

Análisis y visualización de Datos sobre Niveles de Obesidad en México, Perú y Colombia

Andrés Salcedo Vera, Andrés Ricardo Cáceres, Daniela Alejandra Benavidez, Leidy Katherine Vela
Fundación Universitaria Konrad Lorenz

Resumen—Este documento presenta un análisis exploratorio de un conjunto de datos sobre los niveles de obesidad en México, Perú y Colombia. Se aborda la problemática del incremento de la obesidad en la población, utilizando datos obtenidos a través de encuestas en línea. Se detallan las variables analizadas, el preprocesamiento de los datos y los métodos empleados para el análisis. Finalmente, se presentan los próximos pasos a seguir en la investigación.

I. INTRODUCCIÓN

La obesidad es un problema de salud pública que afecta a millones de personas en todo el mundo. En América Latina, países como México, Perú y Colombia han registrado un incremento en los índices de obesidad, lo que genera preocupaciones en términos de salud y bienestar. El presente estudio tiene como objetivo analizar un conjunto de datos que estima los niveles de obesidad en individuos de estos países, basado en hábitos alimenticios y condiciones físicas.

El conjunto de datos utilizado proviene de encuestas anónimas realizadas a individuos de entre 14 y 61 años. Contiene 2111 registros y 17 atributos relacionados con la alimentación, la actividad física y el estado de salud de los participantes.

II. CONJUNTO DE DATOS

El conjunto de datos utilizado está disponible en formato CSV y contiene información detallada sobre hábitos alimenticios y condiciones físicas de los participantes. Puede consultarse en el siguiente enlace:

- **Fuente:** <https://www.kaggle.com/datasets/aravindpcoder/obesity-or-cvd-risk-classifyregressorcluster>
- **Variables principales:**
 - Consumo frecuente de alimentos calóricos (FAVC)
 - Frecuencia de consumo de verduras (FCVC)
 - Número de comidas principales (NCP)
 - Consumo de agua (CH2O)
 - Actividad física (FAF)
 - Tiempo de uso de dispositivos tecnológicos (TUE)
 - Transporte utilizado (MTRANS)
- **Clasificación de obesidad:**
 - Bajo peso (<18.5)
 - Normal (18.5 - 24.9)
 - Sobrepeso (25.0 - 29.9)
 - Obesidad I (30.0 - 34.9)
 - Obesidad II (35.0 - 39.9)
 - Obesidad III (>40)

III. ANÁLISIS EXPLORATORIO DE DATOS

Para comprender mejor la distribución de los datos, se realizó un análisis exploratorio, el cual incluyó:

- Limpieza de datos: detección y manejo de valores nulos y atípicos.
- Análisis estadístico descriptivo: media, mediana, moda y desviación estándar de cada variable.
- Visualización de datos: gráficos de distribución de variables clave.

Se encontró que la mayoría de los participantes tienen un IMC dentro de las categorías de normalidad y sobrepeso, con una menor proporción en las categorías de obesidad. También se observó una correlación entre la actividad física y el índice de obesidad.

IV. PREGUNTAS DE INVESTIGACIÓN

Para comprender mejor la distribución y los factores asociados a la obesidad en México, Perú y Colombia, se plantean las siguientes preguntas clave:

- ¿Qué patrones se pueden identificar en el consumo de alimentos y su relación con los niveles de obesidad?
- ¿Cómo influyen factores como la edad, el género y el medio de transporte en los niveles de obesidad?
- ¿Qué tendencias generales se pueden observar a partir de la visualización de los datos en Power BI?

El análisis exploratorio y la generación de visualizaciones permitirán responder estas preguntas y proporcionar una visión más clara sobre la obesidad en los países estudiados.

V. EXPLICACIÓN DEL AVANCE Y PRÓXIMOS PASOS

Hasta el momento, se han realizado las siguientes acciones:

- Carga y limpieza del conjunto de datos.
- Análisis exploratorio mediante gráficos y estadísticas descriptivas.
- Identificación de patrones en los hábitos alimenticios y la actividad física en relación con los niveles de obesidad.

Los siguientes pasos incluyen:

- Creación de visualizaciones en Power BI para analizar la distribución de la obesidad en México, Perú y Colombia.
- Exploración de tendencias y correlaciones entre hábitos alimenticios, actividad física y niveles de obesidad.
- Generación de un informe con hallazgos y análisis detallado de la situación de la obesidad en los países estudiados.

VI. CONCLUSIÓN

Hasta el momento, se ha realizado la limpieza y exploración inicial del conjunto de datos sobre obesidad en México, Perú y Colombia. Esto ha permitido comprender la distribución de la obesidad en los países analizados y la estructura de los datos disponibles.

A futuro, se trabajará en la creación de visualizaciones interactivas en Power BI para identificar patrones y tendencias más detalladas. Estas visualizaciones facilitarán el análisis de la relación entre hábitos de vida y niveles de obesidad, proporcionando una base para estudios más profundos en el tema.