

Introducción

En este proyecto tomamos el conjunto de datos [Sleep Health and Lifestyle Dataset | Kaggle](#) en donde se encuentran los datos de distintas personas como su profesión, su actividad física y la calidad de sueño. Tomaremos este conjunto de datos para conocer qué cosas son las que afectan al sueño como por ejemplo la cantidad de actividad física, la profesión y la presión a la que cada uno de ellos se somete.

Las preguntas que buscaremos responder con los datos de este conjunto son las siguientes:

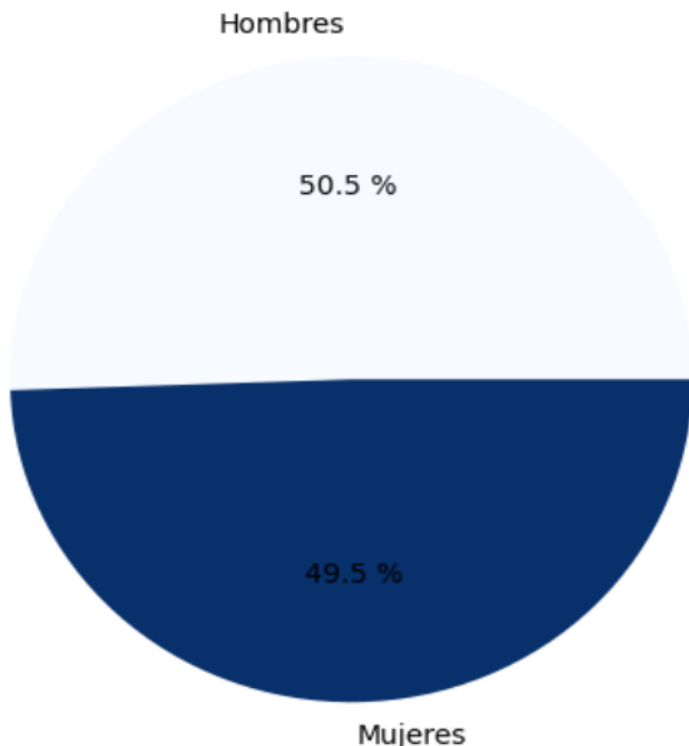
- ¿Quiénes tienen peor calidad del sueño, hombres o mujeres?
- ¿Existe una relación entre la calidad del sueño de las personas y su profesión?
- ¿La actividad física afecta al sueño?
- ¿Qué profesión presenta la peor calidad de sueño?
- ¿Cuál profesión tienen a las personas con el mayor nivel de estrés y el mayor índice de masa corporal?
- ¿En qué rango de edades se encuentran la mayor cantidad de trastornos del sueño?
- La cantidad de pasos al día, ¿afecta la calidad del sueño? ¿al índice de masa corporal?

Estas preguntas nos servirán como punto de partida en nuestro de datos.

Análisis del conjunto

Análisis General

Sabemos que en el conjunto de datos contamos con 374 filas y 13 columnas de datos, de las cuales sabemos que en general los datos se dividen de la siguiente forma:



Género	Cantidad
Hombres	189
Mujeres	185

Esto nos da un vistazo general del conjunto y podemos decir que el conjunto se encuentra en proporción, respecto a la cantidad de hombre y mujeres.

Clasificación dada por edad

En el conjunto podemos encontrar las edades de las personas, con respecto a las edades podemos encontrar lo siguiente:

- La edad más baja es de 27 años
- La edad más alta es de 59 años

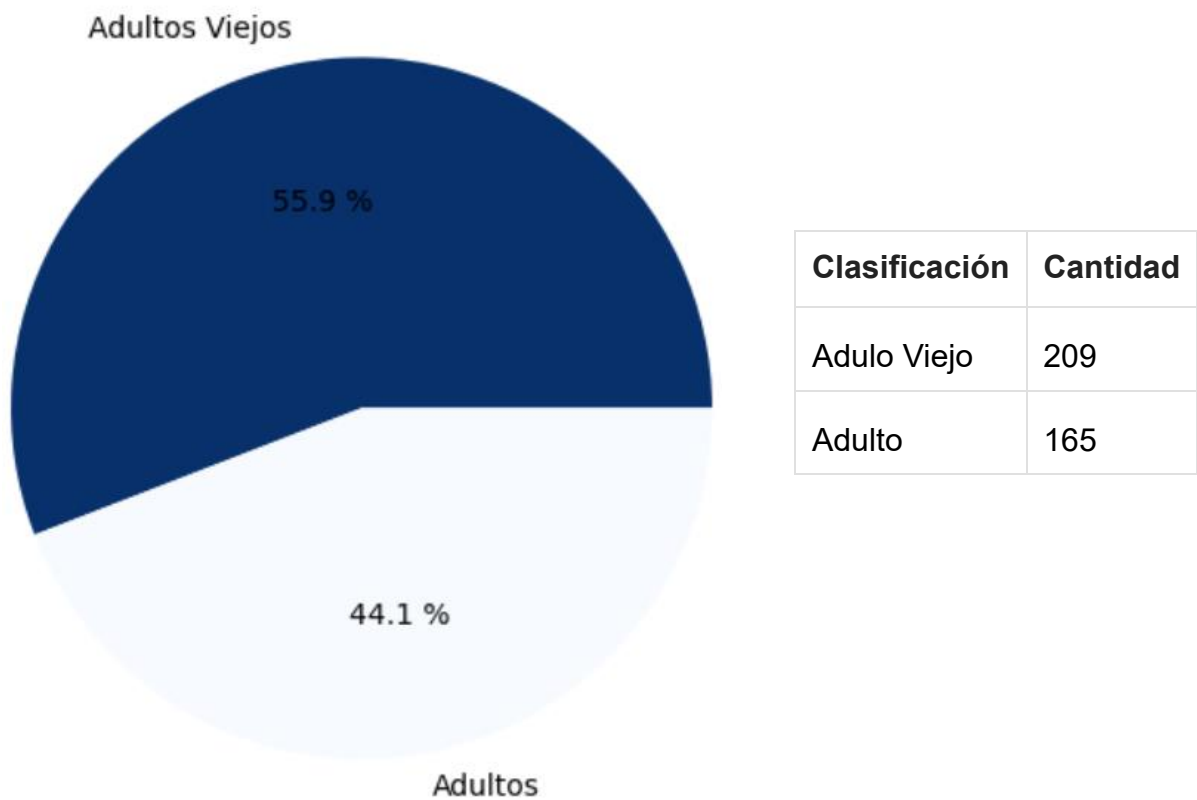
Lo que podemos decir que el conjunto de datos toma en cuenta a personas que ya se encuentran en la edad laboral iniciando en el trabajo o con pocos años en el mismo, o de personas cercanas a jubilarse.

La edad promedio del conjunto de datos es de 42 años y la desviación estándar es de 8.67, lo que quiere decir que en este conjunto podemos encontrar datos muy dispersos lo que puede deberse a los pocos datos recabados.

Para apoyarnos en este análisis se agregó una nueva columna al conjunto de datos a la que se le llamó `Age_group` en la cuál se clasifico la edad de las personas de la siguiente manera:

- Edad menor o igual a 40: `Adulto`
- Edad mayor a 40: `Adulto viejo`

Tomando en cuenta este ordenamiento de los datos tenemos lo siguiente:

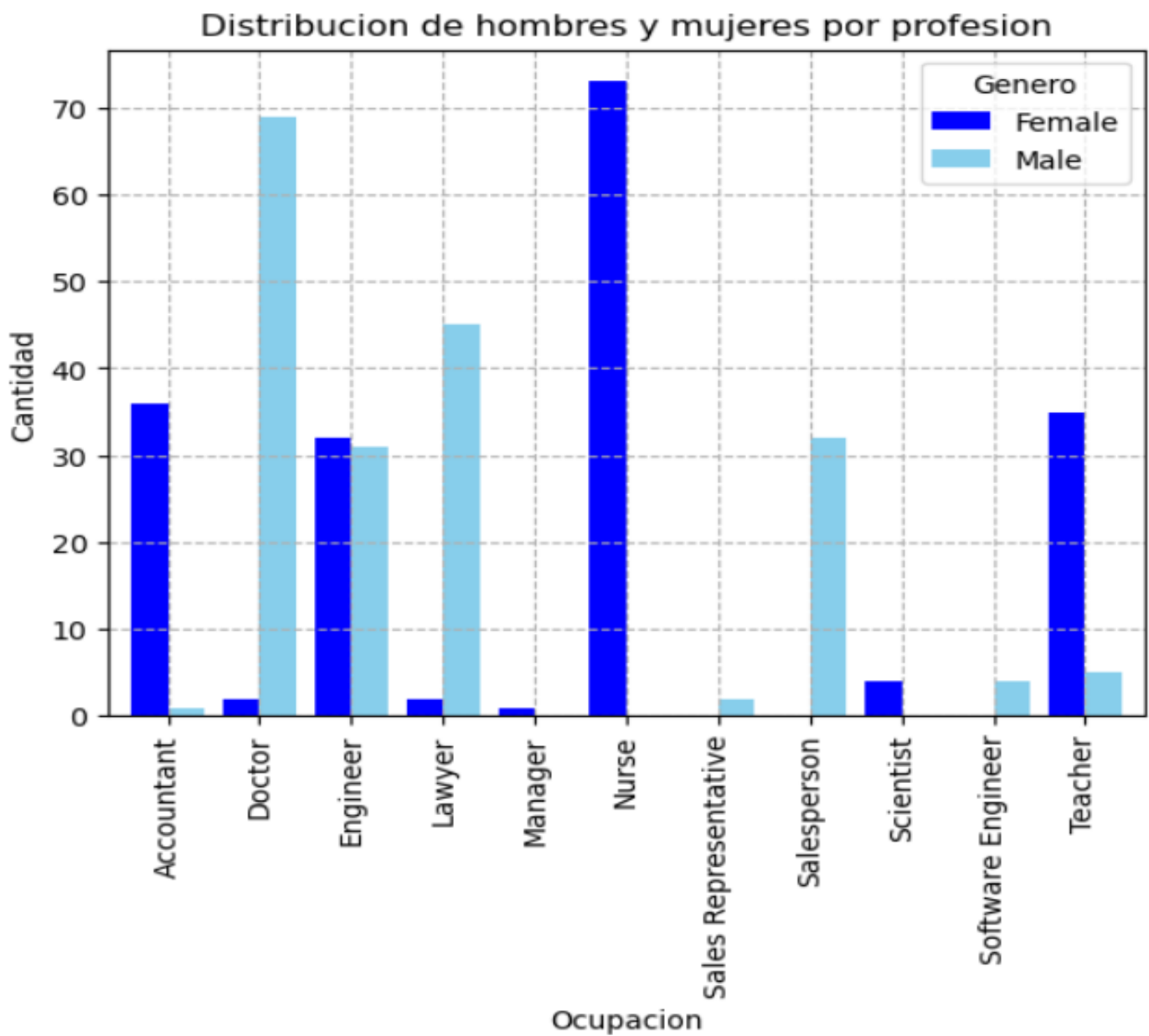


Finalmente, el promedio de edades por cada género queda de la siguiente forma:

Gender	mean	Std
Female	47.40	8.10
Male	37.07	5.66

Podemos ver que existe una diferencia de 10 años en el promedio de edades de ambos géneros, y también notamos una diferencia grande en la desviación estándar, siendo ésta de 3 años.

Ocupaciones



En esta gráfica notamos la siguiente:

Las ocupaciones en donde hay más presencia de mujeres son:

- Contador
- Ingeniería
- Gerente
- Enfermería
- Ciencia
- Maestro

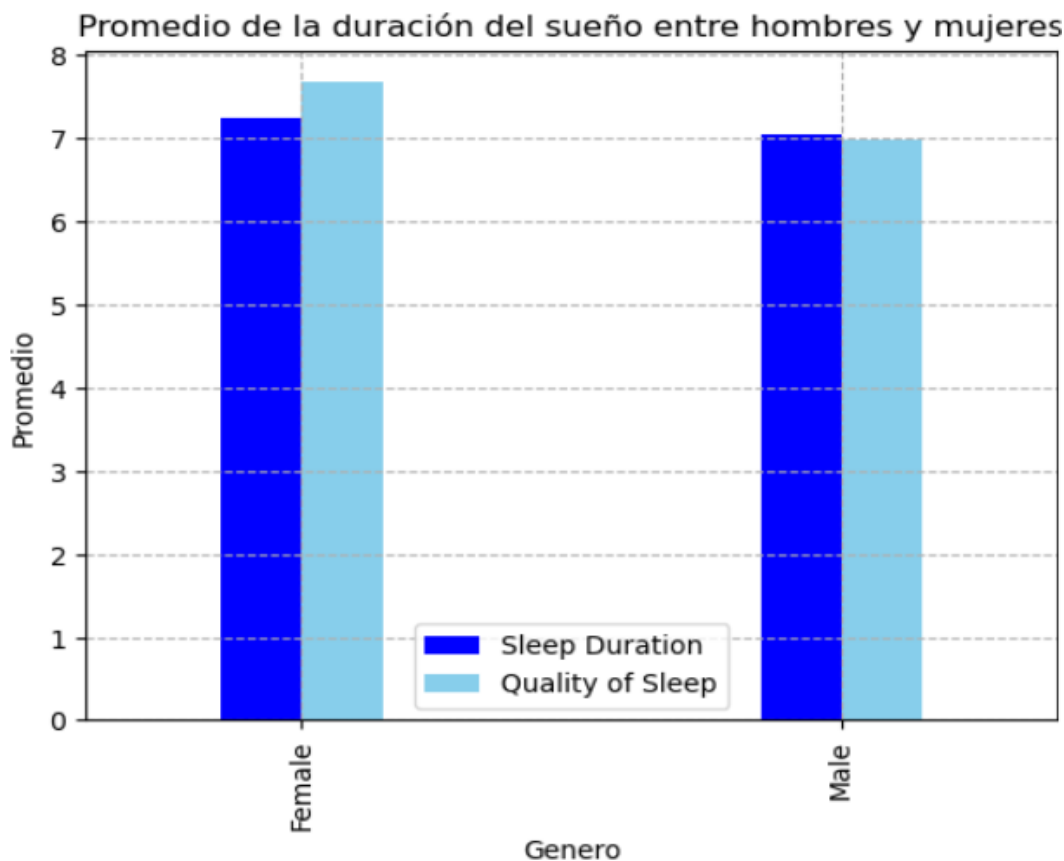
Las ocupaciones en donde los hombres tienen más presencia:

- Doctor
- Representante de ventas
- Ventas
- Abogado
- Ingeniero de Software

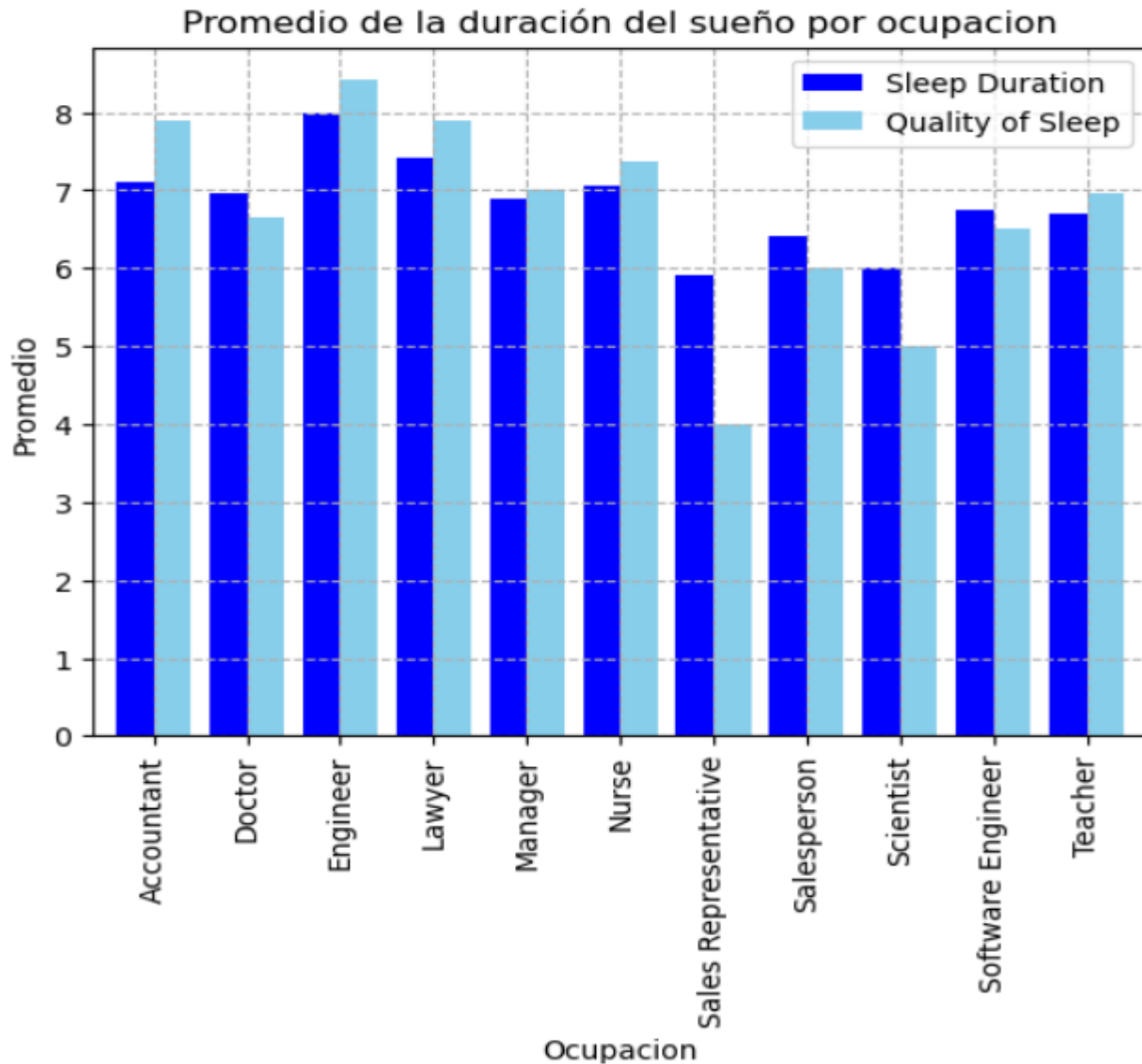
Es importante mencionar que solamente en *Ingeniería* la diferencia es de solamente una persona.

Calidad del Sueño

El promedio de la duración del sueño entre hombres y mujeres es la siguiente:



Podemos ver que en general, las mujeres tienen un mejor sueño que los hombres no solamente en cantidad, sino que, en calidad, para profundizar más, obtendremos los mismos datos, pero ahora en cada una de las profesiones del conjunto.



En estas tablas podemos ver que las peores ocupaciones que tienen la peor calidad y cantidad de sueño son:

- Representante de Ventas
- Científico
- Ventas

Y aquellas que tienen mejor calidad y tiempo de sueño son:

- Ingeniero
- Abogado
- Contador

En las profesiones donde peor calidad y cantidad de sueño hay se encuentran más hombres que mujeres.

Problemas de salud

Finalmente buscaremos que tan buena salud tienen cada una de las personas en las distintas profesiones, con este dato podremos sacar nuestras conclusiones y podremos responder las preguntas que nos hicimos en un inicio al comenzar con nuestro análisis.

Occupation	BMI Category	Count
Lawyer	Normal Weight	1
Software Engineer	Overweight	1
	Obese	1
Manager	Overweight	1
Teacher	Obese	1
Doctor	Normal Weight	2
Lawyer	Obese	2
	Overweight	2
Software Engineer	Normal Weight	2
Sales Representative	Obese	2
Engineer	Overweight	3
	Normal Weight	4
Doctor	Obese	4
Scientist	Overweight	4
Accountant	Normal Weight	5
	Overweight	6
Teacher	Normal	6
Nurse	Normal Weight	7
Accountant	Normal	26
Salesperson	Overweight	32
Teacher	Overweight	33
Lawyer	Normal	42
Engineer	Normal	56
Doctor	Normal	65

Nurse	Overweight	66
-------	------------	----

Otro punto importante a tener en cuenta es saber si la cantidad de pasos al día afecta a la salud y a la calidad del sueño, para conocer esto tendremos el promedio de los pasos diarios, el nivel de actividad física, duración del sueño y la calidad del sueño en cada una de las profesiones.

Occupation	Daily Steps	Physical Activity Level	Sleep Duration	Quality of Sleep
Sales Representative	3000.0	30.0	5.9	4.0
Scientist	5350.0	41.0	6.0	5.0
Manager	5500.0	55.0	6.9	7.0
Software Engineer	5800.0	48.0	6.75	6.5
Teacher	5957.5	45.62	6.69	6.97
Engineer	5980.95	51.86	7.99	8.41
Salesperson	6000.0	45.0	6.40	6.0
Doctor	6808.45	55.35	6.97	6.65
Accountant	6881.08	58.11	7.11	7.89
Lawyer	7661.70	70.42	7.41	7.89
Nurse	8057.53	78.59	7.06	7.37

Conclusiones

Este conjunto de datos nos sirve para darnos un vistazo general de cada una de las profesiones y como éstas afecta a la salud de las personas. Digo general porque un problema de este conjunto es la cantidad de datos que tiene, ya que al contar solamente con 307 filas son muy pocas para todo el margen de edad que se tomó. Eso lo podemos notar al obtener el promedio de las edades y la desviación estándar por lo que, al terminar con nuestro análisis, podemos decir que hacen falta más datos o en su caso reducir el margen de edades de los encuestados.

Pasando al sueño de cada una de las personas tenemos lo siguiente:

Las mujeres en términos generales tienen una mejor calidad y cantidad de sueño respecto a los hombres ya que, las mujeres pueden tener un mejor control de estrés o porque las mujeres a las que se les hizo la encuesta parecen tener profesiones en las que se habla con menos personas que las profesiones de los hombres.

Esto lo podemos acompañar si tomamos solamente las primeras cinco profesiones que presentan la mejor calidad y cantidad de sueño:

- Ingeniería
- Abogado
- Contador
- Enfermería
- Gerente

De estas profesiones en tres se encuentran más mujeres (Contador, Enfermería y Gerente) en una hay más hombres (Abogado) y en Ingeniería se encuentra la diferencia de una persona entre ambos.

Muchas de estas profesiones también se repiten cuando revisamos la cantidad de pasos al día, podemos ver que las cinco profesiones donde más pasos al día se dan son:

- Enfermería
- Abogado
- Contador
- Doctor
- Vendedor

En la mayoría de estas personas podemos ver que caminar forma una parte fundamental de ellas por lo que no es de extrañar que éstas presenten un promedio alto de pasos al día y esto también se ve en la salud del sueño, ya que si vemos la tabla de los pasos al día podemos ver que en ella existe una relación que entre más pasos se den mejor va a ser el sueño.

Aunque esta misma relación no se nota cuando hablamos del índice de masa corporal, porque de la lista anterior podemos ver que en las siguientes se encuentran muchas personas con sobrepeso:

- Enfermería
- Vendedor

Finalmente, las preguntas hechas alrededor de la edad de los encuestados no se respondieron ya que se encontró que el rango de edades es muy alto y la cantidad de datos muy pocos por lo que las respuestas que se puedan encontrar en este ámbito no pueden llegar a ser relevantes con la realidad.

ISAIAS DAVID PEREZ SARMIENTO.

ECONOMISTA / ANALISTA DE DATOS