Dataset Natural History Museum

INITIAL STEPS

Start by implementing this instruction into CoLab

1. import numpy as np

import pandas as pd

Naming this CSV file in CoLab

chd=pd.read_csv("resource.csv")

This is in place in order to use "chd" instead of typing out the long address.

To get the first 4 rows

Type out this in CoLab

chd [0:4]

The chart will show

-order -phylum -scientificName -specificEpithet -subfamily -subgenus -suborder -superfamily -taxonRank -type

You can also get down to one category you're focusing on
_This is important to narrowing down specific interests _

For example, if you just want to find PHYLUMS:

chd ["phylum"][0:4]

---

*You **can** combine them*

Making a broader search

Sticking with only 4 rows as results

- chd [["phylum" , "scientificName" , "specificEpithet" , "subfamily"]] [0:4]

Only finding NaN (specific column)

1. chd.iloc[:,3]
2. This will ONLY narrow down the columns, there are many rows still
3. 0 NaN 1 NaN 2 NaN 3 NaN 4 NaN ... 824749 NaN 824750 NaN 824751 NaN 824752 NaN 824753 NaN

Finding Taxon Rank

chd[chd["taxonRank"]== "TR"]

This is another specific search option.

This is important because it adds another aspect in the process of narrowing the data down and identifies the #s associated with just TR

Links (for closer inspection on the process in CoLab)

https://colab.research.google.com/drive/1MXLZo8RTayuao2tWTdAyKFj_8XLNjMq2#scrollTo=aZJmjCQHxxQd