**Task 1: Analyse user engagement across internship domains**

**code:**

import pandas as pd

import matplotlib.pyplot as plt

# Load internship data

internship\_df = pd.read\_csv("/content/DOC-20250611-WA0003. - DOC-20250611-WA0003..csv")

# Group by internship domain

domain\_summary = internship\_df.groupby("Internship\_Domain").agg(

Applications=('Applied', 'sum'),

Participations=('Participated', 'sum')

).reset\_index()

# Plotting

plt.figure(figsize=(10, 6))

bar\_width = 0.35

index = range(len(domain\_summary))

most\_applied\_domain = domain\_summary.loc[domain\_summary['Applications'].idxmax()]

print("Most Applied Domain:")

print(most\_applied\_domain)

# Find most popular domain by number of participations

most\_participated\_domain = domain\_summary.loc[domain\_summary['Participations'].idxmax()]

print("\nMost Participated Domain:")

print(most\_participated\_domain)

plt.bar(index, domain\_summary['Applications'], bar\_width, label='Applications', color='skyblue')

plt.bar([i + bar\_width for i in index], domain\_summary['Participations'], bar\_width, label='Participations', color='orange')

plt.xlabel('Internship Domain')

plt.ylabel('Number of Students')

plt.title('Internship Applications and Participations by Domain')

plt.xticks([i + bar\_width/2 for i in index], domain\_summary['Internship\_Domain'])

plt.legend()

plt.tight\_layout()

plt.show()

**output:**



