…………………………………………………………….Assignment………………………………………………………………………

1. Write a python program using pandas Interpolation to fill in missing values in the data frame.

Input:

dfpd.DataFrame(("Math":[12, 4, 7, None, 2],

"English" (4, 3, 57, 3, None).

"Hindi": (20, 16, None, 3, 8).

"Science" [14, 3, None, None, 6]})

//code

import pandas as pd

import numpy as np

# Create the DataFrame

data = {

"Math": [12, 4, 7, None, 2],

"English": [4, 3, 57, 3, None],

"Hindi": [20, 16, None, 3, 8],

"Science": [14, 3, None, None, 6]

}

df = pd.DataFrame(data)

# Display the original DataFrame

print("Original DataFrame:")

print(df)

# Interpolate the missing values

df\_interpolated = df.interpolate()

# Display the DataFrame after interpolation

print("\nDataFrame after interpolation:")

print(df\_interpolated)

output:

Original DataFrame:

Math English Hindi Science

0 12.0 4.0 20.0 14.0

1 4.0 3.0 16.0 3.0

2 7.0 57.0 NaN NaN

3 NaN 3.0 3.0 NaN

4 2.0 NaN 8.0 6.0

DataFrame after interpolation:

Math English Hindi Science

0 12.0 4.0 20.0 14.0

1 4.0 3.0 16.0 3.0

2 7.0 57.0 12.0 4.5

3 4.5 3.0 3.0 5.25

4 2.0 30.0 8.0 6.0

………………………………………………………………………………………………………………………………………………….

2. Write a python program using Interpolation to fill in missing values in the data frame. Then generate a subject wise highest score in the form of a bar chart.

Input:

data = {

'Subject': ['Math', 'Physics', 'Chemistry', 'Biology'].

'Student A': [80, 85, np.nan, 70],

'Student B: [90, np.nan, 75, 85],

'Student C: [80, 88, np.nan, 78]

//code

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

# Create the DataFrame

data = {

'Subject': ['Math', 'Physics', 'Chemistry', 'Biology'],

'Student A': [80, 85, np.nan, 70],

'Student B': [90, np.nan, 75, 85],

'Student C': [80, 88, np.nan, 78]

}

df = pd.DataFrame(data)

# Display the original DataFrame

print("Original DataFrame:")

print(df)

# Interpolate the missing values

df\_interpolated = df.interpolate()

# Display the DataFrame after interpolation

print("\nDataFrame after interpolation:")

print(df\_interpolated)

# Find the highest score for each subject

max\_scores = df\_interpolated.set\_index('Subject').max(axis=1)

# Generate the bar chart

max\_scores.plot(kind='bar', color='skyblue')

plt.title('Highest Scores by Subject')

plt.xlabel('Subject')

plt.ylabel('Score')

plt.xticks(rotation=45)

plt.tight\_layout()

plt.show()

output:

Original DataFrame:

Subject Student A Student B Student C

0 Math 80.0 90.0 80.0

1 Physics 85.0 NaN 88.0

2 Chemistry NaN 75.0 NaN

3 Biology 70.0 85.0 78.0

DataFrame after interpolation:

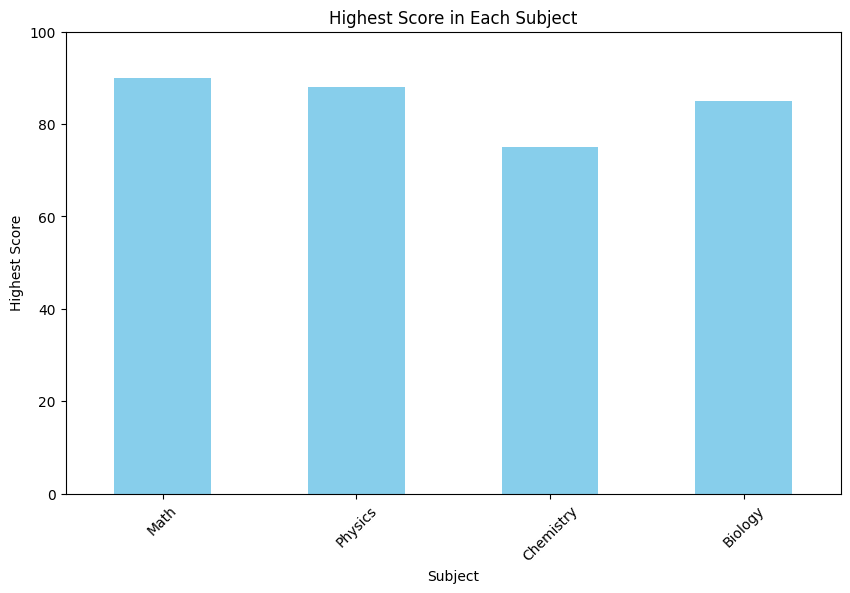
Subject Student A Student B Student C

0 Math 80.0 90.0 80.0

1 Physics 85.0 81.5 88.0

2 Chemistry 82.5 75.0 81.5

3 Biology 70.0 85.0 78.0



…………………………………………………………………………………………………………………………………………………………….