

ENCUESTA CUATRIENAL DE ESTRUCTURA SALARIAL 2014

Categoría: Bachillerato

Grupo: PEPE

Alumno: José M. García Arcas

Tutora: M. José Calvo Leva

IES Rey Carlos III



OBJETIVOS. TÉCNICAS USADAS. PROGRAMAS INFORMÁTICOS

OBJETIVOS	TÉCNICAS Y METODOLOGÍA. Software
<p>1. Comprobar si existen diferencias entre las medias del salario bruto anual para hombres y mujeres con el fin de ver si el sexo influye en el salario bruto anual.</p>	<p>-Haremos un contraste de medias con un test de hipótesis. Ayudándonos del programa Microsoft Excel para calcular los parámetros.</p> <p>-Previamente haremos un test de normalidad de las variables salario bruto de hombres y salario bruto de mujeres mediante el programa Statgraphics.</p> <p>- Hemos usado la función “var” de la hoja EXCEL y funciones como comparar dos muestras simples independientes; análisis de una variable; o ajuste a una distribución del programa Statgraphics.</p>
<p>2. Comprobar si la antigüedad influye en el salario bruto anual. Es decir, si a más antigüedad en el puesto de trabajo se cobra más.</p> <p>3. Estudiar los histogramas de las variables cuantitativas continuas “antigüedad” y “salario bruto anual”; así como los parámetros estadísticos.</p>	<p>-Haremos un ajuste lineal entre la antigüedad y el salario bruto anual; hemos considerado oportuno agrupar la variable meses de antigüedad y la variable años de antigüedad. En este caso la variable antigüedad será la independiente. Usaremos el icono de regresión simple del programa Statgraphics.</p> <p>- Realizaremos un análisis de los parámetros estadísticos más usuales como son la media, desviación típica, coeficiente de variación, máximo, mínimo y los cuartiles.</p> <p>- También realizaremos sendos intervalos de confianza para las medias de las dos variables anteriores.</p>
<p>4. Comprobar si las gratificaciones extraordinarias están relacionadas de forma lineal con el salario bruto anual.</p> <p>5. Estudiar si los datos referentes a gratificaciones extraordinarias están muy dispersos en la muestra de los 200 encuestados.</p> <p>6. Observar si existe brecha salarial en las diferentes regiones.</p> <p>7. Estudiar el tipo de mercado más habitual.</p>	<p>- Haremos una recta de regresión considerando el salario bruto anual la variable independiente. Calcularemos el coeficiente de regresión entre las dos variables. Usaremos el programa Statgraphics.</p> <p>-Realizaremos un gráfico de líneas para la variable gratificaciones extraordinarias. Usaremos la hoja Excel.</p> <p>-Dividiremos los datos por regiones y sexo y haremos un gráfico de área para ver si existe brecha salarial entre los salarios medios de cada región. Lo realizaremos con la hoja Excel.</p> <p>-Realizamos un diagrama de sectores para ver los porcentajes de encuestados que trabajan en los diversos tipo de mercado.</p>
<p>8. Analizar diversas variables estratificando por sexo: tipo de jornada, duración de contrato, nivel de estudios y salario medio según nivel de estudios.</p>	<p>- Hemos realizado diagramas de barras y de columnas agrupadas para ver si existen diferencias significativas entre sexos y tipo de jornada y entre tipo de jornada y nivel de estudios. Usaremos gráficos de la Hoja Excel.</p>

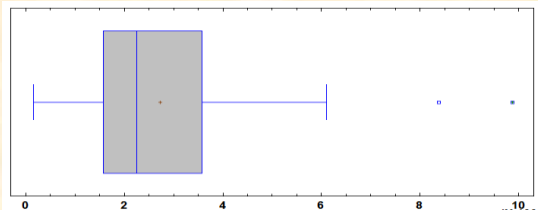
Observación: Hemos considerado oportuno hacer uso del programa Statgraphics por la rapidez en cálculos y gráficos que ofrece, así como la posibilidad de comprar rápidamente muestras independientes o realizar histogramas o ajustar rectas de regresión.

Objetivo 1. Contraste de medias. Paso 1: Test de normalidad

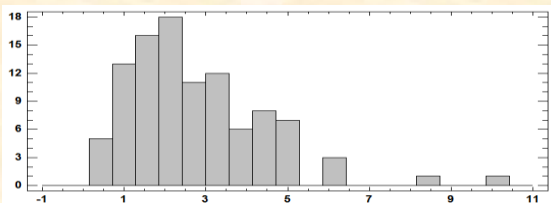
Haremos un contraste de medias entre dos muestras, para ello queremos comprobar si la variable salario anual bruto es normal; lo haremos por sexo. Por tanto, veamos el paso 1: "Comprobar la normalidad de los datos".

Realizamos el histograma, diagrama de caja y bigotes y prueba de normalidad con el programa Statgraphics; con ello intentaremos ver si los datos se pueden ajustar de forma gráfica a una distribución normal:

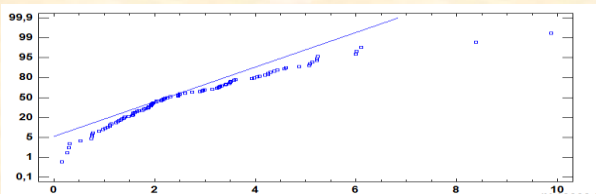
SEXO: HOMBRE



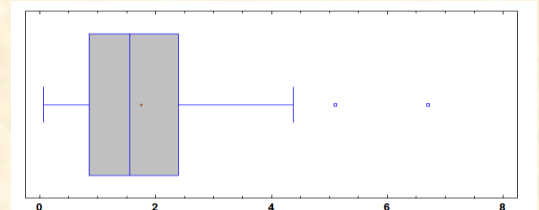
- Se observan dos valores atípicos tanto en el Histograma como en el gráfico de caja y bigotes.



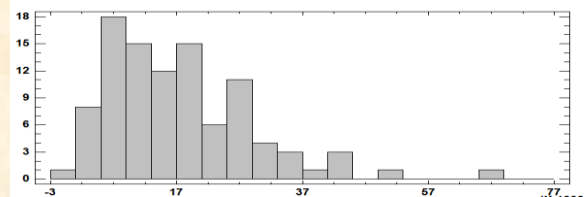
La siguiente gráfica nos indica que cuanto más se acercan los puntos a la línea más cercana es nuestra distribución a una Normal



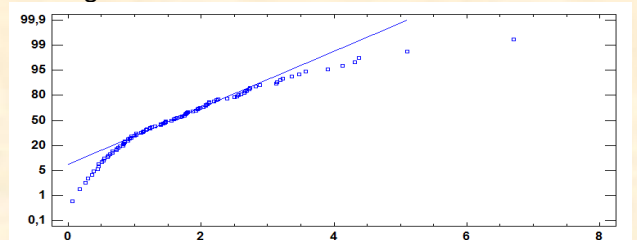
SEXO: MUJER



- También se observan dos valores atípicos y la mediana está más centrada en la caja.

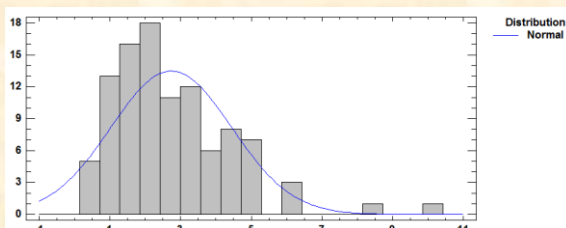


Se ven claramente los datos atípicos en el histograma a la derecha del mismo.



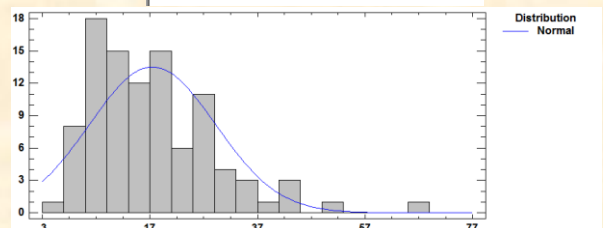
Por último hacemos un ajuste teórico para comprobar que efectivamente podemos considerar con un alto porcentaje de acierto que los datos se distribuyen de forma normal. Hacemos un test de normalidad que nos afirma que podemos afirmar que los datos se ajustan con un nivel de confianza del 95% a una variable normal cuyos parámetros son:

Fitted Distributions	
Normal	
mean =	27364,9
standard deviation =	17073,9



Gráficamente se observa que los datos referidos a mujeres parece que ajustan incluso mejor que los datos de los hombres a una normal, pero vamos a realizar un test de normalidad al igual que antes. El resultado es que también podemos afirmar con un 5% de error que los datos se distribuyen de e forma normal.

Fitted Distributions	
Normal	
mean =	17471,9
standard deviation =	11688,2



Objetivo 1. Contraste de medias. Paso 2

Bajo el supuesto que los salarios brutos anuales se distribuyen según distribuciones normales, tenemos:

$$\bar{x}_b = 27364\text{€} \quad \sigma_b = 17073 \quad n_b = 101$$

$$\bar{x}_m = 17471\text{€} \quad \sigma_m = 11688 \quad n_m = 99$$

Bajo el supuesto que los salarios brutos anuales se distribuyen según distribuciones normales, tenemos:

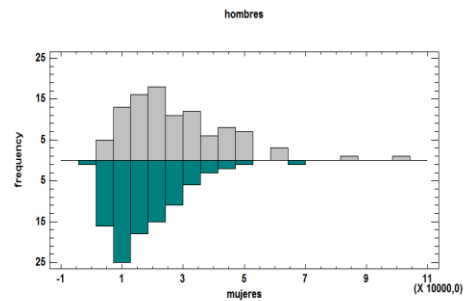
$$\bar{X}_M - \bar{X}_B \approx N\left(\mu_M - \mu_B, \sqrt{\frac{\sigma_M^2}{n_M} + \frac{\sigma_B^2}{n_B}}\right)$$

El intervalo de confianza a un nivel de confianza de $1-\alpha$ para la diferencia de medias viene dado por la expresión:

$$(\bar{X}_M - \bar{X}_B) \pm t(\min\{n_M - 1, n_B - 1\}, \alpha/2) \sqrt{\frac{S_M^2}{n_M} + \frac{S_B^2}{n_B}}$$

Calculamos el estadístico de contraste:

$$t^* = \frac{(\bar{X}_M - \bar{X}_B) - (\mu_M - \mu_B)}{\sqrt{\frac{S_M^2}{n_M} + \frac{S_B^2}{n_B}}} = -4,79$$



La t-Student que necesitamos será mínimo (98,100); es decir de 98 grados de libertad y a un nivel de confianza del 95% --> El valor será menor que 0,005 que es el área por debajo del valor "t" y por tanto el p-valor si el contraste es unilateral será menor que 0,01. Por tanto rechazamos la hipótesis nula y por tanto existen diferencias significativas en las medias de los salarios brutos de hombres y mujeres.

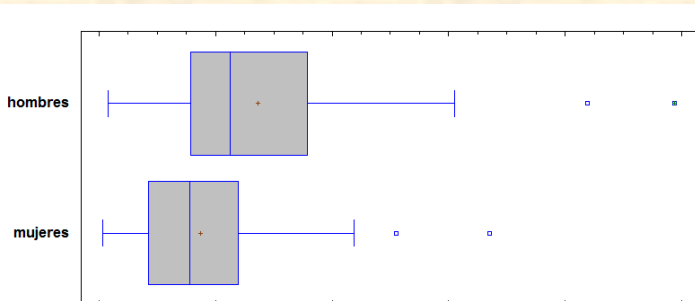
Hemos usado el programa Statgraphics para conseguir intervalos de confianza:

Al 95,0% el intervalo de confianza para la media de los hombres es: [23994,3; 30735,5]

Al 95,0% e intervalo de confianza para la media de las mujeres es [15140,7; 19803,0]

Al 95,0% el intervalo de confianza para la diferencia de medias es [5804,96; 13981,0]

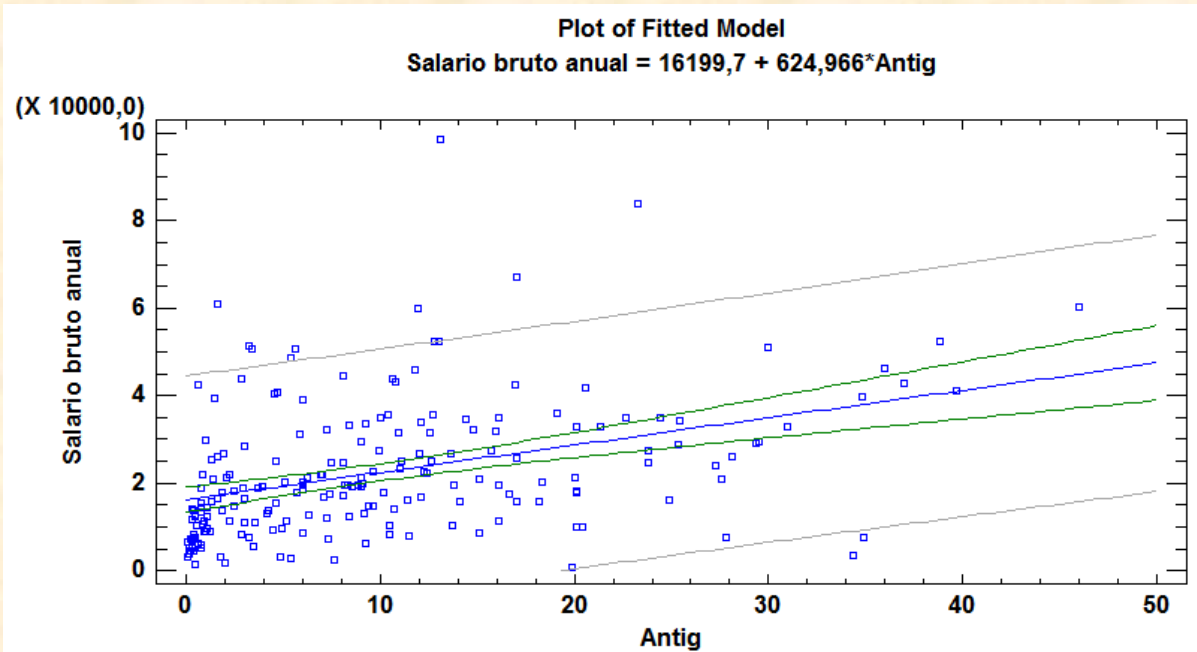
Se puede observar que el intervalo de confianza para la media de los hombres tiene los valores extremos muy por encima del intervalo de confianza para la media de las mujeres.



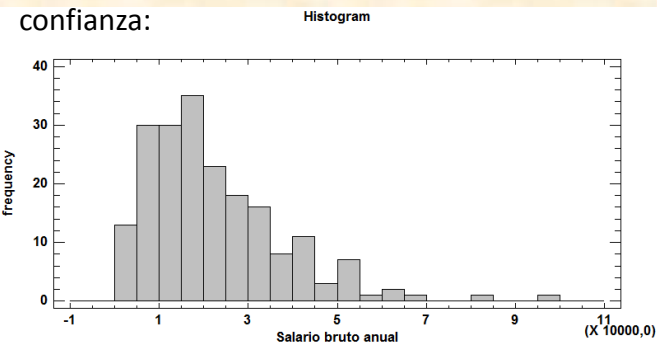
En el gráfico de la izquierda podemos ver los dos diagramas de cajas y bigotes para ambas distribuciones, y se puede ver cómo quedan los datos de mujeres por debajo de los de hombres.

Objetivos 2 y 3. Salario bruto anual y antigüedad

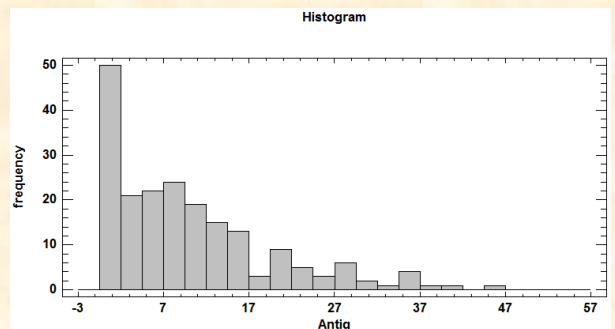
❖ Nos planteamos una recta de regresión para estudiar la nube de puntos o diagrama de dispersión. De todas las posibles rectas que se podrían ajustar a los datos el programa Statgraphics considera que la mejor es $\text{Salario bruto anual} = 161997 + 624,966 \cdot \text{Antig}$. El coeficiente de correlación calculado es de 0.38, lo que indica que existe una relación lineal positiva pero no es muy fuerte porque queda lejos del 1. Este modelo ajustado explicaría tan solo el 15% de los datos de salario bruto anual.



❖ Hemos realizado un histograma para las dos variables que hemos relacionado en la regresión lineal de arriba, puesto que ambas son variables cuantitativas continuas; a la izquierda se observa el histograma de salario bruto anual, y debajo se observan los principales parámetros estadísticos de dicha variable junto con un intervalo de confianza para la media; a la derecha se observa el histograma de la variable antigüedad y debajo los principales parámetros estadísticos junto con un intervalo de confianza:



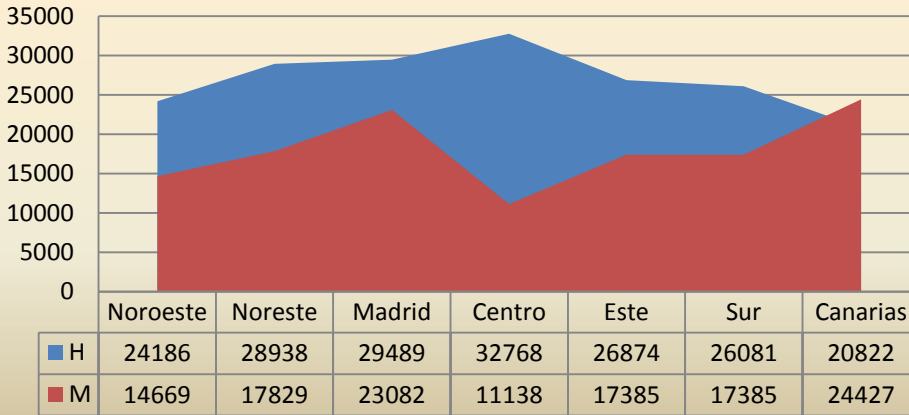
Media: 22467€
 Desviación típica: 15438€
 CV: 68%
 Mínimo: 683€ ; Máximo: 98661€
 Q1: 11226€ ; Q2: 19254€ ; Q3: 31343€
 Intervalo de confianza para la media al 95% (20315;24620)



Media: 10 años
 Desviación típica: 9 años
 CV: 95%
 Mínimo: 0,08 años ; Máximo : 46 años
 Q1: 2 años ; Q2: 7,8 años ; Q3: 14 años
 Intervalo de confianza para la media al 95% (8,6 ; 11,3)

Objetivos 4, 5 y 6. Salario por regiones y sexo. Pagas extras respecto a salario. Tipo de mercado.

Brecha salarial por regiones



• Hemos dividido los datos por regiones y por sexo y hemos realizado un gráfico donde se reflejan las medias aritméticas. Se observa que la mayor brecha salarial ocurre en la región centro siendo casi de 20000€ dicha brecha.

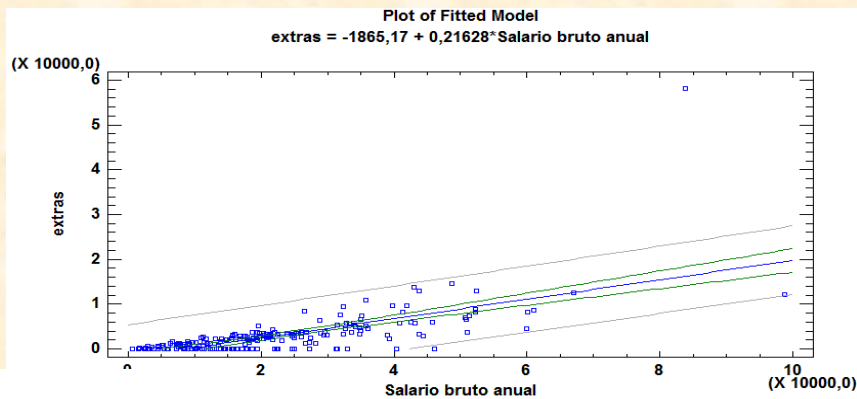
• También se observa que en la única región donde las mujeres cobran más es Canarias.

• En todas las regiones se observa la conocida brecha salarial (diferencias entre los salarios medios de hombres y mujeres).

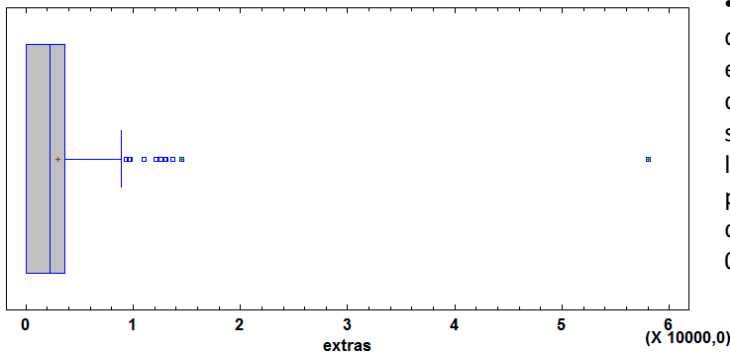
• Ahora nos vamos a plantear si existe relación lineal positiva entre el salario bruto anual y las gratificaciones extraordinarias. El programa Statgraphics nos ofrece un coeficiente de correlación bastante elevado (0,67) y nos ofrece la mejor recta que ajusta a los datos:

$$\text{Gratificaciones extraordinarias} = -1865,17 + 0,21 \cdot \text{salario bruto anual}$$

Box-and-Whisker Plot



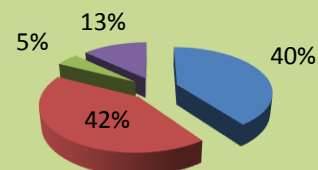
• Hemos realizado un diagrama de caja y bigotes para ver cómo quedarían los datos de “gratificaciones extraordinarias” distribuidas. Se observa cómo los valores quedan agrupados en torno al cero y los valores atípicos son los que sobrepasan los 1000€. Hay un valor atípico que llega casi a los 6000€. La media que se obtiene es de 2994€ pero lógicamente el CV es del 165% puesto que hay mucha dispersión en torno a la media ya que muchos valores son 0€.



• Observamos en el gráfico de la derecha que el 42% de los individuos trabajan a nivel local o regional; y un 40% a nivel nacional. Lo que significa que un 82% trabaja a nivel local, regional o nacional y tan sólo un 18% lo hace a nivel europeo o mundial.

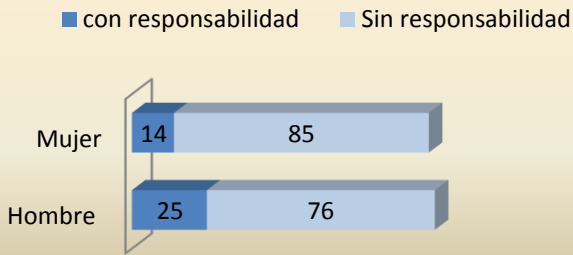
Tipo de mercado

■ Nacional ■ Local/regional ■ Unión Europea ■ Mundial

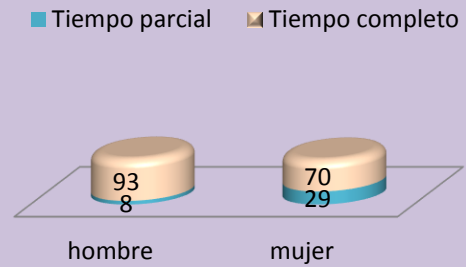


Objetivo 8. Estratificar por la variable "Sexo"

Responsabilidad por sexos

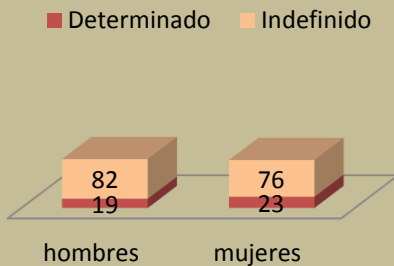


Tipo de jornada por sexo

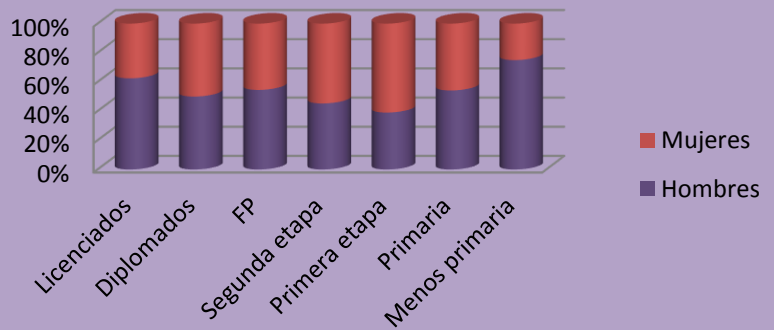


- ❖ Hemos estratificado la variable responsabilidad por sexo, así como la variable tipo de jornada y podemos observar que el nº de mujeres con puesto de responsabilidad en el trabajo es casi la mitad que de hombres.
- ❖ Hemos estudiado también el tipo de jornada según el sexo, observando que hay casi 4 veces más mujeres que hombres a tiempo parcial, pero la mayoría de hombres y mujeres están a tiempo completo, sólo 18% de encuestados está a tiempo parcial.

Tipo de contrato por sexos



Nivel de estudios por sexo

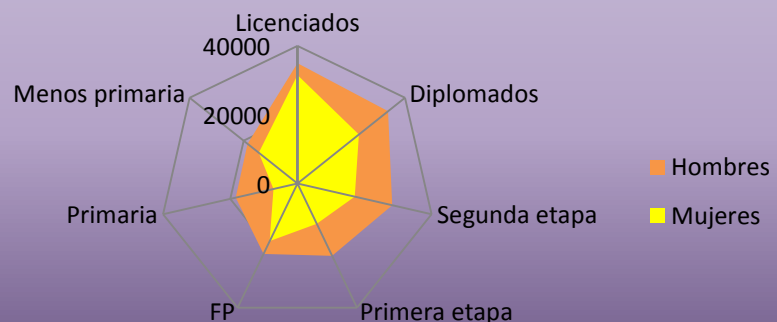


- ❖ Hemos estratificado la variable tipo de contrato por sexo, observando que el número de contratos de duración determinada es un poco superior en mujeres que en hombres.
- ❖ En la gráfica de arriba a la derecha se observa que el número de hombres es superior al de mujeres en las categorías: "licenciados y similares"; "menos que primaria".
- ❖ En la gráfica de "sueldo medio por sexos y nivel de estudios se observa que en todas las categorías de nivel de estudios el sueldo medio de los hombres es superior al de las mujeres.

Tabla de sueldos medios:

Estudios	Hombre	Mujer
Licenciados	34980	31573
Diplomados	33725	22897
Segunda etapa	28256	17116
Primera etapa	23276	12971
FP	22648	18589
Primaria	18383	7110
Menos primaria	18516	14730

Sueldo medio por sexos y nivel de estudios



CONCLUSIONES

- ❖ El salario medio bruto anual para hombres es significativamente distinto que el de mujeres, es decir, existe la conocida "brecha salarial" en los datos estudiados.
- ❖ El 95% de los datos de salario bruto anual para hombres están en el intervalo [23994,30735]; y el intervalo para las mujeres es [15140,19803]; es decir, el límite inferior es 8000€ más para hombres que para mujeres y de 11000 euros para el límite superior. También se observa dicha diferencia en las medias de salarios brutos anuales de hombres y mujeres: la media en los hombres es de 27364€ frente a 17471€ que es la media de mujeres; lo que suponen unos 10000€ de diferencia.
- ❖ La antigüedad y el salario bruto anual tienen una relación positiva pero no es significativa.
- ❖ La media de salario bruto anual tanto en hombres como en mujeres es de 22467€ y hay una variabilidad del 68% de los datos con respecto a esta media.
- ❖ La media de antigüedad en los datos es de 10 años; estando la mayoría de los datos por debajo de 17 años de antigüedad. De hecho, el 75% de los datos tienen menos de 14 años de antigüedad.
- ❖ La brecha salarial por sexo se acentúa más en la región centro donde alcanza unos 20000€ de diferencia entre los salarios medios de hombres y mujeres. Y la única región donde el salario medio de mujeres supera al de hombres es en Canarias.
- ❖ Existe una relación positiva fuerte entre el salario bruto y las gratificaciones extraordinarias. Estas últimas están en torno al cero en muchos casos y se consideran valores atípicos los que superan los 1000€. Hay mucha variabilidad entre los datos.
- ❖ El 82% de los encuestados trabajan a nivel local/regional o nacional frente a tan sólo un 18% que lo hace a nivel europeo o mundial.
- ❖ Existen diferencias significativas en torno a los puestos de responsabilidad que ocupan hombres y mujeres, siendo el porcentaje de hombres un 10% superior al de mujeres.
- ❖ El número de mujeres que trabaja a tiempo parcial cuadruplica al número de hombres a tiempo parcial; sin embargo la diferencia es poco significativa cuando hablamos del tipo de contrato.
- ❖ Si estudiamos el nivel de estudios entre hombres y mujeres, los datos están bastante igualados excepto en las dos etapas de secundaria: hay más mujeres cuyos estudios son la primera etapa y segunda etapa de secundaria. El dato de menos de primaria es poco significativo porque sólo hay 4 personas.
- ❖ Como venimos diciendo en conclusiones anteriores, se observan también diferencias de salario por sexos y nivel de estudios. La diferencia menor se encuentra en la categoría de licenciados y de menos de primaria; en el resto la diferencia está alrededor de los 10000€.