**DOCUMENTATION**

This program has 3 tasks, and the 2nd and 3rd tasks are built on the 1st one. So the function that will be used in the 1st task will be used in 2nd and 3rd as well, similarly functions used in task 2 will be used in task 3.

* In Task one, there are functions like **readCSVfiles()** which reads the data from 10 csv files into tables (3 tables- one for sensor A, one for sensor B and one which has combined data of both sensors); **drop\_tables()** which drop tables created every time program runs to recreate tables; create\_tables() that creates tables; three **insert\_tables()** functions to inserts data into 3 tables created and last one **select\_merged\_tables()** to display all the rows from the merged table.
* In Task two, there are functions like **create\_frames()** that reads data from tables in the data base and creates data frames and subset of dataframe to have one for sensor A and one for sensor B and calculates average temperature and humidity across the time period; **operations\_perday()** that groups the data frame by date and calculates average temperature and humidity for each group for sensor A as well as sensor B; **user\_time\_period()** that takes a start and end time period from user and if valid input and within the range then calculates average temperature and humidity for both sensors in that time period; **operations\_whole()** calculates average temperature and humidity throughout the entire period including the data for both sensors together.
* In Task three, there are functions like **create\_frames()** that reads data from tables in the data base and creates data frames and subset of dataframe to have one for sensor A and one for sensor B; **operations\_perday()** that groups the data frame by date and calls the function user\_plot\_time\_period() that takes a start and end time period from user and if valid input and within the range, asks users for what sensor should visualization be performed and what kind of visualization should be performed, and keeps asking for a valid input till users enters one and then plots respective graph for the selected sensor within the selected time period.

This is how the program holistically performs, and all the functions used have been explained.

* Now, this program can be used in a lot of different scenarios. Since, this program crates a new database, inserts data from 10 csv files into the tables and performs statistical analysis on the structured data, this program can be used for analyzing and kind of structured data, create database and tables and insert data into the tables. Also, it can be used to perform different kind of analysis on structured data and also do visualization analysis on the structured data.
* This program also can be used to create subsets of the data extracted from the tables into the data frames and perform analysis on a selected sub set of data. Further this program uses panda library to conduct analysis and matplotlib to do visualization, it has a scope to incorporate numpy as well to do analysis on data.