

Presentación Final

Métodos cuantitativos y análisis de datos grandes

Lic. Maximiliano Cacace Bach

5 de julio 2021

Algunas características de los becarios de la ECyT-UNSAM

En este trabajo se describe al grupo de becarios de grado de la Escuela de Ciencia y Tecnología de la Universidad Nacional de San Martín del año 2020. Se basa en datos obtenidos a través de formularios de encuesta. Son datos de uso institucional. Se propone una hipótesis para ponerla a prueba a los fines de este seminario.

Es un estudio de tipo exploratorio-descriptivo. La población es finita, son lxs estudiantes con beca de apoyo económico, la cual a su vez es una muestra y representación de la población total estudiantil de la Escuela de Ciencia y Tecnología de la UNSAM.

Datos

Los datos reúnen las respuestas obtenidas a partir de encuestas a becarixs de la ECyT, en período 2020. Se han obtenido 164 respuestas de un total de 184 becarios.

La encuesta tiene 32 + - 14 preguntas/variables, en función de la opción que eligieron. Para este trabajo se seleccionaron 5 variables que combinan variables nominales (3) y de ratio (2).

Variables analizadas

Edad de los becarios

Esta variable es:

- ▶ De ratio y
- ▶ Discreta

Tiempo de beca

Esta variable es:

- ▶ De ratio
- ▶ de 5 niveles y
- ▶ Continua

Variables analizadas II

Costeo de estudios

Esta variable es:

- ▶ Nominal
- ▶ de 5 niveles y
- ▶ Discreta

Trabajo

Esta variable es:

- ▶ Nominal
- ▶ de 2 niveles y
- ▶ Discreta

Con quien vive

Esta variable es:

- ▶ Nominal
- ▶ de 4 niveles y
- ▶ Discreta

Tipo de Estudio: Descriptivo

Se hace un análisis descriptivo de los datos. Se eligen medidas de:

- ▶ Tendencia central -> Media para Edad y Tiempo de Estudio
- ▶ Dispersión -> Rango, Rango intercuartil en Edad y Desviación Estándar en Edad y Tiempo de Estudio
- ▶ Como no se puede sacar en R, no está la Moda de las demás variables

Hipótesis

- ▶ H_0 : La variable Edad tiene distribución normal
- ▶ H_1 : La variable Edad no tiene distribución normal

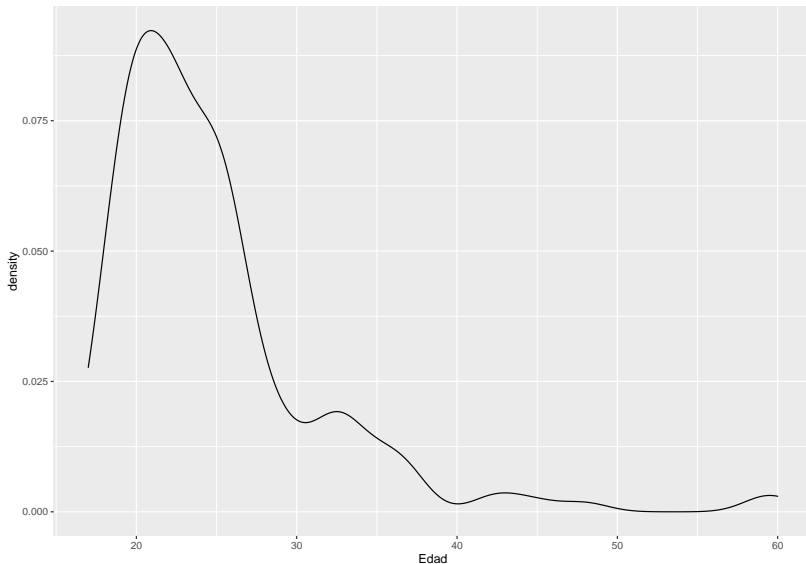
Hipótesis: Test aplicable

Para refutar o no la hipótesis nula se utilizará

Prueba paramétrica:

*El test de Shapiro-Wilks

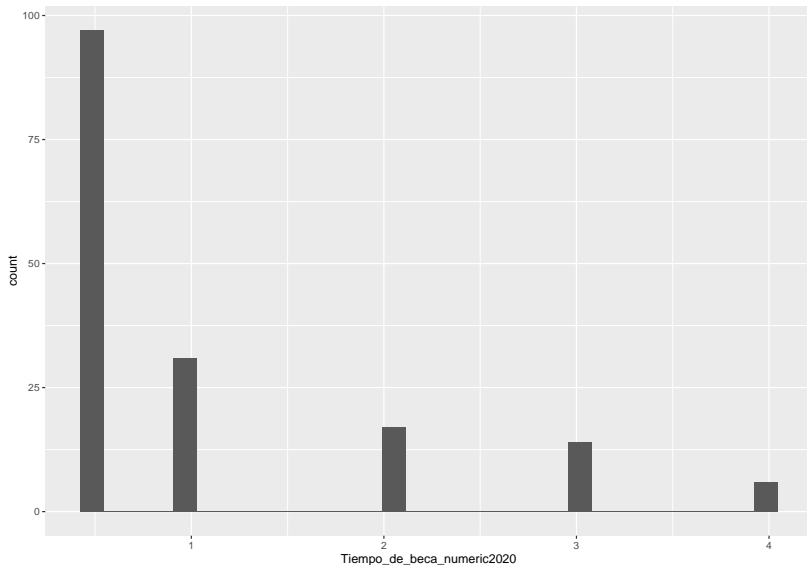
Análisis descriptivo: Edad



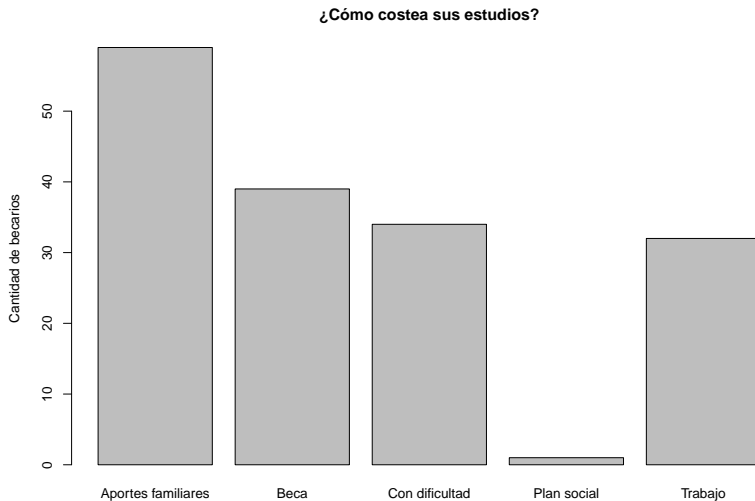
No tiene distribución normal. Tiene sesgo positivo, mayor cantidad de becarios entre los “17 y 28 años”.

Análisis descriptivo: Tiempo de beca

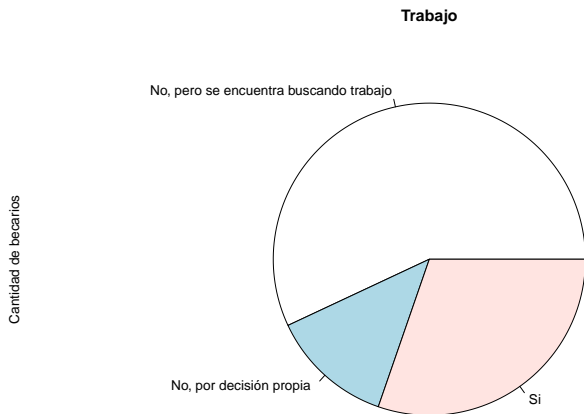
`stat_bin()` using `bins = 30`. Pick better value with `



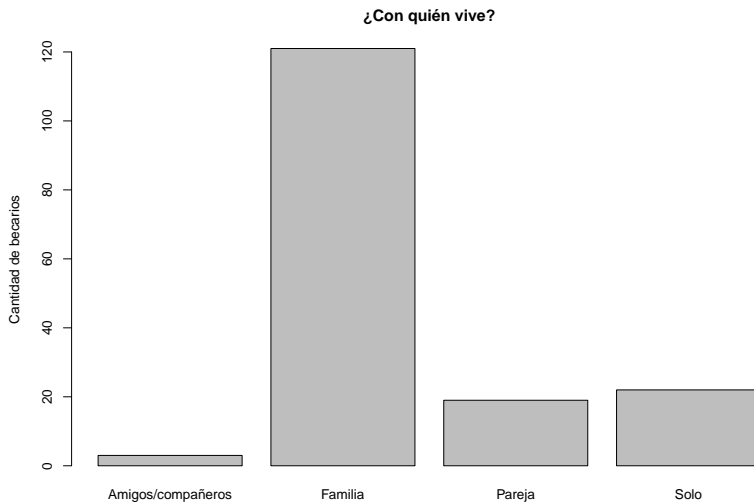
Análisis descriptivo: Costeo de estudios



Análisis descriptivo: Trabajo

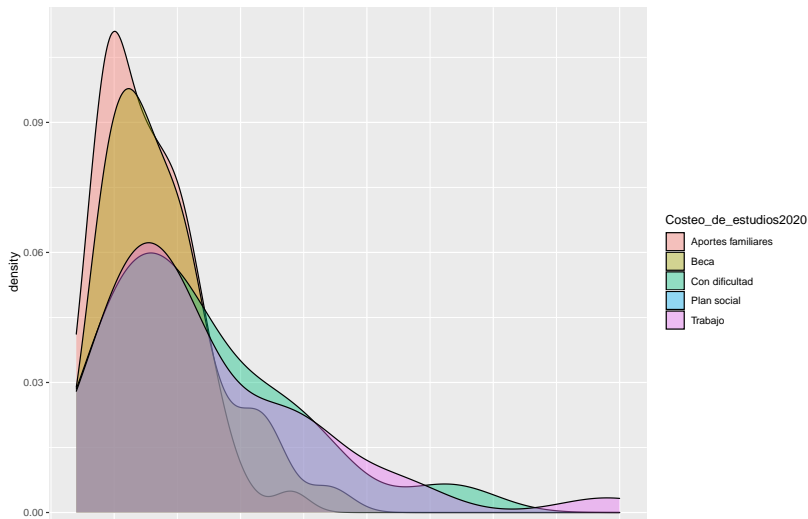


Análisis descriptivo: Con quien vive

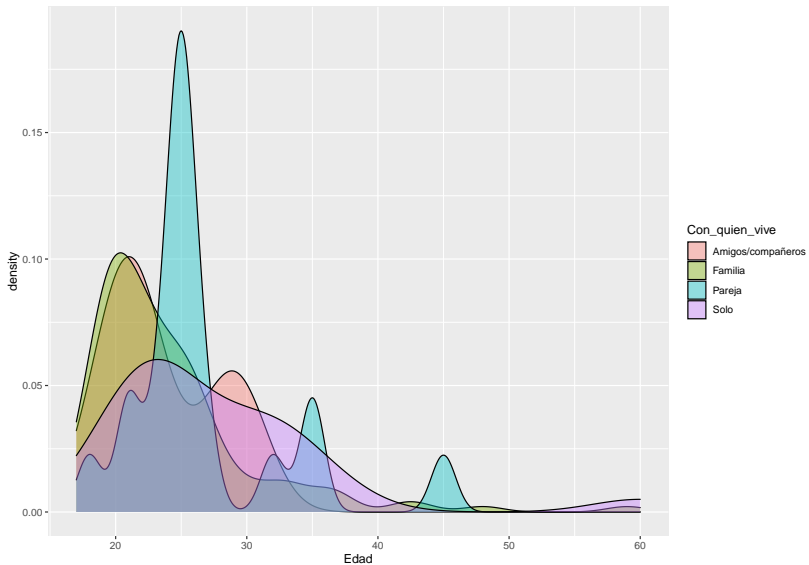


Edad x Costeo de Estudios

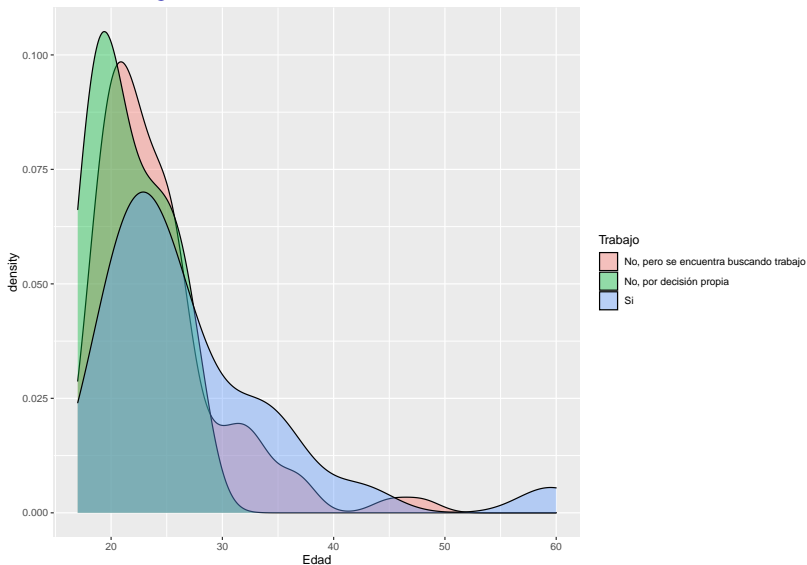
```
## Warning: Groups with fewer than two data points have been removed
## Warning in max(ids, na.rm = TRUE): ningun argumento finito
## -Inf
```



Edad x Con quien vive



Edad x Trabajo



A partir de los 28/30 años disminuyen los becarios que no están trabajando o que no están buscando trabajo.

Centralización y Dispersión

Media y mediana de Edad

```
## [1] 24.86061
```

```
## [1] 23
```

```
## [1] 17 60
```

```
## [1] 6.821681
```

##	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
##	17.00	20.00	23.00	24.86	26.00	60.00

Media y Desviación Estándar de Tiempo de beca

```
## [1] 1.087879
```

```
## [1] 0.9562878
```

Análisis: Hipótesis

- ▶ H_0 : La variable Edad tiene distribución normal
- ▶ H_1 : La variable Edad no tiene distribución normal

Nivel de significanza = 0.5

```
##
```

```
## Shapiro-Wilk normality test
```

```
##
```

```
## data: Becarios_2020$Edad
```

```
## W = 0.78464, p-value = 0.000000000000002625
```

El valor de p es inferior al nivel alfa, por ende, se rechaza la hipótesis nula. La variable Edad no tiene distribución normal.