1.Simple sales report using Google Colab

1. Open Google Colab

Go to: https://colab.research.google.com

Click on “New Notebook”

2. Program:

import pandas as pd

import matplotlib.pyplot as plt

data = {

'Product': ['A', 'B', 'A', 'C', 'B', 'C', 'A'],

'Quantity': [10, 5, 8, 6, 9, 4, 7],

'Price': [100, 200, None, 300, None, 350, 100]

}

df = pd.DataFrame(data)

df['Price'] = df.groupby('Product')['Price'].transform(lambda x: x.fillna(x.mean()))

df['Total\_Sales'] = df['Quantity'] \* df['Price']

sales\_by\_product = df.groupby('Product')['Total\_Sales'].sum()

top\_product = sales\_by\_product.idxmax()

print(f"Product with highest total sales: {top\_product}")

sales\_by\_product.plot(kind='bar', color='skyblue')

plt.title('Total Sales by Product')

plt.xlabel('Product')

plt.ylabel('Total Sales')

plt.grid(axis='y')

plt.show()

\_\_\_\_\_\_\_\_\_\_

2.Daily Temperature Tracker analysis in Google Colab

1. Open Google Colab

Go to https://colab.research.google.com

Click "New Notebook"

2. Program:

import pandas as pd

import matplotlib.pyplot as plt

data = {

'Date': ['2025-05-01', '2025-05-02', '2025-05-03', '2025-05-04', '2025-05-05'],

'Min\_Temp': [22, 24, None, 23, 25],

'Max\_Temp': [34, None, 36, 35, 37]

}

df = pd.DataFrame(data)

df['Date'] = pd.to\_datetime(df['Date'])

df['Min\_Temp'].fillna(df['Min\_Temp'].mean(), inplace=True)

df['Max\_Temp'].fillna(df['Max\_Temp'].mean(), inplace=True)

df['Average\_Temp'] = (df['Min\_Temp'] + df['Max\_Temp']) / 2

hottest\_day = df.loc[df['Average\_Temp'].idxmax()]

print(f"Date with highest average temperature: {hottest\_day['Date'].date()}")

plt.figure(figsize=(10,5))

plt.plot(df['Date'], df['Average\_Temp'], marker='o', linestyle='-', color='orange')

plt.title('Average Temperature Over Time')

plt.xlabel('Date')

plt.ylabel('Average Temperature')

plt.grid(True)

plt.xticks(rotation=45)

plt.tight\_layout()

plt.show()

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

3.COVID-19 Daily Cases analysis using Google Colab

1. Open Google Colab

Go to https://colab.research.google.com

Click on “New Notebook”

2. Program:

import pandas as pd

import matplotlib.pyplot as plt

data = {

'Date': ['2025-05-01', '2025-05-02', '2025-05-03', '2025-05-04', '2025-05-05'],

'Cases': [100, None, 250, 180, None]

}

df = pd.DataFrame(data)

df['Date'] = pd.to\_datetime(df['Date'])

df['Cases'].fillna(0, inplace=True)

total\_cases = df['Cases'].sum()

average\_cases = df['Cases'].mean()

print(f"Total Cases: {total\_cases}")

print(f"Average Daily Cases: {average\_cases}")

highest\_day = df.loc[df['Cases'].idxmax()]

print(f"Date with highest cases: {highest\_day['Date'].date()}, Cases: {highest\_day['Cases']}")

plt.figure(figsize=(10,5))

plt.plot(df['Date'], df['Cases'], marker='o', linestyle='-', color='red')

plt.title('COVID-19 Daily Cases')

plt.xlabel('Date')

plt.ylabel('Cases')

plt.grid(True)

plt.xticks(rotation=45)

plt.tight\_layout()

plt.show()

\_\_\_\_\_\_\_\_\_\_\_\_

4.Movie Ratings Dataset in Google Colab

1. Open Google Colab

Visit: https://colab.research.google.com

Click "New Notebook"

2. Program

import pandas as pd

import matplotlib.pyplot as plt

data = {

'Movie\_Name': ['Movie A', 'Movie B', 'Movie C', 'Movie D', 'Movie E', 'Movie F'],

'Viewer\_Rating': [4.5, 4.0, None, 3.5, 5.0, 4.8]

}

df = pd.DataFrame(data)

average\_rating = df['Viewer\_Rating'].mean()

print(f"Average Viewer Rating: {average\_rating:.2f}")

top\_3 = df.sort\_values(by='Viewer\_Rating', ascending=False).head(3)

print("Top 3 Movies:")

print(top\_3)

top\_5 = df.sort\_values(by='Viewer\_Rating', ascending=False).head(5)

plt.figure(figsize=(8,5))

plt.bar(top\_5['Movie\_Name'], top\_5['Viewer\_Rating'], color='purple')

plt.title('Top 5 Movie Ratings')

plt.xlabel('Movie Name')

plt.ylabel('Viewer Rating')

plt.ylim(0, 5)

plt.grid(axis='y')

plt.tight\_layout()

plt.show()

\_\_\_\_\_\_\_\_\_\_\_\_\_\_

5.Online Course Completion Data in Google Colab

1. Open Google Colab

Go to https://colab.research.google.com

Click “New Notebook”

2. Program:

import pandas as pd

import matplotlib.pyplot as plt

data = {

'Student\_ID': [101, 102, 103, 104, 105],

'Completion\_Status': ['Yes', None, 'No', 'Yes', None]

}

df = pd.DataFrame(data)

df['Completion\_Status'].fillna('No', inplace=True)

completed\_count = df[df['Completion\_Status'] == 'Yes'].shape[0]

print(f"Number of students who completed the course: {completed\_count}")

status\_counts = df['Completion\_Status'].value\_counts()

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

plt.pie(status\_counts, labels=status\_counts.index, autopct='%1.1f%%', startangle=90, colors=['green', 'red'])

plt.title('Course Completion Distribution')

plt.show()