

V OLIMPIADA ESTADÍSTICA

CATEGORÍA: BACHILLERATO Y CICLOS FORMATIVOS

Participantes: Sofía Quintanilla y Andrés Timón

Tutor: Leonardo Martín Herrero

Grupo: Lemon

Centro educativo: British Council School

Provincia: Madrid

METODOLOGÍA

1. HERRAMIENTAS:

- Microsoft Excel para realizar gráficos.
- Power Point para representar los datos y gráficos.
- Encuesta Cuatrienal de estructura salarial 2014.

3. PLANIFICAR:

- Realizar una hipótesis con sentido para cada uno de los objetivos.

5. RESULTADOS:

- A partir de los resultados obtenidos creamos conclusiones para cada objetivo, donde analizamos si coinciden con nuestras hipótesis.

2. EXPLORAR Y DECIDIR:

- A partir de los datos proporcionados, tuvimos que pensar nuestros 5 objetivos.
- Estos objetivos debían ser coherentes a la par que interesantes.
- Además, tuvimos que decidir los métodos matemáticos a seguir, todos ellos los hemos estudiado en estadística durante este curso.

4. DESARROLLO DEL TRABAJO:

- Analizamos los datos mediante la realización de tablas estadísticas y gráficos con el programa "Excel".
- Calculamos medias, coeficientes de correlación y probabilidades.
- Decidimos además desarrollar cada objetivo con un método estadístico diferente:
 - Objetivo 1: mediante la representación gráfica de los datos.
 - Objetivo 2: utilizamos un estudio bidimensional para encontrar correlaciones.
 - Objetivo 3 y 4: seguimos un estudio unidimensional.
 - Objetivo 5: nos basamos en un estudio binomial de los datos.

6. PRESENTACIÓN:

- Por último creamos una presentación con el programa informático "Power Point" con toda la información.
- Luego exportamos la presentación en formato .pdf.

OBJETIVOS

1. Desglosar la información obtenida en la encuesta sobre los tipos de actividades económicas, centrándonos en las actividades administrativas, de industria manufacturera y las actividades sanitarias y servicios sociales. Los parámetros que hemos tenido en cuenta son género, regiones y tramo de edades.

2. Analizar posibles correlaciones entre salario y otros factores como tramo de edad, gratificaciones extraordinarias, jornada anual pactada, estudios y diferencias entre género. Nuestro objetivo era encontrar relaciones directas que pudieran ayudar a predecir datos.

3. Encontrar la media del salario bruto anual en cada mercado para poder comparar y analizar qué mercado proporciona un sueldo más elevado. Para ello hemos estudiados los datos en función de la varianza, la desviación típica y el coeficiente de variación; así podemos comprobar si las medias obtenidas son fiables según la dispersión de los datos.

4. Hallar cuál es el tipo de persona predominante en esta encuesta para poder ver qué valores son los más comunes y poder así realizar una conclusión sobre la situación laboral y demográfica de España en 2014. Para llevar esto a cabo hemos encontrado la moda y la media de cada aspecto de la encuesta para más tarde proceder a analizarlos.

5. Comparar el promedio español de contratos indefinidos y determinados en las diferentes regiones del país. Para realizar esta comparación hemos realizado un estudio binomial: a partir del promedio nacional calculamos el porcentaje de personas con cada tipo de contrato por región y comparamos esa previsión con la realidad de los encuestados.

HIPÓTESIS

Creemos que atendiendo a las labores que se llevan a cabo, aquellos que trabajan en administración serán los más jóvenes, seguidos de los de manufacturación y de los de servicios sociales. En cuestión de regiones, consideramos que la manufacturación será más numerosa en el norte, la administrativa predominante en la Comunidad de Madrid y los servicios sociales en el Este y Sur. Finalmente, pensamos que habrá mayoría de hombres en la industria manufacturera, y mayoría de mujeres en servicios sociales y administración.

Sabemos que el salario posee una importancia crítica en la actividades económicas, por lo que imaginamos que tendrá una correlación fuerte con los estudios ya que determina las posibilidades de trabajo. A la vez, queremos demostrar la diferencia de salarios entre géneros con la misma formación. Así mismo, intentamos encontrar la edad crítica en la que los sueldos son mayores. Finalmente, creemos que tanto las jornadas anuales como las gratificaciones pueden potencialmente afectar al salario y por ende, tener una correlación con él.

Pensamos que el mercado que pagará los salarios más bajos será el regional o local ya que los mercados locales poseen una demanda inferior. Además, un cargo local tendrá menos responsabilidades, y por lo tanto esto se verá reflejado en sus sueldos. De manera contraria, suponemos que la media de los sueldos más altos se encontrará en el mercado mundial, porque su influencia económica será mayor al tener mayor participación de agentes económicos. El mercado nacional tendrá la segunda media más baja seguida por la media de la Unión Europea.

Asumimos, partiendo del hecho de que en España existe un predominio del sector terciario, que la actividad económica más común será una relacionada con los servicios. En 2014, España se encontraba todavía en una fase de recuperación de la recesión económica, consecuentemente estimamos que habrá un mayor número de contratos de duración determinada a tiempo parcial entre los encuestados. Por último, sabemos que la población de España está envejeciendo y por lo tanto predecimos que el tramo de edad modal se situará entre los 30 y 50 años de edad.

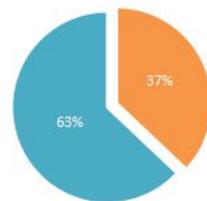
Estimamos que el promedio español será parecido a aquel de la comunidad de Madrid ya que al ser la capital del país presenta, junto a Barcelona la mayor parte de la actividad económica. Por otra parte, es cierto que la zona centro y norte de la península son las que más aportan al PIB de España, y por lo tanto, tienen una economía de grandes dimensiones. Esto hace que tengan una influencia importante sobre el promedio nacional, y por lo tanto, estas regiones seguirán el promedio. Por otro lado, creemos que la región sur no seguirá la media nacional y será inferior.

Tipos de actividades económicas

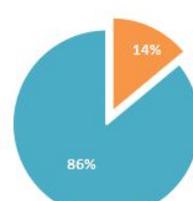


OBJETIVO 1

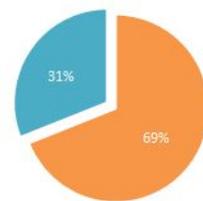
Actividades administrativas por género



Actividades sanitarias y servicios sociales por género



Industria manufacturera por género



Códigos de actividades económicas:

- B0** : Industrias extractivas
- C1 - C8**: Industrias manufactureras
- D0**: Suministro de energías
- E0**: Suministro de agua
- F0**: Construcción
- G1 - G2**: Comercio
- H1 - H2**: Transporte y almacenamiento
- I0**: Hostelería
- J0**: Información y comunicaciones
- K0**: Actividades financieras
- L0**: Actividades inmobiliarias
- M0**: Actividades profesionales
- N0**: Actividades administrativas
- O0**: Administración pública y defensa
- P0**: Educación
- Q0**: Actividades sanitarias y servicios sociales
- R0**: Actividades artísticas
- S0**: Otros servicios

Actividades administrativas por edades



Actividades sanitarias y de servicios sociales por edades



Industria manufacturera por edades



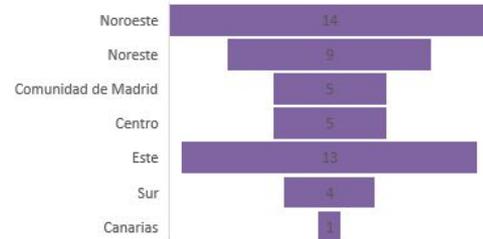
Actividades administrativas por regiones



Actividades sanitarias y de servicios sociales por regiones



Industria manufacturera por regiones

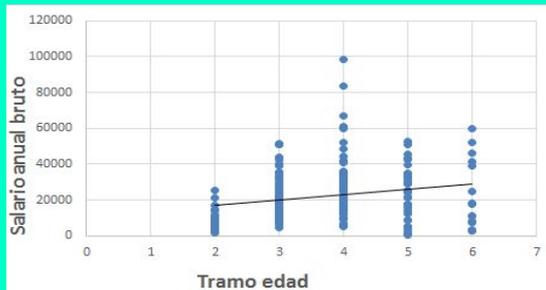


RELACIÓN ENTRE TRAMO DE EDAD Y SALARIO BRUTO ANUAL

LEYENDA:

- 1 - Menos de 19
- 2 - 20 a 29
- 3 - 30 a 39
- 4 - 40 a 49
- 5 - 50 a 59
- 6 - Más de 59

MEDIA X	3,760
MEDIA Y	22467,83
VARIANZA X	1,21
VARIANZA Y	237093077
D.TÍPICA X	1,10
D.TÍPICA Y	15397,83
COVARIANZA	3672,4
PEARSON	0,21681926



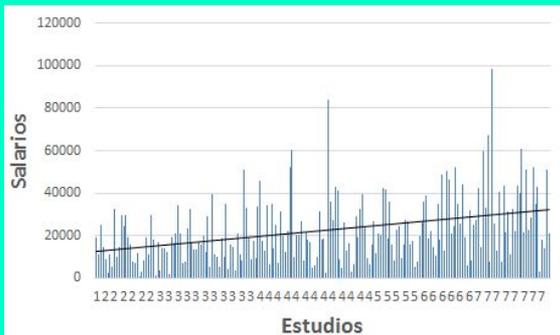
OBJETIVO 2

RELACIÓN ENTRE ESTUDIOS Y SALARIO BRUTO ANUAL

LEYENDA:

- 1 - Menos que primaria
- 2 - Educación primaria
- 3 - ESO
- 4 - Bachillerato
- 5 - Formación profesional
- 6 - Universitarios
- 7 - Licenciados

MEDIA X	4,330
MEDIA Y	22467,83
VARIANZA X	2,79
VARIANZA Y	237093077
D.TÍPICA X	1,67
D.TÍPICA Y	15397,83
COVARIANZA	10008,1
PEARSON	0,389125864



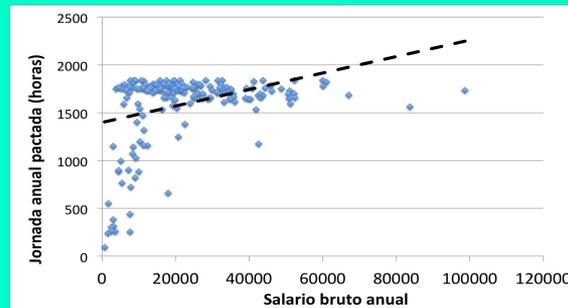
COMPARACIÓN EN EL SALARIO BRUTO ANUAL ENTRE GÉNEROS TENIENDO LA MISMA FORMACIÓN EDUCATIVA

Diferencias de salarios entre hombres y mujeres



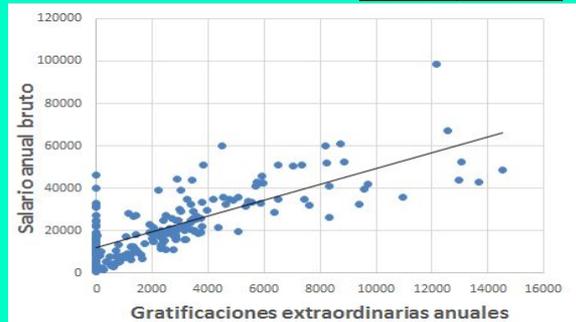
RELACIÓN ENTRE JORNADA ANUAL PACTADA Y SALARIO BRUTO ANUAL

MEDIA X	22467,832
MEDIA Y	1591,51
VARIANZA X	237165621,5
VARIANZA Y	140732,5599
D.TÍPICA X	15400,18252
D.TÍPICA Y	375,1433858
COVARIANZA	2042404,069
PEARSON	0,353523677



RELACIÓN ENTRE GRATIFICACIONES EXTRAORDINARIAS Y SALARIO BRUTO ANUAL

MEDIA X	2703,96
MEDIA Y	22049,15
VARIANZA X	9253903,3
VARIANZA Y	255731445,6
D.TÍPICA X	3042,02
D.TÍPICA Y	15991,61
COVARIANZA	34645954,2
PEARSON	0,712193461





CONCLUSIONES OBJETIVO 1 & 2



OBJETIVO 1:

TRAMO DE EDAD

Los resultados fueron, en su mayoría, bastante dispares con nuestras hipótesis. En las actividades de administración nos sorprendió que la moda fuese entre los **40-49 años**, ya que supusimos que estos trabajos requerían menor experiencia laboral. En actividades sanitarias y de servicios sociales nuestra hipótesis sí se confirmó ya que la moda se encontraba en los **40-49 años**, posiblemente porque estos trabajos suelen realizarse a partir de una edad más madura, y al final de una carrera más demandante. Y, los resultados sobre las actividades manufactureras también nos sorprendieron ya que la moda se situaba en los **30-39 años**, quizás por la necesidad de requerir un alto nivel físico.

REGIONES:

Los resultados en la representación gráfica de las regiones tampoco fueron mal encaminados. En la industria manufacturera el **Norte (Oeste + Este)** pudo ser la zona más desarrollada con un **45%** debido a la cantidad de áreas montañosas. En las actividades sanitarias y de servicios sociales, un **21%** de ellas tenían lugar en el **Este**, probablemente a causa de mayor pobreza. Y, en referencia a las actividades administrativas, los resultados sí que fueron sorprendentes ya que no esperábamos que el Centro, y la Comunidad de Madrid (al ser ciudades tan cosmopolitas) no tuvieran representación. Sin embargo, el **Este** acaparó casi un **48%** de ellas, a pesar de que no encontremos una razón para estos resultados.

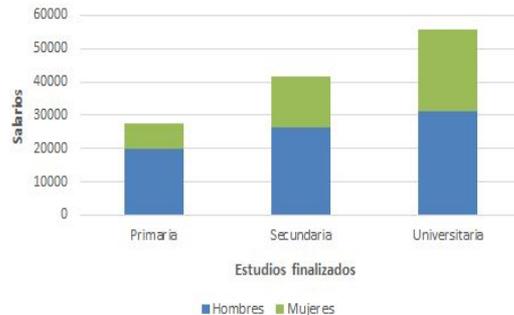
GÉNERO

Los resultados confirmaron nuestras hipótesis, ya que en las actividades sanitarias y de servicios sociales las mujeres claramente eran la mayoría (**86%**), debido a que la carrera de enfermería es más común en mujeres. En las actividades administrativas, el resultado es más parecido ya que estos trabajos son realizados por ambos sexos indistintamente, aunque la mujer predomine en este caso (**63%**). Y, en la industria manufacturera nos sorprendió el porcentaje tan pequeño de mujeres (**31%**), pese a que intuimos que los hombres fuesen a predominar (**69%**).

OBJETIVO 2:

En el estudio bidimensional analizamos diferentes factores con el salario bruto anual, aunque la única correlación fuerte que encontramos (con una coeficiente de correlación de Pearson de **0,71**) fue gratificaciones extraordinarias con salario. Al ver los resultados consideramos que la razón por la cual las personas con mayores salarios obtienen además mayores gratificaciones anuales, es porque sus puestos de trabajo son más altos por lo que las comisiones extras les corresponden a ellos.

Diferencias de salarios entre hombres y mujeres



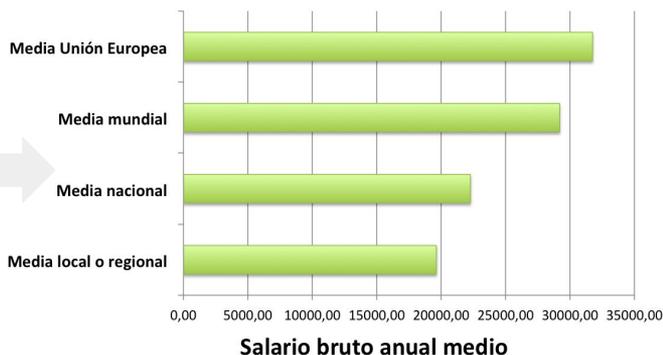
Al hacer la correlación de estudios con el salario llegamos a una correlación débil de tan solo **0,38**. Pero, pensamos que las excepciones fueron las que evitaron que hubiera una relación más fuerte y fiable. Por lo tanto, hicimos más investigación con el gráfico de comparación de salarios por género, y llegamos a la conclusión de que nuestra hipótesis era cierta ya que había una tendencia clara. Este gráfico también coincide con nuestra hipótesis de que los hombres ganan más que las mujeres pese a tener la misma formación.

Finalmente, observando los resultados llegamos a la conclusión de que pese a que no hubiera una correlación entre edades y salarios, si hay que una edad crítica en la que se concentran las personas con mayores salarios, siendo el tramo entre los **40-49 años**.

Y, aunque pensábamos que la jornada anual pactada sí que afectaba al salario, al buscar la correlación hallamos una relación bastante débil de **0,35**, lo que significa que nuestra estimación estaba lejos de la realidad.

OBJETIVO 3

Para el tercer objetivo hemos encontrado el salario medio de cada sector. Además, hemos calculado el coeficiente de variación para poder ver la fiabilidad de los resultados. A la derecha presentamos una representación gráfica para facilitar la comparación de los resultados.



Media local o regional	19.619,55 €
Media nacional	22.265,16 €
Media mundial	29.192,79 €
Media Unión Europea	31.742,11 €

Varianza local	179755014,72
Varianza nacional	198729101,07
Varianza mundial	965645166,15
Varianza union europea	810302613,22

Desviación típica local	13407,27
Desviación típica nacional	14097,13
Desviación típica mundial	31074,83
Desviación típica UE	28465,81

Coefficiente de variación local	0,68
Coefficiente de variación nacional	0,63
Coefficiente de variación mundial	1,06
Coefficiente de variación UE	0,90

OBJETIVO 4

Promedio de persona según los datos de la encuesta proporcionada. En este objetivo encontramos la moda de cada categoría salvo en aquellas marcadas con un asterisco (*) donde tuvimos que realizar la media ya que no existía una moda como tal. El objetivo era encontrar el dato más común en cada categoría creando así el "prototipo de persona" que realizó la encuesta. En la siguiente tabla se presentan los resultados.

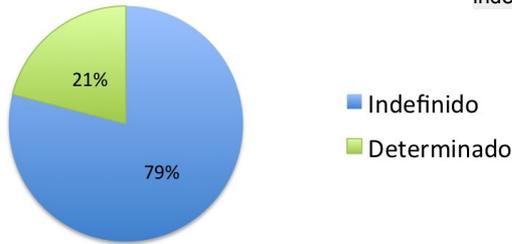
	Sin formato	Con formato
Región	5	Este
Actividad económica	Q0	Actividades sanitarias y servicios sociales
Tamaño del centro de trabajo	3	200 y más trabajadores
Mercado	1	Local o regional
Sexo	1	Hombre
Nacionalidad	1	España
Ocupación	D0	Técnicos; profesionales de apoyo
Responsabilidad en la organización o supervisión	0	No
Nivel de estudios	4	Segunda etapa de educación secundaria
Años de antigüedad	0	0
Meses de antigüedad	1	1
Tipo de jornada	1	Tiempo completo
Duración del contrato	1	Duración indefinida
Jornada anual pactada (horas) *	1591,5	1591 horas 30 minutos
Jornada semanal pactada (horas) *	35,66	35 horas 40 minutos
Salario bruto anual *	22467,83	22467,83
Gratificaciones extraordinarias abonadas en el año *	2994,17	2994,17
Valoración en especie *	84,47	84,47
Tramo edad	3	De 30 a 39 años

CONCLUSIÓN: A partir de los resultados obtenidos, podemos ver que nuestra hipótesis iba bien encaminada. Sin embargo, cabe destacar que la media del sueldo bruto anual es superior en la **Unión Europea** a la del mercado **mundial**. Sin embargo, a partir del coeficiente de variación podemos observar que la categoría con los datos más dispersos es la del mercado **mundial**, lo cual podría explicar porque la media es inferior. El coeficiente de variación para el mercado **europeo** es también bastante alto, por lo tanto podemos concluir que la media quizás no ilustra la realidad.

CONCLUSIÓN: Los resultados obtenidos han sido, en general, muy sorprendentes en relación con nuestra hipótesis. Nos ha llamado la atención que la mayoría de las personas tuviesen tan solo **estudios secundarios**. Además, consideramos que con una moda de **1 mes de antigüedad**, es extraño que la mayoría tuviesen tanto una **jornada completa** como un **contrato indefinido**. También nos sorprendió que las actividades más comunes fuesen **sanitarias**. Eso sí, también cabe destacar que nuestra hipótesis acertó con el tramo de edad (comenzando en los **30 años**).

OBJETIVO 5

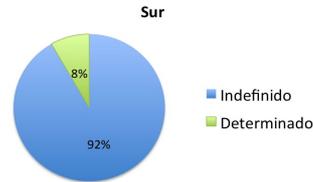
Total



Para realizar un estudio binomial, hemos decidido coger como probabilidad los tipos de contrato existentes en España. A partir de aquí, realizamos un estudio binomial con una probabilidad de éxito de 0,79 (la probabilidad de tener un contrato indefinido).

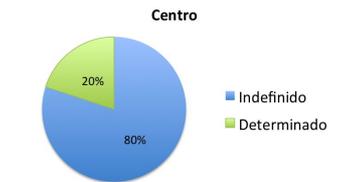
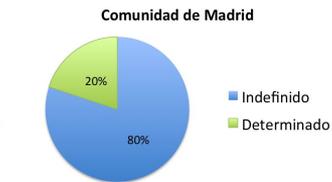
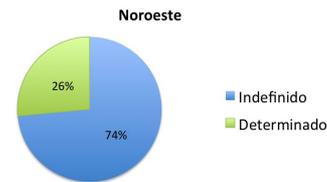
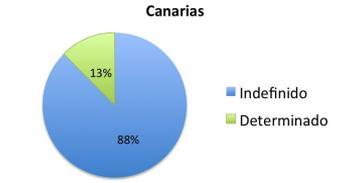
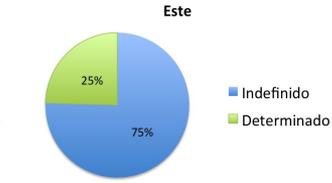
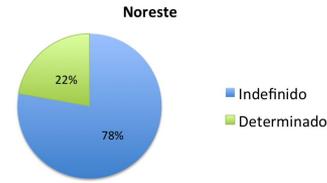
Aquí podemos ver los resultados de toda España a partir de las 200 personas encuestadas. Las dos variables son duración de contrato indefinido o determinado.

	Indef	Deter	TOTAL
ESPAÑA	158	42	200
Probabilidad	0,79	0,21	1
Porcentaje	79%	21%	100%



En este estudio binomial, tomamos como "n" el número total de encuestados de cada región. A continuación realizamos una "Previsión", calculando la media de personas que debería tener cada tipo de contrato, aplicando como probabilidad de éxito la total de España; es decir, multiplicamos el número de encuestados por 0,79 y redondeamos al número entero más cercano. La "Previsión" por lo tanto es el número de encuestados que debería tener un contrato indefinido en cada región siguiendo el promedio de España.

Por último, miramos los datos reales y podemos comprobar si alguna de las regiones sigue el promedio español. Para facilitar la comparación hemos realizado varios gráficos de sectores:



En la esquina superior derecha se encuentra el promedio español de los tipos de contrato según los datos de la encuesta, y a la derecha están los gráficos con el número real de encuestados con contratos indefinidos y determinados por regiones.

Encuestados (n)	PREVISIÓN		REALIDAD	
	Indef	Deter	Indef	Deter
Región Noroeste	34	7	25	9
Región Noreste	27	6	21	6
Comunidad de Madrid	30	6	24	6
Región Central	20	4	16	4
Región Este	57	12	43	14
Región Sur	24	5	22	2
Canarias	8	2	7	1

CONCLUSIÓN: Una vez realizado el estudio, podemos observar que la hipótesis inicial era correcta. Como habíamos previsto, la **Región Noreste**, la **Comunidad de Madrid** y la **Región Central** siguen el promedio nacional; mientras que en la **Región Sur** y en **Canarias** hay un mayor número de contratos indefinidos que la media. Nos llama la atención el hecho de haber en la **Región Este** y **Noroeste** más contratos determinados que el promedio nacional, lo cual puede deberse a que provincias como la Comunidad Valenciana y Murcia tienen un PIB por debajo de la media española, lo cual representa una economía más floja y consecuentemente, un mayor número de contratos determinados. En la **Región Sur** y en **Canarias** encontramos un número mayor de contratos indefinidos que en la media española, lo cual nos sorprendió. Esto se debe a que en una economía fuerte tiende a haber más estabilidad y por lo tanto hay un mayor número de trabajadores con contratos indefinidos. Sin embargo, en la **Región Sur** y en **Canarias** hay más contratos de duración indefinida que la media; se nos ocurre que quizás el número de encuestados no sea lo suficientemente representativo.