1)

import pandas as pd

import matplotlib.pyplot as plt

import numpy as np

data = pd.read\_csv(r'C:\Users\Darth Rahul\Documents\DsResearch\Digital Marketing\online\_advertising\_performance\_data.csv', encoding='ISO-8859-1')

data['Date'] = pd.to\_datetime(data['day'].astype(str) + '-' + data['month'], format='%d-%B')

engagement\_mapping = {'Low': 1, 'Medium': 2, 'High': 3}

data['n\_user\_engagement'] = data['user\_engagement'].map(engagement\_mapping)

avg\_engagement = data.groupby('Date')['n\_user\_engagement'].mean()

plt.figure(figsize=(14, 8))

plt.plot(avg\_engagement.index, avg\_engagement.values, marker='o', linestyle='-', color='b', label='Average Engagement')

x = np.arange(len(avg\_engagement))

y = avg\_engagement.values

z = np.polyfit(x, y, 1)

p = np.poly1d(z)

plt.plot(avg\_engagement.index, p(x), "r--")

plt.title('Trend in User Engagement Throughout the Campaign Period')

plt.xlabel('Date')

plt.ylabel('Average User Engagement')

plt.grid(True)

plt.show()

print(avg\_engagement)

2)

import pandas as pd

import matplotlib.pyplot as plt

data = pd.read\_csv(r'C:\Users\Darth Rahul\Documents\DsResearch\Digital Marketing\online\_advertising\_performance\_data.csv', encoding='ISO-8859-1')

bc = data.groupby('banner')['clicks'].agg(['sum', 'mean']).reset\_index()

bc.columns = ['banner', 'total\_clicks', 'average\_clicks']

print(bc)

plt.figure(figsize=(14, 8))

plt.bar(bc['banner'], bc['total\_clicks'], color='red')

plt.xlabel('Banner Size')

plt.ylabel('Total Clicks')

plt.title('Total Clicks by Banner Size')

plt.show()

plt.figure(figsize=(14, 8))

plt.bar(bc['banner'], bc['average\_clicks'], color='blue')

plt.xlabel('Banner Size')

plt.ylabel('Average Clicks')

plt.title('Average Clicks by Banner Size')

plt.show()

3)

import pandas as pd

import matplotlib.pyplot as plt

data = pd.read\_csv(r'C:\Users\Darth Rahul\Documents\DsResearch\Digital Marketing\online\_advertising\_performance\_data.csv', encoding='ISO-8859-1')

placement\_stats = data.groupby('placement')[['displays', 'clicks']].sum().reset\_index()

placement\_sb\_displays = placement\_stats.sort\_values(by='displays', ascending=False)

placement\_sb\_clicks = placement\_stats.sort\_values(by='clicks', ascending=False)

print("Placements with the highest number of displays:")

print(placement\_sb\_displays.head(1))

print("\nPlacements with the highest number of clicks:")

print(placement\_sb\_clicks.head(1))

plt.figure(figsize=(14, 8))

plt.bar(placement\_sb\_displays['placement'][:5], placement\_sb\_displays['displays'][:10], color='red')

plt.xlabel('Placement')

plt.ylabel('Total Displays')

plt.title('Top 5 Placements by Total Displays')

plt.show()

plt.figure(figsize=(14, 8))

plt.bar(placement\_sb\_clicks['placement'][:5], placement\_sb\_clicks['clicks'][:10], color='blue')

plt.xlabel('Placement')

plt.ylabel('Total Clicks')

plt.title('Top 5 Placements by Total Clicks')

plt.show()

4)

import pandas as pd

data = pd.read\_csv(r'C:\Users\Darth Rahul\Documents\DsResearch\Digital Marketing\online\_advertising\_performance\_data.csv', encoding='ISO-8859-1')

correlation = data['cost'].corr(data['revenue'])

print("Correlation between cost and revenue:",correlation)

5)

import pandas as pd

data = pd.read\_csv(r'C:\Users\Darth Rahul\Documents\DsResearch\Digital Marketing\online\_advertising\_performance\_data.csv', encoding='ISO-8859-1')

total\_revenue = data['revenue'].sum()

total\_clicks = data['clicks'].sum()

avg\_revenue\_per\_click = total\_revenue / total\_clicks

print("Average revenue per click:",avg\_revenue\_per\_click)

6)

import pandas as pd

data = pd.read\_csv(r'C:\Users\Darth Rahul\Documents\DsResearch\Digital Marketing\online\_advertising\_performance\_data.csv', encoding='ISO-8859-1')

data['conversion\_rate'] = (data['post\_click\_conversions'] / data['clicks'].replace(0, 1)) \* 100

campaign\_cr = data.groupby('campaign\_number')['conversion\_rate'].mean().reset\_index()

print("Top campaigns with the highest post-click conversion rates:")

print(campaign\_cr)

7)

import pandas as pd

import matplotlib.pyplot as plt

import numpy as np

data = pd.read\_csv(r'C:\Users\Darth Rahul\Documents\DsResearch\Digital Marketing\online\_advertising\_performance\_data.csv', encoding='ISO-8859-1')

data['Date'] = pd.to\_datetime(data['day'].astype(str) + '-' + data['month'], format='%d-%B')

sales = data.groupby('Date')['post\_click\_sales\_amount'].sum().reset\_index()

sales['Date\_ordinal'] = sales['Date'].map(pd.Timestamp.toordinal)

correlation = sales['Date\_ordinal'].corr(sales['post\_click\_sales\_amount'])

print ("Correlation between date and post click sales: ",correlation)

plt.figure(figsize=(14, 8))

plt.plot(sales['Date'], sales['post\_click\_sales\_amount'], marker='o', linestyle='-', label='Post-Click Sales Amount')

z = np.polyfit(sales.index, sales['post\_click\_sales\_amount'], 1)

p = np.poly1d(z)

plt.plot(sales['Date'], p(sales.index), "r--", label='Trendline')

plt.title('Post-Click Sales Amount Over Time')

plt.xlabel('Date')

plt.ylabel('Post-Click Sales Amount')

plt.grid(True)

plt.legend()

plt.show()

8)

import pandas as pd

import matplotlib.pyplot as plt

data = pd.read\_csv(r'C:\Users\Darth Rahul\Documents\DsResearch\Digital Marketing\online\_advertising\_performance\_data.csv', encoding='ISO-8859-1')

engagement\_mapping = {'Low': 1, 'Medium': 2, 'High': 3}

data['user\_engagement\_numeric'] = data['user\_engagement'].map(engagement\_mapping)

avg\_engagement\_by\_banner = data.groupby('banner')['user\_engagement\_numeric'].mean()

plt.figure(figsize=(14, 8))

avg\_engagement\_by\_banner.plot(kind='bar', color='green')

plt.title('Average User Engagement by Banner Size')

plt.xlabel('Banner Size')

plt.ylabel('Average User Engagement')

plt.show()

print(avg\_engagement\_by\_banner)

9)

import pandas as pd

data = pd.read\_csv(r'C:\Users\Darth Rahul\Documents\DsResearch\Digital Marketing\online\_advertising\_performance\_data.csv', encoding='ISO-8859-1')

placement\_grouped = data.groupby('placement').sum()

placement\_grouped['post\_click\_conversion\_rate'] = placement\_grouped['post\_click\_conversions'] / placement\_grouped['clicks']

sorted\_placements = placement\_grouped.sort\_values(by='post\_click\_conversion\_rate', ascending=False)

print(sorted\_placements[['post\_click\_conversion\_rate']]))

10)

import pandas as pd

import matplotlib.pyplot as plt

data = pd.read\_csv(r'C:\Users\Darth Rahul\Documents\DsResearch\Digital Marketing\online\_advertising\_performance\_data.csv', encoding='ISO-8859-1')

data['Date'] = pd.to\_datetime(data['day'].astype(str) + '-' + data['month'], format='%d-%B')

data.set\_index('Date', inplace=True)

monthly\_data = data.resample('ME').sum()

plt.figure(figsize=(14, 8))

plt.plot(monthly\_data.index, monthly\_data['displays'], marker='o', linestyle='-', color='b', label='Displays')

plt.title('Displays Over Time')

plt.xlabel('Date')

plt.ylabel('Displays')

plt.show()

plt.figure(figsize=(14, 8))

plt.plot(monthly\_data.index, monthly\_data['clicks'], marker='o', linestyle='-', color='g', label='Clicks')

plt.title('Clicks Over Time')

plt.xlabel('Date')

plt.ylabel('Clicks')

plt.show()

11)

import pandas as pd

data = pd.read\_csv(r'C:\Users\Darth Rahul\Documents\DsResearch\Digital Marketing\online\_advertising\_performance\_data.csv', encoding='ISO-8859-1')

engagement\_mapping = {'Low': 1, 'Medium': 2, 'High': 3}

data['user\_engagement\_numeric'] = data['user\_engagement'].map(engagement\_mapping)

correlation = data['user\_engagement\_numeric'].corr(data['revenue'])

print("Correlation between user engagement and revenue:",correlation)

12)

import pandas as pd

data = pd.read\_csv(r'C:\Users\Darth Rahul\Documents\DsResearch\Digital Marketing\online\_advertising\_performance\_data.csv', encoding='ISO-8859-1')

def outliers(data, column):

Q1 = data[column].quantile(0.25)

Q3 = data[column].quantile(0.75)

IQR = Q3 - Q1

l = Q1 - 1.5 \* IQR

u = Q3 + 1.5 \* IQR

out = data[(data[column] < l) | (data[column] > u)]

return out

outliers\_cost = outliers(data, 'cost')

outliers\_clicks = outliers(data, 'clicks')

outliers\_revenue = outliers(data, 'revenue')

out\_summary = {

'cost': outliers\_cost.shape[0],

'clicks': outliers\_clicks.shape[0],

'revenue': outliers\_revenue.shape[0]

}

print("Summary of Outliers Detected:")

print(out\_summary)

print("\nDetails for Cost Outliers:")

print(outliers\_cost['cost'].describe())

print("\nDetails for Clicks Outliers:")

print(outliers\_clicks['clicks'].describe())

print("\nDetails for Revenue Outliers:")

print(outliers\_revenue['revenue'].describe())

13)

import pandas as pd

data = pd.read\_csv(r'C:\Users\Darth Rahul\Documents\DsResearch\Digital Marketing\online\_advertising\_performance\_data.csv', encoding='ISO-8859-1')

banner\_size = data.groupby('banner').agg({

'revenue': 'mean',

'clicks': 'sum',

'post\_click\_conversions': 'sum'

}).reset\_index()

banner\_size['post\_click\_conversion\_rate'] = banner\_size['post\_click\_conversions'] / banner\_size['clicks']

placement\_type = data.groupby('placement').agg({

'revenue': 'mean',

'clicks': 'sum',

'post\_click\_conversions': 'sum'

}).reset\_index()

placement\_type['post\_click\_conversion\_rate'] = placement\_type['post\_click\_conversions'] / placement\_type['clicks']

print("Effectiveness of campaign based on Banner Size:")

print(banner\_size[['banner', 'revenue', 'post\_click\_conversion\_rate']])

print("\nEffectiveness of campaign based on Placement Type:")

print(placement\_type[['placement', 'revenue', 'post\_click\_conversion\_rate']])

14)

import pandas as pd

data = pd.read\_csv(r'C:\Users\Darth Rahul\Documents\DsResearch\Digital Marketing\online\_advertising\_performance\_data.csv', encoding='ISO-8859-1')

data['ROI'] = ((data['revenue'] - data['cost']) / data['cost'].replace(0, 1))\*100

campaign\_roi = data.groupby('campaign\_number')['ROI'].mean().reset\_index()

campaign\_roi = campaign\_roi.sort\_values(by='ROI', ascending=False)

banner\_roi = data.groupby('banner')['ROI'].mean().reset\_index()

banner\_roi = banner\_roi.sort\_values(by='ROI', ascending=False)

print("Campaigns sorted by ROI(%):")

print(campaign\_roi)

print("\nBanner Sizes sorted by ROI(%):")

print(banner\_roi)

15)

import pandas as pd

import matplotlib.pyplot as plt

data = pd.read\_csv(r'C:\Users\Darth Rahul\Documents\DsResearch\Digital Marketing\online\_advertising\_performance\_data.csv', encoding='ISO-8859-1')

placement\_conversions = data.groupby('placement')['post\_click\_conversions'].mean().reset\_index()

print(placement\_conversions)

plt.figure(figsize=(14, 8))

plt.bar(placement\_conversions['placement'], placement\_conversions['post\_click\_conversions'], color='green')

plt.xlabel('Placement Type')

plt.ylabel('Average Post-Click Conversions')

plt.title('Distribution of Post-Click Conversions Across Different Placement Types')

plt.show()

16)

import pandas as pd

import matplotlib.pyplot as plt

def what\_day(day):

if day < 5:

return 'Weekday'

else:

return 'Weekend'

data = pd.read\_csv(r'C:\Users\Darth Rahul\Documents\DsResearch\Digital Marketing\online\_advertising\_performance\_data.csv', encoding='ISO-8859-1')

data['Date'] = pd.to\_datetime(data['day'].astype(str) + '-' + data['month'], format='%d-%B')

data['DayofWeek'] = data['Date'].dt.dayofweek

data['Weekday\_Weekend'] = data['DayofWeek'].apply(what\_day)

engagement = data.groupby('Weekday\_Weekend')[['displays', 'clicks', 'post\_click\_sales\_amount']].mean()

engagement.plot(kind='bar', figsize=(14, 8))

plt.title('Average User Engagement Levels: Weekdays vs. Weekends')

plt.xlabel('Day Type')

plt.ylabel('Average Values')

plt.xticks(rotation=0)

plt.show()

print(engagement)

17)

import pandas as pd

data = pd.read\_csv(r'C:\Users\Darth Rahul\Documents\DsResearch\Digital Marketing\online\_advertising\_performance\_data.csv', encoding='ISO-8859-1')

data\_nz\_clicks = data[data['clicks'] > 0]

data\_nz\_clicks['CPC'] = data\_nz\_clicks['cost'] / data\_nz\_clicks['clicks']

campaign\_cpc = data\_nz\_clicks.groupby('campaign\_number')['CPC'].mean().reset\_index()

banner\_cpc = data\_nz\_clicks.groupby('banner')['CPC'].mean().reset\_index()

print("Campaigns by cost per click(CPC):")

print(campaign\_cpc)

print("\nBanner Sizes by cost per click(CPC):")

print(banner\_cpc)

18)

import pandas as pd

data = pd.read\_csv(r'C:\Users\Darth Rahul\Documents\DsResearch\Digital Marketing\online\_advertising\_performance\_data.csv', encoding='ISO-8859-1')

data\_nz\_conversions = data[data['post\_click\_conversions'] > 0]

data\_nz\_conversions['cost\_per\_conversion'] = data\_nz\_conversions['cost'] / data\_nz\_conversions['post\_click\_conversions']

campaign\_cpc = data\_nz\_conversions.groupby('campaign\_number')['cost\_per\_conversion'].mean().reset\_index()

campaign\_cpc = campaign\_cpc.sort\_values(by='cost\_per\_conversion')

placement\_cpc = data\_nz\_conversions.groupby('placement')['cost\_per\_conversion'].mean().reset\_index()

placement\_cpc = placement\_cpc.sort\_values(by='cost\_per\_conversion')

print("Campaigns after sorting by Cost per Post-Click Conversion:")

print(campaign\_cpc)

print("\nPlacements after sorting by Cost per Post-Click Conversion:")

print(placement\_cpc)

19)

import pandas as pd

import matplotlib.pyplot as plt

import numpy as np

data = pd.read\_csv(r'C:\Users\Darth Rahul\Documents\DsResearch\Digital Marketing\online\_advertising\_performance\_data.csv', encoding='ISO-8859-1')

data['Date'] = pd.to\_datetime(data['day'].astype(str) + '-' + data['month'], format='%d-%B')

data['DayOfWeek'] = data['Date'].dt.day\_name()

week = ['Sunday','Monday', 'Tuesday', 'Wednesday', 'Thursday', 'Friday', 'Saturday']

data['DayOfWeek'] = pd.Categorical(data['DayOfWeek'], categories=week, ordered=True)

data['post\_click\_conversion\_rate'] = (data['post\_click\_conversions'] / data['clicks'].replace(0, 1)) \* 100

pccc = data.groupby('DayOfWeek')['post\_click\_conversion\_rate'].mean().reindex(week)

plt.figure(figsize=(14, 8))

plt.bar(pccc.index, pccc.values, color='skyblue')

x = np.arange(len(pccc))

y = pccc.values

z = np.polyfit(x, y, 1)

p = np.poly1d(z)

plt.plot(pccc.index, p(x), "r--")

plt.title('Average Post-Click Conversion Rate by Day of the Week')

plt.xlabel('Day of the Week')

plt.ylabel('Average Conversion Rate (%)')

plt.show()

print(pccc)

20)

import pandas as pd

import matplotlib.pyplot as plt

data = pd.read\_csv(r'C:\Users\Darth Rahul\Documents\DsResearch\Digital Marketing\online\_advertising\_performance\_data.csv', encoding='ISO-8859-1')

engagement\_mapping = {'Low': 1, 'Medium': 2, 'High': 3}

data['user\_engagement\_numeric'] = data['user\_engagement'].map(engagement\_mapping)

ec\_data = data.groupby(['user\_engagement', 'campaign\_number']).agg({

'post\_click\_conversions': 'mean'

}).reset\_index()

pt= ec\_data.pivot(index='campaign\_number', columns='user\_engagement', values='post\_click\_conversions')

pt.plot(kind='bar', figsize=(14, 8))

plt.xlabel('Campaign Number')

plt.ylabel('Average Post-Click Conversions')

plt.title('Average Post-Click Conversions by Campaign and User Engagement Type')

plt.xticks(rotation=0)

plt.legend(title='User Engagement', loc='upper right')

plt.show()

print(pt)