

OLIMPIADA DE ESTADÍSTICA 2020

Encuesta de presupuestos
familiares 2018

Grupo: SALLIBAC2

► *Categoría A*

2º Bachillerato

Colegio Salliver, Málaga



Componentes:
Laura Segura Morales
María Peinado Toledo
Noemí Arias Campano

Tutor: Pedro Arturo Sánchez Meneses

IN
e

Instituto
Nacional de
Estadística

Actualmente se plantean grandes retos para los jóvenes, puesto que cada vez tienen más dificultades para emanciparse, que tienen que afrontar un futuro energético en el que los combustibles fósiles deben tener menos preponderancia y en el que las zona rurales tienden a despoblarse.

Somos tres jóvenes a punto de terminar nuestros estudios de Bachillerato y acceder a la universidad y como tales nos preocupa el futuro que nos espera.

Este año, con la **Encuesta de Presupuestos Familiares 2018**, hemos querido estudiar varios aspectos que están a la orden del día.

Nuestros objetivos principales con este proyecto son:

1.

Estudiar la dificultad de los hijos para **emanciparse**. Hemos analizado la relación existente entre el **número de personas en el hogar y aquellos el número de personas ocupadas**.

2.

El uso de **energías para el consumo en el hogar** cada vez son mayores. Dentro de las fuentes no renovables que nos muestra la recogida de datos, observar cuáles son las **más comunes** y cómo se relacionan entre ellas.

3.

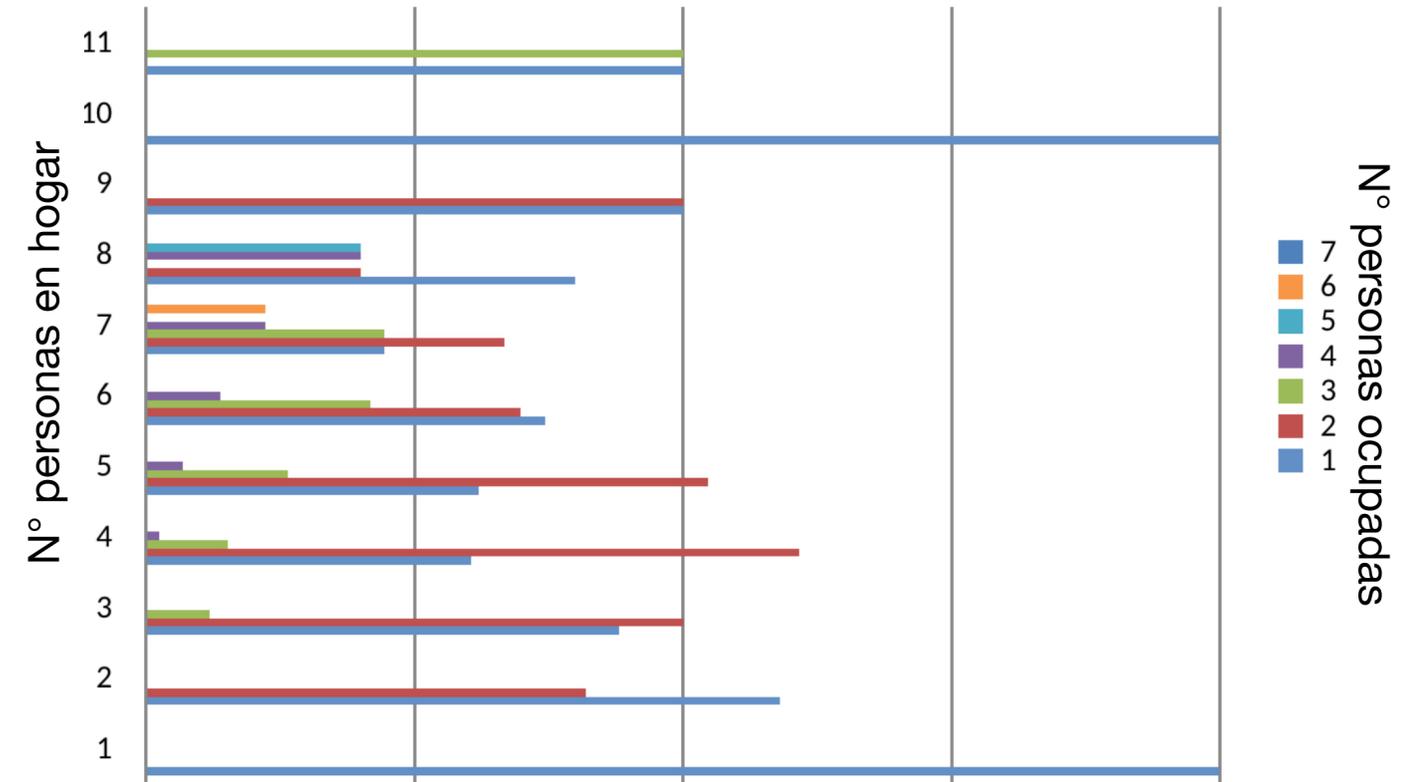
La **despoblación de los pueblos** cada vez es más usual. Queremos, analizar el **nivel de estudios** presente en cada uno de ellos y a su vez, el **número estudiantes** perteneciente a cada hogar ya que suponemos que este ambos van en detrimento.

Para ello hemos utilizado:

- Keynote para editar la presentación
- Fórmulas para facilitar y completar nuestros estudios
- Las aplicaciones de Numbers y Excel para realizar gráficas
- Funciones de Excel para filtrar datos como
 - “CONTAR.SI”
 - “CONTAR.SI.CONJUNTO”

Ocupados en el hogar

N° personas en hogar	N° personas ocupadas						
	1	2	3	4	5	6	7
1	376	0	0	0	0	0	0
2	484	336	0	0	0	0	0
3	443	503	60	0	0	0	0
4	270	542	68	11	0	0	0
5	54	91	23	6	0	0	0
6	16	15	9	3	0	0	0
7	2	3	2	1	0	1	0
8	2	1	0	1	1	0	0
9	1	1	0	0	0	0	0
10	1	0	0	0	0	0	0
11	1	0	1	0	0	0	0



Intervalo de confianza: Los siguientes cálculos realizados hemos utilizado los datos para hogares con 4 o más personas, para afirmar nuestra hipótesis nula que es que la edad de las personas ocupadas en el hogar es de dos o más personas.

$$z \approx \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = -8,39$$

- Media = 1,8287
- Varianza = 0,6857
- n = 1127
- 95%

Como la muestra es bastante grande supongo que la varianza de la población es la varianza de la muestra.

$$Z_{1-\alpha} = 2.33$$

$$Z \geq -Z_{1-\alpha} \quad H_0 \quad \mu \leq 2$$

$$-8,39 \not\geq -2.33 \quad H_1 \quad \mu > 2$$

No se acepta la hipótesis alternativa. Por lo tanto, y al 95% de confianza, podemos suponer que el número de personas ocupadas en el hogar es menor que dos

Estudio de la correlación:

$$\rho_{xy} = \frac{Cov_{xy}}{\sigma_x \sigma_y} = 0,39$$

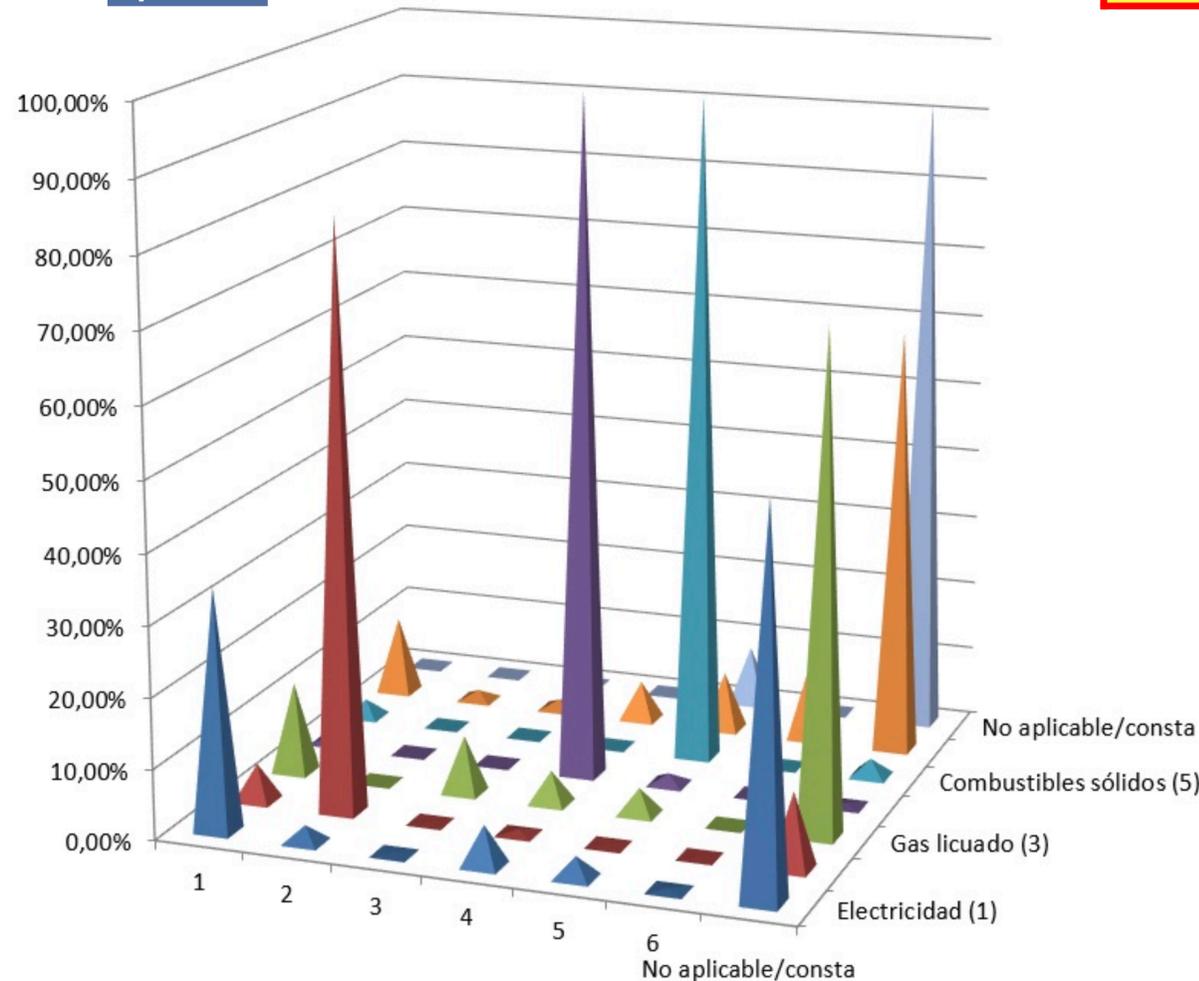
$$\left. \begin{aligned} Cov_{xy} &= 0,28 \\ \sigma_x &= 1,21 \\ \sigma_y &= 0,62 \end{aligned} \right\}$$

Por lo tanto, la relación entre las personas viviendo en el hogar y las personas ocupadas en el hogar es débil.

Relación entre fuentes de energía

Calefacción

	1	2	3	4	5	6	No aplicable/consta
1	34,34 %	2,38 %	0,16 %	5,72 %	3,02 %	0,32 %	54,05 %
2	5,26 %	83,09 %	0,05 %	0,69 %	0,10 %	0,05 %	10,77 %
3	12,87 %	0,02 %	8,14 %	4,64 %	3,69 %	0,00 %	70,48 %
4	1,14 %	0,00 %	0,00 %	96,96 %	1,52 %	0,00 %	0,38 %
5	2,50 %	0,00 %	0,00 %	0,00 %	95,00 %	0,00 %	2,50 %
6	11,11 %	1,39 %	1,39 %	5,56 %	8,33 %	11,11 %	61,11 %
No aplicable	0,00 %	0,00 %	0,00 %	0,00 %	8,33 %	0,00 %	91,67 %



Además vamos a realizar un segundo estudio centrándonos en la electricidad y el gas licuado como fuentes de energía para el agua caliente con el fin de comprobar si la siguiente hipótesis es correcta:

- Sabiendo que la electricidad y el gas licuado son las fuentes de energía más utilizadas para el agua caliente y donde la electricidad es la mayoritaria, ¿lo es también para la calefacción?

	1	3
1	432	136
2	30	2
3	2	86
4	72	49
5	38	39
6	4	0
No aplicable/consta	680	745
	n1 = 1258	n2 = 1057
	p1 = 0,3434	p2 = 0,1287

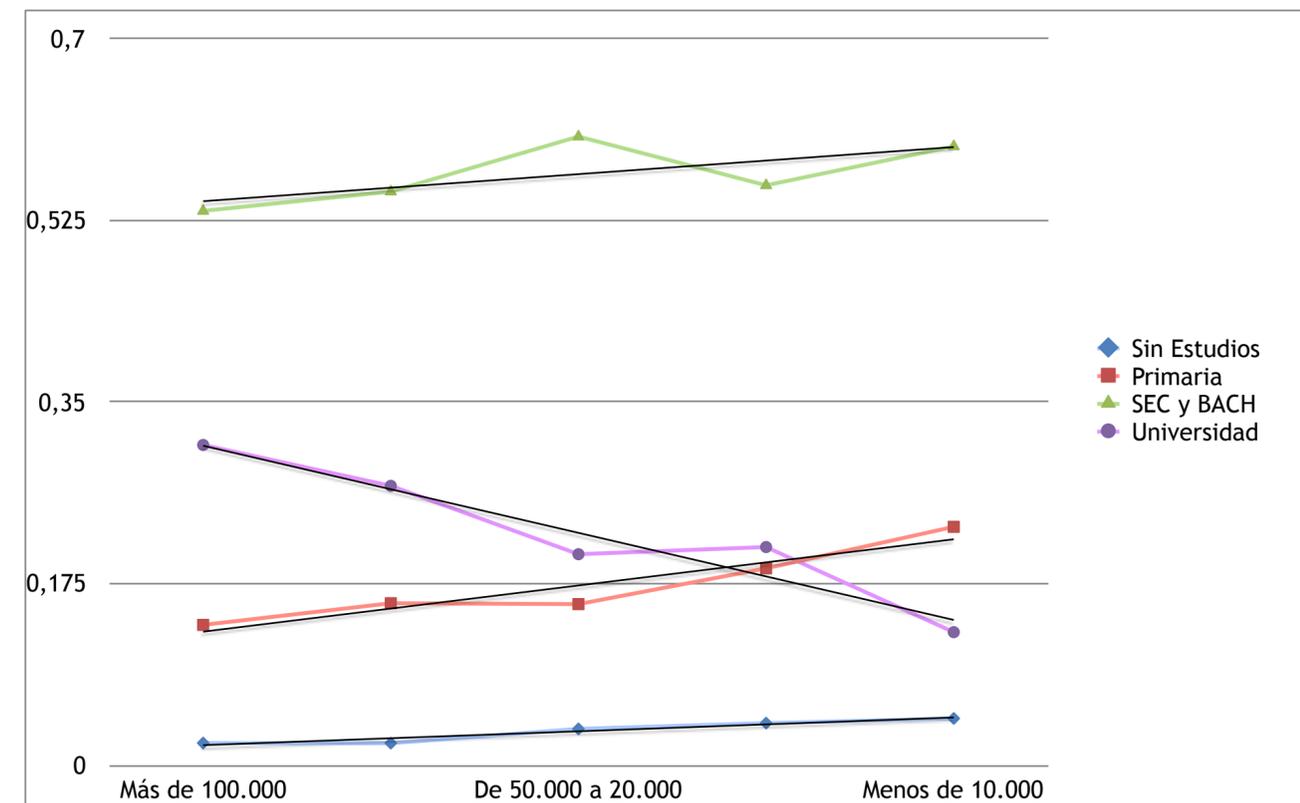
$$Z = \frac{p_1 - p_2}{\sqrt{\frac{n_1 p_1 + n_2 p_2}{n_1 \cdot n_2} \left(1 - \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}\right)}} = 11,9202$$

Nivel de confianza Z(1-alfa/2)	Región de aceptación
0.95	Cumple la hipótesis
	11,9202 > -1,65

Aprobamos la hipótesis puesto que cumple la región de aceptación.

Pueblos despoblados

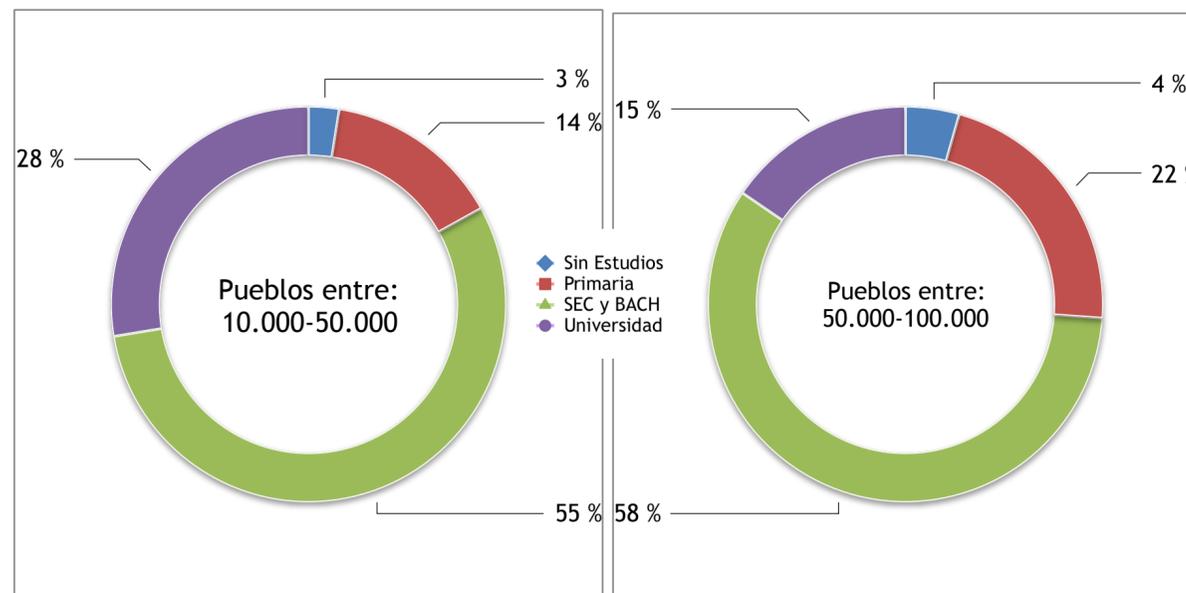
	Sin Estudios	Primaria	SEC y BACH	Universidad
Más de 100.000	0,0217	0,1354	0,5340	0,3087
De 50.000 a 100.000	0,0216	0,1563	0,5526	0,2693
De 50.000 a 20.000	0,0353	0,1554	0,6055	0,2035
De 20.000 a 10.000	0,0409	0,1899	0,5586	0,2104
Menos de 10.000	0,0454	0,2298	0,5961	0,1284



Estadísticas del contraste:

Dividimos la población entre menos de 50.000 personas y entre más de 50.000 para saber si la proporción de personas **sin estudios** en ambas poblaciones es parecido, porque la gráfica muestra un recorrido lineal. Para ello utilizamos una diferencia de proporciones y suponemos una hipótesis nula de que ambas proporciones son iguales.

$$Z = \frac{p_1 - p_2}{\sqrt{\frac{n_1 p_1 + n_2 p_2}{n_1 \cdot n_2} \left(1 - \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}\right)}} = -3,6419$$



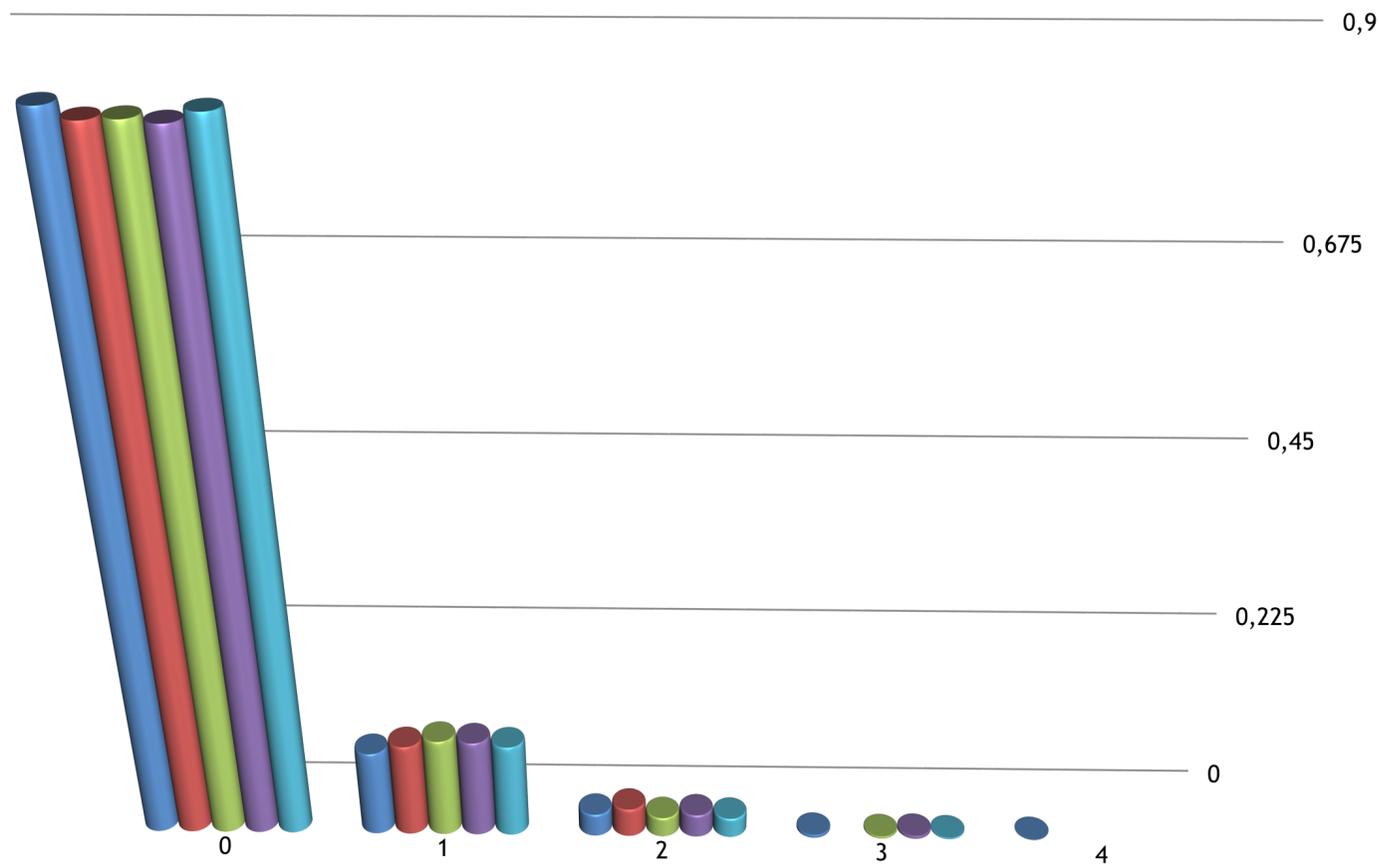
Nivel de confianza $Z(1-\alpha/2)$

0.95 Rechazamos porque no cumple $-1.96 < -3.6419 < 1.96$

Al **rechazarla** suponemos que la población de más de 50.000 es mayor que la otra y *volvemos a emplear la fórmula.*

Nivel de confianza $Z(1-\alpha)$

0.95 Aceptamos porque cumple $-3.6419 < 1.645$



■ Más de 100.000
■ De 50.000 a 100.000
■ De 20.000 a 50.000
■ De 10.000 a 20.000
■ Menos de 10.000

Número de estudiantes por hogar

	0	1	2	3	4
Más de 100.000	0,8544	0,1121	0,0292	0,0037	0,0005
De 50.000 a 100.000	0,8405	0,1222	0,0371	0	0
De 20.000 a 50.000	0,8419	0,1302	0,0252	0,0025	0
De 10.000 a 20.000	0,8379	0,1284	0,0297	0,0037	0
Menos de 10.000	0,8496	0,1232	0,0253	0,0017	0

En esta gráfica podemos apreciar que el número de estudiantes que hay en cada hogar tampoco influye con respecto al tamaño del municipio. Por lo que las personas que viven en los pueblos pequeños no están condicionadas a no estudiar.

- **Covarianza** = -0,5483
- **D.T(x)** = 1,5990
- **D.T(y)** = 1,7497

$$\rho_{xy} = \frac{Cov_{xy}}{\sigma_x \sigma_y} = -0.1959$$

Con la correlación queríamos analizar si existe considerable relación entre ambas variables ya que la gráfica mostraba lo contrario a lo que habíamos supuesto desde un principio. El resultado nos ha resultado muy bajo, por lo que no hay relación entre el número de estudiantes por hogar y los pueblos donde habitan.

Conclusión Objetivo 1:

Tuvimos que *rechazar la hipótesis* de que existían más de 2 personas ocupadas en el hogar para demostrar que había muchas dificultades de emancipación.

Entonces entendemos que la dificultad para emanciparse puede venir dada por empleos con remuneración baja o dificultades para encontrar trabajo; el que la media sea menor que dos puede interpretarse también como dificultades para encontrar trabajo.

Además, la relación entre la cantidad de personas en el hogar y la cantidad de personas ocupadas en el hogar resultó ser débil, lo que nos indica que no tiene porque haber más personas trabajando en el hogar a pesar de haber más personas en el mismo.

Conclusión 1 Objetivo 2: Observamos cómo en los hogares se utilizan mayoritariamente la misma fuente de energía tanto para la calefacción como para el agua caliente.

Conclusión 2 Objetivo 2: Podemos afirmar que la electricidad es la fuente de energía que destaca sobre las demás utilizadas para la calefacción.

Conclusión 2 Objetivo 3:

No influye el tamaño de municipio con el número de estudiantes por hogar, por lo tanto la hipótesis que teníamos al principio del trabajo era errónea.

Conclusión 1 Objetivo 3:

Los pueblos de más de 50.000 habitantes tiene una mayor proporción de personas sin estudios. Por lo que a pesar de que hoy en día sea noticia que la gente abandona los pueblos por falta de trabajos y estudios, se demuestra que en las ciudades hay una mayor cantidad de personas sin estudios.