**Welcome!**

Welcome to "Intermediate Pandas Python Library for Data Science". 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 these learning objectives:

1. Splitting of data into subsets.

2. Imputation and interpolation of missing values.

3.  Sorting by columns and ranges.

4. Selecting and slicing data.

5. Joining, Concatenation & Advanced Joins

6. Grouping, calculation of aggregate and binning.

By the end of this course, you will improve your efficiency when using Pandas**.**

**Course Structure**

This course is divided into 4 parts:

1. Course Overview: This introductory reading material.
2. About the dataset " What Does Aid to Africa Finance? " : This section contains information about the real world dataset that you will be working with. To save time, during the guided project, I only explain the variables that we work with. However, you can explore all the variables using the variable description file linked within.
3. "Intermediate Pandas Python Library for Data Science" **:** This is the hands on project that we will work on in Rhyme.
4. 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 "Intermediate Pandas Python Library for Data Science" is divided into the following tasks:

Task 1: Splitting of data into subsets.

Task 2: Imputation and interpolation of missing values.

Task 3: Sorting by columns and ranges.

Task 4: Selecting and slicing data.

Task 5: Joining, Concatenation & Advanced Joins

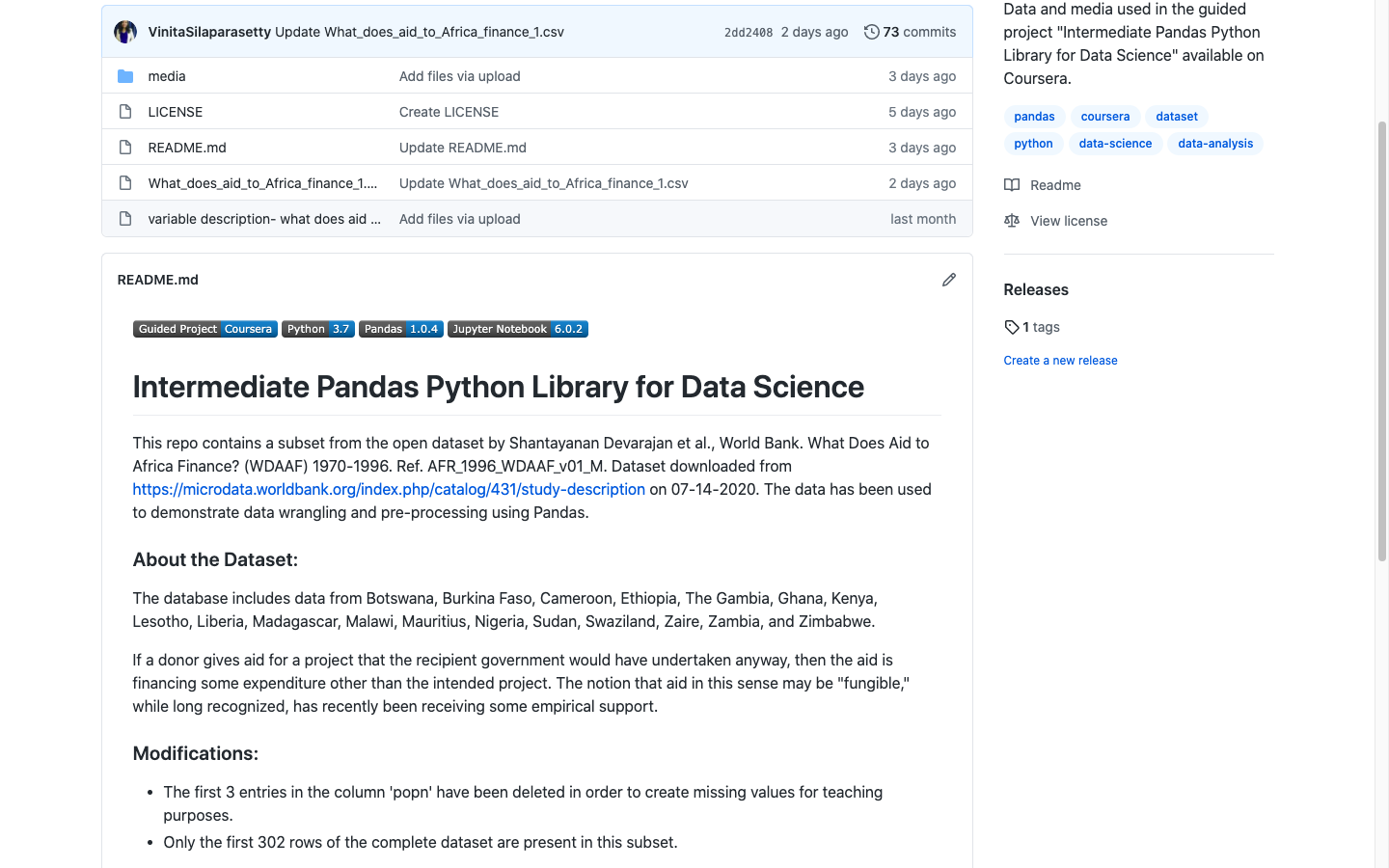
Task 6: Grouping, calculation of aggregate and binning

**Meet the Instructor**

Hi! I am Vinita Silaparasetty, a certified data science instructor specializing in Python and R for machine learning as well as Tensorflow 2 and Keras for Deep Learning. I am the author of the book " Deep Learning Projects using TensorFlow 2, Neural Network Development using Python and Keras " published by Apress, New York.

**About the Dataset**

***Note:*** *During the guided project, I explain only the variables that we work with. However, feel free to explore all the variables using the variable description file which is available* [*here.*](https://github.com/VinitaSilaparasetty/Coursera-Intermediate-Pandas/blob/master/variable%20description-%20what%20does%20aid%20to%20africa%20finance.pdf)



This guided project uses a subset from the open dataset by Shantayanan Devarajan et al., World Bank. What Does Aid to Africa Finance? (WDAAF) 1970-1996. Ref. AFR\_1996\_WDAAF\_v01\_M. Dataset downloaded from <https://microdata.worldbank.org/index.php/catalog/431/study-description> on 07-14-2020. The data has been used to demonstrate data wrangling and pre-processing using Pandas.

**About the Dataset:**

The database includes data from Botswana, Burkina Faso, Cameroon, Ethiopia, The Gambia, Ghana, Kenya, Lesotho, Liberia, Madagascar, Malawi, Mauritius, Nigeria, Sudan, Swaziland, Zaire, Zambia, and Zimbabwe.

If a donor gives aid for a project that the recipient government would have undertaken anyway, then the aid is financing some expenditure other than the intended project. The notion that aid in this sense may be "fungible," while long recognized, has recently been receiving some empirical support.

**Modifications:**

* Three entries in the column 'popn' have been deleted at random, in order to create missing values for teaching purposes.
* Only the first 302 rows of the complete dataset are present in this subset.