30/01/2020

ENDSPE project preparation

Github repo : https://github.com/ddonatien/endspe

Data sources

We picked datas about endangered species from different sources :

Nature Serve's datasets are about biodiversity in North America. We can submit a form to be granted an access so their datasets, which gathers information about localization of high biodiversity regions and imperiled species.
https://www.natureserve.org/conservation-tools/national-species-dataset?fbclid=IwAR3JqHRPLvo5jGIPQdImiJqr8SlYNPySZhiz4Y6krtfzDBG4T-uh4GtXtuE

Endangered Species Act of US DATA.GOV, gathering various datasets about our topic.
https://catalog.data.gov/dataset/endangered-species-act?fbclid=IwAR38W4n7gARy91fpHnMfM9z7z14pjQUqGetNL-M1oa4qSeTJxyao6Qbzvgs

US Department of interior, "Critical Habitat for Threatened and Endangered Species Dataset"
https://data.doi.gov/dataset/fws-critical-habitat-for-threatened-and-endangered-species-datasetd55fc

European « Red List », that is the dataset we used for our tests.
https://data.europa.eu/euodp/en/data/dataset/data_european-red-lists-4

Data cleaning

We exported European Red List's dataset to .csv, and we read it with Pandas (Python library), in a Jupiter environnement (cf Data Exploration.ipynb in our Github repository). We explored our dataset by ploting some charts to visualize data attributes.