



IPN
Instituto Politécnico Nacional



UPIIZ
*Unidad Profesional Interdisciplinaria de Ingeniería
Campus Zacatecas*

Ingeniería en:
Sistemas Computacionales

Materia:

Data Mining

Docente:

Héctor Alejandro Acuña Cid

Actividad:

*“Práctica 02. Visualización y Estructuras
Multidimensionales”*

Alumno:

Alejandro Tamayo Castro

Grupo:

7CM1

Zacatecas, Zac. septiembre 19 de 2023

Código

Para crear el data set de 500 se implemento un llenado aleatorio de datos partiendo de diccionarios establecidos, en este caso se usaron 10 nombres, 6 materias y 8 semestres, así como un llenado aleatorio de 0 hasta 10 en la calificación.

```
# Crear listas con valores posibles para cada columna
nombres = ['Domingo', 'Eloisa', 'Carlos', 'Maria', 'Eusebio', 'India', 'Claudio', 'Odalys', 'Regina', 'Isabel']
materias = ['Geografia', 'Historia', 'Fisica', 'Mecanografia', 'Matematicas', 'Confiteria']
semestres = ['Primer', 'Segundo', 'Tercer', 'Cuarto', 'Quinto', 'Sexto', 'Septimo', 'Octavo']
```

Se genero el dataset de 500 filas y se guardo en un DataFrame para su uso en el código.

```
# Generar 500 filas de datos
for _ in range(500):
    nombre = random.choice(nombres)
    materia = random.choice(materias)
    semestre = random.choice(semestres)
    calificacion = random.randint(0, 10)

    # Añadir los datos generados a las listas
    lista_nombre.append(nombre)
    lista_materia.append(materia)
    lista_semestre.append(semestre)
    lista_calificaciones.append(calificacion)

# Crear un DataFrame de pandas con los datos
df = pd.DataFrame({
    'Nombre': lista_nombre,
    'Materia': lista_materia,
    'Semestre': lista_semestre,
    'Calificacion': lista_calificaciones
})
```

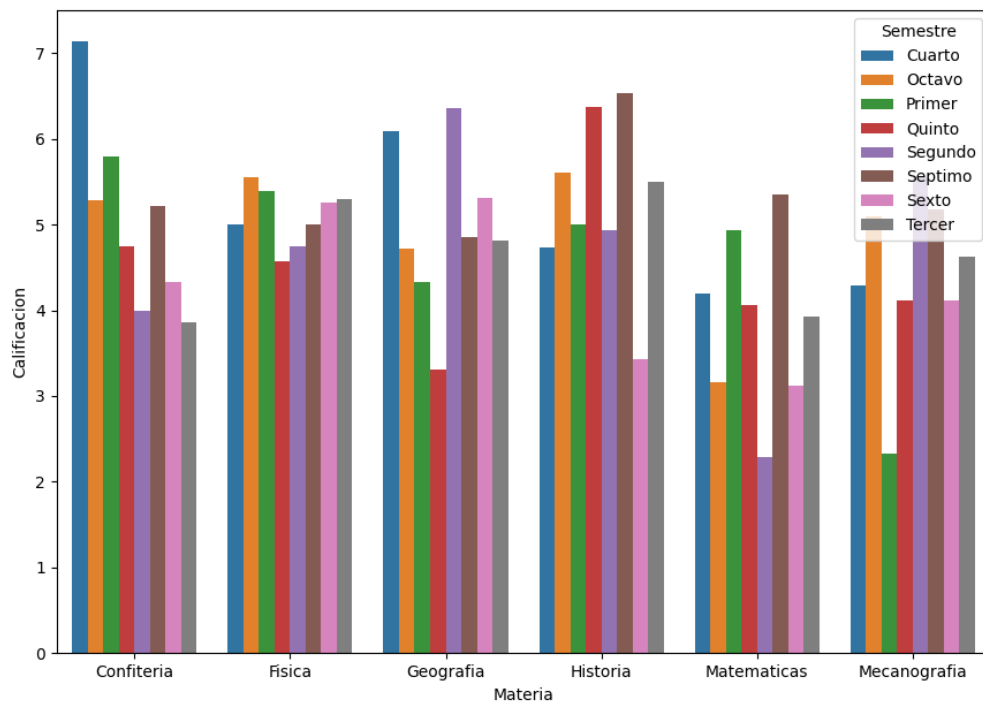
Para crear el cubo de datos, empleando la función groupby, se realizó de la siguiente manera:

```
#Crear cubo de datos con group.by
cubo = df.groupby(['Materia', 'Semestre'])['Calificacion'].mean().reset_index()
print("Cubo: ")
print(cubo)
```

Obteniendo de resultado lo siguiente.

Cubo:				20	Geografia	Segundo	6.363636
0	Confiteria	Cuarto	7.142857	21	Geografia	Septimo	4.857143
1	Confiteria	Octavo	5.285714	22	Geografia	Sexto	5.307692
2	Confiteria	Primer	5.800000	23	Geografia	Tercer	4.818182
3	Confiteria	Quinto	4.750000	24	Historia	Cuarto	4.727273
4	Confiteria	Segundo	4.000000	25	Historia	Octavo	5.600000
5	Confiteria	Septimo	5.222222	26	Historia	Primer	5.000000
6	Confiteria	Sexto	4.333333	27	Historia	Quinto	6.375000
7	Confiteria	Tercer	3.857143	28	Historia	Segundo	4.933333
8	Fisica	Cuarto	5.000000	29	Historia	Septimo	6.538462
9	Fisica	Octavo	5.555556	30	Historia	Sexto	3.428571
10	Fisica	Primer	5.384615	31	Historia	Tercer	5.500000
11	Fisica	Quinto	4.571429	32	Matematicas	Cuarto	4.200000
12	Fisica	Segundo	4.750000	33	Matematicas	Octavo	3.166667
13	Fisica	Septimo	5.000000	34	Matematicas	Primer	4.928571
14	Fisica	Sexto	5.250000	35	Matematicas	Quinto	4.062500
15	Fisica	Tercer	5.300000	36	Matematicas	Segundo	2.285714
16	Geografia	Cuarto	6.090909	37	Matematicas	Septimo	5.357143
17	Geografia	Octavo	4.714286	38	Matematicas	Sexto	3.117647
18	Geografia	Primer	4.333333	39	Matematicas	Tercer	3.923077
19	Geografia	Quinto	3.307692	40	Mecanografia	Cuarto	4.285714
20	Geografia	Segundo	6.363636	41	Mecanografia	Octavo	5.090909
				42	Mecanografia	Primer	2.333333
				43	Mecanografia	Quinto	4.111111
				44	Mecanografia	Segundo	5.545455
				45	Mecanografia	Septimo	5.181818
				46	Mecanografia	Sexto	4.111111
				47	Mecanografia	Tercer	4.625000

Para la visualización de los datos implementado el código del pdf se obtuvieron los siguientes resultados:



- ¿Cuál materia tiene el promedio más alto y en qué semestre?
 - La materia de Confiteria muestra la calificación de 7 que es la mas alta en 4to semestre.
- ¿Hay alguna materia que mejore o empeore significativamente entre semestres?
 - La materia de Geografia muestra buenas calificaciones en los semestres 4to, 2do, sexto, mientras que los demás semestres se muestran calificaciones peores a estas.
- ¿Cómo podrían los educadores utilizar estos insights para mejorar el currículo?
 - Implementando estrategias de aprendizaje diferentes en los semestres donde se muestran bajas calificaciones, integrando y mejorando sus rubricas de calificación así como invirtiendo en material de las materias con peores calificaciones para corregirlas.