**Akhilesh Kumar - Pandas Exam Paper 2**

**Only for Answer**

**Section A : Data Manipulation**

1. import pandas as pd

data = {'Product': ['A', 'B', 'C'], 'Price': [100, 200, 300]}

df = pd.DataFrame(data)

df['Price'] = df['Price'].apply(lambda x: x \* 2)

print(df)

Output -

Product Price

0 A 200

1 B 400

2 C 600

1. import pandas as pd

data = {'Name': ['John', 'Alice', 'Bob'],

'Passed': ['Yes', 'No', 'Yes']}

df = pd.DataFrame(data)

df['Passed'] = df['Passed'].map({'Yes': True, 'No': False})

print(df)

Output -

Name Passed

0 John True

1 Alice False

2 Bob True

1. import pandas as pd

data = {'Names': ['John', 'ALICE', 'Bob']}

df = pd.DataFrame(data)

df['Names'] = df['Names'].str.lower()

print(df)

Output -

Names

0 john

1 alice

2 bob

1. import pandas as pd

data = {'City': ['Tokyo', 'mumbai', 'New York']}

df = pd.DataFrame(data)

df['City'] = df['City'].str.upper()

print(df)

Output -

City

0 TOKYO

1 MUMBAI

2 NEW YORK

1. import pandas as pd

data = {'FullName': ['John Doe', 'Alice Smith', 'Bob Johnson']}

df = pd.DataFrame(data)

df[['FirstName', 'LastName']] = df['FullName'].str.split(' ', expand=True)

print(df)

Output -

FullName FirstName LastName

0 John Doe John Doe

1 Alice Smith Alice Smith

2 Bob Johnson Bob Johnson

1. import pandas as pd

data = {'Name': ['Alice', 'Bob', 'Charlie'],

'Email': ['alice@gmail.com', 'bob@yahoo.com', 'charlie@gmail.com']}

df = pd.DataFrame(data)

filtered\_df = df[df['Email'].str.contains('@gmail.com', na=False)]

print(filtered\_df)

Output -

Name Email

0 Alice alice@gmail.com

2 Charlie [charlie@gmail.com](mailto:charlie@gmail.com)

1. import pandas as pd

data = {'Name': ['Alice', 'Bob', 'Charlie'],

'Email': ['alice@example.com', 'bob@example.com', 'charlie@example.com']}

df = pd.DataFrame(data)

df['Email'] = df['Email'].str.replace('@example.com', '@newdomain.com')

print(df)

Output -

Name Email

0 Alice alice@newdomain.com

1 Bob bob@newdomain.com

2 Charlie [charlie@newdomain.com](mailto:charlie@newdomain.com)

**Section B: Grouping and Aggregation (8 Questions)**

1. import pandas as pd

data = {'Department': ['HR', 'Finance', 'HR', 'IT', 'Finance', 'IT'],

'Salary': [50000, 60000, 55000, 70000, 65000, 72000]}

df = pd.DataFrame(data)

mean\_salary = df.groupby('Department')['Salary'].mean()

print(mean\_salary)

Output -

Department

Finance 62500.0

HR 52500.0

IT 71000.0

Name: Salary, dtype: float64

1. import pandas as pd

data = {'Region': ['East', 'West', 'East', 'West', 'North', 'South'],

'Sales': [200, 300, 250, 400, 150, 500]}

df = pd.DataFrame(data)

aggregated\_sales = df['Sales'].agg(['mean', 'max'])

print(aggregated\_sales)

Output -

mean 300.0

max 500.0

Name: Sales, dtype: float64

1. import pandas as pd

data = {'Name': ['Alice', 'Bob', 'Charlie', 'David'],

'Marks': [85, 90, 78, 92]}

df = pd.DataFrame(data)

aggregated\_marks = df['Marks'].aggregate(['sum', 'count'])

print(aggregated\_marks)

Output -

sum 345

count 4

Name: Marks, dtype: int64

1. import pandas as pd

data = {'Name': ['Alice', 'Bob', 'Charlie', 'David'],

'City': ['New York', 'Los Angeles', 'Chicago', 'New York']}

df = pd.DataFrame(data)

filtered\_df = df[df['City'].isin(['New York', 'Los Angeles'])]

print(filtered\_df)

Output -

Name City

0 Alice New York

1 Bob Los Angeles

3 David New York

1. import pandas as pd

data = {'Gender': ['Male', 'Female', 'Male', 'Female', 'Male'],

'Marks': [85, 90, 78, 88, 92]}

df = pd.DataFrame(data)

grouped\_data = df.groupby('Gender')['Marks'].sum()

print(grouped\_data)

Output-

Gender

Female 178

Male 255

Name: Marks, dtype: int64

1. import pandas as pd

data = {'Age': [25, 30, 35, 28, 40],

'Salary': [50000, 60000, 65000, 58000, 70000]}

df = pd.DataFrame(data)

aggregated\_data = df.agg({'Age': ['min', 'max', 'mean'],

'Salary': ['min', 'max', 'mean']})

print(aggregated\_data)

Output -

Age Salary

min 25.0 50000.0

max 40.0 70000.0

mean 31.6 60600.0

1. import pandas as pd

data = {'City': ['Mumbai', 'Delhi', 'Mumbai', 'Chennai', 'Delhi', 'Chennai', 'Mumbai'],

'Name': ['Akhil', 'Raj', 'Priya', 'John', 'Anita', 'Sam', 'Karan']}

df = pd.DataFrame(data)

grouped\_data = df.groupby('City').size()

print(grouped\_data)

Output -

City

Chennai 2

Delhi 2

Mumbai 3

dtype: int64

1. import pandas as pd

data = {'Department': ['HR', 'IT', 'IT', 'HR', 'Finance', 'Finance', 'IT'],

'Salary': [50000, 70000, 65000, 52000, 60000, 75000, 68000]}

df = pd.DataFrame(data)

def salary\_range(series):

return series.max() - series.min()

salary\_ranges = df.groupby('Department')['Salary'].apply(salary\_range)

print(salary\_ranges)

Output -

Department

Finance 15000

HR 2000

IT 5000

Name: Salary, dtype: int64

**Section C: Merging, Joining and Concatenating (5 Questions)**

1. import pandas as pd

data1 = {'Name': ['Akhil', 'Priya'],

'Age': [25, 22]}

df1 = pd.DataFrame(data1)

data2 = {'Name': ['Raj', 'Anita'],

'Age': [30, 28]}

df2 = pd.DataFrame(data2)

result = pd.concat([df1, df2], ignore\_index=True)

print(result)

Output -

Name Age

0 Akhil 25

1 Priya 22

2 Raj 30

3 Anita 28

1. import pandas as pd

data1 = {'ID': [1, 2, 3],

'Name': ['Akhil', 'Priya', 'Raj']}

df1 = pd.DataFrame(data1)

data2 = {'ID': [1, 2, 4],

'Age': [25, 22, 30]}

df2 = pd.DataFrame(data2)

merged\_data = pd.merge(df1, df2, on='ID', how='inner')

print(merged\_data)

Output -

ID Name Age

0 1 Akhil 25

1 2 Priya 22

1. import pandas as pd

data1 = {'EmployeeID': [101, 102, 103],

'Name': ['Akhil', 'Priya', 'Raj']}

df1 = pd.DataFrame(data1)

data2 = {'ID': [101, 104, 102],

'Age': [25, 30, 22]}

df2 = pd.DataFrame(data2)

merged\_data = pd.merge(df1, df2, left\_on='EmployeeID', right\_on='ID', how='inner')

print(merged\_data)

Output -

EmployeeID Name ID Age

0 101 Akhil 101 25

1 102 Priya 102 22

1. import pandas as pd

data1 = {'Name': ['Akhil', 'Priya', 'Raj'],

'Age': [25, 22, 30]}

df1 = pd.DataFrame(data1)

data2 = {'City': ['Hyderabad', 'Chennai', 'Delhi'],

'Salary': [50000, 60000, 70000]}

df2 = pd.DataFrame(data2)

result = pd.concat([df1, df2], axis=1)

print(result)

Output -

Name Age City Salary

0 Akhil 25 Hyderabad 50000

1 Priya 22 Chennai 60000

2 Raj 30 Delhi 70000

1. import pandas as pd

# Assuming df1 and df2 are your DataFrames

result = pd.merge(df1, df2, on="ID", how="outer")

print(result)

**Section D: Reshaping and Input/Output (10 Questions)**

1. import pandas as pd

data = {'A': [1, 2], 'B': [3, 4]}

df = pd.DataFrame(data)

transposed\_df = df.T

print(transposed\_df)

Output -

0 1

A 1 2

B 3 4

1. transposed\_df = df.T

print(transposed\_df)

Output -

0 1

A 1 2

B 3 4

1. df.to\_csv('output.csv', index=False)

print("DataFrame saved to output.csv")

1. df.to\_excel('output.xlsx', index=False)

print("DataFrame exported to output.xlsx")

1. df.to\_json('output.json', orient='records')

print("DataFrame converted to output.json")

1. df.to\_html('output.html')

print("DataFrame rendered and saved as output.html")

1. import pandas as pd

df = pd.read\_csv('student\_data.csv')

print(df.head())

1. import pandas as pd

df = pd.read\_excel('sales\_data.xlsx')

print(df.head())

1. df[['Name', 'Salary']].to\_csv('employees.csv', index=False)

print("Selected columns saved to employees.csv")

1. df[['Name', 'Department']].to\_json('output.json', orient='records')

print("Selected columns saved to output.json")