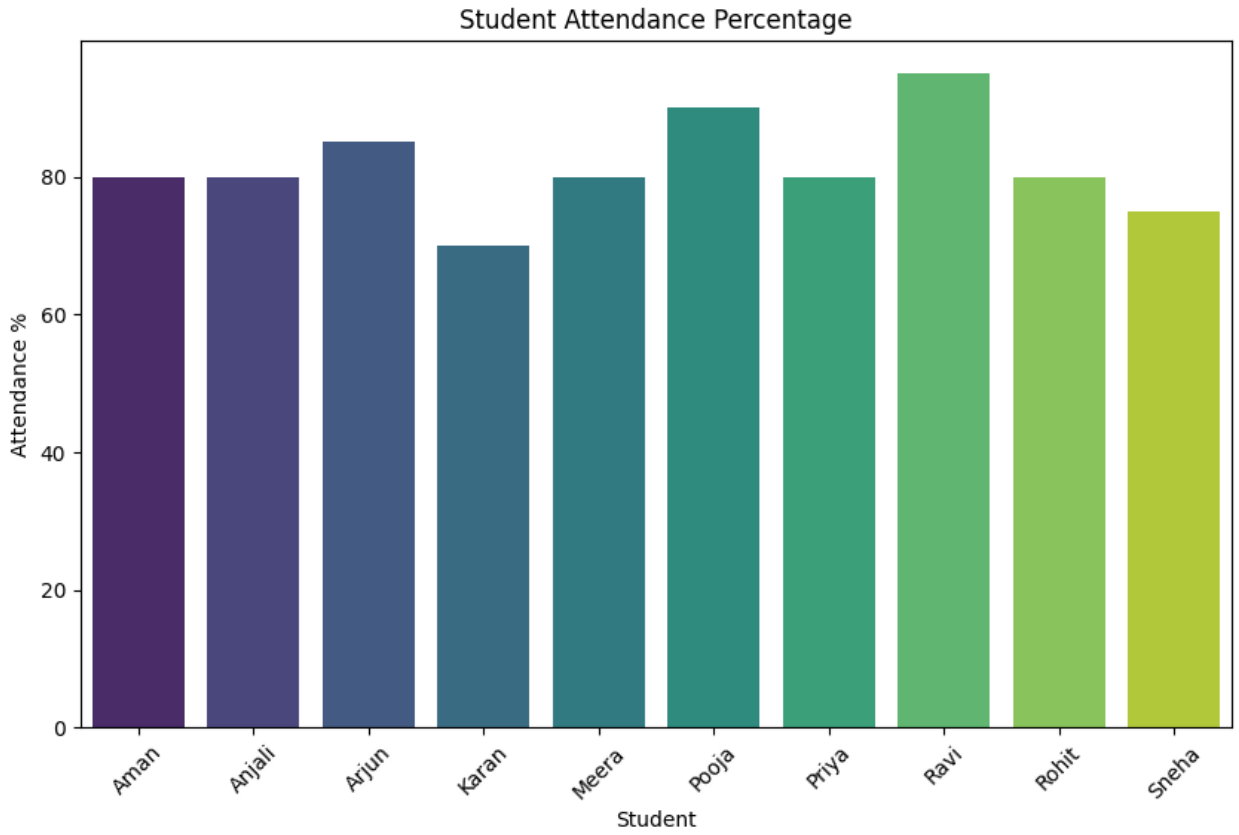
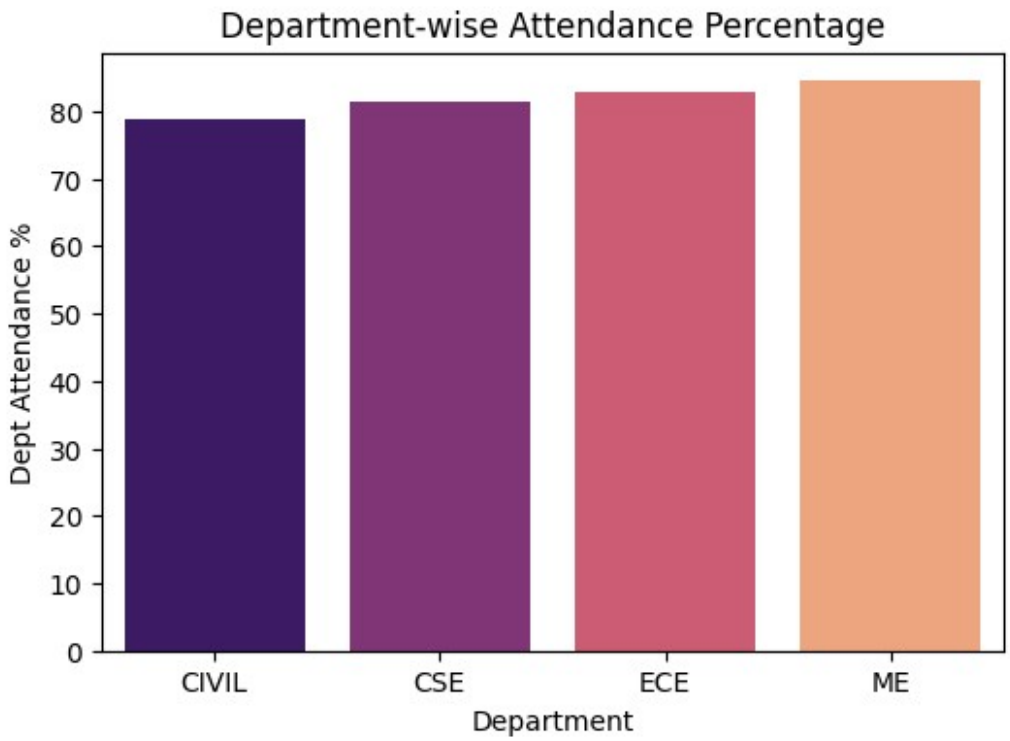
```python
# Student Attendance Analysis in Google Colab

# Step 1: Import Libraries
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns

# Step 2: Create a Sample Dataset
# (You can replace this with your own CSV upload)
np.random.seed(42)

students = ["Aman", "Priya", "Rohit", "Sneha", "Arjun", "Meera",
"Karan", "Anjali", "Ravi", "Pooja"]
departments = ["CSE", "ECE", "ME", "CIVIL"]
subjects = ["Maths", "Physics", "Computer", "English"]

data = []
for day in pd.date_range(start="2025-01-01", end="2025-01-20"):
    for s in students:
        record = {
            "Date": day,
            "Student": s,
            "Department": np.random.choice(departments),
            "Subject": np.random.choice(subjects),
            "Status": np.random.choice(["Present", "Absent"], p=[0.8,
0.2])
        }
        data.append(record)

df = pd.DataFrame(data)

# Step 3: Basic Dataset Info
print("First 10 rows of dataset:")
print(df.head(10))
print("\nTotal Records:", len(df))

# Step 4: Attendance Percentage per Student
attendance_summary = df.groupby("Student")["Status"].apply(
    lambda x: (x == "Present").mean() * 100
).reset_index(name="Attendance %")

print("\nStudent-wise Attendance Percentage:")
print(attendance_summary)

# Step 5: Department-wise Attendance %
dept_summary = df.groupby("Department")["Status"].apply(
    lambda x: (x == "Present").mean() * 100
).reset_index(name="Dept Attendance %")
```

```
print("\nDepartment-wise Attendance Percentage:")
print(dept_summary)

# Step 6: Subject-wise Attendance %
subject_summary = df.groupby("Subject")["Status"].apply(
    lambda x: (x == "Present").mean() * 100
).reset_index(name="Subject Attendance %")

print("\nSubject-wise Attendance Percentage:")
print(subject_summary)

# Step 7: Visualization - Student Attendance
plt.figure(figsize=(10,6))
sns.barplot(x="Student", y="Attendance %", data=attendance_summary,
palette="viridis")
plt.xticks(rotation=45)
plt.title("Student Attendance Percentage")
plt.show()

# Step 8: Visualization - Department Attendance
plt.figure(figsize=(6,4))
sns.barplot(x="Department", y="Dept Attendance %", data=dept_summary,
palette="magma")
plt.title("Department-wise Attendance Percentage")
plt.show()

# Step 9: Visualization - Subject Attendance
plt.figure(figsize=(6,4))
sns.barplot(x="Subject", y="Subject Attendance %",
data=subject_summary, palette="coolwarm")
plt.title("Subject-wise Attendance Percentage")
plt.show()

# Step 10: Attendance Heatmap (Date vs Student)
pivot = df.pivot_table(index="Date", columns="Student",
values="Status", aggfunc=lambda x: (x=="Present").mean())
plt.figure(figsize=(12,6))
sns.heatmap(pivot, cmap="Greens", cbar_kws={'label': 'Attendance
Rate'})
plt.title("Daily Attendance Heatmap (1=Present, 0=Absent)")
plt.show()

# Step 11: Overall Attendance Pie Chart
labels = ["Present", "Absent"]
sizes = df["Status"].value_counts().values
colors = ["#4CAF50", "#F44336"]
plt.pie(sizes, labels=labels, autopct='%1.1f%%', colors=colors,
startangle=90)
plt.title("Overall Attendance Distribution")
plt.show()
```

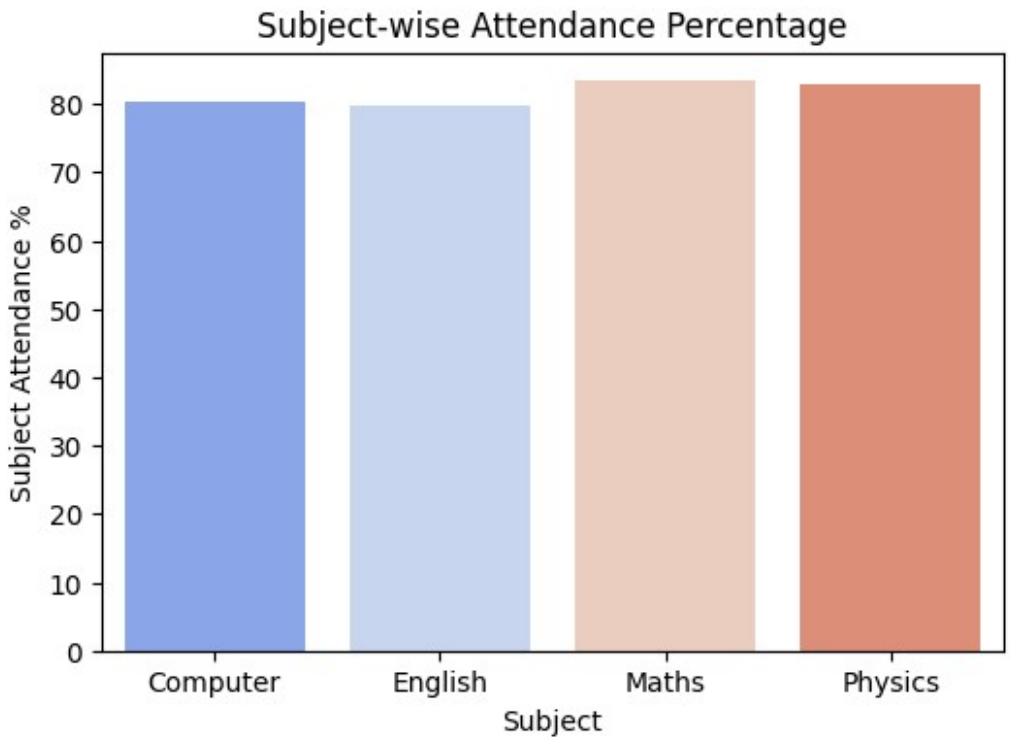
```
First 10 rows of dataset:
        Date Student Department   Subject   Status
0 2025-01-01    Aman         ME   English   Absent
1 2025-01-01   Priya         ME   English  Present
2 2025-01-01   Rohit         ME   Physics  Present
3 2025-01-01   Sneha         ME  Computer   Absent
4 2025-01-01   Arjun      CIVIL   English  Present
5 2025-01-01   Meera        ECE     Maths   Absent
6 2025-01-01   Karan      CIVIL   Physics  Present
7 2025-01-01  Anjali      CIVIL   English  Present
8 2025-01-01    Ravi      CIVIL   Physics  Present
9 2025-01-01   Pooja      CIVIL     Maths  Present

Total Records: 200

Student-wise Attendance Percentage:
  Student  Attendance %
0    Aman          80.0
1  Anjali          80.0
2   Arjun          85.0
3   Karan          70.0
4   Meera          80.0
5   Pooja          90.0
6   Priya          80.0
7    Ravi          95.0
8   Rohit          80.0
9   Sneha          75.0

Department-wise Attendance Percentage:
  Department  Dept Attendance %
0      CIVIL          78.688525
1        CSE          81.355932
2        ECE          82.857143
3         ME          84.444444

Subject-wise Attendance Percentage:
    Subject  Subject Attendance %
0  Computer             80.434783
1   English             79.629630
2     Maths             83.333333
3   Physics             82.758621

/tmp/ipython-input-3876286149.py:62: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be
removed in v0.14.0. Assign the `x` variable to `hue` and set
`legend=False` for the same effect.

  sns.barplot(x="Student", y="Attendance %", data=attendance_summary,
palette="viridis")
```

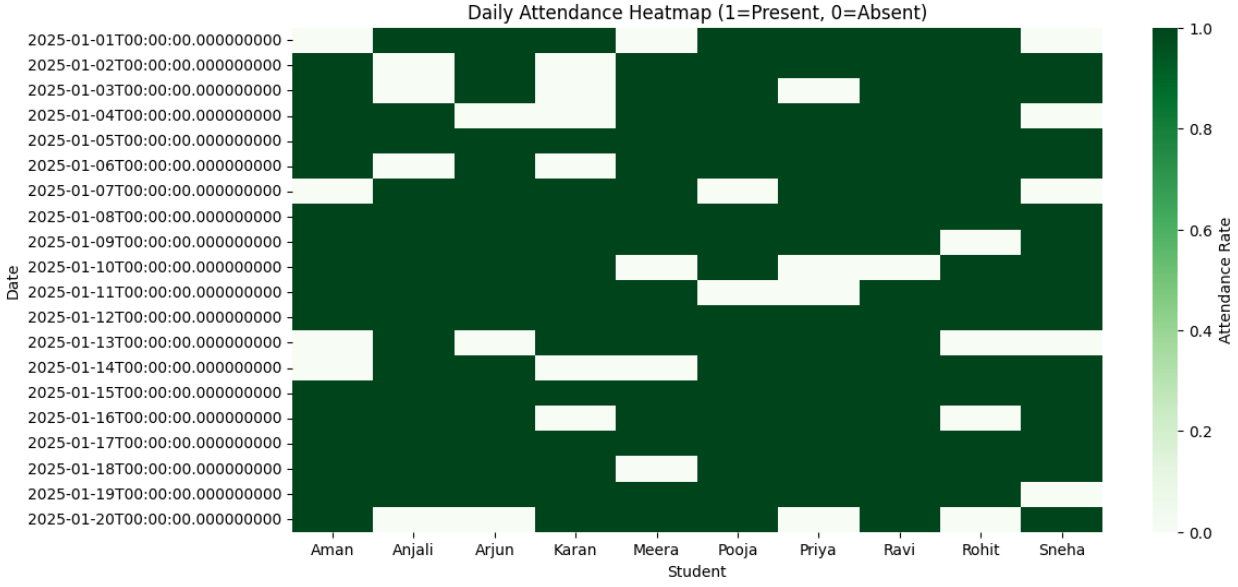


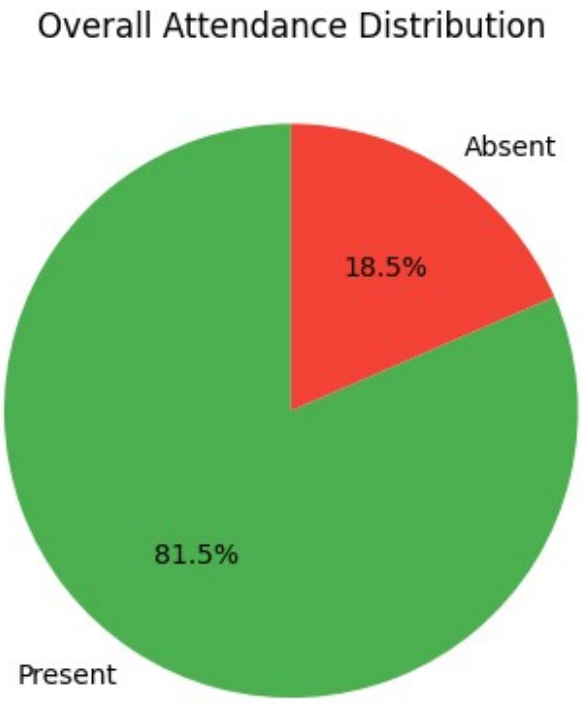

```
/tmp/ipython-input-3876286149.py:69: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be
removed in v0.14.0. Assign the `x` variable to `hue` and set
`legend=False` for the same effect.

  sns.barplot(x="Department", y="Dept Attendance %",
data=dept_summary, palette="magma")
```




```
/tmp/ipython-input-3876286149.py:75: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `x` variable to `hue` and set `legend=False` for the same effect.

  sns.barplot(x="Subject", y="Subject Attendance %", data=subject_summary, palette="coolwarm")
```


Subject-wise Attendance Percentage
Subject Attendance %
0
10
20
30
40
50
60
70
80
Computer
English
Maths
Physics
Subject


Daily Attendance Heatmap (1=Present, 0=Absent)
Date
2025-01-01T00:00:00.000000000
2025-01-02T00:00:00.000000000
2025-01-03T00:00:00.000000000
2025-01-04T00:00:00.000000000
2025-01-05T00:00:00.000000000
2025-01-06T00:00:00.000000000
2025-01-07T00:00:00.000000000
2025-01-08T00:00:00.000000000
2025-01-09T00:00:00.000000000
2025-01-10T00:00:00.000000000
2025-01-11T00:00:00.000000000
2025-01-12T00:00:00.000000000
2025-01-13T00:00:00.000000000
2025-01-14T00:00:00.000000000
2025-01-15T00:00:00.000000000
2025-01-16T00:00:00.000000000
2025-01-17T00:00:00.000000000
2025-01-18T00:00:00.000000000
2025-01-19T00:00:00.000000000
2025-01-20T00:00:00.000000000
Aman
Anjali
Arjun
Karan
Meera
Pooja
Priya
Ravi
Rohit
Sneha
Student
Attendance Rate
1.0
0.8
0.6
0.4
0.2
0.0


Overall Attendance Distribution
Absent
18.5%
81.5%
Present