



LABORATORIO #3. TEA



Estudiante: Valeria Paola Araujo Cedeño

Identificador de Github: valeriaaraujo03

Nombre del Grupo del Proyecto: Golden Health

Nombre del Proyecto: Nutri-Analytics

Organización en GitHub: GoldentHealth

<https://github.com/GoldHealth>

Repositorio en GitHub: Nutri-Analytics

<https://github.com/orgs/GoldHealth/repositories>

Parte 1: Repositorio del Proyecto

- **¿Cuál es el URL/Link/Enlace del Proyecto?**

R: <https://github.com/orgs/GoldHealth/repositories>

- **¿Cuál es su usuario de github?**

R: **User:** valeriaraujo03

Link: <https://github.com/valeriaaraujo03>

Parte 2: Fuente de Datos

- **¿Cuál es el formato de la fuente de datos (CSV/JSON/SQL/etc)?**

R: El formato que utilizaremos para nuestra fuente de datos es .CSV

- **¿Cuántos Registros/Líneas de información se esperan?**

Se espera tener más de 30 datos para poder decir que tenemos una buena muestra de la población. Se obtuvieron finalmente 84 encuestas, con opción a que continúen aumentando las mismas con el pasar del tiempo.

- **¿Cuáles son las propiedades/atributos de la fuente de datos?**

R: Las propiedades/atributos son los siguientes:

Edad, Estatura, Sexo, Año

- En edad, se tomaron 3 rangos distintos, encapsulamos en Menos de 18 años, 18 a 22 años y más de 22 años, esto nos permitirá establecer si existe alguna relación respecto a la edad y el IMC.
- En estatura, se toma la medida en metros (m), para utilizarlo posteriormente con la formula del IMC.
- Sexo, va de la mano del IMC, pues tambien es un parametro indicador del estado corporal.
- Año, se toma el año para buscar una relación entre el IMC, y el año que se encuentra cursando el estudiante.

Como se mencionó, hasta el momento se han logrado coleccionar 84 registros de encuestas, con opción a que continúe creciendo hasta el punto tal del inicio del análisis de datos, la primera tanda de los mismos se encuentra en el siguiente subdirectorio, parte del repositorio del proyecto, enlace: <https://github.com/GoldHealth/Nutri-Analytics/tree/main/data>

- **¿Cuál es el enlace o directorio en el que se almacenó la fuente de datos en GitHub?**

R: <https://github.com/GoldHealth/Nutri-Analytics/tree/main/data> como aun nos encontramos en proceso de recolección, hemos subido una primera parte que incluye 84 elementos de nuestra muestra.

El link de la encuesta aún activa es el siguiente: <https://forms.gle/ZJAKQo5J5HXdr1h27>

- **¿Qué análisis estadísticos u operaciones se pueden aplicar a esta fuente de datos? ¿Cuál es el propósito?**

Se planea utilizar la estadística descriptiva, así mismo la media de los datos, y utilizar la correlación lineal con el propósito de observar gráficamente la relación entre el IMC, edad y el año que se cursa.