sample_compute_interpolation_error

March 15, 2022

The Coastal Grain Size Portal (C-GRASP) dataset Will Speiser, Daniel Buscombe, Evan Goldstein > Calculate Interpolation Error from Known Sample Values

The purpose of this notebook
This notebook will output a CSV containing all of the data from a chosen C-GRASP dataset with new fields containing an estimated percent error for interpolation of distribution percentiles. This will only be calculated for samples where distribution percentile values are included in the source dataset, as that is the only way to establish a "known" value. As C-Grasp file sizes vary completion of this task will vary with internet connectivity and computer processing power. This notebook provides simple code that estimates the percent error for various interpolated distribution values in the C-Grasp dataset.
To do so, a user choose a CGRASP dataset of choice . The notebook then runs loops through each sample with known distribution percentile values, recalculates that value and calculates an estimate for percent error of the scipy interpolation function (see the "sample_compute_percentile" notebook).

[1]:
```
import pandas as pd
import scipy
from scipy.interpolate import interp1d
import requests
import ipywidgets
import math
import numpy as np
import matplotlib.pyplot as plt
```

Select a dataset

[2]:
```
#Dataset collection widget
zen=ipywidgets.Select(
    options=['Entire Dataset', 'Estimated Onshore Data', 'Verified Onshore␣
 ↪Data', 'Verified Onshore Post 2012 Data'],
    value='Entire Dataset',
    # rows=10,
    description='Dataset:',
    disabled=False
)

display(zen)
```

```
Select(description='Dataset:', options=('Entire Dataset', 'Estimated Onshore␣
↪Data', 'Verified Onshore Data', '…
```

Download the dataset

[3]:
```
url = 'https://zenodo.org/record/6099266/files/'
if zen.value=='Entire Dataset':
    filename='dataset_10kmcoast.csv'
if zen.value=='Estimated Onshore Data':
    filename='Data_EstimatedOnshore.csv'
if zen.value=='Verified Onshore Data':
    filename='Data_VerifiedOnshore.csv'
if zen.value=='Verified Onshore Post 2012 Data':
    filename='Data_Post2012_VerifiedOnshore.csv'
print("Downloading {}".format(url+filename))
```

```
Downloading https://zenodo.org/record/6099266/files/Data_VerifiedOnshore.csv
```

The next cell will download the CGRASP dataset and read it in as a pandas dataframe with variable name `df`

[4]:
```
url=(url+filename)
print('Retrieving Data, Please Wait')
#retrieve data
df=pd.read_csv(url)
print('Sediment Data Retrieved!')
```

```
Retrieving Data, Please Wait
Sediment Data Retrieved!
```

Let's take a quick look at the file

[5]:
```
df
```

[5]:

	ID	Sample_ID	Sample_Type_Code	\
0	876	SPIbeach5	1	
1	878	SPI6	1	
2	877	SPI6	1	
3	1429	SPIbeach4	1	
4	1430	SPIbeach3	1	
…	…	…	…	
5348	591703	NaN	1	
5349	591704	NaN	1	
5350	591705	NaN	1	
5351	607152	3WellsBeach	1	
5352	607153	Wells Beach	1	

	Project	dataset	Date	\
0	SandSnap, image taken by:	sandsnap	2021-11-08	
1	SandSnap, image taken by:	sandsnap	2021-11-08	

```
2                            SandSnap, image taken by:   sandsnap  2021-11-08
3                            SandSnap, image taken by:   sandsnap  2021-11-08
4                            SandSnap, image taken by:   sandsnap  2021-11-08
…                                                   …       …         …
5348                                                  NaN     UMASS  2016-03-10
5349                                                  NaN     UMASS  2015-08-27
5350                                                  NaN     UMASS  2016-03-10
5351  Brian McFall, sandsnap training round 1 March …  sandsnap  2019-09-23
5352            SandSnap, image taken by: rose.dopsovic  sandsnap  2021-09-14

     Location_Type  latitude  longitude                  Contact  …  \
0          Beach?Y  26.12871  -97.16718           Sandsnap, USACE  …
1          Beach?Y  26.12899  -97.16713           Sandsnap, USACE  …
2          Beach?Y  26.12899  -97.16713           Sandsnap, USACE  …
3          Beach?Y  26.16883  -97.17248           Sandsnap, USACE  …
4          Beach?Y  26.16885  -97.17284           Sandsnap, USACE  …
…              …       …         …                        …  …
5348    berm crest  42.87166  -70.81624  Jonathan Woodruff, UMASS  …
5349          dune  42.87166  -70.81624  Jonathan Woodruff, UMASS  …
5350          dune  42.87166  -70.81630  Jonathan Woodruff, UMASS  …
5351       Beach?Y  43.30194  -70.56639           Sandsnap, USACE  …
5352       Beach?Y  43.30194  -70.56639           Sandsnap, USACE  …

           d16       d25       d30       d50       d65       d75       d84  \
0     0.565657  0.624976  0.657068  0.785439  0.889342  1.016927  1.131754
1     0.565657  0.624976  0.657068  0.785439  0.889342  1.016927  1.131754
2     0.565657  0.624976  0.657068  0.785439  0.889342  1.016927  1.131754
3     0.565657  0.624976  0.657068  0.785439  0.889342  1.016927  1.131754
4     0.565657  0.624976  0.657068  0.785439  0.889342  1.016927  1.131754
…        …         …         …         …         …         …         …
5348  0.358041  0.400608  0.423038  0.490549       NaN  0.557579  0.580922
5349  0.239851  0.259365  0.267726  0.300685       NaN  0.325147  0.333953
5350  0.385868  0.431429  0.449862  0.516696       NaN  0.581913  0.604390
5351  0.181053  0.189600  0.194784  0.215519  0.232742  0.253463  0.272112
5352  0.731995  0.806752  0.846725  1.006613  1.139310  1.301861  1.448158

           d90       d95                                                Notes
0     1.276942  1.397932                                                  NaN
1     1.276942  1.397932                                                  NaN
2     1.276942  1.397932                                                  NaN
3     1.276942  1.397932                                                  NaN
4     1.276942  1.397932                                                  NaN
…        …         …                                                      …
5348  0.596342  0.609027  Season: winter; Beach Face: 0.068; Transect: A
5349  0.339824  0.344716  Season: summer; Beach Face: 0.075; Transect: A
5350  0.619024  0.631220  Season: winter; Beach Face: 0.068; Transect: A
5351  0.300978  0.325033                                  Coin: quarter;
```

```
5352  1.632341  1.785827                                                NaN

[5353 rows x 34 columns]
```

Lets take a look at what distributions are provided from source data:

[6]:
```
given_values=np.array2string(df['Measured_Distributions'].unique()) #Find each␣
 ↪distribution in entire dataset that was provided provided in source data for␣
 ↪at least one sample
given_values= given_values[:].replace(" ",",").replace("'","").replace("[","").
 ↪replace("]","") #convert to string and remove array artefacts
given_values=(list(set(given_values.split(',')))) #extract delete duplicates (i.
 ↪e. when multiple source datasets provide the same  distribution)
given_values.remove('nan') #remove nan from list
given_values=np.array(given_values) #Turn it into an array for use later
print('Given distribution values from source data in dataset: ', given_values)
```

```
Given distribution values from source data in dataset:  ['d65' 'd10' 'd84' 'd90'
'd50' 'd75' 'd25' 'd16']
```

Create a new, blank calculated interpolation value and percent error columns for each of those distributions

[7]:
```
for d in range (0,len(given_values)):
            calc_column=str(given_values[d]+'_calc')
            error_column=str(given_values[d]+'_error')
            df[calc_column]=''
            df[error_column]=''
            d=d+1
```

0.1 This next cell is where the calculations will occur:

- In the outer most for loop, the function is iterating over each sample and accounting for the number of distributions provided in its source data.
- For the next loop within the previous one, the value and name of each distribution provided in the source data is being collected
- In the next loop,the function is one by one "hiding" a distribution from the dataset and is re-interpolated from the other distributions from the source data. This distribution is re-introduced in the next iteration, and another distribution is hidden/re-interpolated. These re-interpolated values go in the "_calc" columns.
- After that, the percent error of each re-interpolated distribution value is calculated with the distribution value from the source data

0.2 *The output will be the addition of 2 new columns distribution provided in a sample's source data, the re-interpolated value (in '_calc') and the calculated percent error (in '_error')*

[8]:
```
for z in range (0,len(df)): #loop on each sample
    if df['num_orig_dists'].iloc[z] < 3: #if the number of given distributions␣
↪is less than 3 skip the sample
        pass
    else:
        #try:
            num_orig_dists=len(df['Measured_Distributions'].iloc[z].
↪split(','))#extract amount of known distributions per sample
            given_dist_names=[]
            given_dist_vals=[]
            i=0
            for i in range(0,num_orig_dists): #find distribution values␣
↪provided in source data for each sample
                a=(df['Measured_Distributions'].iloc[z]) #extract sample's␣
↪provided distributions
                a=a.split(',')[i] #extract distribution focused on in this␣
↪iteration
                b=a.split('d')[1] #pull number value from name
                val=int(b)/100 #turn value to decimal (e.g. 90 to .9)
                given_dist_names.append(a) #collect given distribution names␣
↪from each sample
                given_dist_vals.append(val) #collect given distribution values␣
↪from each sample
            i=0

            for n in range (0,num_orig_dists): #Repeats for each distribution
                    # "deleting" the distribution name and value to be␣
↪recalculated
                    new_dist_names=np.delete(given_dist_names, n)
                    new_dist_vals=np.delete(given_dist_vals, n)
                    # "preserving" the distribution name to be recalculated as␣
↪another variable
                    focus_column=given_dist_names[n]
                    focus_column_value=given_dist_vals[n]
                    calc_column=str(given_dist_names[n]+'_calc')
                    error_column=str(given_dist_names[n]+'_error')
                    #new columns for recalculated value and error



                    grain_size_bins=[]
                    ia=0
                    for ia in range(0,(num_orig_dists-1)):
```

```
                        bin_size=df[new_dist_names[ia]].iloc[z]
                        grain_size_bins.append(bin_size)

                    grain_size_frequencies=new_dist_vals
                     #This interpolates the value using the gathered "original"␣
↪distributions from above
                    distribution = scipy.interpolate.
↪interp1d(grain_size_frequencies, grain_size_bins, bounds_error=False,␣
↪fill_value='extrapolate')
                    #This adds them to the new calculated column
                    df.loc[z,[calc_column]] =␣
↪float(distribution(given_dist_vals[(n)]))
                    df.loc[z, error_column]=abs(((df[calc_column].
↪iloc[z]-df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
        #except:
print('Error Calculation Successful!')
```

```
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
```

```
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
```

```
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
```

```
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
```

```
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
```

```
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
```

```
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
```

```
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
```

```
Error Calculation Successful!
```

```
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
```

```
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
```

```
/tmp/ipykernel_888389/2652930518.py:43: RuntimeWarning: invalid value
encountered in double_scalars
  df.loc[z, error_column]=abs(((df[calc_column].iloc[z]-
df[focus_column].iloc[z])/df[focus_column].iloc[z])*100)
```

Lets see if that worked

[9]:
```
df
```

[9]:
```
           ID    Sample_ID  Sample_Type_Code  \
0         876    SPIbeach5                 1
1         878         SPI6                 1
2         877         SPI6                 1
3        1429    SPIbeach4                 1
4        1430    SPIbeach3                 1
…         …           …                 …
5348   591703          NaN                 1
5349   591704          NaN                 1
5350   591705          NaN                 1
5351   607152  3WellsBeach                 1
5352   607153  Wells Beach                 1

                                                Project   dataset        Date  \
0                          SandSnap, image taken by:   sandsnap  2021-11-08
1                          SandSnap, image taken by:   sandsnap  2021-11-08
2                          SandSnap, image taken by:   sandsnap  2021-11-08
3                          SandSnap, image taken by:   sandsnap  2021-11-08
4                          SandSnap, image taken by:   sandsnap  2021-11-08
…                                                  …       …          …
5348                                             NaN     UMASS  2016-03-10
5349                                             NaN     UMASS  2015-08-27
5350                                             NaN     UMASS  2016-03-10
5351  Brian McFall, sandsnap training round 1 March …  sandsnap  2019-09-23
5352            SandSnap, image taken by: rose.dopsovic  sandsnap  2021-09-14

     Location_Type  latitude  longitude                  Contact  …  \
0          Beach?Y  26.12871  -97.16718          Sandsnap, USACE  …
1          Beach?Y  26.12899  -97.16713          Sandsnap, USACE  …
2          Beach?Y  26.12899  -97.16713          Sandsnap, USACE  …
3          Beach?Y  26.16883  -97.17248          Sandsnap, USACE  …
4          Beach?Y  26.16885  -97.17284          Sandsnap, USACE  …
…            …       …        …                        …  …
5348    berm crest  42.87166  -70.81624  Jonathan Woodruff, UMASS  …
5349          dune  42.87166  -70.81624  Jonathan Woodruff, UMASS  …
5350          dune  42.87166  -70.81630  Jonathan Woodruff, UMASS  …
5351       Beach?Y  43.30194  -70.56639          Sandsnap, USACE  …
5352       Beach?Y  43.30194  -70.56639          Sandsnap, USACE  …
```

```
      d90_calc d90_error  d50_calc  d50_error  d75_calc d75_error  d25_calc  \
0     1.208305   5.37509  0.790204   0.606737  1.016927       0.0  0.623835
1     1.208305   5.37509  0.790204   0.606737  1.016927       0.0  0.623835
2     1.208305   5.37509  0.790204   0.606737  1.016927       0.0  0.623835
3     1.208305   5.37509  0.790204   0.606737  1.016927       0.0  0.623835
4     1.208305   5.37509  0.790204   0.606737  1.016927       0.0  0.623835
…        …        …        …         …        …        …
5348
5349
5350
5351  0.284545  5.460065  0.216564   0.484626  0.253463       0.0  0.190177
5352  1.545688  5.308484  1.014601   0.793493  1.301861       0.0  0.804688

     d25_error  d16_calc d16_error
0      0.18256  0.561042  0.815872
1      0.18256  0.561042  0.815872
2      0.18256  0.561042  0.815872
3      0.18256  0.561042  0.815872
4      0.18256  0.561042  0.815872
…        …        …        …
5348
5349
5350
5351   0.30392  0.176464  2.534656
5352  0.255917  0.724966  0.960215

[5353 rows x 50 columns]
```

0.2.1 Write to file

Finally, define a csv file name for the output dataframe

[10]:
```
output_csvfile='../data_interp_error.csv'
```

write the data to that csv file

[11]:
```
df.to_csv(output_csvfile) #convert data to CSV
```

[ ]:
```

```