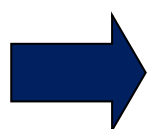


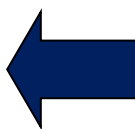
SERVICIO NACIONAL DE ADIESTRAMIENTO EN TRABAJO INDUSTRIAL

FORMACIÓN PROFESIONAL

CURSO DE PRÁCTICA INTENSIVA



CUADERNO DE INFORMES



FORMACIÓN PROFESIONALCFP/UCP/ESCUELA: _____ **SENATI**ESTUDIANTE: _____ **Jhon Erick Rodríguez Monja**ID: _____ **1535824** _____ BLOQUE: _____ **Dual, Tec.Ind, Adm, T Ing 2024-2**CARRERA: **Ingeniería de Software con Inteligencia Artificial**INSTRUCTOR: **Eduardo Alfonso Torres Luna**SEMESTRE: _____ **IV** _____ DEL: **23/09/24** AL: **26/09/24**.



INSTRUCCIONES PARA EL USO DEL CUADERNO DE INFORMES

1. PRESENTACIÓN.

El Cuaderno de Informes es un documento de auto control, en el cual el estudiante, registra diariamente, durante la semana, las tareas, operaciones que ejecuta en su aprendizaje, es un medio para desarrollar la Competencia de Redactar Informes.

2. INSTRUCCIONES PARA EL USO DEL CUADERNO DE INFORMES.

2.1 En la hoja de informe semanal, el estudiante registrará diariamente los trabajos que ejecuta, indicando el tiempo correspondiente. El día de asistencia registrará los contenidos que desarrolla. Al término de la semana totalizará las horas.

De las tareas ejecutadas durante la semana, el ESTUDIANTE seleccionará la tarea más significativa (1) y él hará una descripción del proceso de ejecución con esquemas, diagramas y dibujos correspondientes que aclaren dicho proceso.

2.2 Semanalmente, el Instructor revisará y calificará el Cuaderno de Informes haciendo las observaciones y recomendaciones que considere convenientes, en los aspectos relacionados a la elaboración de un Informe Técnico (letra normalizada, dibujo técnico, descripción de la tarea y su procedimiento, normas técnicas, seguridad, etc.

2.3 Escala de calificación vigesimal:

CUANTITATIVA	CUALITATIVA	CONDICIÓN
16,8 – 20,0	Excelente	Aprobado
13,7 – 16,7	Bueno	
10,5 – 13,6	Aceptable	
00 – 10,4	Deficiente	Desaprobado

INFORME SEMANAL

.....IV....SEMESTRE SEMANA N° 7

DÍA	TAREAS EFECTUADAS	HORAS
LUNES	<p><u>EXPLICACION DE ESTADISTICA BASICA I</u></p> <ul style="list-style-type: none">• Una disciplina matemática que permite recolectar, organizar y analizar datos según sea necesario. Esto puede incluir obtener resultados, comparar información y tomar decisiones más informadas, entre muchas otras aplicaciones.• Al emplear métodos, procedimientos y fórmulas, la estadística se considera la ciencia del análisis de datos.	5
JUEVES	<p><u>PRACTICA CON LIBRERÍA PANDAS Y NUMPY</u></p> <ul style="list-style-type: none">• Realizar operaciones sobre Estadística• Calcular la media, mediana y moda de un archivo csv	7
TOTAL		12

INFORME DE TAREA MÁS SIGNIFICATIVA

Tarea Significativa: Crear un análisis y visualización de los datos de población, en la que calculaste la edad promedio ponderada por departamento.

Descripción del Proceso:

Importación de Bibliotecas:

- Se importan las bibliotecas necesarias: pandas para el manejo de datos y matplotlib.pyplot para la visualización.

Carga de Datos:

- Se carga un archivo CSV que contiene datos de población. Se especifica la ruta del archivo, el encoding (UTF-8), y el separador (punto y coma).

Renombrar Columnas:

- Las columnas del DataFrame se renombran para facilitar su comprensión y manipulación.

Definición de Pesos por Edad:

- Se define un diccionario age_weights que asigna un peso a cada grupo de edad, lo que permitirá calcular la edad promedio ponderada.

Cálculo de Edad Promedio Ponderada:

- Se calcula la edad promedio ponderada utilizando los grupos de edad y sus respectivos pesos. Este cálculo se realiza para cada fila del DataFrame.

Agrupación de Datos:

- Los datos se agrupan por DEPARTAMENTO, y se suman las poblaciones de cada grupo de edad, así como se obtiene la edad promedio para cada departamento.

Conversión de Tipo:

- La columna de edad promedio se convierte a tipo entero para simplificar la visualización y el análisis.

Visualización:

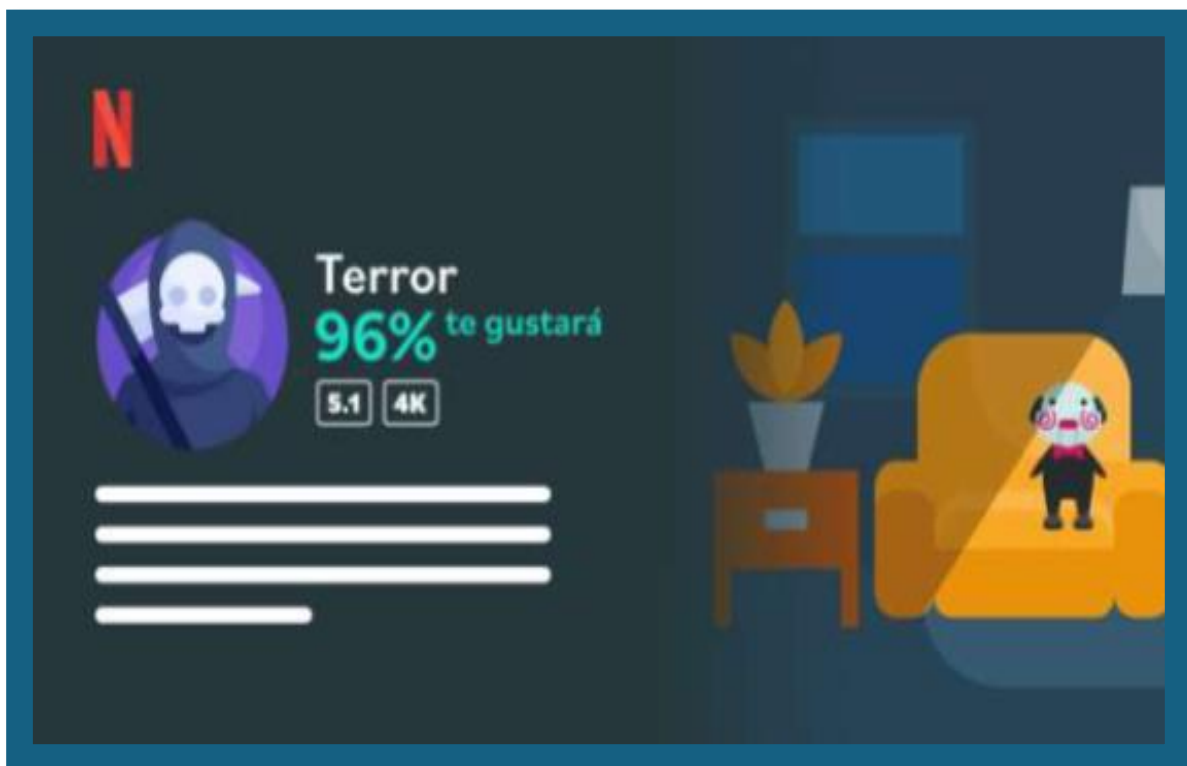
- Se crea un gráfico de barras que muestra la edad promedio por departamento. Se personaliza el gráfico con títulos, etiquetas y rotación de las etiquetas del eje x para mejorar la legibilidad.

LUNES:

Ejercicios planteados en clase:

ESTADISTICA BASICA I

La estadística es una disciplina matemática que permite recolectar, organizar y analizar datos según sea necesario. Esto puede incluir obtener resultados, comparar información y tomar decisiones más informadas, entre muchas otras aplicaciones. Al emplear métodos, procedimientos y fórmulas, la estadística se considera la ciencia del análisis de datos. Su principal objetivo es ayudar a comprender lo que ocurre en el entorno a partir de la información disponible.



Ventajas de la Estadística.

- Permite resumir grandes cantidades de información, realizar comparaciones y predecir resultados.
- Es muy útil para tomar decisiones acertadas y responder preguntas sobre cualquier objeto de estudio.
- Apoya la justificación de ideas o afirmaciones con pocos argumentos.
- Es una herramienta valiosa para obtener conclusiones objetivas investigaciones.

TIPOS DE ESTADISTICA

- **Estadística Descriptiva:** Ayuda a organizar grandes cantidades de datos utilizando métodos, tablas y gráficos para presentar los resultados de manera ordenada. En este curso, te mostramos algunas de estas herramientas.
- **Estadística Inferencial:** Se enfoca en realizar conclusiones y deducciones a partir de una muestra de datos. Es útil para tomar decisiones o determinar tendencias en un grupo de información.

Media aritmética.

- La media aritmética, también conocida simplemente como promedio, es una medida de tendencia central que se utiliza para representar un conjunto de datos
- mediante un solo valor. Este valor se obtiene sumando todos los datos y dividiendo el resultado entre el número total de datos.
- Fórmula
- Para un conjunto de datos x_1, x_2, \dots, x_n , la media aritmética se calcula de la siguiente manera:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{x_1 + x_2 + \dots + x_n}{n}$$

Donde:

- \bar{x} es la media aritmética.
- x_i representa cada uno de los valores del conjunto de datos.
- n es el número total de datos.

Ejemplo

- Supongamos que tenemos los siguientes datos: 4, 8, 6, 5, 3, 7.
- Sumamos todos los valores:
- $4 + 8 + 6 + 5 + 3 + 7 = 33$
- Dividimos la suma entre el número total de datos:

$$\bar{x} = \frac{33}{6} = 5.5$$

- Por lo tanto, la media aritmética de estos datos es 5.5

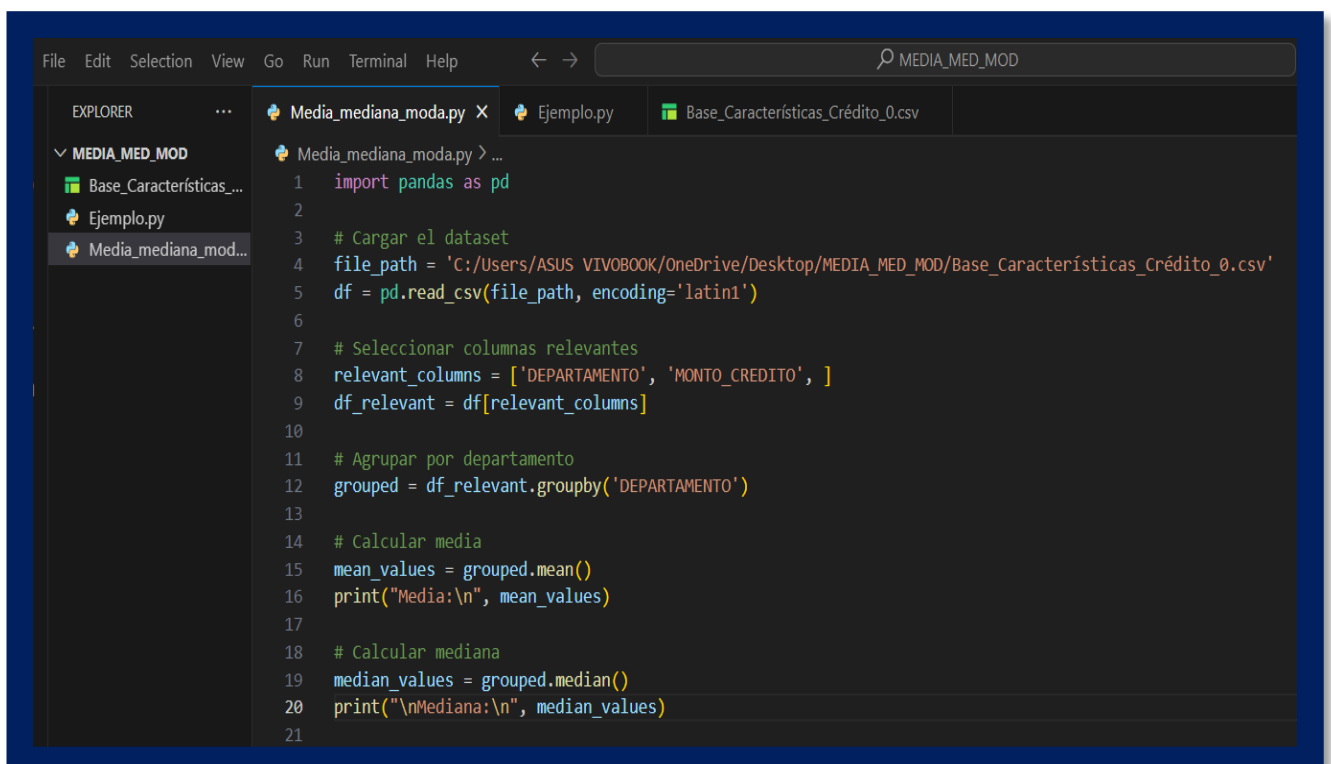
JUEVES

TRABAJO INDIVIDUAL PYTHON

Importamos la librería NUMPY y PANDAS PARA calcular la Media, Mediana y Moda de un archivo CSV.

EJEMPLO 1:

1. Observar y analizar dataset: Base_Características_Credito_0.csv
2. Seleccionar campos relevantes
3. Importar csv con Pandas
4. Calcular media, mediana y moda (agrupar por departamento)



```
File Edit Selection View Go Run Terminal Help  < -> MEDIA_MED_MOD

EXPLORER  ...
  MEDIA_MED_MOD
    Base_Características_...
    Ejemplo.py
    Media_mediana_mod...

Media_mediana_moda.py X  Ejemplo.py  Base_Características_Crédito_0.csv

Media_mediana_moda.py > ...
1  import pandas as pd
2
3  # Cargar el dataset
4  file_path = 'c:/Users/ASUS VIVOBOK/OneDrive/Desktop/MEDIA_MED_MOD/Base_Características_Crédito_0.csv'
5  df = pd.read_csv(file_path, encoding='latin1')
6
7  # Seleccionar columnas relevantes
8  relevant_columns = ['DEPARTAMENTO', 'MONTO_CREDITO', ]
9  df_relevant = df[relevant_columns]
10
11 # Agrupar por departamento
12 grouped = df_relevant.groupby('DEPARTAMENTO')
13
14 # Calcular media
15 mean_values = grouped.mean()
16 print("Media:\n", mean_values)
17
18 # Calcular mediana
19 median_values = grouped.median()
20 print("\nMediana:\n", median_values)
21
```


FileEditSelectionViewGoRunTerminalHelp

EXPLORER

MEDIA_MED_MOD

Base_Características_...

Ejemplo.py

Media_mediana_mod...

PROBLEMS

OUTPUT

DEBUG CONSOLE

TERMINAL

PORTS

PS C:\Users\ASUS VIVOBOOK\OneDrive\Desktop\MEDIA_MED_MOD> & "C:/Users

Media:

DEPARTAMENTO	MONTO_CREDITO
AMAZONAS	85513.942174
ANCASH	71456.767674
APURIMAC	128057.344211
AREQUIPA	141449.826977
AYACUCHO	93723.829502
CAJAMARCA	91148.459393
CALLAO	162306.957414
CUSCO	155745.778803
HUANCVELICA	93678.817347
HUANUCO	108219.531532
ICA	70997.526893
JUNIN	104709.120165
LA LIBERTAD	104088.593846
LAMBAYEQUE	107647.957125
LIMA	167014.254224
LORETO	79043.333089
MADRE DE DIOS	43734.828767
MOQUEGUA	80683.768259
PASCO	57397.728644
PIURA	86822.305301
PUNO	71910.081204
PUNO	71910.081204
SAN MARTIN	64220.042084
TACNA	86034.718166
TUMBES	62394.332692
UCAYALI	42115.512850

Mediana:

DEPARTAMENTO	MONTO_CREDITO
AMAZONAS	83500.0
ANCASH	65970.0
APURIMAC	127706.0
AREQUIPA	137037.0
AYACUCHO	80000.0
CAJAMARCA	85000.0
CALLAO	156800.0
CUSCO	148000.0
HUANCVELICA	91200.0
HUANUCO	97500.0
ICA	59852.0
JUNIN	99000.0
LA LIBERTAD	99400.0
LAMBAYEQUE	103599.0
LIMA	151400.0
LORETO	65600.0
MADRE DE DIOS	46100.0
MOQUEGUA	74700.0
PASCO	50000.0
PIURA	78100.0
PUNO	63050.0
SAN MARTIN	49950.0
TACNA	85937.5
TUMBES	56500.0
UCAYALI	25232.0

PS C:\Users\ASUS VIVOBOOK\OneDrive\Desktop\MEDIA_MED_MOD>

> OUTLINE

> TIMELINE

Ejemplo 2:

- Crea un gráfico y muestra la edad promedio por Departamento y año del (2007-2023) importa el archivo

```
Edad_promedio.py > ...
1 import pandas as pd
2 import matplotlib.pyplot as plt
3
4 file_path = 'C:/Users/ASUS VIVOBOK/OneDrive/Desktop/MEDIA_MED_MOD/poblacion2 (1).csv'
5 data = pd.read_csv(file_path, encoding='UTF-8', sep=';')
6
7 data.columns = ['DEPARTAMENTO', 'AÑO', '0-19', '20-39', '40-59', '60+']
8
9 age_weights = {'0-19': 10, '20-39': 30, '40-59': 50, '60+': 70}
10
11 # edad promedio ponderada por departamento y año
12 data['edad_promedio'] = (
13     data['0-19'] * age_weights['0-19'] +
14     data['20-39'] * age_weights['20-39'] +
15     data['40-59'] * age_weights['40-59'] +
16     data['60+'] * age_weights['60+']
17 ) / (data['0-19'] + data['20-39'] + data['40-59'] + data['60+'])
18
19 data['edad_promedio'] = data['edad_promedio'].astype(int)
20
21 grouped_data = data.groupby('DEPARTAMENTO').agg({
22     'AÑO': 'first',
23     '0-19': 'sum',
24     '20-39': 'sum',
25     '40-59': 'sum',
26     '60+': 'sum',
27     'edad_promedio': 'mean'
28 }).reset_index()
29
30 grouped_data['edad_promedio'] = grouped_data['edad_promedio'].astype(int)
31
32 print(grouped_data)
33
34
35 plt.figure(figsize=(10, 6))
36 plt.bar(grouped_data['DEPARTAMENTO'], grouped_data['edad_promedio'], color='skyblue')
37 plt.title('Edad Promedio por Departamento')
38 plt.xlabel('Departamento')
39 plt.ylabel('Edad Promedio (años)')
40 plt.xticks(rotation=90) # Rotar de etiquetas
41 plt.tight_layout() # Asegurarme qque se muestre correcamente el grafico
42 plt.show()
43
```



EVALUACIÓN DEL INFORME DE TRABAJO SEMANAL

NOTA

OBSERVACIONES Y RECOMENDACIONES

DEL INSTRUCTOR:

FIRMA DEL ESTUDIANTE:

FIRMA DEL INSTRUCTOR:



PROPIEDAD INTELECTUAL DEL SENATI. PROHIBIDA SU
REPRODUCCIÓN Y VENTA SIN LA AUTORIZACIÓN
CORRESPONDIENTE