



UNIVERSIDAD DE COLIMA – TRANSICIÓN

DIPLOMADO DE ANALÍTICA Y CIENCIA DE DATOS

PROYECTO FINAL

ANÁLISIS DE LA SALUD DE SUEÑO Y ESTILOS DE VIDA

PRESENTA:

CINTHIA LEONORA MURILLO AVALOS

OCTUBRE 2024.

Índice

1. Selección y Justificación de la Base de Datos.....	3
<i>Introducción.....</i>	<i>3</i>
<i>Justificación de la selección</i>	<i>3</i>
<i>Objetivos.....</i>	<i>3</i>
2. Preparación y Limpieza de los Datos	4
<i>Descripción de la Base de Datos:.....</i>	<i>4</i>
<i>Limpieza de Datos:.....</i>	<i>4</i>
3. Análisis Exploratorio de Datos (EDA)	5
4. Visualización de Datos.....	6
5. Interpretación y Conclusiones.....	8
<i>Resultados del análisis</i>	<i>8</i>
<i>Conclusiones</i>	<i>9</i>

1. Selección y Justificación de la Base de Datos

Introducción

La base de datos seleccionada ofrece información relevante en relación a la salud del sueño y los estilos de vida, siendo estos dos aspectos de suma importancia para la salud en general. Se incluyen las variables de género, edad, ocupación, duración del sueño, calidad del sueño, nivel de actividad física, niveles de estrés, categoría de IMC, presión arterial, frecuencia cardíaca, pasos diarios y presencia o ausencia de trastornos del sueño. El análisis de este conjunto de variables permite identificar patrones y tendencias que influyen en la salud en general, con lo cual se pueden prevenir enfermedades y promover hábitos saludables, mejorando así la calidad de vida de las personas.

La base de datos seleccionada fue: Sleep_health_and_lifestyle_dataset.csv, la cual se obtuvo del URL: <https://www.kaggle.com/datasets/uom190346a/sleep-health-and-lifestyle-dataset/data>.

Justificación de la selección

El motivo de selección de la base de datos fue la relevancia e impacto que tiene tanto la calidad del sueño como el estilo de vida en la salud en general. En este sentido, el analizar patrones o tendencias ayuda a la mejor toma de decisiones, en este caso, decisiones relacionadas con la salud de la población. Por lo tanto, el conocer cómo ciertos factores del estilo de vida afectan el sueño, permite prevenir trastornos relacionados y enfermedades crónicas, promover hábitos saludables y por ende mejorar la calidad de vida de las personas.

Objetivos

- Caracterización demográfica de los individuos con relación a género, edad y ocupación.
- Comparativa de la cantidad de horas de sueño por género.
- Comparativa de la calidad de sueño por género.
- Identificar si el nivel de estrés tiene asociación con los minutos de actividad física realizados al día.

- Identificar si la categoría de IMC tiene asociación con el nivel de estrés.

2. Preparación y Limpieza de los Datos

Para el tratamiento de los datos se utilizó excel así como para el análisis exploratorio, y para la visualización de datos mediante gráficos descriptivos se usó Tableau.

Descripción de la Base de Datos:

La base de datos “salud del sueño y estilo de vida” está conformada por 374 registros y 13 variables, de las cuales nueve son numéricas y cuatro categóricas.

La descripción a detalle de cada una de las variables de estudio es la siguiente:

- Identificación de la persona: # identificador para cada individuo
- Género: con opción de respuesta masculino o femenino.
- Edad: años de vida de cada individuo.
- Ocupación: ocupación o profesión de cada individuo con opción de respuesta abogado, científico, contador, doctor, enfermera, gerente, ingeniero, ingeniero de software, maestro, representante de ventas y vendedor.
- Duración del sueño (horas): número de horas que cada individuo duerme al día.
- Calidad del sueño: calificación subjetiva de la calidad del sueño, con escala de 1 a 10.
- Nivel de actividad física (minutos/día): número de minutos de actividad física que cada individuo realiza al día.
- Nivel de estrés: calificación subjetiva del nivel de estrés experimentado por cada individuo, con escala de 1 a 10.
- Categoría de IMC: categoría de IMC de cada individuo, con opción de respuesta bajo peso, normal, sobrepeso y obeso.
- Presión arterial (sistólica/diastólica): medición de la presión arterial de cada individuo.
- Frecuencia cardíaca (bpm): frecuencia cardíaca en reposo de la persona en latidos por minuto.
- Pasos diarios: número de pasos que cada individuo da por día.
- Trastorno del sueño: presencia o ausencia de un trastorno de sueño en cada individuo con opción de respuesta: ninguno, insomnio y apnea del sueño.

Limpieza de Datos:

Para la identificación de valores faltantes se utilizó en Excel la función de CONTAR.BLANCO, lo que nos arrojó 0% de registros nulos (NULL), por tal motivo se conservó para el análisis, el total de los registros (374), debido a la captura adecuada no fue necesario hacer corrección de datos ni transformaciones.

3. Análisis Exploratorio de Datos (EDA)

Para el análisis exploratorio, en principio, se le dio formato de tabla al conjunto de datos en Excel, posteriormente, se elaboraron tablas dinámicas con las variables demográficas (género, edad y ocupación), duración de sueño, calidad de sueño, minutos de actividad física, nivel de estrés y categoría de IMC. Para la variable edad se calcularon estadísticas descriptivas, específicamente, la media y desviación estándar. Además de la media para la duración de sueño y calidad de sueño.

Ilustración 1 Análisis exploratorio en Excel – parte 1

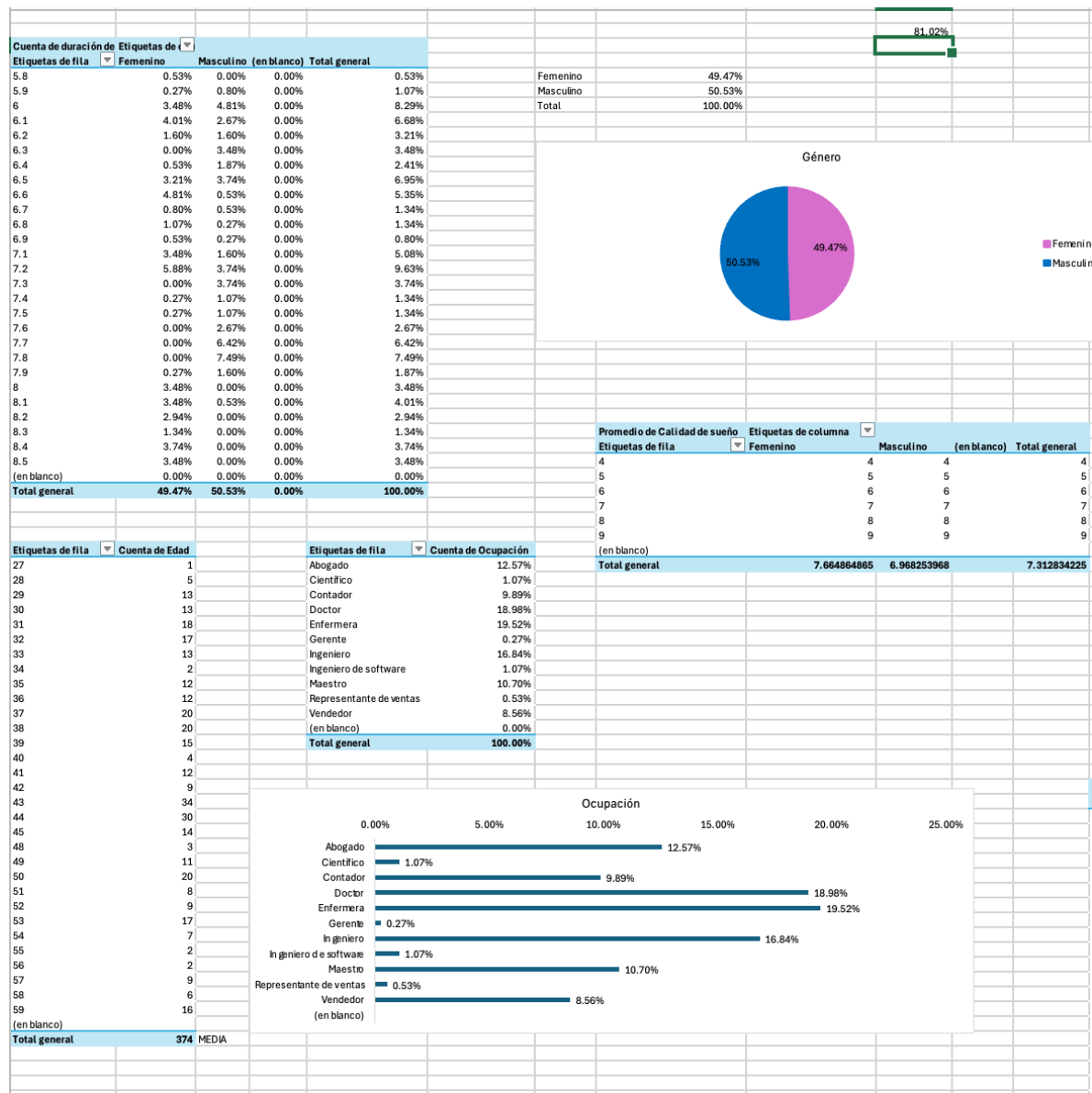


Ilustración 2 Análisis exploratorio en Excel – parte 2



4. Visualización de Datos

Para la visualización interactiva de los datos se utilizó Tableau en el cual se generaron tres historias. La primera se titula “Caracterización de los individuos de estudio” con base a las variables género, edad y ocupación. La segunda se llama “Salud de sueño” en la cual se observa la cantidad de sueño por ocupación, así como la calidad de sueño por género. Y la tercera lleva por título “Tablas cruzadas del nivel de estrés” presentando el nivel de estrés por actividad física y nivel de estrés por la categoría de IMC.

Ilustración 3 Gráficos de la historia 1 en Tableau.

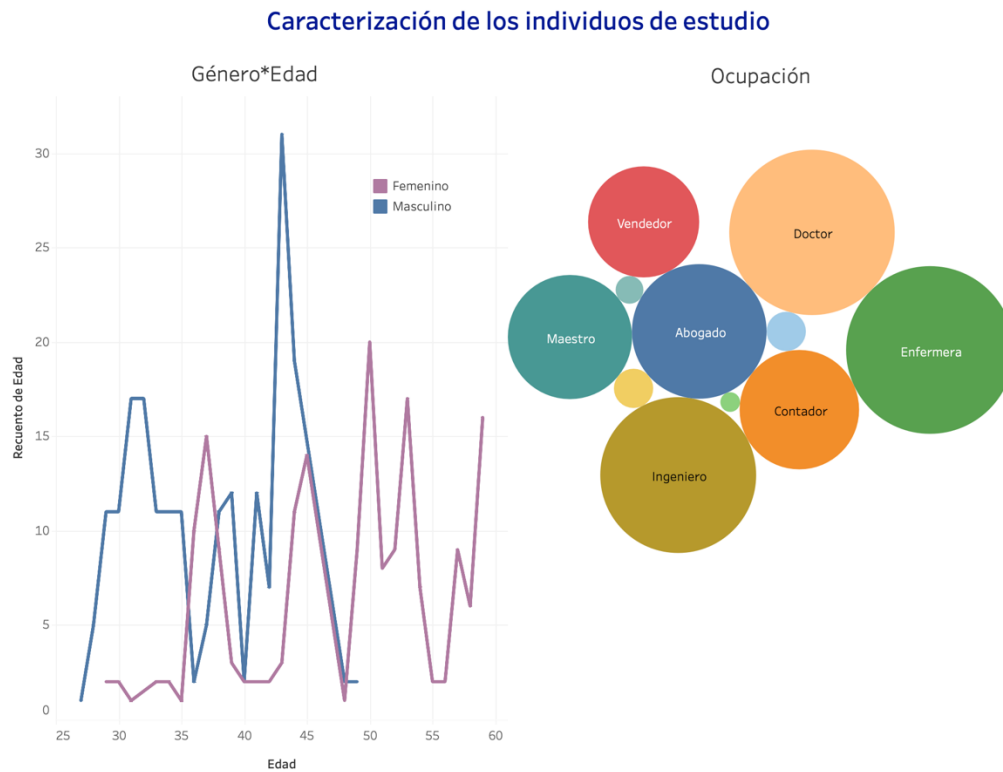


Ilustración 4 Gráficos de la historia 2 en Tableau.

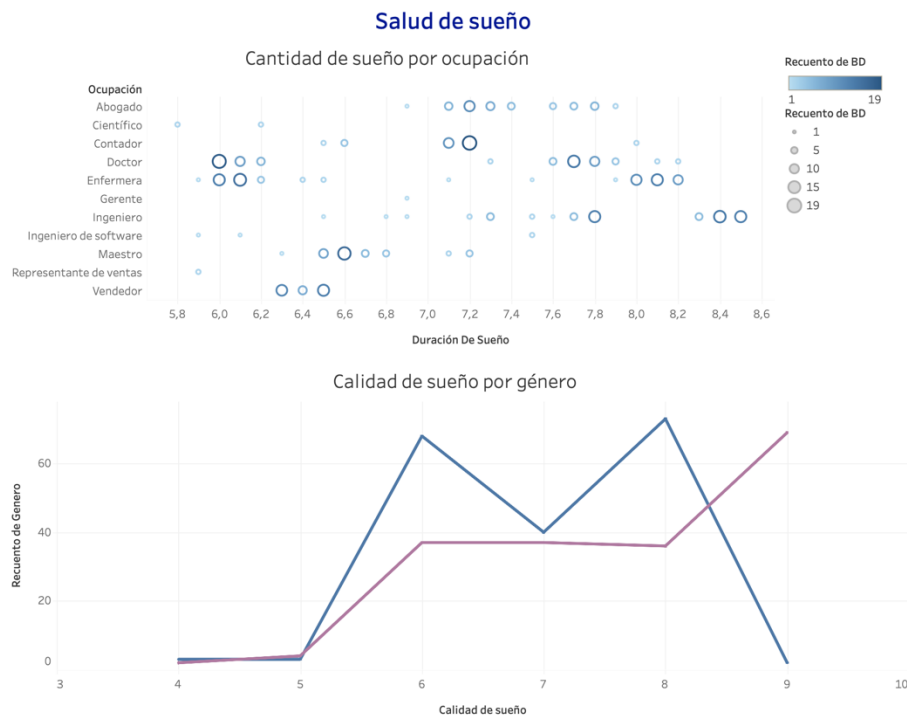
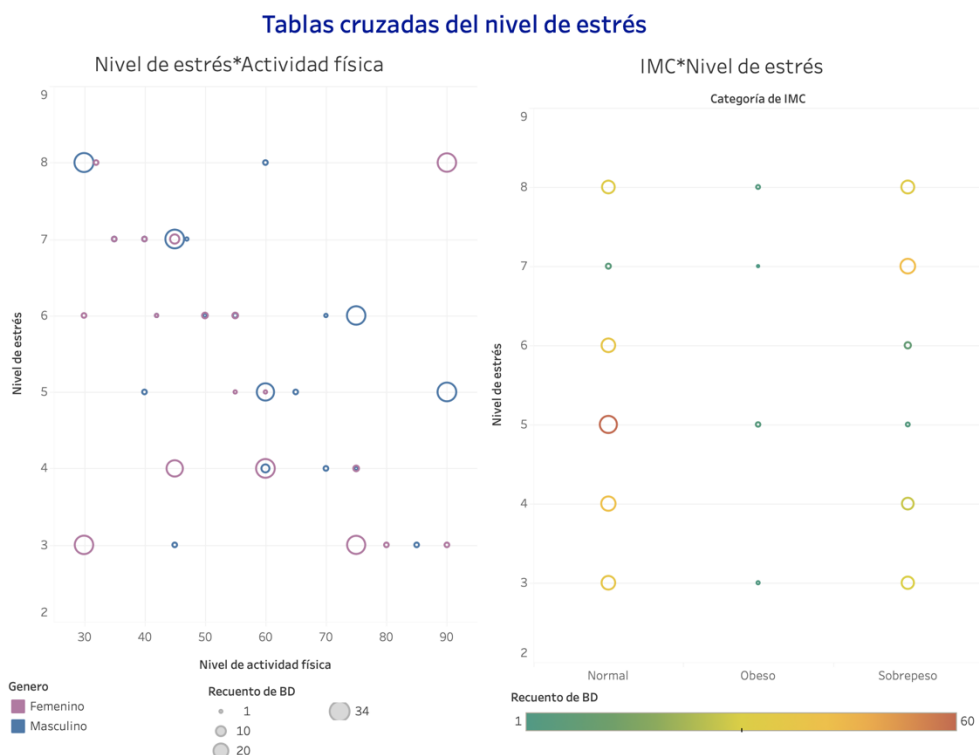


Ilustración 5 Gráficos de la historia 3 en Tableau.



Visualizar gráficos en el link:

https://public.tableau.com/app/profile/cinthia.murillo/viz/ProyectoFinal_17301875265050/Historia?publish=yes

5. Interpretación y Conclusiones

Resultados del análisis

La muestra esta conformada por el 49.47% de sexo femenino y el 50.53% masculino, oscilando entre los 27 y 59 años de edad, con una media de 42.18 años y una desviacion estandar de 8.67.

Las tres ocupaciones más frecuentes entre los individuos de análisis son: enfermera/o con el 19.52%, doctora/o con 18.98 e ingeniera/o con 16.84%.

Se encontró que el 62.7% de las mujeres suele dormir menos de 8 horas diarias contra el 98.9% de los hombres, sin embargo, el promedio de duración de sueño es similar en ambos sexos, ya que en mujeres es de 7.22 horas y en hombres de 7.03 horas.

De acuerdo a la calificación subjetiva mencionada por los individuos de estudio, resultó que las mujeres tienen un promedio más alto en calidad de sueño que los hombres, es decir, el promedio en mujeres es de 7.66 y en hombres de 6.96.

Las tres ocupaciones con el promedio más alto en cantidad de sueño son: ingeniero con 7.98 hrs., abogado con 7.41 hrs. y contador con 7.11 hrs.

Las cantidades de minutos de actividad física más frecuentes en los participantes son: 30" (18.18%), 45" (18.18%), 60" (18.72%), 75" (17.91%) y 90" (17.91%). Los niveles de estrés con mayor frecuencia son: 3 (18.98%), 4 (18.72%) y 8 (18.72%). No se encontró asociación entre ambas variables.

Los resultados indican que el 57.75% de los individuos pertenecen a la categoría de IMC normal, el 39.57% está en sobrepeso y el 2.67% en obesidad.

El cruce de las variables categoría de IMC y nivel de estrés no presentaron asociación.

Conclusiones

El análisis de datos nos permite identificar patrones y tendencias para la mejor toma de decisiones. El presente estudio tuvo por objetivo conocer las características demográficas, salud de sueño y los estilos de vida de 374 participantes. En este sentido, se lograron realizar algunas comparativas de la cantidad y calidad de sueño por género y ocupación, así como el conocer que no existe asociación entre la variable nivel de estrés con actividad física o categoría de IMC. De manera general se sugiere considerar el dormir al menos ocho horas diarias, ya que los resultados indican que solo el 19% suele hacerlo. Asimismo se sugiere atender los niveles de estrés ya que el 18.27% respondió tener un nivel alto (nivel 8 de 10), y su IMC al resultar que alrededor del 40% están en sobrepeso. Los análisis del presente estudio podrán servir como base para futuras investigaciones sobre el tema, realizando un análisis con mayor detalle y aplicando métodos estadísticos avanzados para identificar la asociación de la calidad de sueño y los estilos de vida.