Introducción:

Criterios de depuración:

Al comenzar a explorar la base de datos se encuentra que el campo del ***programa*** al cual pertenecía cada estudiante, algunos los escribían en mayúsculas y otros en minúscula, esto significa algo problemático para el R ya que diferencia mayúsculas y minúsculas dando lugar a duplicación de información; por lo cual el primer criterio es colocar que la carrera comience siempre con mayúsculas.

Nos resultan 6 niveles {Estadística, I Administrativa, Forestal, I Industrial, I Sistemas, Zootecnia}.

Columna ***Problemática*:**

El criterio para unificar esta será colocar en una sola palabra global varias que se podrían englobarse en ella además de unificar algunas que tengan problemas de escritura

**ESTO PODRIA SER EN UNA TABLA.**

*Corrupción*= {'corrupción','mal manejo dinerodel gob','corrupción','corrupción política', 'política'}.

Aunque la corrupción se da en muchos ámbitos, tanto privados como públicos asumiremos esta como una sola.

*Problema agrario*= {'Agricultura','las lluvias','TLC'}.

*Seguridad vial* = {'alta tasa accidentes','conductores ebrios'}.

*Seguridad ciudadana* = {'bacrim','impunidad derechos',' inseguridad ','la justicia','seguridad','violencia','narcotráfico'}.

Aunque el narcotráfico tiene grandes componentes económicos tomaremos la problemática que esta ha generado en el tema de la ciudadanía, recordando hechos tristes de terrorismo como carros bombas, asesinatos en los últimos años.

*Conflicto armado* = {'charlas con guerrilla','conflicto armado','guerra Intena','la guerra','Proceso de paz'}.

*Economía y equidad*= {‘desigualdad’,'desplazamiento','La economía','pobreza','Pobreza','hidroeléctricas','paro Naciona' }.

*Educación*= {'educación','educación Calidad'}.

*Salud* = {salud}.

Algunas personas dejaron en blanco este campo, el cual se decidió cambiar por un “No responde”.

En total quedaron 9 niveles para esta columna.

Columna ***PARTIDO***:

Se pudo observar casi la presencia de todos los partidos, la depuración consistió en solamente lo siguiente.

u = {uribista}.

No responde = {“ ”, no respondo}.

Ninguno= {na,ninguno,n/a}.

Por simplicidad vamos a poner na,n/a y ninguno como uno solo.

En total quedaron 9 niveles: conservador, izquierda, liberal, mira, polo, u , verde , ninguno, no responde.

MIRAR BIEN LO DE IZQUIERA Y LIBERAL.

Columna **POLITICO.ID**

Esta columna se refiere a con que político se siente identificado el estudiante.

La depuración consistió en

No responde={“ ”, no respondo}.

Ninguno ={na,ninguno,n/a,ns/nr}.

En total quedaron 8 niveles: fajardo, jorge r, navarro, ninguno, no responde, petro , uribe y vargas ll.

Columna **VOTAR:**

Esta columna registra si el estudiando tiene pensado votar en las próximas elecciones.

La depuración consistió en

No={no, no puedo}.

Si={si}.

No responde = {“ “, no sabe}.

Columna **VP. VOTAR**:

Esta columna registra si para el estudiante vale la pena votar en las próximas elecciones colombianas.

La depuración consistió en

Si= {si,cvb si }.

No = {no,n}.

No responde = {“ ”, no sabe, Ns/nr,depende}.

En total quedaron 4 niveles: en blanco, No,No responde y Si.

Columna **POSTURA.EUN**

Esta columna registra si como estudiante de La Universidad Nacional de Colombia es necesario tener una postura política clara.

La depuración consistió en

Si = {si, Si}.

No={no}.

No responde = { “ ” , no necesariamente, n/a,posiblmente} .

En total quedaron 4 niveles: en blanco, No, No responde, Si.

Columna **PERT.PARTIDO**

Esta columna registra si el estudiante de La Universidad Nacional de Colombia debe pertenecer a un partido político?.

No = {n o, no}.

No responde= {No sabe,no necesariamente, depende}.

Si = {s,si}.

En total quedaron 3 niveles: No, No responde, Si.

Columna **VOTO.P:**

Esta columna registra si el padre del estudiante voto en las anteriores elecciones presidenciales.

Si={si}.

No={no,fallecido}.

No responde= {“ ”, na}

No sabe={no sabe}.

Las personas que no sabían si su padre había votado quedaron unificadas en las que no respondían al no ser muchas.

En total quedaron 3 niveles: No, No responde, Si.

Columna **CANDIDATO.P**

Esta columna registra por cual candidato voto el padre.

La depuración consistió en

No responde={“ ”, no respondo,’ ’}.

No aplica ={na,n/a }.

En blanco={en blanco }

Santos= ‘santos’.

No sabe = {No sabe}.

Columna **VOTO.M:**

Esta columna registra si la madre del estudiante voto en las anteriores elecciones presidenciales.

Si={si}.

No={no }.

No responde= {“ ”}

No sabe={no sabe}.

Columna **CANDIDATO.M**

Esta columna registra por cual candidato voto la madre.

La depuración consistió en

No responde={“ ”, no respondo,’ ’}.

No aplica ={ n/a }.

Mockus={mockus}.

Fajardo={fajardo}.

Uribe= {uribe}.

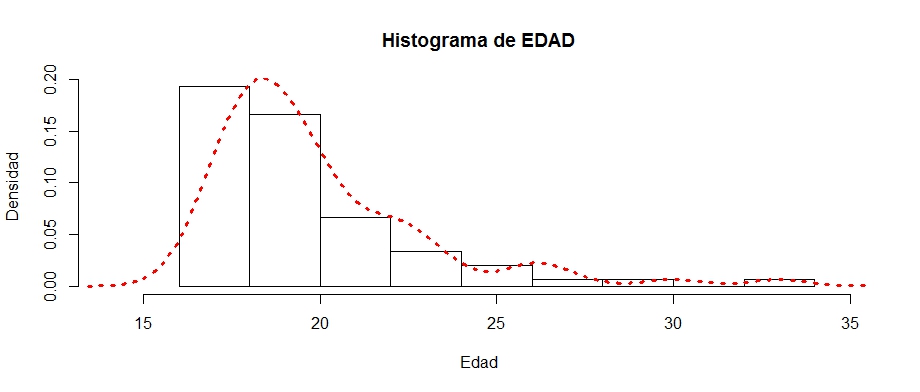
En blanco={en blanco }

* Análisis exploratorio de cada variable.

HACER BIEN EL DE LA EDAD!!!!!.

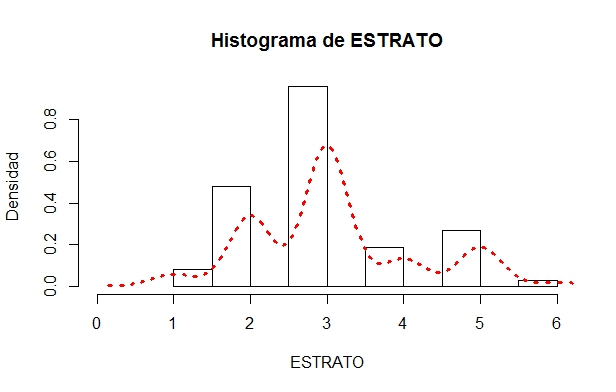
Variable **Edad**:

|  |  |
| --- | --- |
| EDAD | frecuencia |
| 18 | 18 |
| 19 | 15 |
| 17 | 10 |
| 20 | 10 |
| 22 | 6 |
| 21 | 4 |
| 23 | 4 |
| 26 | 3 |
| 16 | 1 |
| 24 | 1 |
| 27 | 1 |
| 30 | 1 |
| 33 | 1 |



Variable **ESTRATO**:

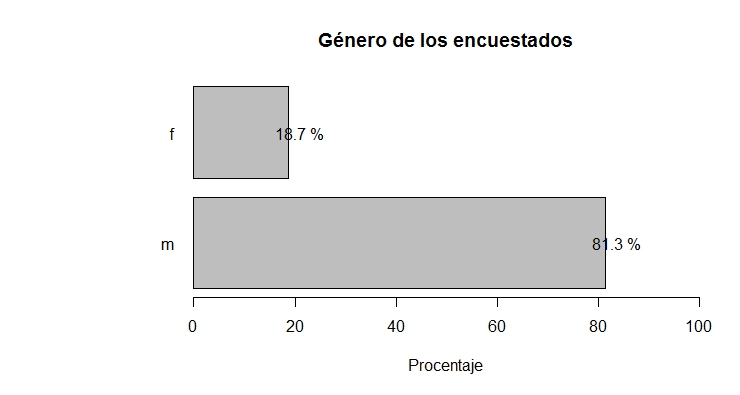
|  |  |
| --- | --- |
| ESTRATO | Frecuencia |
| 1 | 3 |
| 2 | 18 |
| 3 | 36 |
| 4 | 7 |
| 5 | 10 |
| 6 | 1 |



Variable **GENERO**

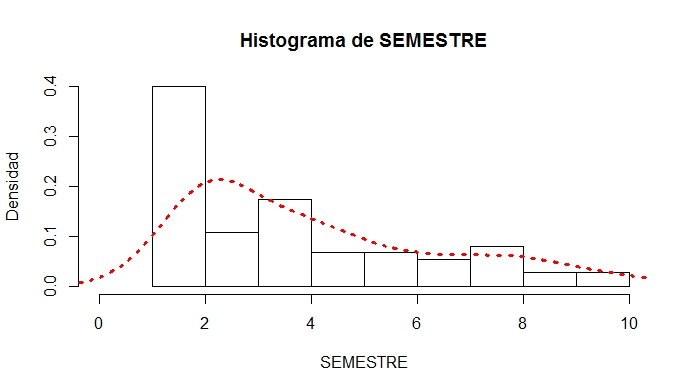
Tabla de frecuencia con su porcentaje:

|  |  |  |
| --- | --- | --- |
| Genero | Frecuencia | Porcentaje (%) |
| m | 61 | 81,3 |
| f | 14 | 18,7 |



Variable **SEMESTRE:**

|  |  |
| --- | --- |
| SEMESTRE | Frecuencia |
| 1 | 1 |
| 2 | 29 |
| 3 | 8 |
| 4 | 13 |
| 5 | 5 |
| 6 | 5 |
| 7 | 4 |
| 8 | 6 |
| 9 | 2 |
| 10 | 2 |



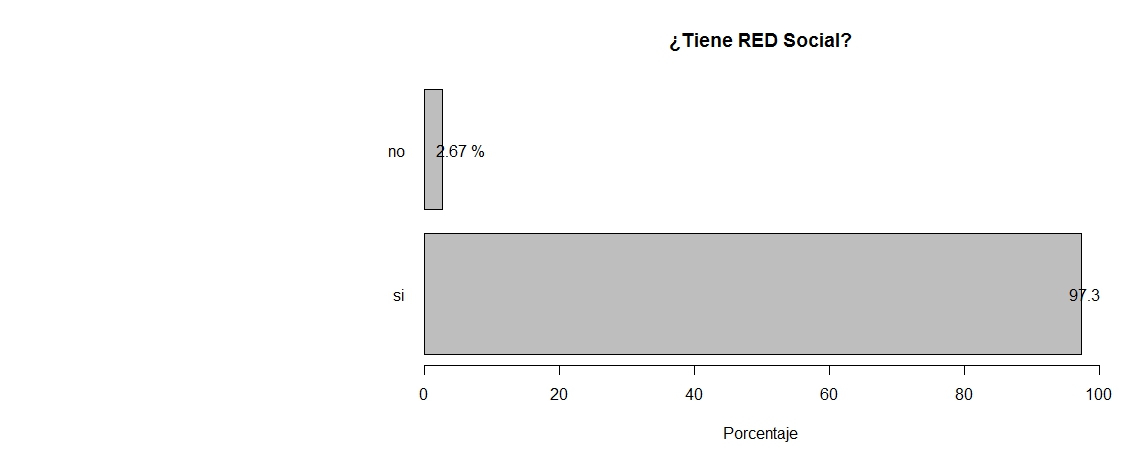
Variable **PROGRAMA**:

Tabla de frecuencia con su porcentaje:

|  |  |  |
| --- | --- | --- |
| PROGRAMA | Frecuencia | Porcentaje (%) |
| I Sistemas | 42 | 56 |
| Estadística | 26 | 34,67 |
| I Industrial | 3 | 4 |
| I Forestal | 2 | 2,67 |
| I Administrativa | 1 | 1,33 |
| C:\Users\EQUIPO JOSE\Dropbox\8 semestre\Estadistica Descriptiva\TrabajoIndividual\1\Entregable\Imagenes\Rplot.jpegZootecnia | 1 | 1,33 |

Variable **RED.SOCIAL.**

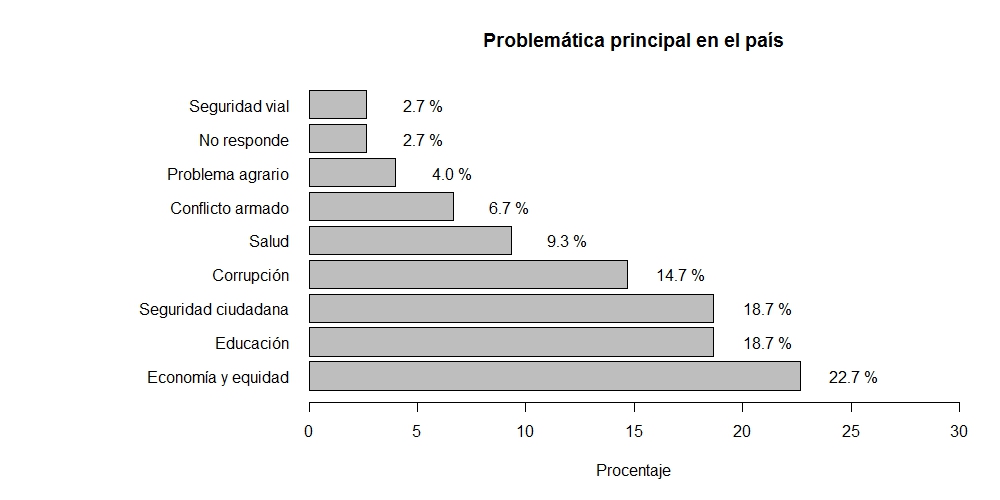
|  |  |  |
| --- | --- | --- |
|  | frecuencia | Porcentaje( % ) |
| si | 73 | 97,33 |
| no | 2 | 2,67 |



Variable**: Problemática**

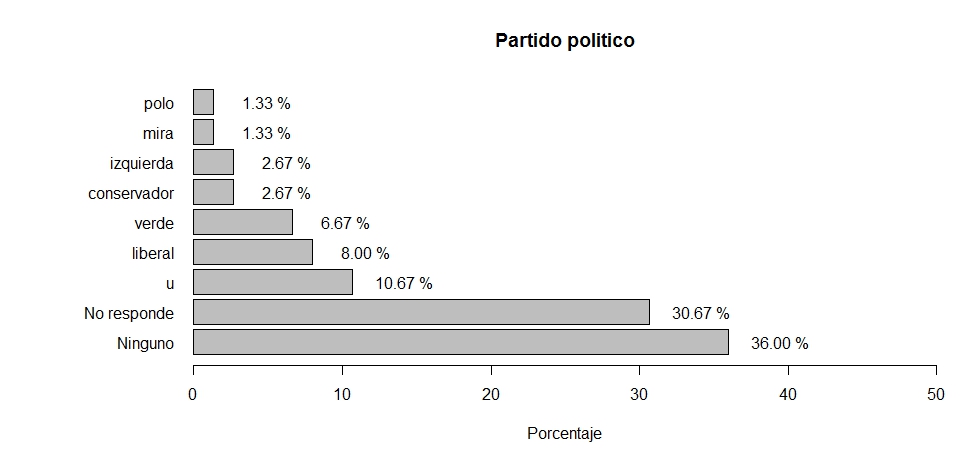
Tabla de frecuencia con su porcentaje:

|  |  |  |
| --- | --- | --- |
| Problemática | Frecuencia | Porcentaje (%) |
| Economía y equidad | 17 | 22,7 |
| Educación | 14 | 18,7 |
| Seguridad ciudadana | 14 | 18,7 |
| Corrupción | 11 | 14,7 |
| Salud | 7 | 9,3 |
| Conflicto armado | 5 | 6,7 |
| Problema agrario | 3 | 4 |
| No responde | 2 | 2,7 |
| Seguridad vial | 2 | 2,7 |

vvvbvbvbvbvbv

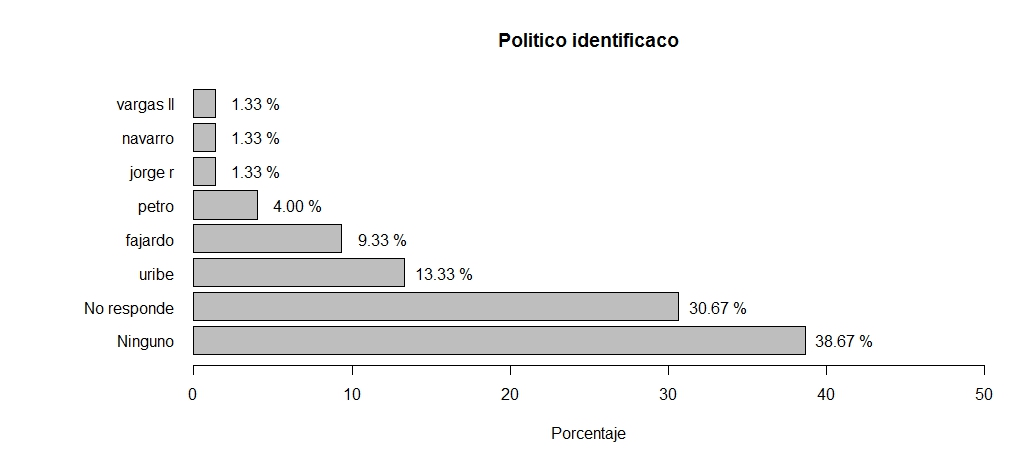
Variable **PARTIDO:**

|  |  |  |
| --- | --- | --- |
| Partido | Frecuencia | Porcentaje |
| Ninguno | 27 | 36,0 |
| No responde | 23 | 30,7 |
| u | 8 | 10,7 |
| liberal | 6 | 8,0 |
| verde | 5 | 6,7 |
| conservador | 2 | 2,7 |
| izquierda | 2 | 2,7 |
| mira | 1 | 1,3 |
| polo | 1 | 1,3 |



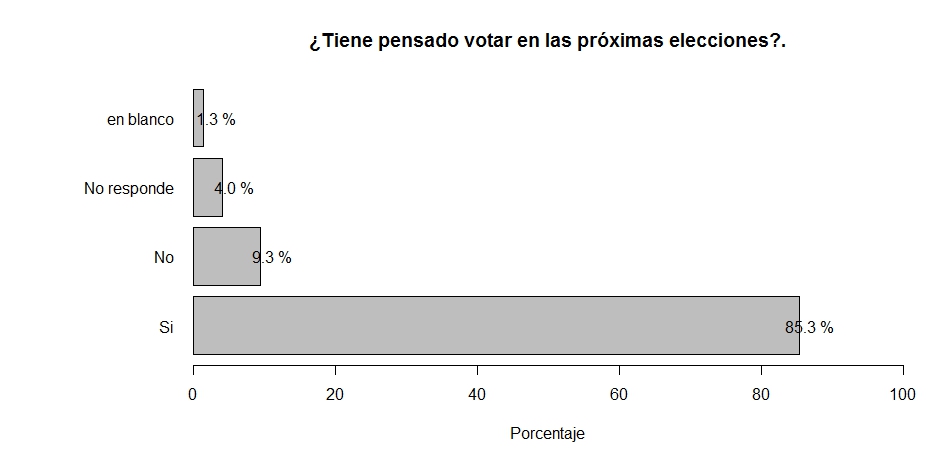
Variable **POLITICO.ID:**

|  |  |  |
| --- | --- | --- |
| POLITICO.ID | Frecuencia | Porcentaje |
| Ninguno | 29 | 38,67 |
| No responde | 23 | 30,67 |
| Uribe | 10 | 13,33 |
| fajardo | 7 | 9,33 |
| Petro | 3 | 4,00 |
| Jorge r | 1 | 1,33 |
| navarro | 1 | 1,33 |
| Vargas ll | 1 | 1,33 |

****

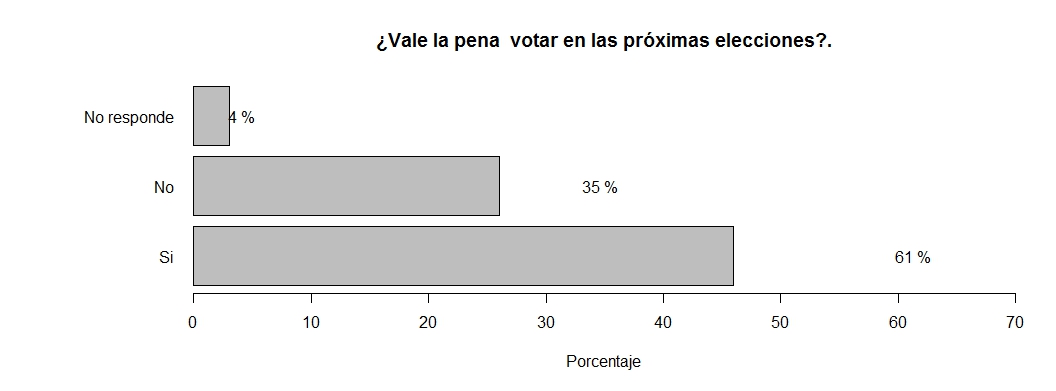
Variable **Votar**

|  |  |  |
| --- | --- | --- |
| VOTAR | Frecuencia | Porcentaje(%) |
| Si | 64 | 85,33 |
| No | 7 | 9,33 |
| No responde | 3 | 4,00 |
| en blanco | 1 | 1,33 |

****

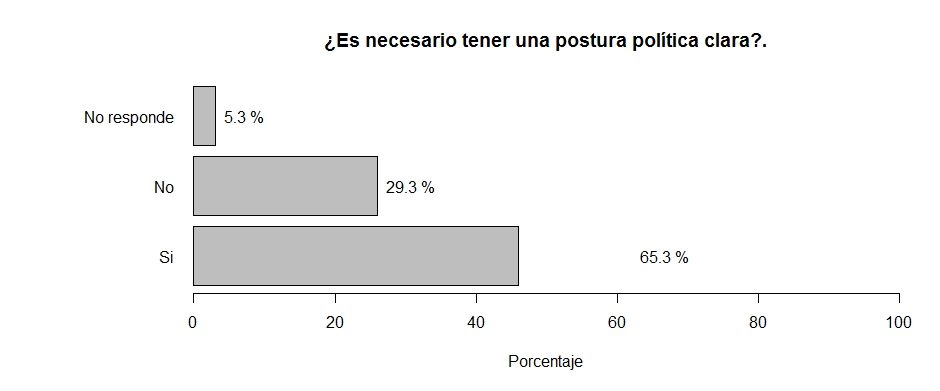
Variable **VP.VOTAR:**

|  |  |  |
| --- | --- | --- |
| VP.VOTAR | Frecuencia | Porcentaje (%) |
| Si | 46 | 61,3 |
| No | 26 | 34,7 |
| No responde | 3 | 4,0 |



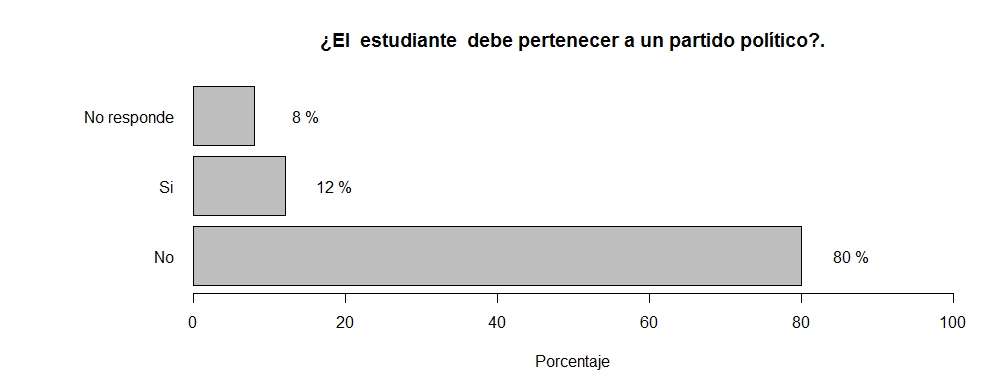
Variable **POSTURA.EUN:**

|  |  |  |
| --- | --- | --- |
| POSTURA.EUN | Frecuencia | Porcentaje (%) |
| Si | 49 | 65,3 |
| No | 22 | 29,3 |
| No responde | 4 | 5,3 |



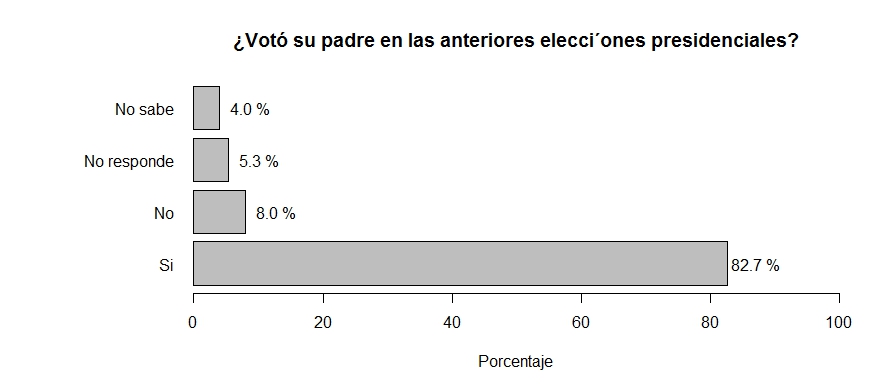
Variable **PERT.PARTIDO:**

|  |  |  |
| --- | --- | --- |
| PERT.PARTIDO | Frecuencia | Porcentaje (%) |
| Si | 49 | 65,3 |
| No | 22 | 29,3 |
| No responde | 4 | 5,3 |



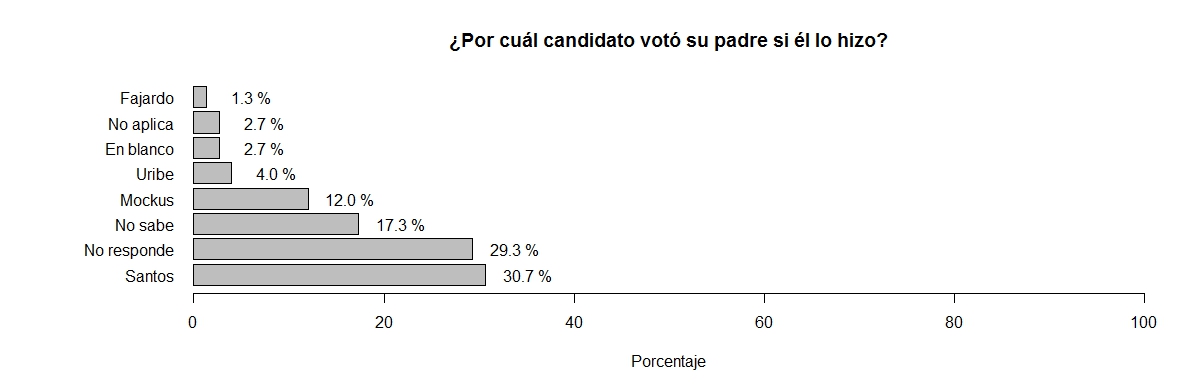
Variable **VOTO.P**

|  |  |  |
| --- | --- | --- |
| VOTO.P | Frecuencia | Porcentaje (%) |
| Si | 62 | 82,7 |
| No | 6 | 8,0 |
| No responde | 4 | 5,3 |
| No sabe | 3 | 4,0 |



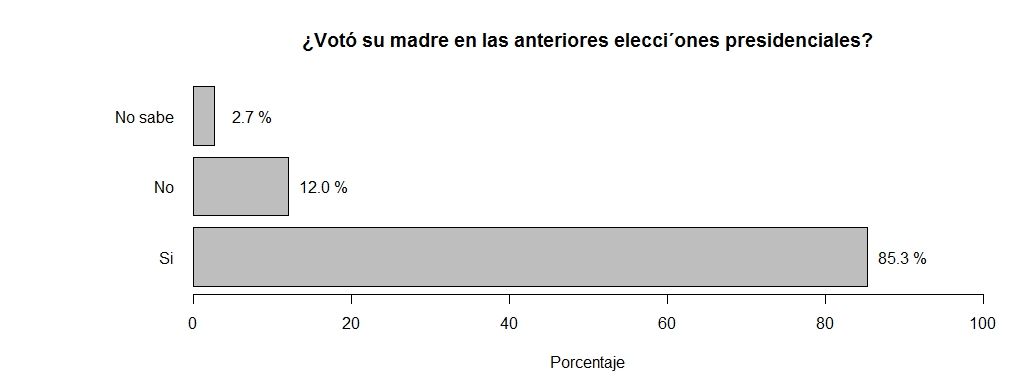
Variable **CANDIDATO.P**

|  |  |  |
| --- | --- | --- |
| CANDIDATO.P | Frecuencia | Porcentaje ( %) |
| Santos | 23 | 30,7 |
| No responde | 22 | 29,3 |
| No sabe | 13 | 17,3 |
| Mockus | 9 | 12,0 |
| Uribe | 3 | 4,0 |
| En blanco | 2 | 2,7 |
| No aplica | 2 | 2,7 |
| Fajardo | 1 | 1,3 |



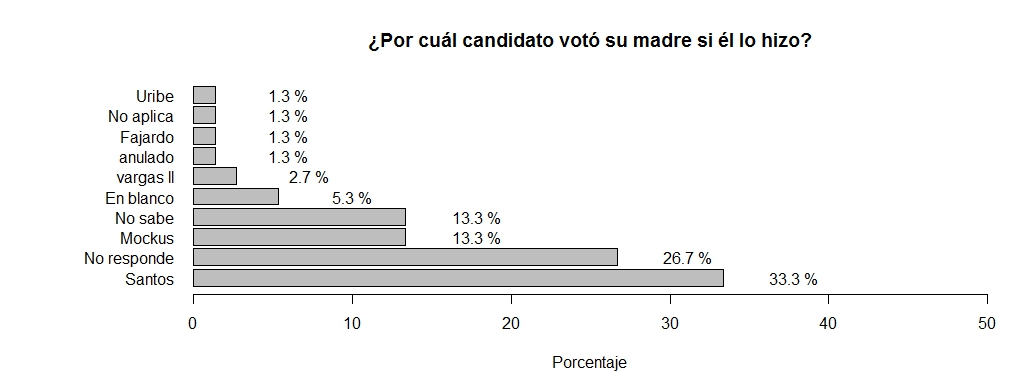
Variable **VOTO.M**

|  |  |  |
| --- | --- | --- |
| VOTO.M | Frecuencia | Porcentaje (% ) |
| Si | 64 | 85,3 |
| No | 9 | 12,0 |
| No sabe | 2 | 2,7 |



Variable **CANDIDATO.M**

|  |  |  |
| --- | --- | --- |
| CANDIDATO.M | Frecuencia | Porcentaje |
| Santos | 25 | 33,3 |
| No responde | 20 | 26,7 |
| Mockus | 10 | 13,3 |
| No sabe | 10 | 13,3 |
| En blanco | 4 | 5,3 |
| vargas ll | 2 | 2,7 |
| anulado | 1 | 1,3 |
| Fajardo | 1 | 1,3 |
| No aplica | 1 | 1,3 |
| Uribe | 1 | 1,3 |



## Hacer análisis de los partidos.

## PONER PIES DE PAGINA

# HACER ANALISIS GENERO POR CARRERA –

Cruce de Variables:

Conclusiones y recomendaciones:

Bibliografia: