Course Objectives

In this course, we are going to focus on **three** learning objectives:

1. How to visualize and understand geographical data in an interactive way with Python.
2. How the K-Means algorithm works, and some of the shortcomings it has.
3. Density-based clustering approaches, and how to deal with any outliers they may classify.

By the end of this course, you will be able to **take raw geographical data, and cluster it effectively using basic or more advanced density-based clustering techniques -- you will also learn about how to determine the strength of a given clustering. You will learn how to plot this data in an interactive way within Python that can be exported easily.**

Course Structure

This course is divided into 3 parts:

1. Course Overview: This introductory reading material.
2. **Clustering Geolocation Data Intelligently in Python:**This is the hands on project that we will work on in Rhyme.
3. Graded Quiz: This is the final assignment that you need to pass in order to finish the course successfully.

Project Structure

The hands on project on **Clustering Geolocation Data** is divided into following tasks:

**Task 1:** An introduction to the problem, as well as basic exploratory data analysis and visualizations

**Task 2:** Visualizing geographical data in a more meaningful and interactive way

**Task 3:** Methods of evaluating the strength of a clustering algorithm

**Task 4:** Theory behind K-Means, and how to use it for our problem

**Task 5:** Introduction to density-based clustering approaches, and how to use DBSCAN

**Task 6:** Introduction to HDBSCAN, to alleviate constraints of classical DBSCAN

**Task 7:** A simple method to address outliers classified by density-based models

### WHAT YOU WILL LEARN

* Clean and preprocess geolocation data for clustering
* Visualize geolocation data interactively using Python
* Cluster this data ranging from simple to more advanced methods, and evaluate these clustering algorithms