**Approach 1**:

**connect\_populate\_db.py** - Script to connect and populate database created on AWS

To compile this script through command prompt- **python connect\_populate\_db.py**

**Approach 2:**

1. First step is to create database shelve

**dbshelve.py**- Script to create a persistent, dictionary-like object to store the images

To compile this script through command prompt **–**

**python dbshelve.py --dataset dataset --shelve db.shelve**

Description:

* + --dataset dataset: this indicates the local folder where the images will be stored
  + --shelve db.shelve: is the output dataset using this Python shelve library

1. Second step is to create the dataset of images from local folder

**create\_dataset.py** – Script to create dataset of images and also resize or adulterate it with filters

To compile this script through command prompt **–**

**python create\_dataset.py --input faces1 --output dataset**

Description:

* --input faces1: faces1 is a local folder which contains images
* –output dataset: dataset is a local folder wherein images will be stored as the output of this script

1. Third and last step is to calculate hash of images present in the dataset and then compare the hash of images with that of query image to search for similar/almost similar images.

**hash.py**- Script to calculate hash of images and retrieve similar images

To compile this script through command prompt **–**

**python hash.py --dataset dataset --query faces1/308.jpg**

Description:

* --dataset dataset: this will let script to access dataset to calculate hash of all the images present in the dataset
* --query faces1/308.jpg: this is the query image for which hashes will be compared to retrieve all the similar images