## Question 3: Finding IDs within Percentage Threshold

import pandas as pd

# Assuming df is the DataFrame created in Question 2

def find\_ids\_within\_ten\_percentage\_threshold(df, reference\_value):

# Filter DataFrame for the given reference value

reference\_df = df[df['id\_start'] == reference\_value]

# Check if the reference value exists in the DataFrame

if reference\_df.empty:

print(f"Reference value {reference\_value} not found in the DataFrame.")

return []

# Calculate the average distance for the reference value

average\_distance = reference\_df['distance'].mean()

# Calculate the threshold values

lower\_threshold = average\_distance - (average\_distance \* 0.1)

upper\_threshold = average\_distance + (average\_distance \* 0.1)

# Filter DataFrame for values within the 10% threshold

within\_threshold\_df = df[(df['distance'] >= lower\_threshold) & (df['distance'] <= upper\_threshold)]

# Get the sorted list of values from id\_start column

sorted\_ids\_within\_threshold = sorted(within\_threshold\_df['id\_start'].unique())

return sorted\_ids\_within\_threshold

# Example usage:

# Assuming df is the DataFrame created in Question 2

reference\_value = 123 # Replace with the desired reference value

result = find\_ids\_within\_ten\_percentage\_threshold(df, reference\_value)

print(result)