

Análisis de BD de los Aspirantes UNAL & UdeA

Daniel Villa - Ronald Palencia

Introducción:

Este archivo se creo para el tratamiento de los datos recolectados los días 23 al 25 de octubre de los aspirantes universitarios de la UNAL sede Medellín y UdeA.

Tipo de muestreo

se aplico un muestreo sistemático de 1 en 20, tomando como valor aleatorio inicial por medio del software R

```
# fijar semilla:
set.seed(123)

# muestreo 1 en 20 en la UdeA
udea <- sample(1:20,1, replace = F)

# muestreo 1 en 20 en la unal
unal <- sample(1:10,1, replace = F)

paste("UdeA = Valor generado aleatoriamente: ", udea)
```

```
## [1] "UdeA = Valor generado aleatoriamente: 15"
```

```
paste("UNAL = Valor generado aleatoriamente: ", unal)
```

```
## [1] "UNAL = Valor generado aleatoriamente: 3"
```

el valor $x = \{unal, udea\}$ fue el que se genero por medio de la función `sample()`.

Fecha de inicio y finalización

Primer día

Nota: Unico día de recolección de datos en la UNAL, el 24 y 25 de Oct fueron tomados en la UdeA.

las encuestas empezaron el día 23 de octubre a las 10 AM de la mañana, terminando a las 2 PM de la tarde

Segundo día

las encuestas empezaron el día 24 de octubre a las 8 AM de la mañana, terminando a las 12 del medio día

Tercer día

las encuestas empezaron el día 25 de octubre a las 8 AM de la mañana, terminando a las 11 AM (esto debido a la dificultad de tomar los datos en la UdeA).

Se leen los datos tal cual se descargan de google forms, por ende utilizamos `clean_names()` para organizar y estandarizar los nombres de las variables

```
data = read_xlsx("unal_udea.xlsx", col_names = T)
```

```
## New names:  
## * ' ' -> '...1'
```

```
data %<>% clean_names()
```

ahora con la función siguiente lo que hacemos es transformar la variable: “año de graduación de ultimo año de secundaria”

este cambio se efectuó dado las múltiples respuestas de los usuarios para declarar un mismo año de graduación

```
#data$en_que_ano_salio_o_saldra_del_colegio %>% unique()  
  
data$en_que_ano_salio_o_saldra_del_colegio =  
  str_replace_all(data$en_que_ano_salio_o_saldra_del_colegio,  
    c("Once"= "2022", "11.0" = "2022",  
      "Saldré este año" = "2022",  
      "Ya salí" = "2021", "En el 2019" = "2019",  
      "comunicación audiovisual y multimedia" = "2019",  
      "1º" = "2022",  
      "2020.0" = "2022", "2022.0" = "2022", "2018.0" = "2018",  
      "2019.0" = "2019", "2017.0" = "2017", "2021.0" = "2021"))
```

Con esto en mente se hace la misma verificación de las demás variables, por ende se entiende que las columnas que no aparecen son porque no se encontraron defectos o algún motivo para su cambio (*esto para referirse a las preguntas de opción múltiple*)

Ahora por medio de una caracterización dada por la SNIES, tenemos las siguientes categorías para la clasificación de las carreras de los aspirantes

- A1 : Agronomía, Veterinaria y afines
- A2: Bellas Artes
- A3: Ciencias de la Educación
- A4 : Ciencias de la Salud
- A5: Ciencias Sociales y Humanas
- A6 : Economía, Administración, Contaduría y afines
- A7 : Ingeniería, Arquitectura, Urbanismo y afines
- A8 : Matemáticas y Ciencias Naturales

Aquí se encuentran repartidas cada carrera que se ofrece en todas las IES del país.
por eso con el siguiente código se convierte cada carrera en una categoría de A del 1 al 8.

```
# Creación de las áreas de la SNIES:
areas = c('A1', 'A2', 'A3', 'A4',
          'A5', 'A6', 'A7',
          'A8')

# Estandarización para una mejor categorización

data$carrera_o_licenciatura_a_la_que_se_presenta %<>%
  tolower() %>% chartr("áéíóú", "aeiou", .)

# Clasificación de las carreras presentes en clases de A(1-8)

data$carrera_o_licenciatura_a_la_que_se_presenta <-
  str_replace_all(data$carrera_o_licenciatura_a_la_que_se_presenta,
    c("(^(ing).*)|(^ign).*)" = "A7", "(^(arq).*)" = "A7",
      "(.*(agr).*)|(^zoo).*)" = "A1", "(.*(vet).*)" = "A1",
      "(.*(esta).*)|(^mate).*)" = "A8",
      "(^(fis).*)|(^astro).*)" = "A8",
      "(^(eco).*)|(^nego).*)" = "A6",
      "(^(admi).*)|(^archivis).*)" = "A6",
      "(^(conta).*)" = "A6",
      "(^(lic).*)" = "A3", "(.*(polit).*)" = "A5",
      "(.*(hist).*)" = "A5",
      "(.*(derec).*)" = "A5", "(.*(filo).*)" = "A5",
      "(.*(teolo).*)" = "A5",
      "(.*(antro).*)|(^perio).*)" = "A5",
      "(^(pedag).*)|(^entren).*)" = "A3",
      "(.*(socio).*)|(^traducc).*)" = "A5",
      "(.*(arte).*)|(^cult).*)" = "A2",
      "(.*(multime).*)|(^maquilla).*)" = "A2",
      "(^(pregrado).*)|(^bioingenieria).*)" = "A7",
      "(.*(softw).*)" = "A7",
      "(.*(medici).*)|(^salud).*)" = "A4",
      "(.*(trumenta).*)|(^enferm).*)" = "A4",
      "(.*(odonto).*)|(^cologia).*)" = "A4",
      "(.*(farma).*)" = "A4", "(.*(sistema).*)" = "A7"))

# Presentación de la nueva variable con sus nuevos valores:

data %<>% rename(., areas = carrera_o_licenciatura_a_la_que_se_presenta)

#data$areas %>% table() %>% as.data.frame() %>% xtable()
```

Ahora vemos que para la siguiente también tenemos problemas seccionando la segunda opción del aspirante, por lo cual además de las ocho áreas de las SNIES tenemos una nueva dado que hubieron estudiantes que decían que no les interesa una segunda carrera.

```
# data$si_su_anterior_respuesta_es_si_indica_cual %>% unique() %>% unique()
```

	.	Freq
1	A1	15
2	A2	11
3	A3	8
4	A4	32
5	A5	27
6	A6	17
7	A7	165
8	A8	15

estandarización de las respuestas:

```
data$si_su_anterior_respuesta_es_si_indica_cual %<>%
  tolower() %>% chartr("áéíóú", "aeiou", .)
```

Clasificación de las carreras como segunda opción de los aspirantes:

```
data$si_su_anterior_respuesta_es_si_indica_cual <-
  str_replace_all(data$si_su_anterior_respuesta_es_si_indica_cual,
    c("(^ (ing).*)|^ (ign).*)" = "A7", "(^ (arq).*)" = "A7",
      "(.* (agr).*)|^ (zoo).*)" = "A1", "(.* (vet).*)" = "A1",
      "(.* (esta).*)|^ (mate).*)" = "A8",
      "(^ (fis).*)|^ (astro).*)" = "A8",
      "(^ (eco).*)|^ (nego).*)" = "A6",
      "(^ (admi).*)|^ (archivis).*)" = "A6",
      "(^ (conta).*)" = "A6",
      "(^ (lic).*)" = "A3",
      "(.* (polit).*)" = "A5", "(.* (hist).*)" = "A5",
      "(.* (derec).*)" = "A5", "(.* (filo).*)" = "A5",
      "(.* (teolo).*)" = "A5",
      "(.* (antro).*)|^ (perio).*)" = "A5",
      "(^ (pedag).*)|^ (entren).*)" = "A3",
      "(.* (socio).*)|^ (traduc).*)" = "A5",
      "(.* (arte).*)|^ (cult).*)" = "A2",
      "(.* (multime).*)|^ (maquilla).*)" = "A2",
      "(^ (pregrado).*)|^ (bioingenieria).*)" = "A7",
      "(.* (softw).*)" = "A7",
      "(.* (medici).*)|^ (salud).*)" = "A4",
      "(.* (trumenta).*)|^ (enferm).*)" = "A4",
      "(.* (odonto).*)|^ (cologia).*)" = "A4",
      "(.* (farma).*)" = "A4", "(.* (una).*)|^ (so).*)" = "no",
      "(.* (interesa).*)|^ (nt).*)" = "no",
      "(.* ([.]).*)|^ (uwu).*)" = "no",
      "(.* (music).*)|^ (audio).*)" = "A2",
      "(.* (cine).*)" = "A2",
      "(.* (biol).*)|^ (nutri).*)" = "A4",
      "(.* (bact).*)" = "A4",
      "(.* (quimi).*)" = "A8",
      "(.* (adminis).*)" = "A6", "(.* (educac).*)" = "A3",
      "(.* (web).*)|^ (mecatro).*)" = "A7",
      "(.* (web).*)|^ (aviaci).*)" = "A7",
      "(.* ([n/a]).*)" = "no"))
```

```
# Visualización de la segunda opción por áreas de conocimiento:

data %<>% rename(., segunda_opcion = si_su_anterior_respuesta_es_si_indica_cual)

# data$segunda_opcion %>% table() %>% as.data.frame() %>% xtable()
```

	.	Freq
1	A1	12
2	A2	8
3	A3	6
4	A4	16
5	A5	14
6	A6	27
7	A7	66
8	A8	8
9	no	108

una vez tenemos nuestras variables depuradas, ahora pasamos a generar un subset, es decir, otra base de datos con las variables que son relevantes para el problema (se quita la marca temporal de la fecha y hora de la encuesta).

Base de datos final

Ahora sacamos las tres primeras columnas que corresponden a datos que no son relevantes:

- X1: numeración de la fila
- marca_temporal: fecha y hora de la realización de la encuesta
- por_favor_acepta_nuestra_politica_de_privacidad: aceptación de los tratamientos de los datos.

```
# Base de datos final:
data <- data[-c(1:3)]
```

ahora verificamos que todo este correcto en nuestro nuevo data set:

```
# Cambiamos fluido por Fluido
data[199,1] <- "Fluido"
data[16,1] <- "Fluido"

# Convertimos en factor las siguientes variables:

data$genero_al_que_pertenece %<>% factor()

data$tipo_de_colegio_que_estudia_o_estudio %<>% factor()

data$en_que_ano_salio_o_saldra_del_colegio %<>% factor()

data$trabaja_actualmente %<>% factor()
```

skim_type	skim_variable	n_missing	complete_rate	factor.ordered	factor.n_unique	factor.top_counts
factor	genero_al_que_pertenece	0	1.0000000	FALSE	4	Fem: 150, Mas: 135, Flu: 4, Bin: 1
factor	tipo_de_colegio_que_estudia_o_estudio	0	1.0000000	FALSE	2	Pub: 203, Pri: 87
factor	en_que_año_salio_o_saldrá_del_colegio	0	1.0000000	FALSE	5	202: 240, 202: 38, 201: 9, 201: 2
factor	trabaja_actualmente	0	1.0000000	FALSE	2	No: 235, Si: 55
factor	si_su_anterior_respuesta_es_si_su_trabajo_esta_relacionado_con_la_carrera_a_la_cual_se_presento	46	0.8413793	FALSE	2	No: 209, Si: 35
factor	areas	0	1.0000000	FALSE	8	A7: 165, A4: 32, A5: 27, A6: 17
factor	escoja_una_de_las_siguietes_razones_por_la_que_va_a_estudiar	0	1.0000000	FALSE	11	Me: 108, Qui: 92, Qui: 20, Con: 16
factor	seleccione_la_principal_razon_para_la_eleccion_de_la_carrera_anterior	0	1.0000000	FALSE	21	Las: 105, Las: 55, Rec: 21, La: 20
factor	le_interesa_otra_carrera_diferente_a_la_que_se_presento	0	1.0000000	FALSE	3	Si: 170, No: 113, Pue: 7
factor	segunda_opcion	25	0.9137931	FALSE	9	no: 108, A7: 66, A6: 27, A4: 16
factor	la_anterior_carrera_esta_disponible_en_esta_u_otra_universidad	0	1.0000000	FALSE	3	Si: 263, No: 20, No: 7
factor	cuenta_con_algun_estudio_curso_tecnico_entre_otros_certificable_que_este_relacionado_con_la_carrera_que_seleccione	0	1.0000000	FALSE	2	No: 194, Si: 96
factor	universidad_a_la_que_aspira	12	0.9586207	FALSE	6	Uni: 163, Uni: 105, Uni: 6, Tod: 2

```
data$si_su_anterior_respuesta_es_si_su_trabajo_esta_relacionado_con_la_carrera_a_la_cual_se_presento %<>%
  factor()

data$areas %<>% factor()

data$escoja_una_de_las_siguietes_razones_por_la_que_va_a_estudiar %<>% factor()

data$sseleccione_la_principal_razon_para_la_eleccion_de_la_carrera_anterior %<>%
  factor()

data$le_interesa_otra_carrera_diferente_a_la_que_se_presento %<>% factor()

data$segunda_opcion %<>% factor()

data$la_anterior_carrera_esta_disponible_en_esta_u_otra_universidad %<>% factor()

data$cuenta_con_algun_estudio_curso_tecnico_entre_otros_certificable_que_este_relacionado_con_la_carrera_que_seleccione %<>%
  factor()

# Cambiamos algunos valores para evitar más categorías.

data[144:145,13] <- rep("Todas las anteriores", 2)

# Convertir a factor la ultima variable

data$universidad_a_la_que_aspira %<>% factor()
```

Para ver un resumen rápido tipo `summary()` con la función `skim` de la librería `skimr` que lleva el mismo nombre tenemos la siguiente tabla:

Análisis de nuestras variables:

antes de empezar el análisis tenemos que en la variable `sexo` esta se registra como “Maculino” y con la función siguiente cambiamos el valor a “Masculino”.

```
## change factor levels
data <- mutate(data,
  genero_al_que_pertenece =
    refactor(genero_al_que_pertenece,
      levs = c("Binario", "Femenino", "Fluido"),
      repl = "Masculino"))
```

Base de datos organizada

```
# filter and sort the dataset
# data %>%
#   select(genero_al_que_pertenece:universidad_a_la_que_aspira) %>% head() %>%
#   xtable()
```

	genero_al_que_pertenece	tipo_de_colegio_que_estudia	ano_saldrá_del_colegio	trabaja_actualmente	su_trabajo
1	Masculino	Publico	2022	No	No
2	Masculino	Publico	2022	Si	No
3	Masculino	Publico	2022	No	No
4	Masculino	Publico	2022	No	No
5	Masculino	Privado	2022	No	No
6	Femenino	Publico	2021	No	No

Aquí vemos el encabezado de nuestros datos, es decir, los primeros 5 datos de nuestra base de datos de 290 observaciones.

Tabla de Frecuencias, con su respectivo gráfico para una mejor visualización

Tabla y gráfico del género alusivo.

```
result <- pivotr(
  data,
  cvars = "genero_al_que_pertenece",
  nr = Inf
)
# summary()
#result$tab_freq %>% xtable()
```

	genero_al_que_pertenece	n_obs
1	Binario	1
2	Femenino	150
3	Fluido	2
4	género fluido	2
5	Maculino	135
6	Total	290

Podemos notar que en la muestra tomada (de forma sistemática de 1-20) pudimos visualizar un mayor número de mujeres que fueron más receptivas al momento de realizar la encuesta, además, contamos con algunos otros géneros que no son significativos como el Binario y el Fluido.

Para el caso del tipo de colegio donde estudió o estudia:

```
result <- pivotr(
  data,
  cvars = "tipo_de_colegio_que_estudia_o_estudio",
  nr = Inf
)
# summary()
#result$tab_freq %>% xtable()
```

	tipo_de_colegio_que_estudia_o_estudio	n_obs
1	Privado	87
2	Publico	203
3	Total	290

Son mayoritarios los colegio que estudian en colegios del estado, revelando que la mayoría de los aspirantes a universidades publicas no vienen de colegios privados.

```
result <- pivotr(
  data,
  cvars = "en_que_ano_salio_o_saldra_del_colegio",
  normalize = "total",
  nr = Inf
)
# summary()
# result$tab_freq %>% xtable()
```

	en_que_ano_salio_o_saldra_del_colegio	n_obs
1	2017	1
2	2018	2
3	2019	9
4	2021	38
5	2022	240
6	Total	290

Las promociones de 2022, siempre presentan un numero mayoritario en los aspirantes esto se debe al segundo corte o la prueba de final de año, para el ingreso en el 2023 esto lo decimos ya que podríamos afirmar que en le primer corte los aspirantes salieron del colegio en épocas diferentes, más homogeneidad.

```
result <- pivotr(
  data,
  cvars = "trabaja_actualmente",
  nr = Inf
)
# summary()
# result$tab_freq %>% xtable()
```

	trabaja_actualmente	n_obs
1	No	235
2	Si	55
3	Total	290

En el segundo corte se esperan aspirantes que aun no han salido de estudiar, por ende la mayoría no cuenta con un trabajo, podríamos afirmar que para el primer corte del 2023 estos estudiantes que no ingresaron posiblemente estén laborando para entonces.

```
result <- pivotr(
  data,
  cvars = "areas",
  tabsort = "areas",
  nr = Inf
)
```



```
)
# summary()
# result$tab_freq %>% xtable()
```

	areas	n_obs
1	A1	15
2	A2	11
3	A3	8
4	A4	32
5	A5	27
6	A6	17
7	A7	165
8	A8	15
9	Total	290

Ahora una de las variables más importantes, es saber cuales son las aspiraciones en áreas de conocimiento de los que hicieron examen de admisión, por lo cual planteamos una vez más las diferentes áreas del saber, en las cuales se aterrizan todas las carreras brindadas en el estado de Colombia:

- A1 = Agronomía, Veterinaria y afines
- A2 = Bellas Artes
- A3 = Ciencias de la Educación
- A4 = Ciencias de la Salud
- A5 = Ciencias Sociales y Humanas
- A6 = Economía, Administración, Contaduría y afines
- A7 = Ingeniería, Arquitectura, Urbanismo y afines
- A8 = Matemáticas y Ciencias Naturales

Ahora, observando la tabla tenemos que las ingenierías, arquitectura, urbanismo y afines, son el área más demandada de nuestra muestra.

En parte era evidente ya que gran parte de las encuestas fueron realizadas en la UNAL sede Medellín, una universidad que la mayoría, por no decir todas las carreras son de esta área.

```
result <- pivotr(
  data,
  cvars = "escoja_una_de_las_siguietes_razones_por_la_que_va_a_estudiar",
  tabsort = "desc(n_obs)",
  nr = Inf
)
# summary()
# result$tab_freq %>% xtable()
```

Una de las razones por la cual los aspirantes desean entrar a una universidad es porque tienen un alto interés en superarse o bien en estudiar.

	escoja_una_de_las_siguientes_razones_por_la_que_va_a_estudiar	n_obs
1	Así lo hacen mis parientes, mis amigos, mi novio (a)	1
2	Así lo quieren mis papás	4
3	Con los estudios que tengo no podré conseguir un buen trabajo	16
4	Es normal	7
5	Me interesa estudiar	108
6	Me va a dar prestigio.	10
7	No tengo otra cosa que hacer	12
8	Quiero ganar bien cuando sea profesionista	14
9	Quiero superarme	92
10	Quiero vivir más independiente de mi familia	20
11	Un (a) maestro (a) me dijo que siguiera estudiando	6
12	Total	290

```
result <- pivotr(
  data,
  cvars = "seleccione_la_principal_razon_para_la_eleccion_de_la_carrera_anterior",
  tabsort = "desc(n_obs)",
  nr = Inf
)
# summary()
# result$tab_freq %>% xtable()
```

	seleccione_la_principal_razon_para_la_eleccion_de_la_carrera_anterior	n_obs
1	Esta relacionado con temáticas que me apasionan	2
2	Influencia familiar	16
3	Interés propio	1
4	La facilidad para el pago de esta carrera o para obtener becas	20
5	Las asignaturas del pensum y las líneas de énfasis son de mi interés.	105
6	Las materias de la carrera me llaman la atención	55
7	llamativa	1
8	Los egresados de esta carrera consiguen empleo fácilmente	17
9	Los egresados de esta carrera son bien remunerados o tienen buenas ofertas laborales salariales.	16
10	Me gusta	1
11	Pasión de infancia	3
12	Pasión desde la infancia	3
13	Porque me apasiona	1
14	Porque me gusta	1
15	Quiero aprender a defender la palabra de Dios con la apologetica	3
16	Quiero entender el mundo	1
17	Recomendación de un amigo	21
18	Seguir mi media tecnica y ganar dinero	1
19	Tendencia	13
20	vocación	3
21	Vocación	6
22	Total	290

Ahora esperamos que sea cierto que la mayoría de los aspirantes hayan leído el pensum, para decidir que esta es su linea de interés.

```
result <- pivotr(
  data,
  cvars = "cuenta_con_algun_estudio_curso_tecnico_entre_otros_certificable_que_este_relacionado_con_la_carrera",
  nr = Inf
)
# summary()
# result$tab_freq %>% xtable()
```

	cuenta_con_algun_estudio_curso_tecnico_certificable	n_obs
1	No	194
2	Sí	96
3	Total	290

Aunque creímos que los aspirantes para el 2023 no tendrían una fundamentación en el programa que eligieron, realmente nos dejaron sorprendidos al darnos cuenta que casi 100 personas tienen estudios que se relacionen con su carrera.

```
result <- pivotr(
  data,
  cvars = "le_interesa_otra_carrera_diferente_a_la_que_se_presento",
  nr = Inf
)
# summary()
# result$tab_freq %>% xtable()
```

	le_interesa_otra_carrera_diferente_a_la_que_se_presento	n_obs
1	No	113
2	Puede ser	7
3	Si	170
4	Total	290

La decisión de tomar una carrera y no una línea de conocimiento, es marcada en Colombia ya que observamos a personas cerradas en una carrera, aunque estos mismos después eran los que esperaban al menos dos universidades para ingresar a la misma carrera.

```
result <- pivotr(
  data,
  cvars = "universidad_a_la_que_aspira",
  tabsort = "desc(n_obs)",
  nr = Inf
)
# summary()
# result$tab_freq %>% xtable()
```

Como dijimos anteriormente, la mayoría de nuestros encuestados fueron en la UNAL sede medellín, esto se debe a que hubo una mayor cobertura al ingresar en la universidad, lamentablemente no fue el caso de la UdeA, la gente fue menos receptiva.

```
result <- pivotr(
  data,
  cvars = c(
```

	universidad_a_la_que_aspira	n_obs
1	Itm	1
2	Todas las anteriores	2
3	Universidad de Antioquia	105
4	Universidad de Medellín	1
5	Universidad Nacional de Colombia	163
6	Universidad Nacional de Colombia, Universidad de Antioquia	6
7	NA	12
8	Total	290

```

    "genero_al_que_pertenece",
    "areas"
  ),
  data_filter = "genero_al_que_pertenece %in% c('Femenino', 'Masculino')",
  tabsort = "desc(Masculino)",
  nr = Inf
)
# summary()
# result$tab_freq %>% xtable()

```

	areas	Masculino	Femenino	Total
1	A1	5	9	14
2	A2	5	5	10
3	A3	4	4	8
4	A4	9	22	31
5	A5	13	13	26
6	A6	6	11	17
7	A7	83	81	164
8	A8	10	5	15
9	Total	135	150	285

Tanto hombres como mujeres, prefieren las ingenierías, seguidas de las Ciencias Sociales y Humanas.

```

result <- pivotr(
  data,
  cvars = c(
    "trabaja_actualmente",
    "cuenta_con_algun_estudio_curso_tecnico_entre_otros_certificable_que_este_relacionado_con_la_carrera"
  ),
  data_filter = "genero_al_que_pertenece %in% c('Femenino', 'Masculino')",
  nr = Inf
)
# summary()
# result$tab_freq %>% xtable()

```

	cuenta_con_algun_estudio_curso_tecnico_certificable	No	Si	Total
1	No	154	35	189
2	Sí	78	18	96
3	Total	232	53	285

Vemos como son muy pocos, exactamente 18 persona de las 290 que trabajan actualmente y este trabajo tiene que ver con la carrera seleccionada el día del examen.

```

result <- pivotr(
  data,
  cvars = c(
    "tipo_de_colegio_que_estudia_o_estudio",
    "escoja_una_de_las_siguientes_razones_por_la_que_va_a_estudiar"
  ),
  tabsort = "desc(Publico)",
  nr = Inf
)
# summary()
# result$tab_freq %>% xtable()

```

	escoja_una_de_las_siguientes_razones_por_la_que_va_a_estudiar	Privado	Publico	Total
1	Así lo hacen mis parientes, mis amigos, mi novio (a)	0	1	1
2	Así lo quieren mis papás	1	3	4
3	Con los estudios que tengo no podré conseguir un buen trabajo	7	9	16
4	Es normal	1	6	7
5	Me interesa estudiar	37	71	108
6	Me va a dar prestigio.	3	7	10
7	No tengo otra cosa que hacer	3	9	12
8	Quiero ganar bien cuando sea profesionista	2	12	14
9	Quiero superarme	27	65	92
10	Quiero vivir más independiente de mi familia	5	15	20
11	Un (a) maestro (a) me dijo que siguiera estudiando	1	5	6
12	Total	87	203	290

Otra comparación que resulta llamativa es comparar el tipo de colegio en el que estudio o estudia y la razón por la cual quiere entrar a una universidad.

Son muy parecidos en el conteo de los estudiantes que quieren superarse y les interesa estudiar que salieron de colegios públicos.