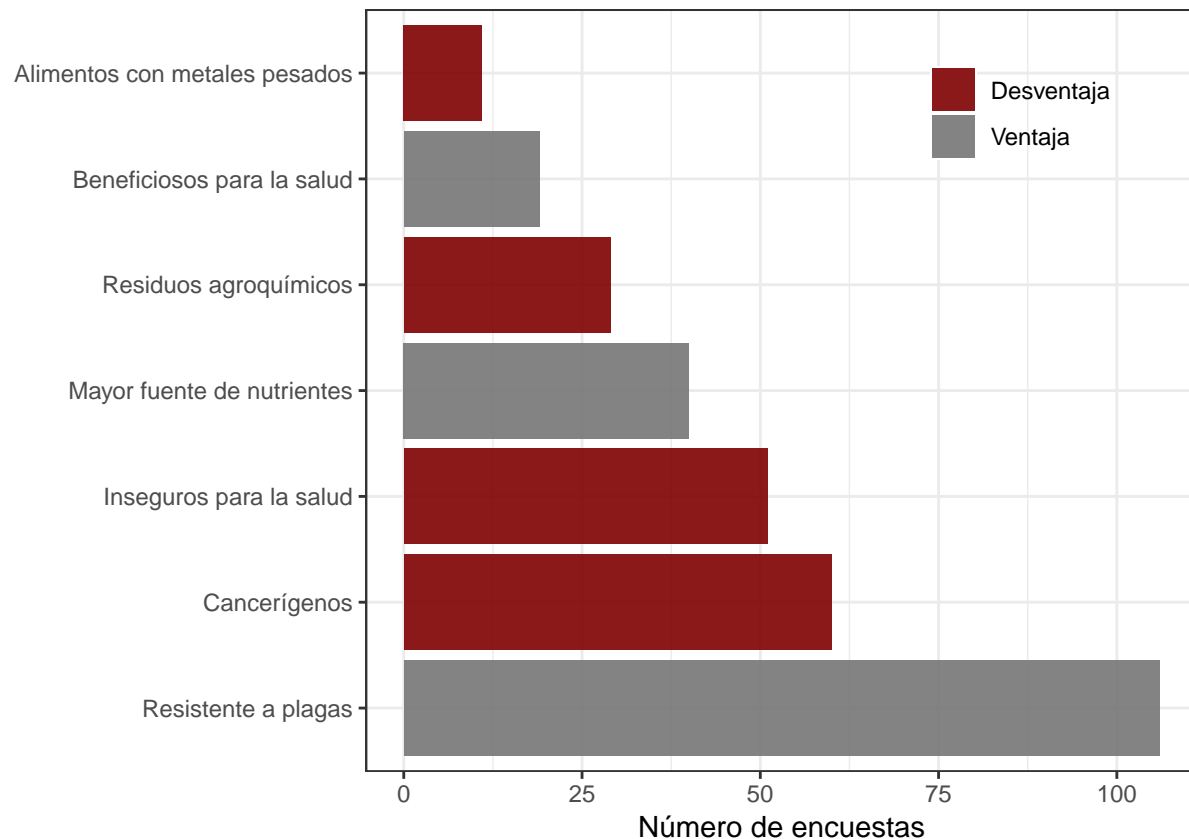
```
risk_benefit <- clean_data %>% select(starts_with("14")) %>%
  gather(key = "Variable", value = "Valor", `14a`:`14h` ) %>%
  group_by(Variable) %>% summarise(N = sum(Valor, na.rm = T) ) %>%
  filter( !(Variable %in% "14h") ) %>%
  mutate(R_B = ifelse(Variable %in% c("14b", "14d", "14g"),
    "Ventaja", "Desventaja") ) %>%
  mutate(R_B = factor(R_B, levels = c("Desventaja", "Ventaja")) ) %>%
  mutate(Variable = ifelse(Variable %in% "14a", "Residuos agroquímicos",
    Variable)) %>%
  mutate(Variable = ifelse(Variable %in% "14b", "Mayor fuente de nutrientes",
    Variable)) %>%
  mutate(Variable = ifelse(Variable %in% "14c", "Inseguros para la salud",
    Variable)) %>%
  mutate(Variable = ifelse(Variable %in% "14d", "Beneficiosos para la salud",
    Variable)) %>%
  mutate(Variable = ifelse(Variable %in% "14e", "Alimentos con metales pesados",
    Variable)) %>%
  mutate(Variable = ifelse(Variable %in% "14f", "Cancerígenos",
    Variable)) %>%
  mutate(Variable = ifelse(Variable %in% "14g", "Resistente a plagas",
    Variable)) %>%
```

```

mutate(Variable = ifelse(Variable %in% "14h", "Mejoramiento genético",
                        Variable)) %>%
arrange(-N) %>% mutate(Variable = factor(Variable, levels = .$Variable))

advantages <- ggplot(risk_benefit, aes(x = Variable, y = N)) +
  geom_bar(aes(fill = R_B), stat = "identity", alpha = 0.9) +
  theme_bw() + labs(fill = "", y = "Número de encuestas", x = "") +
  theme(legend.position=c(0.9,1), legend.justification=c(0.95,1),
        legend.background=element_blank()) +
  coord_flip() + ggsci::scale_fill_uchicago()
advantages

```



Consumo de OGMs

```

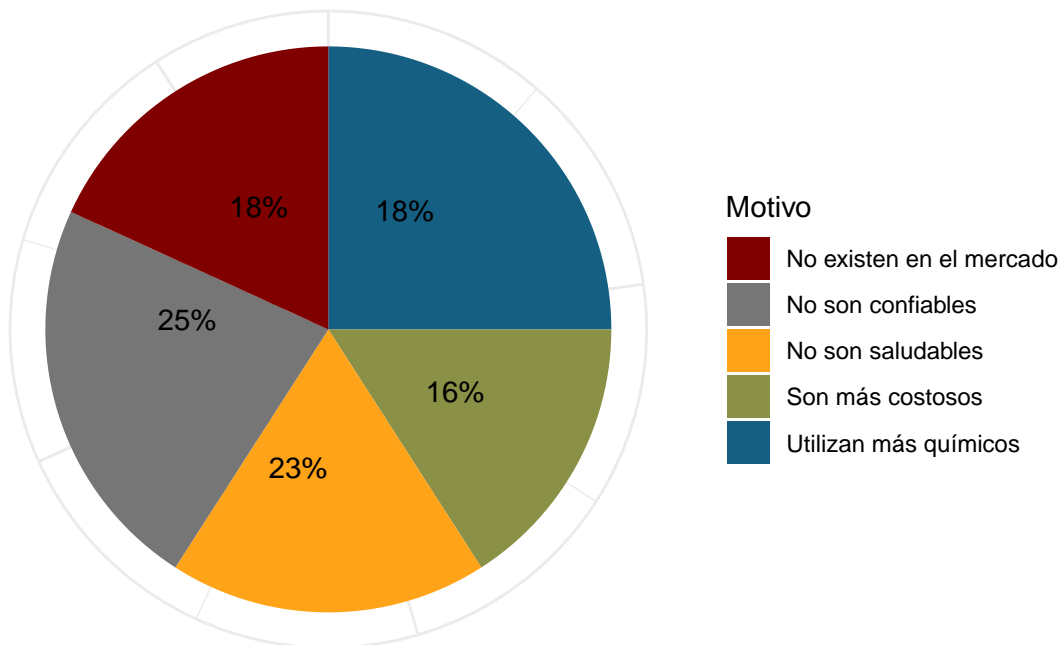
consumption_no <- clean_data %>% filter(`10` == 0) %>% select(`11a`:`11e`) %>%
  gather(key = "Question", value = "Value") %>% group_by(Question) %>%
  summarise(N = sum(Value)) %>% mutate(Per = round(N / sum(N) * 100)) %>%
  mutate(Question = ifelse(Question %in% "11a", "No son saludables", Question)) %>%
  mutate(Question = ifelse(Question %in% "11b", "Son más costosos", Question)) %>%
  mutate(Question = ifelse(Question %in% "11c", "No son confiables", Question)) %>%
  mutate(Question = ifelse(Question %in% "11d", "Utilizan más químicos", Question)) %>%
  mutate(Question = ifelse(Question %in% "11e", "No existen en el mercado", Question))

consumption_pltN <- ggplot(consumption_no, aes(x = "", y = N, fill = Question)) +

```

```
geom_bar(stat = "identity") + coord_polar("y", start = 0) +
theme_minimal() + ggsci::scale_fill_uchicago() +
geom_text(aes(y = N + c(-4, cumsum(N*0.9)[-length(N)]),
              label = paste0(Per, "%") )) +
theme(axis.text.x=element_blank()) +
labs(title = "Consume: No\n", x = '', y = "", fill = "Motivo")
consumtion_pltN
```

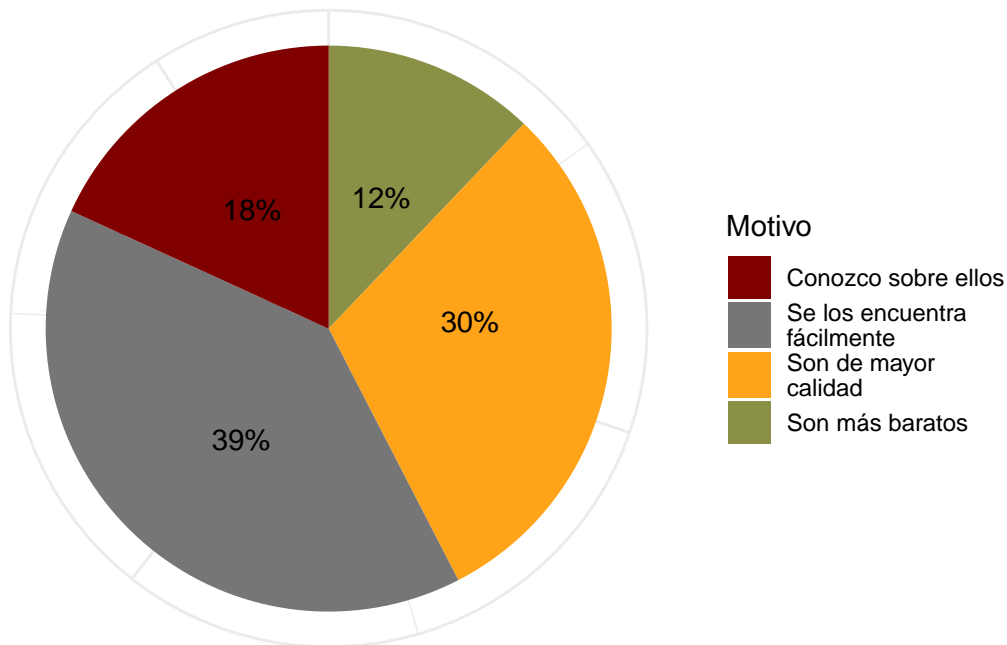
Consume: No



```
consumtion_yes1 <- clean_data %>% filter(`10` == 1) %>% select(`12a`:`12d`) %>%
gather(key = "Question", value = "Value") %>% group_by(Question) %>%
summarise(N = sum(Value, na.rm = T)) %>%
mutate(Per = round(N / sum(N) *100)) %>%
mutate(Question = ifelse(Question %in% "12a",
                        "Se los encuentra\nfácilmente", Question)) %>%
mutate(Question = ifelse(Question %in% "12b",
                        "Son más baratos", Question)) %>%
mutate(Question = ifelse(Question %in% "12c",
                        "Son de mayor\ncalidad", Question)) %>%
mutate(Question = ifelse(Question %in% "12d",
                        "Conozco sobre ellos", Question))

consumtion_pltY <- ggplot(consumtion_yes1, aes(x = "", y = N, fill = Question)) +
geom_bar(stat = "identity") + coord_polar("y", start = 0) +
theme_minimal() + ggsci::scale_fill_uchicago() +
geom_text(aes(y = c(20, 2, 8, 30),
              label = paste0(Per, "%") )) +
theme(axis.text.x=element_blank()) +
labs(title = "Consume: Yes\nPor que?", x = '', y = "", fill = "Motivo")
consumtion_pltY
```

Consume: Yes
Por que?



```
consumption_plot <- ggarrange(consumption_pltN, consumption_pltY,
                              labels = c("A", "B"), ncol = 2, nrow = 1)
```

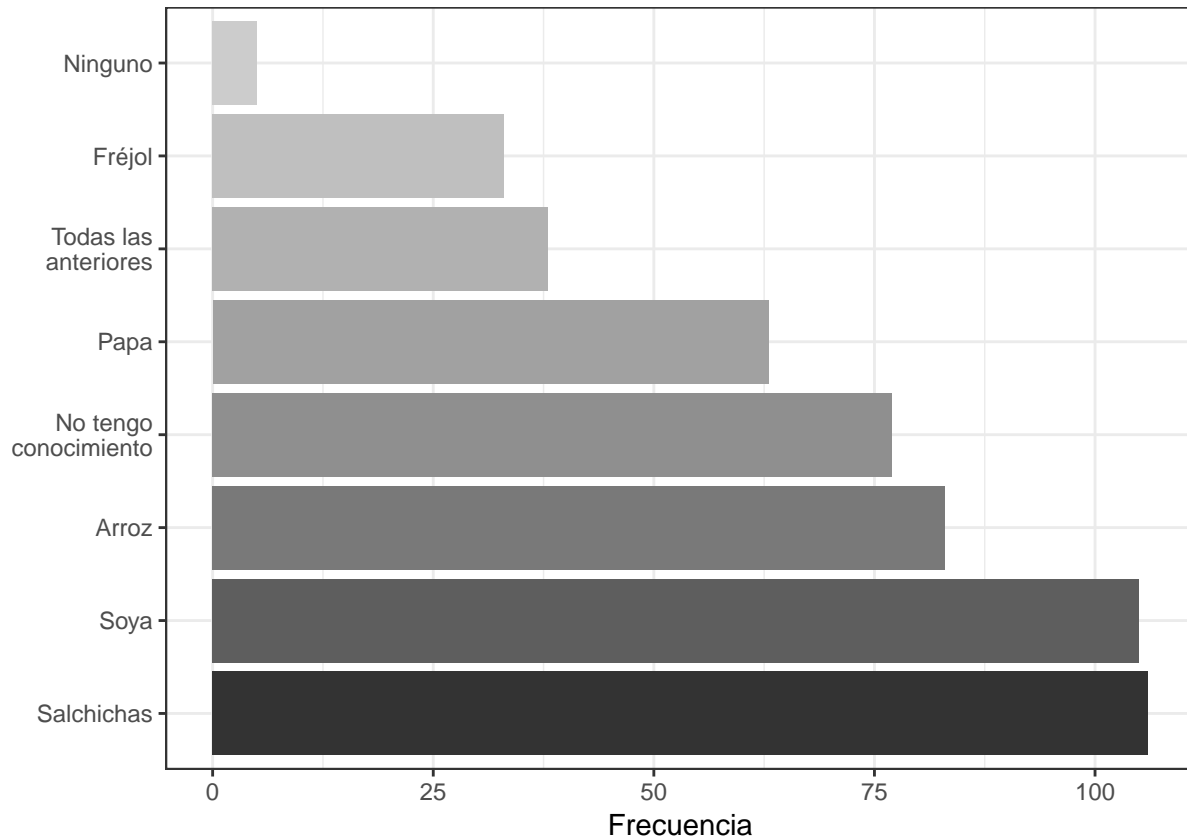
OGM comercializados en EC

Ahora, obtendremos los OGM que las personas encuestadas indican que se consumen en EC.

```
OGM_products <- clean_data %>% select(starts_with("15")) %>%
  gather(key = "Variable", value = "Valor", `15a`:`15h` ) %>%
  group_by(Variable) %>% summarise(N = sum(Valor, na.rm = T) ) %>%
  arrange(-N) %>%
  mutate(Variable = ifelse(Variable %in% "15a", "Soya", Variable) ) %>%
  mutate(Variable = ifelse(Variable %in% "15b", "Arroz", Variable) ) %>%
  mutate(Variable = ifelse(Variable %in% "15c", "Fréjol", Variable) ) %>%
  mutate(Variable = ifelse(Variable %in% "15d", "Papa", Variable) ) %>%
  mutate(Variable = ifelse(Variable %in% "15e", "Salchichas", Variable) ) %>%
  mutate(Variable = ifelse(Variable %in% "15f",
                          "Todas las\n anteriores", Variable) ) %>%
  mutate(Variable = ifelse(Variable %in% "15g", "Ninguno", Variable) ) %>%
  mutate(Variable = ifelse(Variable %in% "15h", "No tengo\nconocimiento",
                          Variable) ) %>%
  mutate(Variable = factor(Variable, levels = .$Variable))

products_plot <- ggplot(OGM_products, aes(Variable, N)) +
  geom_bar(aes(fill = Variable), stat = "identity") +
  scale_fill_grey() + coord_flip() + theme_bw() +
  guides(fill = F) +
  labs(y = "Frecuencia", x = "")
```

products_plot



Percepción de nuevas tecnologías

```
new_tech <- clean_data %>% select(starts_with("20")) %>%
  gather(key = "Variable", value = "Valor", starts_with("20") ) %>%
  count(Variable, Valor) %>% filter(!is.na(Valor)) %>%
  mutate(Variable = ifelse(Variable %in% "20 (Bio)",
                           "Biotecnología", Variable)) %>%
  mutate(Variable = ifelse(Variable %in% "20 (Eolica)",
                           "Generación de\nenergía eólica", Variable)) %>%
  mutate(Variable = ifelse(Variable %in% "20 (Modifica)",
                           "Técnicas de\nmodificación genética", Variable)) %>%
  mutate(Variable = ifelse(Variable %in% "20 (Nano)",
                           "Nanotecnología", Variable)) %>%
  mutate(Variable = ifelse(Variable %in% "20 (NBT)",
                           "Nuevas tecnologías\nde mejoramiento", Variable)) %>%
  mutate(Variable = ifelse(Variable %in% "20 (Quimi)",
                           "Química Agrícola", Variable)) %>%
  mutate(Variable = ifelse(Variable %in% "20 (Solar)",
                           "Generación de\nenergía solar", Variable)) %>%
  mutate(Variable = ifelse(Variable %in% "20 (Transplante)",
                           "Transplate de\nórganos", Variable)) %>%
  filter(Valor <= 3)%>%
```

```
mutate(Valor = ifelse(Valor > 1, "No sé", ifelse(Valor < 1, "No", "Si")))

tech_perception <- ggplot(new_tech, aes(Variable, n )) +
  geom_bar(aes(fill = factor(Valor) ), color = "black",
    stat = "identity", position = "fill") + theme_bw() +
  scale_y_continuous(breaks = seq(0, 1, by = 0.25),
    labels = seq(0, 100, by = 25)) +
  scale_fill_jama() + coord_flip() +
  labs(fill = "Espera un resultado\nbeneficioso", x = "",
    y = "Frecuencia relativa (%)")
tech_perception
```

