



Tendencias e Innovación en Tecnología Agrícola- TEA (CG2335-223E)

Laboratorio y Problemas 08 Proyecto Jupyter Notebooks, Pandas, Matplotlib

Estudiante: Valeria Araujo 24277

Grupo: Gold Health

Proyecto: NUTRI- ANALYTICS

Identificador de Github: valeriaaraujo03

Fecha: 4 de noviembre del 2022.

Parte 1: Problema – Fuente de Datos y Manipulación en DataFrames

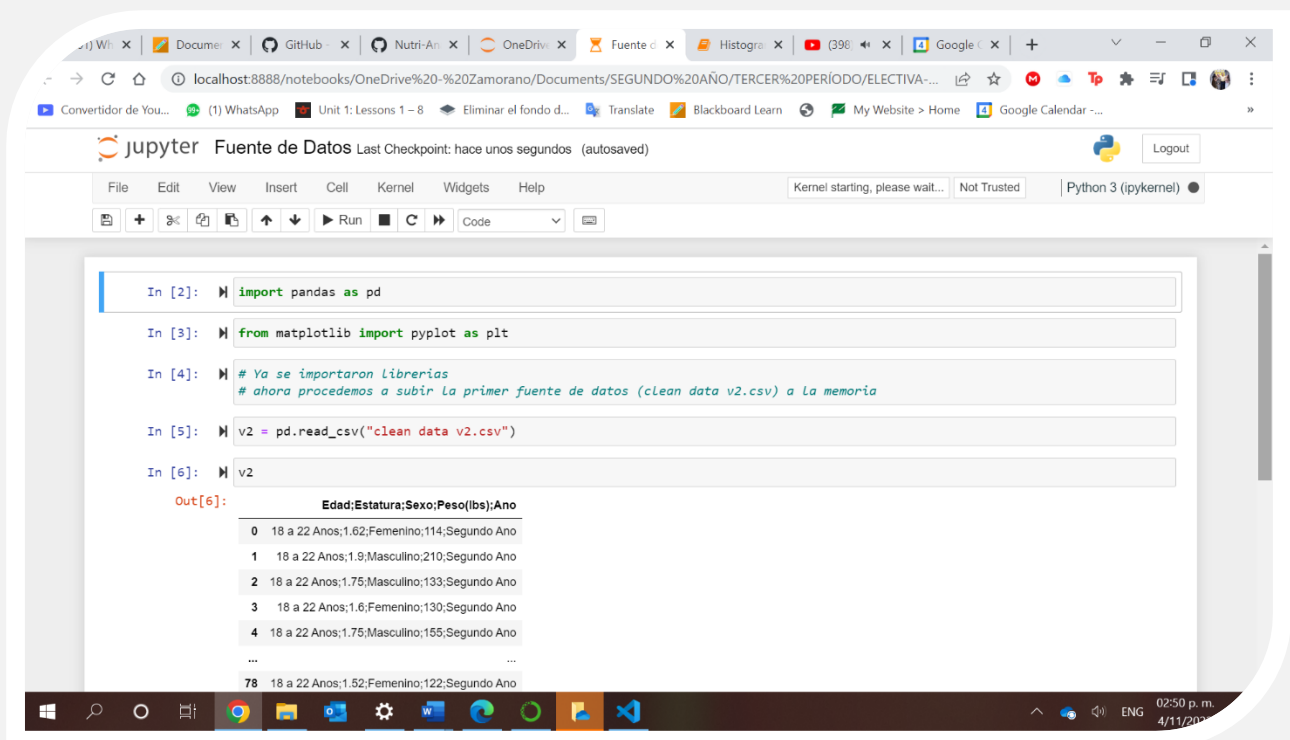
1. Siga las instrucciones del documento de laboratorio y problemas 08.

2. ¿Cuántos registros/líneas tiene su fuente de datos?

Nuestra fuente de datos tiene 259 registros, es decir, se obtuvieron los datos de 259 estudiantes tanto de primero, segundo y tercer año. Hubo un error del 5% y un intervalo de confianza del 95%, que se obtuvo a través de un programa estadístico.

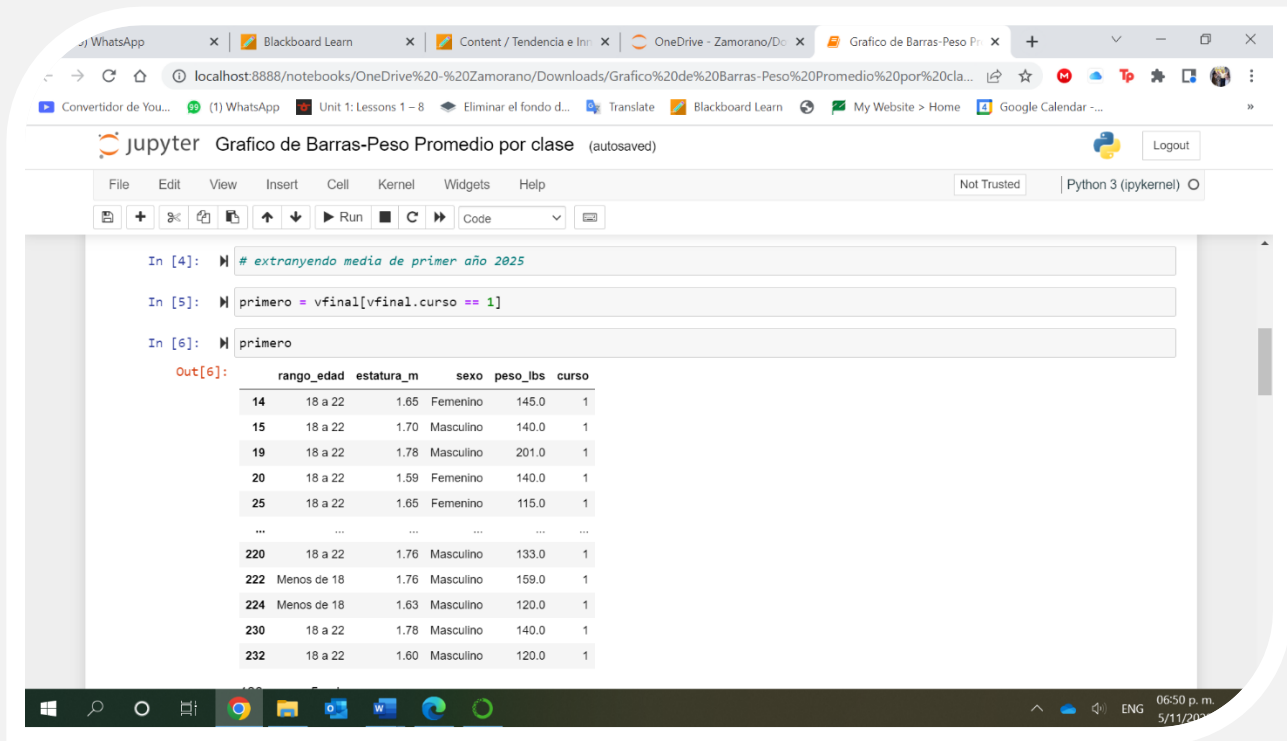
3. Fuente de Datos del Proyecto (Subiendo la fuente de datos a memoria utilizando Pandas)

Incluya la captura de pantalla acá. La captura de pantalla debe mostrar el código, el contenido del DataFrame, y el Jupyter Notebook corriendo en Jupyter (<https://jupyter.org/try-jupyter/lab/>).



4. Filtrando la información en un DataFrame (similar a lo hecho en clase en la cuál filtrábamos los países y creábamos otro DataFrame para Guatemala y Honduras). Utilizando el DataFrame subido en la sección anterior, muestre otro DataFrame que contenga valores filtrados que muestren la preparación de la información para ser analizada y graficada.

Incluya la captura de pantalla acá. La captura de pantalla debe mostrar el código, el contenido del DataFrame, y el Jupyter Notebook corriendo en Jupyter (<https://jupyter.org/try-jupyter/lab/>).



5. ¿Qué columna del DataFrame servirá como el eje X?

Eje "x" será el año que cursan: Aunque esto va a depender del total de gráficas que vamos a realizar, en total, se van a presentar 7 gráficas, siendo 1 histograma por cada año, un gráfico de regresión lineal y de correlación lineal simple Peso-IMC y Peso-Estatura, además 1 gráfico de líneas con el efecto del género y 1 gráfica de barras. Por lo tanto el que se muestra, es el ejemplo que se presentará para este laboratorio.

6. ¿Qué columna del DataFrame servirá como el eje Y?

Eje "y" será el peso promedio: Aunque esto va a depender del total de gráficas que vamos a realizar, en total, se van a presentar 7 gráficas, siendo 1 histograma por cada año, un gráfico de regresión lineal y de correlación lineal simple Peso-IMC y Peso-Estatura, además 1 gráfico de líneas con el efecto del género y 1 gráfica de barras. Por lo tanto el que se muestra, es el ejemplo que se presentará para este laboratorio.

Parte 2: Problema – Graficando los Datos con Matplotlib, Análisis de Datos, Reporte Final, Presentación (Valor: 2 puntos)

1. **¿Revisó la documentación de Matplotlib, Pandas, y los videos de las clases que muestran cómo graficar utilizando un DataFrame y Matplotlib?**

Sí/No

2. **Si la respuesta anterior es No, ¿Por qué?**

3. **¿Qué análisis se está efectuando sobre los datos?**

El análisis que se está efectuando es la creación de gráficos sobre el índice de masa corporal de los estuantes, para luego efectuar la correlación de los datos de estatura con el peso, y así, poder determinar el estado nutricional en que se encuentran los estudiantes de la Universidad Zamorano.

Estamos utilizando múltiples gráficos, específicamente para análisis estadísticos, tenemos prueba de regresión, correlacion (r), determinacion de coeficiente de correlacion (r^2) lineal, graficos de barras comparativos, histogramas para observar el comportamiento normal de los datos, y por ultimo grafico de lineas para ver el comportamiento en general de los datos.

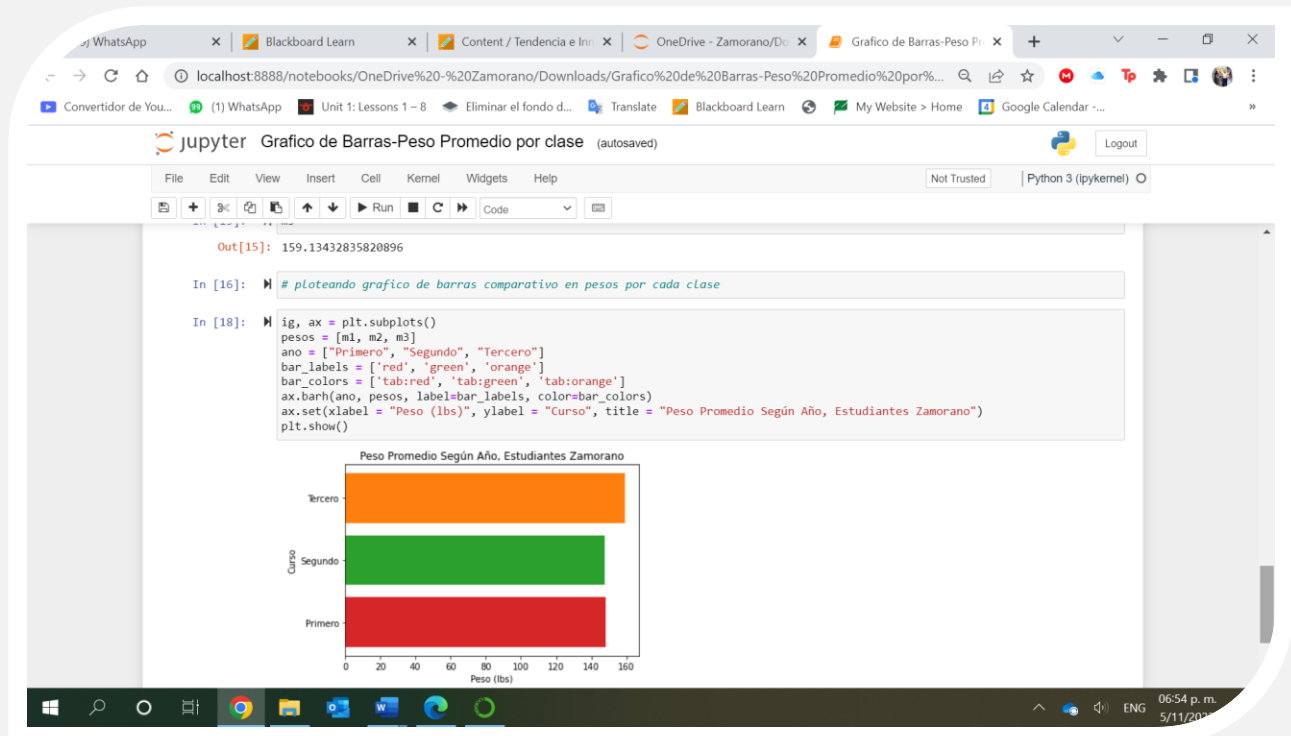
4. **Utilizando el DataFrame que preparó para graficar la Sección anterior, ¿Qué tipo de gráfica están utilizando para el análisis de los datos y por qué?**

El tipo de gráficas que estamos utilizando para el análisis de los datos obetenidos, son los gráficos de barras, debido a que estos ayudan a comparar dos o más variables de una forma clase de entender. Que en el caso de nuestra investigación, se comparan el peso respecto al año.

Para el dataframe presentado, específicamente el de la captura páginas arriba, es usado para observar los pesos, de manera comparativa, entre los tres años.

5. **Analizando la Fuente de Datos del Proyecto (utilizando Matplotlib). Muestre la gráfica del DataFrame preparado anteriormente.**

Incluya la captura de pantalla acá. La captura de pantalla debe mostrar el código, el contenido del DataFrame, la gráfica generada desde Matplotlib, el Jupyter Notebook corriendo en Jupyter (<https://jupyter.org/try-jupyter/lab/>).



El análisis efectuado en este laboratorio debe servir para preparar la documentación del reporte final.

- Con sus compañeros de grupo, creen la documentación del reporte final utilizando los resultados encontrados en Jupyter Notebooks. Expliquen cada una de las gráficas generadas y los DataFrames. Explique el Problema, la solución, la parte técnica, el código fuente, e incluyan referencias.

Incluyan el borrador de la documentación como parte de la entrega en Blackboard.

https://alumnizamorano-my.sharepoint.com/:w:/g/personal/melissa_gonzalez_est_zamorano_edu/EQ-KGktrHB5BmAN4K4MkyLgBDAjrvCB0MfUb5vozNhdx6g?e=jMnKep

URL PARA EL BORRADOR DEL DOCUMENTO, SUJETO A MODIFICACIONES PARA EL REPORTE FINAL.

El análisis efectuado en este laboratorio debe servir para preparar la documentación de la Presentación final.

- Con sus compañeros de grupo, creen la presentación del proyecto final (PowerPoint, KeyNote, Google Slides) utilizando los resultados encontrados en Jupyter Notebooks.

Expliquen cada una de las gráficas generadas y los DataFrames. Explique el Problema, la solución, la parte técnica, el código fuente, e incluyan referencias.

Incluyan el borrador de la documentación cómo parte de la entrega en Blackboard.

https://www.canva.com/design/DAFRAzkmees/tyn74m7V7613lzIAPfGVKg/edit?utm_content=DAFRAzkmees&utm_campaign=designshare&utm_medium=link2&utm_source=s harebutton

URL PARA EL BORRADOR DEL DOCUMENTO, SUJETO A MODIFICACIONES PARA LA PRESENTACIÓN FINAL.

Parte 3: Laboratorio

¿Revisó, estudió, y analizó los videos de la clase y la información adicional de Pandas y Matplotlib?

Sí/No

(Marque su respuesta)

¿Cuál es el url/link/enlace del directorio creado en Github? ¿A dónde se subieron los archivos de este laboratorio?

<https://github.com/GoldHealth/Nutri-Analytics.git>

<https://github.com/GoldHealth/Nutri-Analytics/tree/main/plot%20drafts>

TODOS deben subir este documento a Blackboard. En adición, subir el documento del reporte final y la presentación.