**Welcome!**

Welcome to **Mining Data to Extract and Visualize Insights in Python**. This is a project-based course which should take approximately 1 hour to finish. Before diving into the project, please take a look at the course objectives and structure:

**Course Objectives**

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

1. By the end of Task 1, you will be able to load a dataset and extract basic information using Python.
2. By the end of Task 2, you will be able to learn various ways to clean your dataset.
3. By the end of Task 3, you will be able to visualize patterns and outliers that may be present in your dataset.
4. By the end of Task 4, you will be able to calculate and visualize the correlation between different numeric columns.
5. By the end of Task 5, you will be able to cluster your dataset to identify similar groups.
6. By the end of Task 6, you will be able to visualize your dataset using principal component analysis (PCA).

By the end of this course, you will be able to **clean your dataset, identify similar groups, and create several types of visualizations to present insights on your dataset using Python.**

**Course Structure**

This course is divided into **3** parts:

1. Course Overview: This introductory reading material.
2. **Mining Data to Extract and Visualize Insights 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 **Mining Data to Extract and Visualize Insights in Python** is divided into following tasks:

**Task 1: Load a dataset and extract basic information using Python**

**Task 2: Learn various ways to clean your dataset**

**Task 3: Visualize patterns and outliers that may be present in your dataset**

**Task 4: Calculate and visualize the correlation between different numeric columns**

**Task 5: Cluster your dataset to identify similar groups**

**Task 6: Visualize your dataset using principal component analysis (PCA).**