**UNIVERSIDAD DE PIURA**

FACULTAD DE INGENIERÍA



**“Análisis de datos con Python”**

**PRESENTAN:**

Arizola Rangel, Luis David

Bussalleu Vicente, Gabriel Enrique

Ludeña Campos, Johan Manuel

Medina Coronado, Fernando Jose

Perales Yuyes, Ismael Alejandro

Sosa Silupú, Milagros Fiorella

**DOCENTE:**

Ing. Pedro Rotta

Piura, 23 de enero de 2022

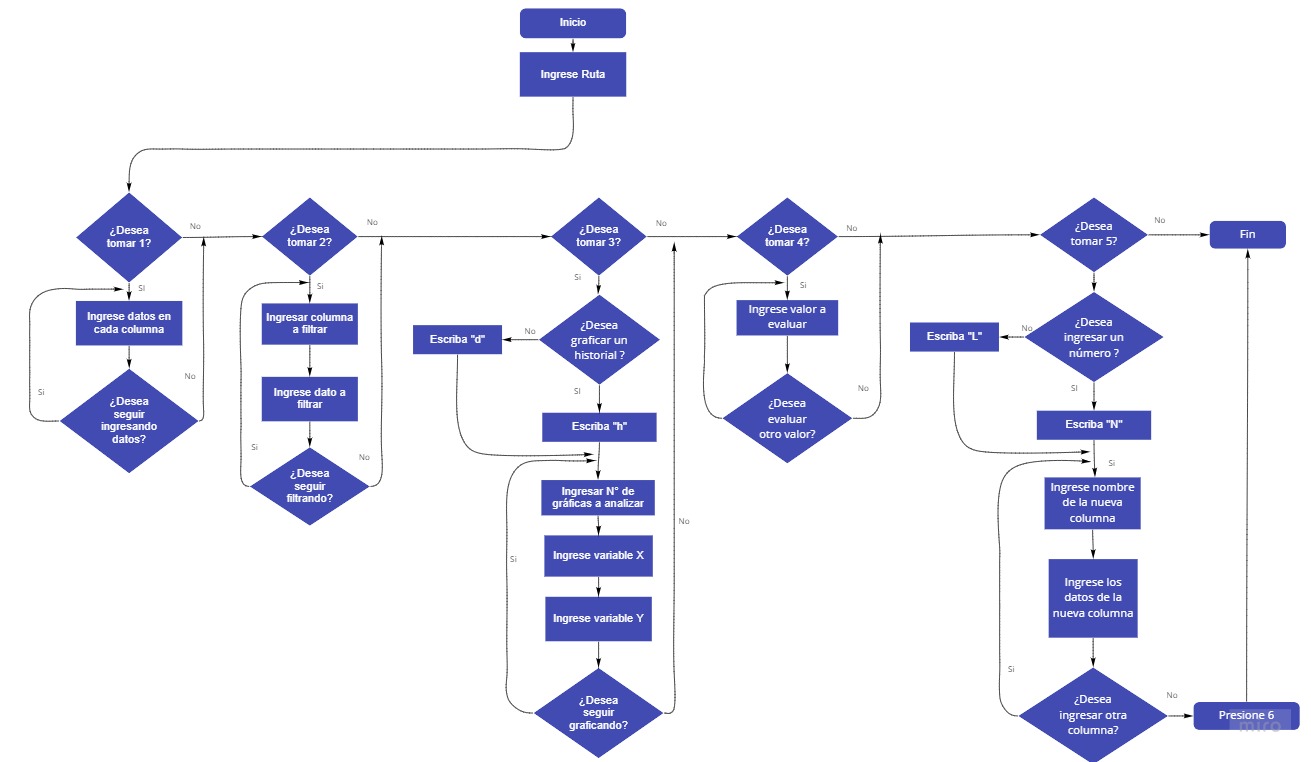
**Análisis de datos con Python-Trabajo Tipo 1**

1. **Introducción:**

Hoy en día donde los recursos tecnológicos se han ido desarrollándose a una velocidad inimaginable, resulta de gran utilidad el poder implementar sistemas que logren solucionar problemas de forma automática, lo que da pie a hablar sobre la automatización de sistemas. Actualmente se ve en el país situaciones en donde es necesaria una solución de automatización como por ejemplo en la energía eléctrica, en donde, anteriormente se destinaba personal para realizar la medición del consumo mensual y hoy en día se cuenta con un sistema interconectado mediante antenas que muestra el consumo detallado en tiempo real de las tarifas eléctricas para evaluar consumos en baja, media y alta tensión y como varia el tipo de costo.

Otro caso es la refrigeración de alimentos para la industria de la agroexportación, en donde la automatización se ve reflejada en el ahorro de energía, además cubre de una mejor manera las exigencias del mercado respecto a la refrigeración de estos productos.

En el presente informe se analizará las **variables:** tiempo, la demanda eléctrica ejecutada, la demanda programada diaria y semanal.

1. Análisis del sistema:

**Ilustración 1. Diagrama de flujo del funcionamiento del programa**

1. *Inicio:*

Como paso inicial se importan las librerías: Pandas, numpy y matplotlib.pyplot. Seguidamente se define la función Columns que me permitirá visualizar la lista de columnas.

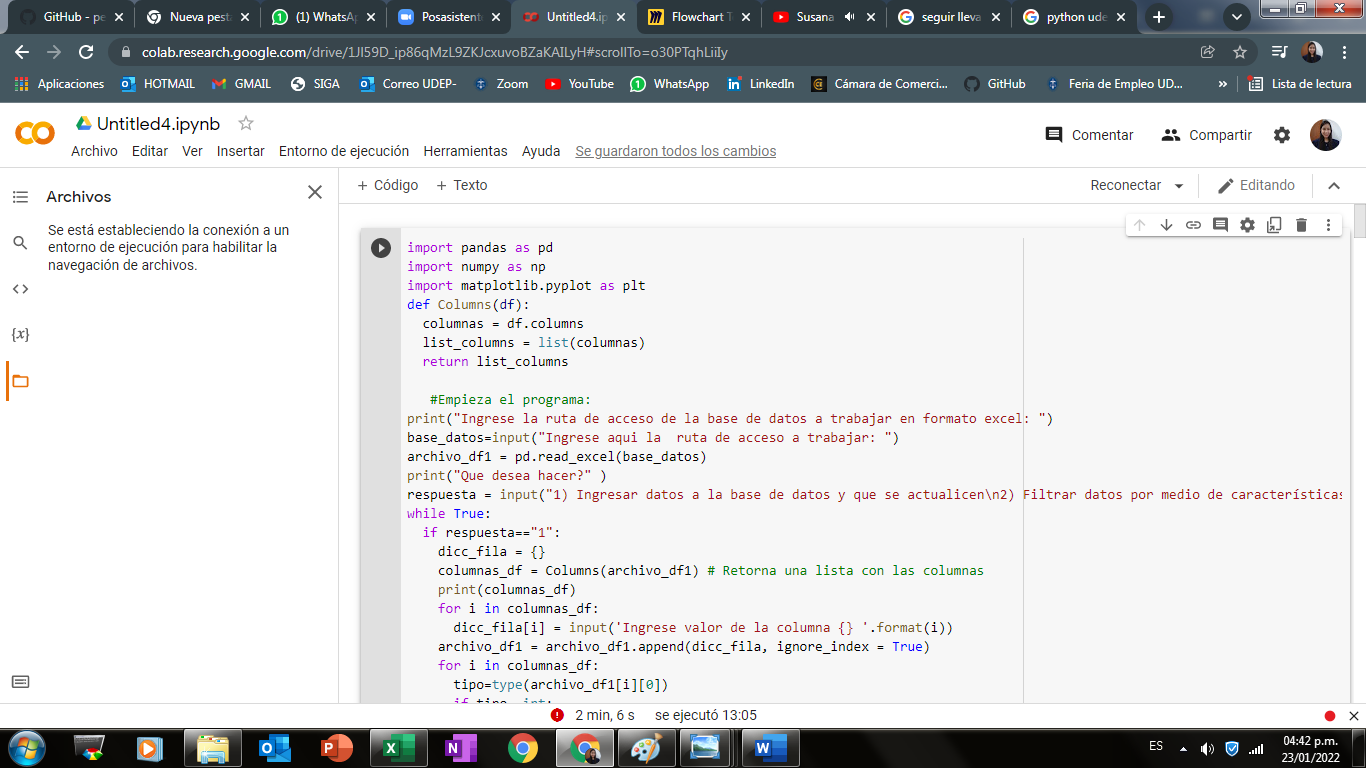


Ilustración 2. Importar librerías y definir una función

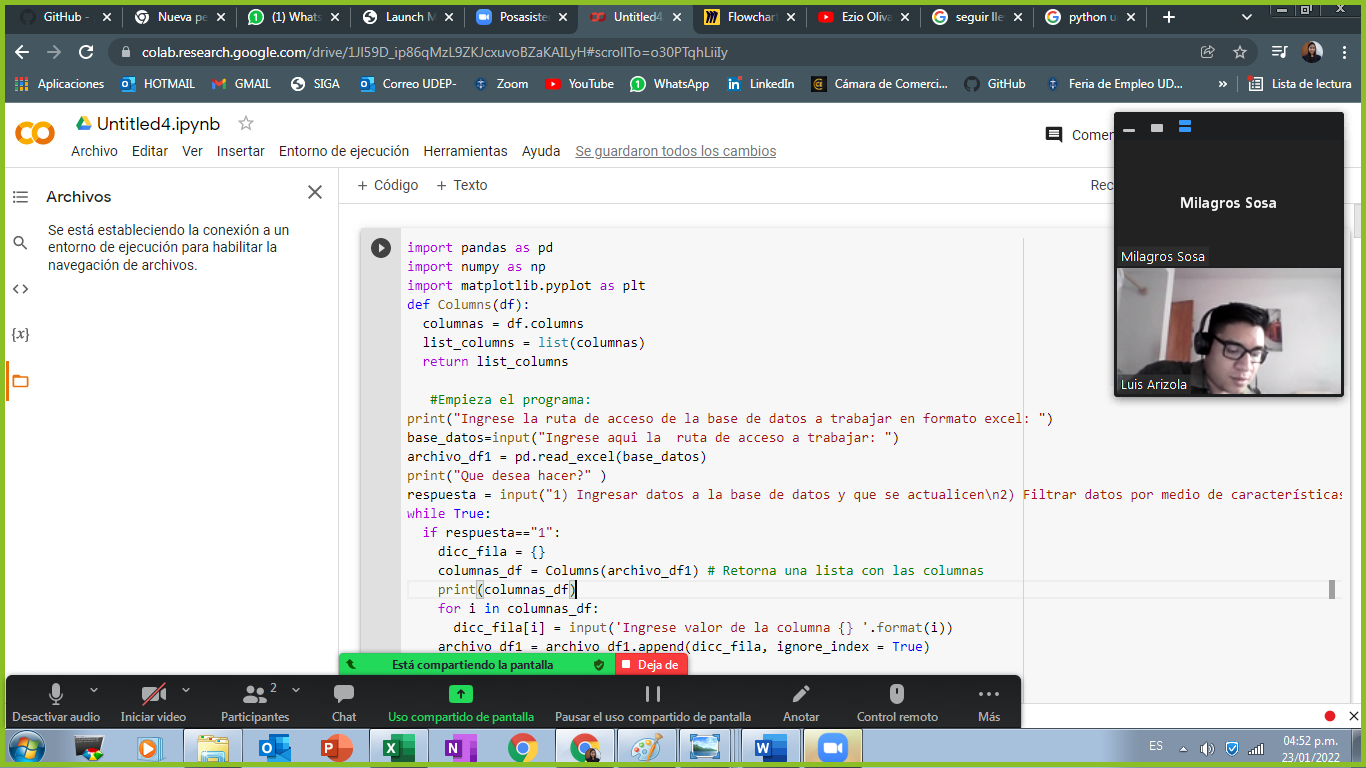
Como segundo paso el programa pedirá al usuario ingresar la ruta de acceso a trabajar (archivo Excel), a continuación se presentan una lista de opciones a realizar dentro del programa:

Ilustración 3. Ingreso de decisión y ingreso de ruta del archivo Excel

1. *Ingresar datos a la base de datos y que se actualicen.*

Si se ha escogido la opción 1, el programa entrará dentro de un bucle While el cual se encargará de añadir valores a cada columna definida en el archivo Excel. Los valores ingresados por el usuario son cadena de texto, por ello se convierten mediante el condicional if a enteros (int)

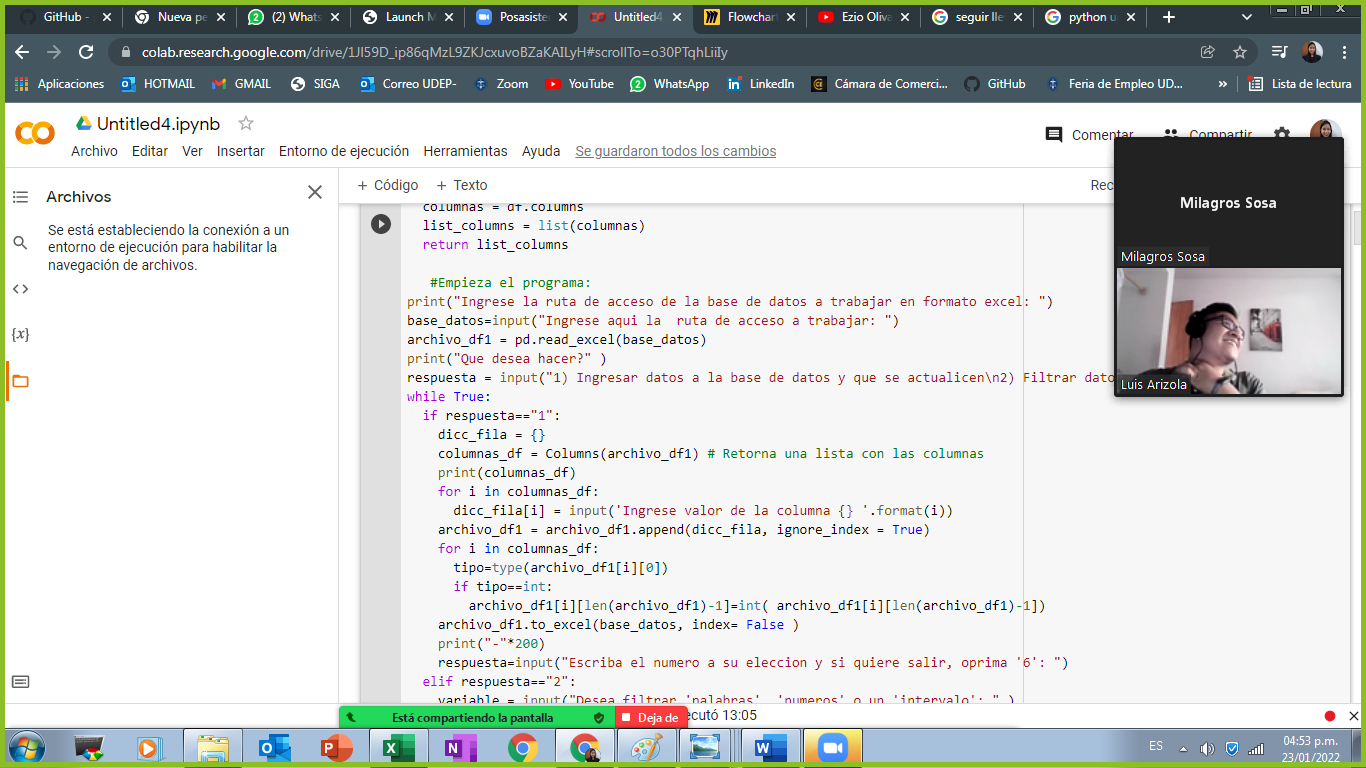


Ilustración 4. Código de la opción 1

1. *Filtrar datos por medio de características:*

Si se elige la opción 2, esta cumple con la función de filtrar datos del archivo EXCEL tales como, palabras, números o intervalos. Se le pide al usuario que desea filtrar (palabras, números o intervalos), luego se escribe la columna que se desea filtrar y seguidamente el dato de dicha columna a filtrar. Si la columna a filtrar son números se debe establecer un float para poder convertirlos a números racionales ya que en Excel están como cadena de texto. En el caso de querer ingresar intervalos hay 3 opciones, si se escribe R se ingresará un Rango, donde seguidamente le usuario deberá escribir un valor mínimo y máximo.

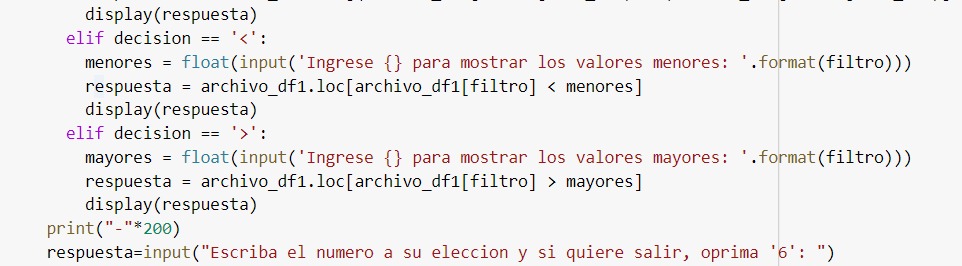
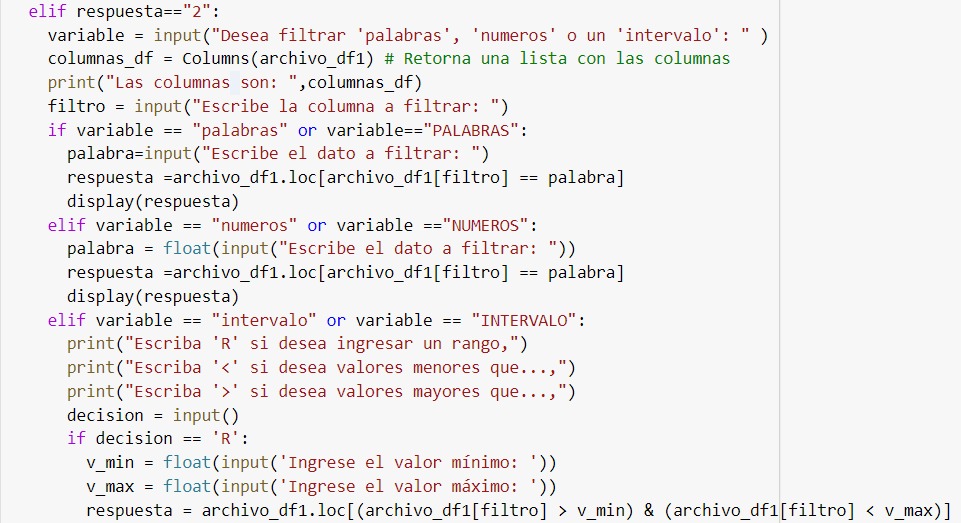


Ilustración 5. Código de la opción 2

1. *Solicitar gráficos de análisis entre algunas variables.*

Si la respuesta es 3, primero el usuario tendrá que elegir qué tipo de gráfico desea, si es de dispersión tendrá que ingresar la letra d y si es de histograma deberá ingresar la letra h. Seguidamente se deberá ingresar la cantidad de gráficos que se van a analizar, para ello se ha establecido un contador inicializado en cero, con el objetivo de graficar el número de gráficos solicitados, posteriormente si la opción elegida es “d” se ingresa con que variables del Excel se desea trabajar tanto para el eje “X” y el eje “Y” y en el caso que la opción sea “h” se elige una sola variable a analizar bajo un histograma el cual le presenta con qué frecuencia se dan los valores de una determinada variable.

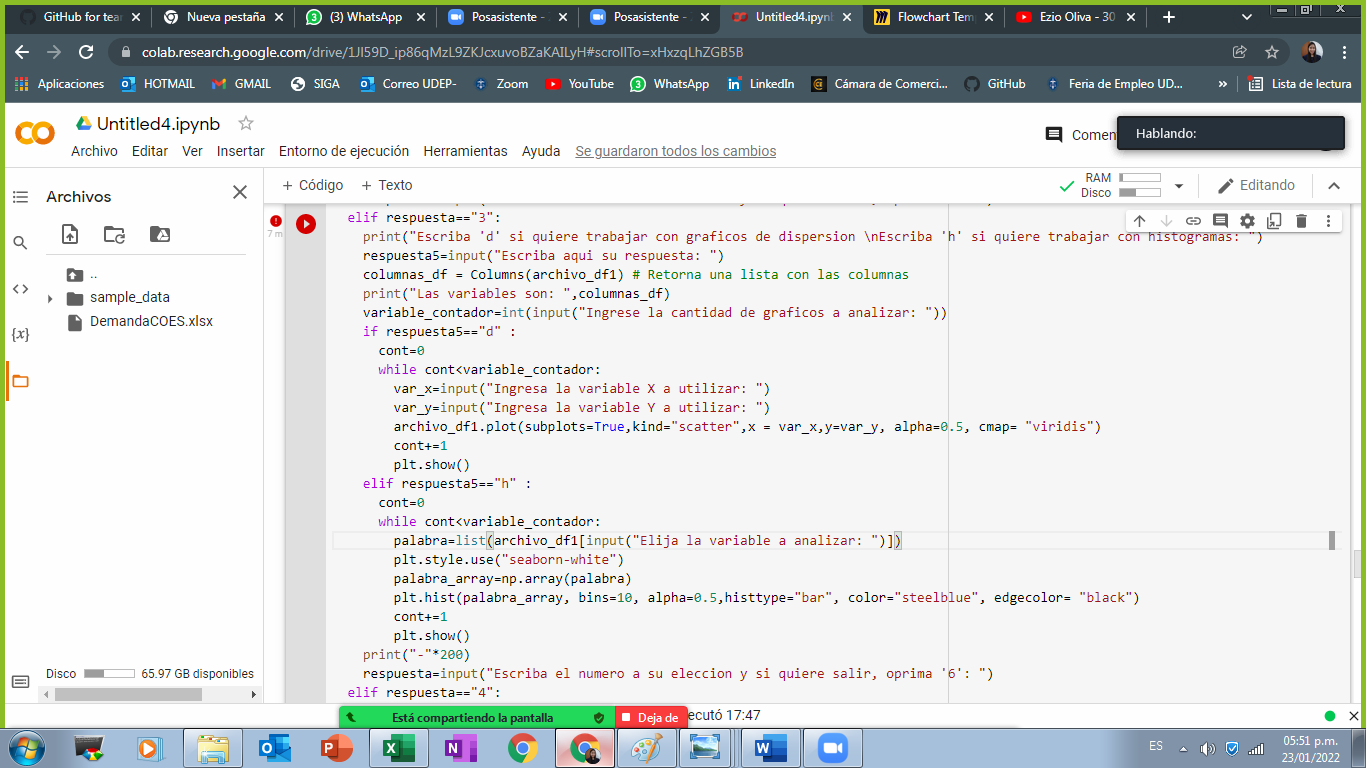


Ilustración 6. Código de la opción 3

1. *Solicitar datos estadísticos de los datos.*

Si se elige la opción 4, el usuario debe ingresar la variable que desea evaluar estadísticamente y el programa le devolverá una tabla con dichos datos estadísticos.

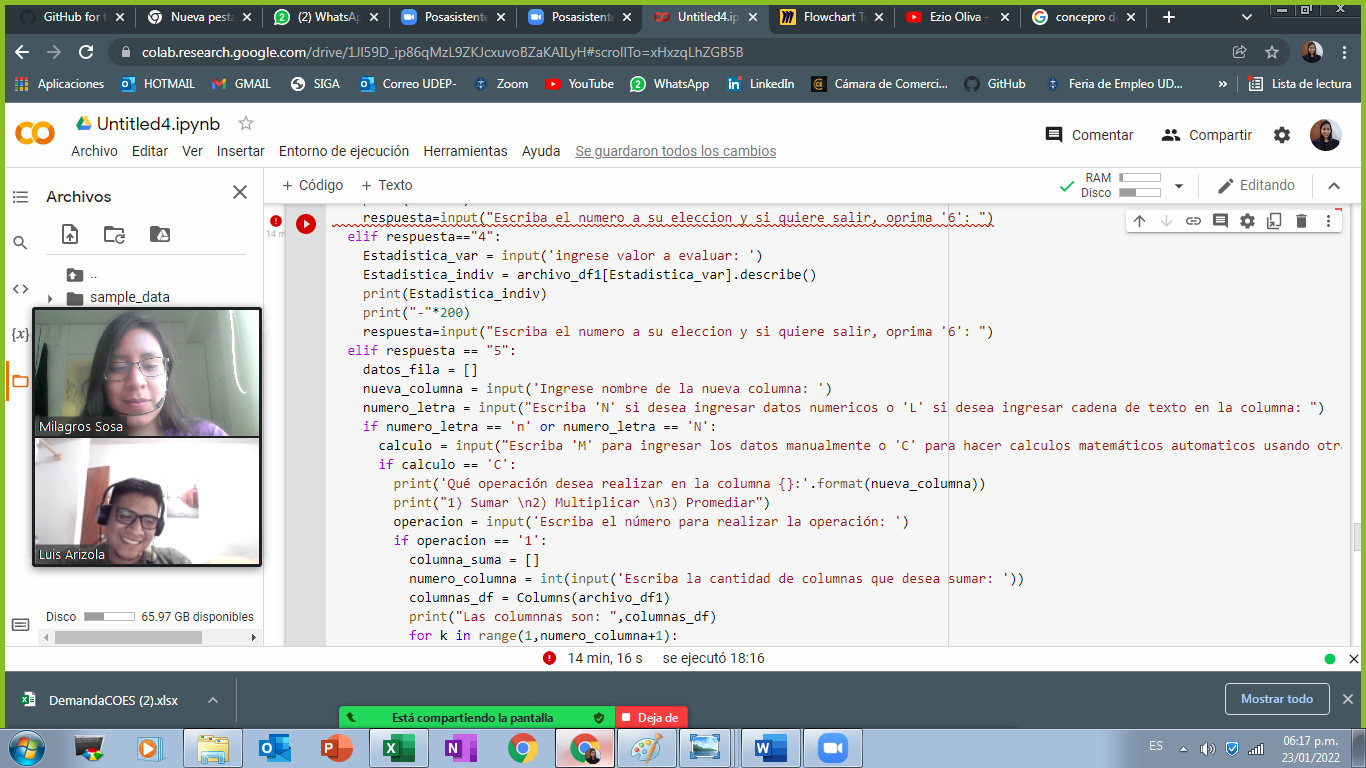


Ilustración 7. Código de la opción 4

1. *Generar nuevas características para los datos*

Si se elige la opción 5, el programa le solicita al usuario el añadir una nueva columna. Posteriormente, si es un dato numérico se ingresa “N” y si es una cadena de texto se ingresa “L”. Para hacer cálculos matemáticos entre las columnas, se coloca la opción “C”, de no ser así, se coloca la opción “M”, esto para ingresar los datos manualmente.

Si se ha colocado la opción C, se debe ingresar la operación a realizar (1 para resta, 2 para multiplicación y 3 para cálculo de promedio). Si (if) se decide por la opción 1, se ha establecido una lista vacía, seguidamente, el usuario ingresará el número de columnas que desea sumar. Para el caso de las operaciones 2 y 3 se realiza el mismo proceso con sus respectivas operaciones matemáticas.

En el caso de que la opción sea “M”, se agregará manualmente los datos de la columna.

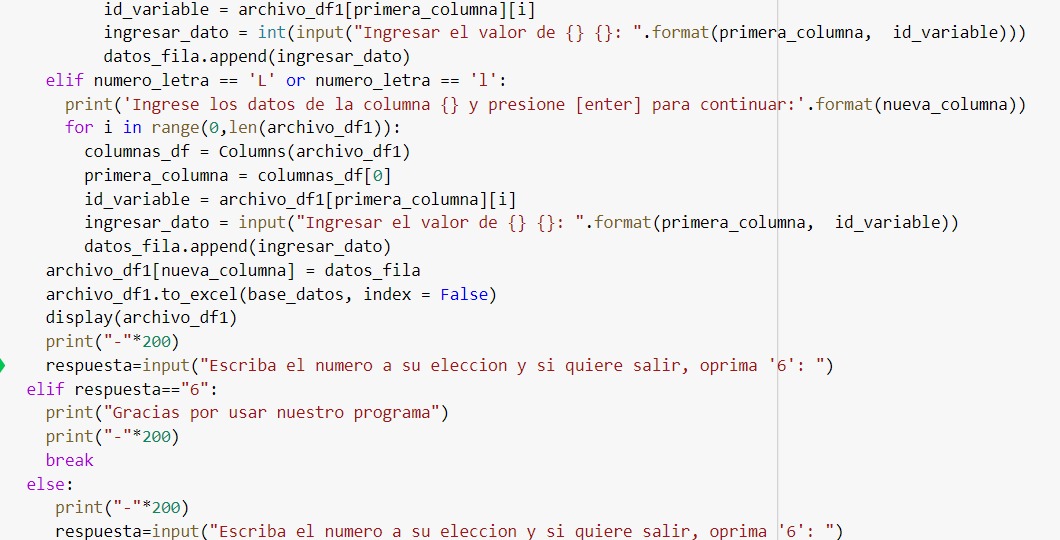
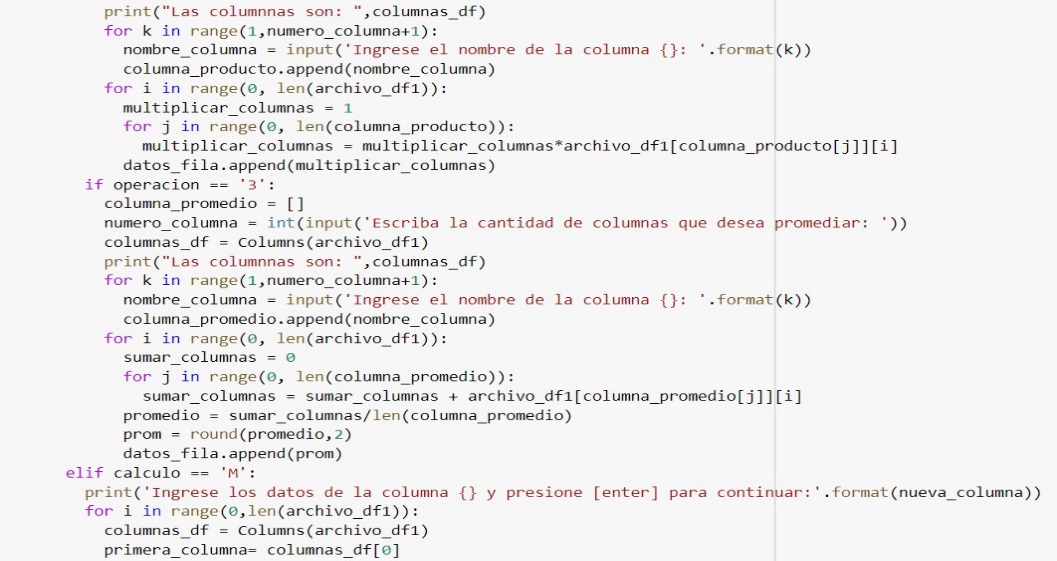
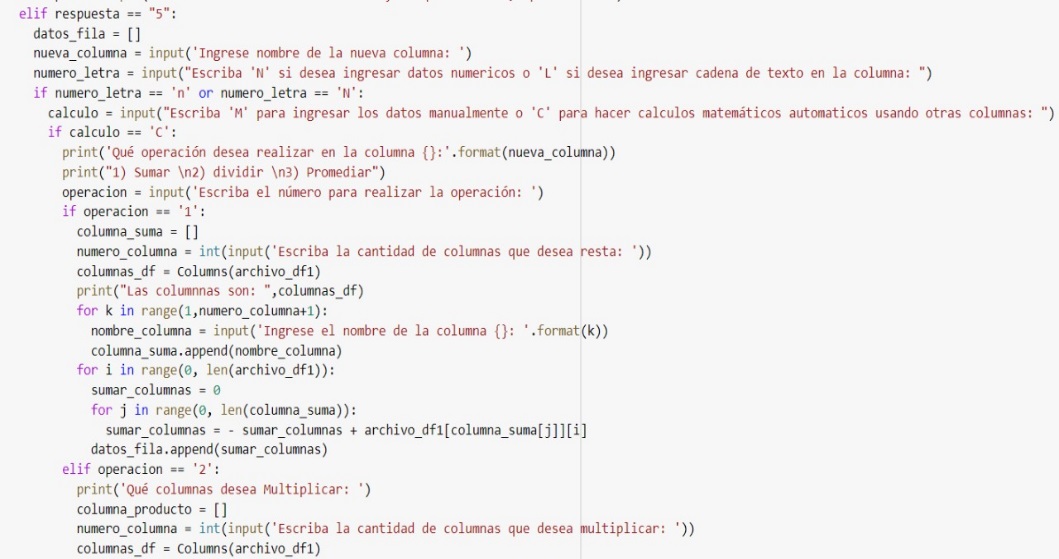
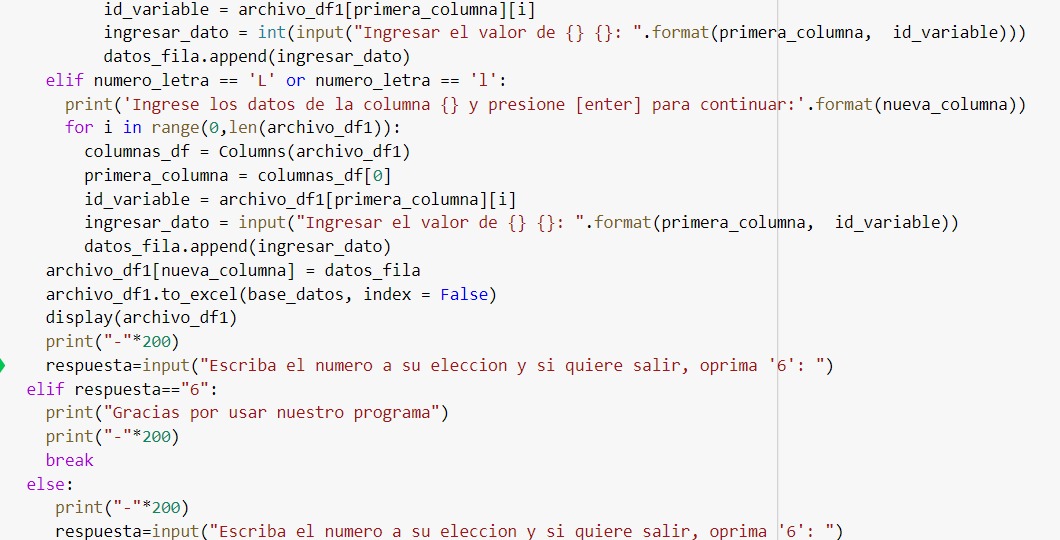


Ilustración 8. Código de la opción 5

1. *Salir del programa.*

Para salir del programa, se debe optar por la opción 6, el cual mostrará el siguiente mensaje: "Gracias por usar nuestro programa".



1. **Ejemplos al correr el programa:**

El programa se puede correr con cualquier archivo Excel, en este caso las variables que analizaremos para el ejemplo son:

1) Tiempo

2) Ejecutado (demanda ejecutada)

3) Prog. Diaria

4) Prog Semanal

* 1. *Introducción de datos:*

Ejemplo de ingresar datos a el archivo Excel adjunto, en este caso se copio la ruta del archivo Excel a utilizar donde había 4 variables de decisión: 1) Tiempo, 2) Ejecutado (demanda ejecutada), 3) Prog. Diaria, 4) Prog Semanal y en cada una de estas columnas se han ido ingresando datos para su posterior análisis.

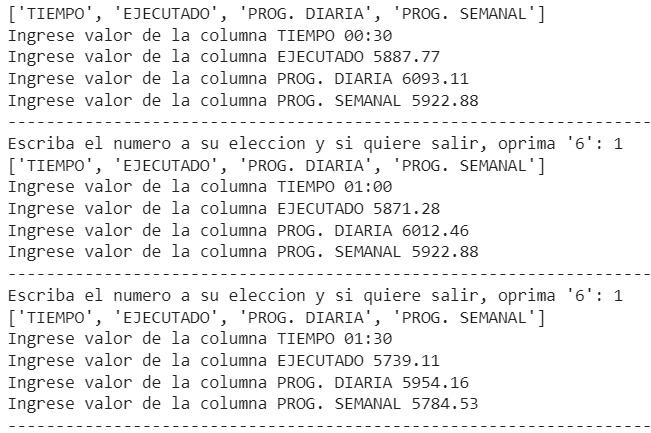
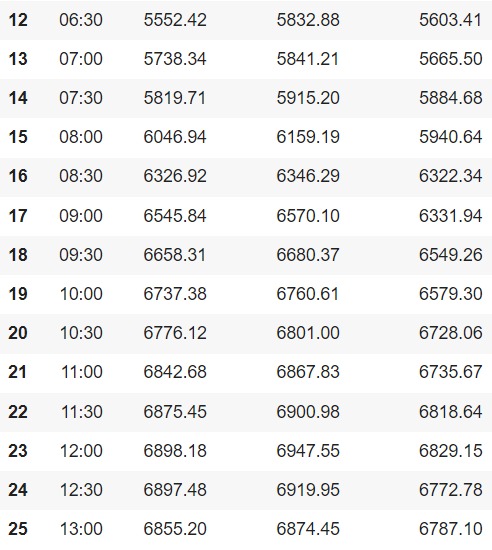


Ilustración 9. Ejemplo de la opción 1

* 1. *Filtrar datos*

Se ha filtrado valores mayores a 6272.16MW para poder tener una tabla de los valores de mayor demanda de energía eléctrica, analizando dicha tabla se obtiene que la mayor demanda ejecutada se da a las 20:00 con el valor de 6931.55 MW

Ilustración 10. Ejemplo aplicativo opción 2



* 1. *Solicitar gráficos de análisis*

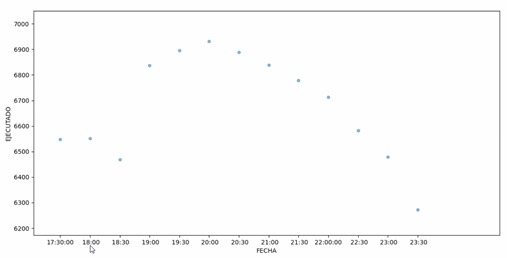
Para este caso se elige por la opción 3, a continuación, el programa solicita el tipo de gráfico que se desea realizar, por lo que se elige por un gráfico de dispersión. Luego en el eje X irá la columna fecha (o Tiempo), mientras que para el eje Y se elige la columna de “EJECUTADO”. Con esta gráfica será posible observar cómo es el comportamiento en la zona de máxima demanda de electricidad que se da entre las 18:00 y las 23:300. Cabe aclarar que estos datos son correspondientes al día 22/01/2022, sacados la página oficial del COES.

Ilustración 11. Ejemplo de aplicación de la opción 3

* 1. *Solicitar datos estadísticos de los datos.*

Para este caso se ha realizado el ejemplo seleccionando la opción 4. Aquí como primer paso se solicita la variable a evaluar, la cual será la columna de “EJECUTADO”.

El valor promedio de la demanda del día 22 de enero de este año fue de 6371.79 MW, con lo cual se puede concluir el valor máximo de la demanda diaria está en un poco más del 90% del promedio.

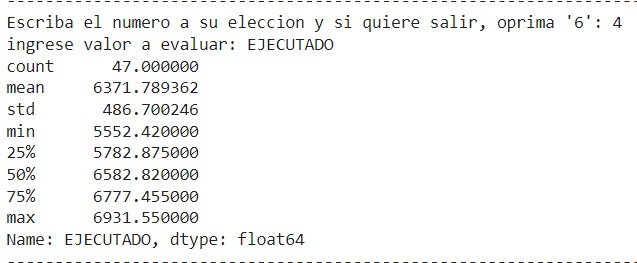


Ilustración 12. Ejemplo de aplicación de la opción 4

* 1. *Ingresar nueva columna:*

Para el caso de seleccionar la opción 5, el programa le solicitará al usuario escribir como es que se nombrará a la nueva columna que se agregará que para este caso se llamará “DESVIACION”. Posteriormente elegimos la opción N para operar con datos numéricos, a continuación, seleccionamos “C” para realizar operaciones matemáticas entre columnas aquí se selecciona la opción 1 de resta. Se elige las columnas de “EJECUTADO” y “PROG. DIARIA”. Lo que representa esta nueva columna es que tanto se aleja la demanda ejecutada, la cual es la real, de la programación diaria que es lo que se estima de la demanda en el día.

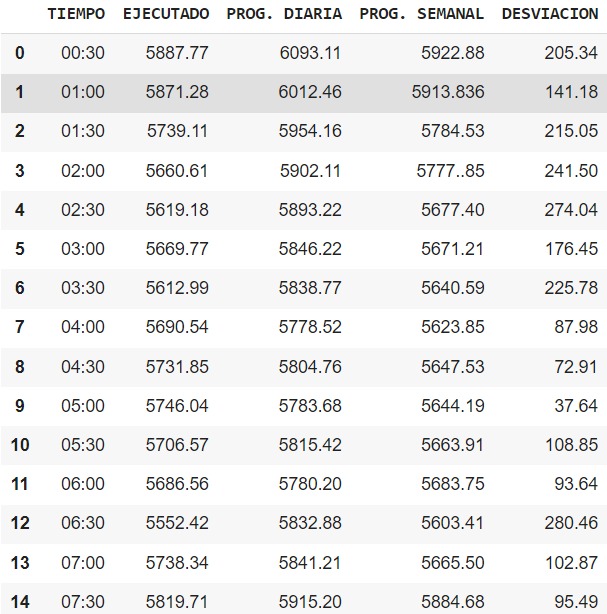
**

Ilustración 13. Ejemplo de aplicación de la opción 5

1. **Conclusiones**

* Del presente proyecto se concluye que para el ejemplo n°1 se observa mediante la tabla que se obtuvo que la demanda máxima de energía eléctrica se da a las 20:00, esto se debe a que alrededor de esta hora la mayoría de las personas usan el fluido eléctrico en iluminación nocturna, publica; artefactos tales como televisores, computadoras. Por lo que tiene sentido que la máxima demanda se de en horas de la noche.
* El empleo del programa permitió optimizar el manejo de gran cantidad de datos y poder analizar cada variable independientemente y así poder sacar conclusiones.
* Dentro del curso de análisis de datos con Python se concluye que resulta de gran utilidad dicho software ya que mejora los sistemas y los automatiza. Además que se puede dar mejores alternativas de solución que contribuyen al desarrollo de una asociación o grupo de trabajo.
* Del ejemplo de la opción 5, se afirma que de tener una mayor data agregada al Excel, se mejora la precisión en cuanto al grafico de dispersión y en consecuencia saber con mayor precisión cual es la tendencia que siguen los datos y posteriormente comenzar una predicción de la demanda para los días posteriores.