

Workclass 01

Departamento de Economía

Carlos Andres Yanes Guerra

Table of contents

1	Presentación	1
1.1	Objetivo	1
2	Propuesta General	2
2.0.1	Ayudas	2
2.0.2	Una forma de importar directamente los datos	3

1 Presentación

Este Workclass pretende en todo momento servir de línea de inicio con el ambiente de aprendizaje de **Python**. Para este primer trabajo le solicitamos responder lo que se le pide.

1.1 Objetivo

En el proceso de familiarización con las herramientas académicas, es esencial adquirir habilidades en el uso de **Google Colab** y comprender la estructura y funcionalidades de los documentos tipo **Markdown**. Además, se debe aplicar un análisis descriptivo, aprovechando los formatos y métricas discutidas en las lecciones.

En esta etapa, se presentará una base de datos de ejemplo denominada “DEMO”. Se sugiere inicialmente la creación de un archivo Excel donde los datos puedan ser trasladados para una primera interacción. A continuación, se recomienda importar temporalmente los datos al entorno de Colab, lo que permitirá iniciar el trabajo y la visualización del proceso completo.

Es importante destacar que este ejercicio no solo facilitará la comprensión práctica, sino también proporcionará una plataforma para aplicar los conceptos teóricos en un contexto real. Mediante la combinación de estas herramientas, se maximizará el aprendizaje y la capacidad para realizar análisis efectivos en el futuro.

2 Propuesta General

Sea la siguiente tabla con información:

Obs	Nombre	Edad	Ingresos	Gastos	Estrato
1	Carlos Sanchez	24	5694258	935000	5
2	Claudia Girrud	29	2753698	645000	3
3	Emilio Vesga	32	1548681	396900	3
4	Camila Sanchez	29	4220699	2230000	4
5	Amaury Suarez	24	3829716	3400000	4
6	Mauricio Silva	25	915761	278000	1
7	Juana Renat	27	4659060	1790000	2
8	Maria Stevez	21	2352287	1587000	3
9	Juan Parniot	23	1446200	989670	2
10	Milton Pípol	25	2676853	1597200	3
11	Matias Guerra	25	1289810	978000	2
12	Adriana Perez	29	2728612	1450000	3
13	Luisa Rodriguez	23	1079456	845200	2
14	Olga Cañon	27	2941315	156900	3
15	Carlos Velas	25	3347724	987300	3

1. Establezca primero con las variables (*Gasto* y *Edad*) un diagrama de dispersión, luego haga lo mismo pero ahora con las variables (*Gasto* e *Ingreso*) . **Realice un análisis comparativo con ambas gráficas.**

2. Establezca un análisis de los Ingresos por estrato socioeconómico. *Pista:* Puede ayudarse de **Pandas** para Python.

3. Cuál es la edad mas común dentro de las observaciones?

4. Realice un histograma con la variable *Ingreso*, establezca un pequeño análisis de eso y apóyese de los estadísticos de *curtosis* y de *asimetría*.

5. Cree una variable dicotoma o **dummy** para el genero de la persona y establezca una comparación entre el gasto por genero de las personas de la base. *Pista:* Debe responder finalmente que genero gasta mas.

2.0.1 Ayudas

Para el ambiente de trabajo no olvide

```

import pandas as pd

# Ruta del archivo
dir = 'ruta/al/archivo/tu_archivo.xlsx'

# Cargar los datos desde el archivo original
datos = pd.read_excel()

# Mirar los resultados
print(datos.head())

```

2.0.2 Una forma de importar directamente los datos

```

import pandas as pd
datos = {
    'Obs': [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15],
    'Nombre': ['Carlos Sanchez', 'Claudia Girrud', 'Emilio Vesga',
               'Camila Sanchez', 'Amaury Suarez', 'Mauricio Silva',
               'Juana Renat', 'Maria Stevez', 'Juan Parniot', 'Milton Pipol',
               'Matias Guerra', 'Adriana Perez', 'Luisa Rodriguez',
               'Olga Cañon', 'Carlos Velas'],
    'Edad': [24, 29, 32, 29, 24, 25, 27, 21, 23, 25, 25, 29, 23, 27, 25],
    'Ingresos': [5694258, 2753698, 1548681, 4220699, 3829716, 915761,
                 4659060, 2352287, 1446200, 2676853, 1289810, 2728612, 1079456,
                 2941315, 3347724],
    'Gastos': [935000, 645000, 396900, 2230000, 3400000, 278000,
               1790000, 1587000, 989670, 1597200, 978000, 1450000, 845200,
               156900, 987300],
    'Estrato': [5, 3, 3, 4, 4, 1, 2, 3, 2, 3, 2, 3, 2, 3, 3]
}
df = pd.DataFrame(datos)
# Mirar el DataFrame
print(df)

```