ECS 116
Databases for Non-Majors

Discussion 5
(5/9/24) |Spring '24

Today's Agenda

1. Quick Updates
2. Using Jupyter notebook to manipulate data using Py Pandas
3. Creating functions, benchmarking, comparative analysis

Quick Updates

Programming Assignment 2

- Part 1 and 2 are online.
- Go to files->Prog-Assmt-2

File naming convention

- Focus on the file naming for submission

Files

Today's Files:

1. Download the folder: *Files->Juypter Notebooks->DISC_5_FILES*
2. Extract the files and open on Anaconda
3. We will use the Boilerplate: *DISC-5-BOILERPLATE-v01.ipynb*
4. Completed notebook: *DISC-5-MAIN-v01.ipynb*

Benchmarking and Comparative Analysis

An index offers an efficient way to quickly access the records from the database files stored on the disk drive. It optimizes the database querying speed by serving as an organized lookup table with pointers to the location of the requested data.

Database benchmarking is a well-defined, proven technique for comparing and analyzing how databases or database management systems (DBMS) perform.

Average Performance Comparison With and Without Benchmarking
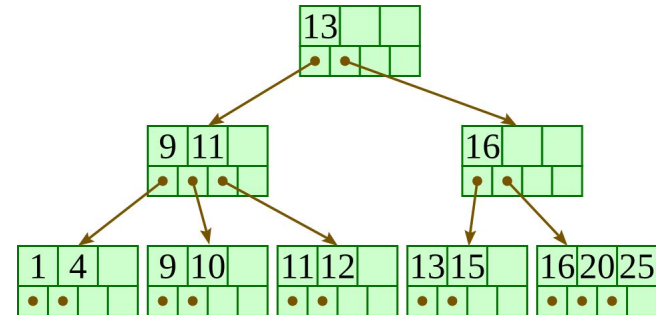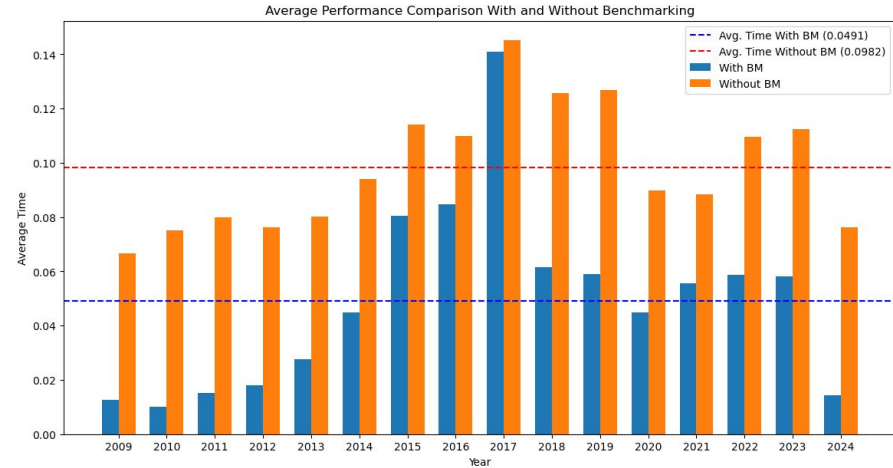
How does it all work?

1. Open Jupyter Notebook

2. Import required files/packages

3. Create connections (env)

4. Manipulate Data

5. Create functions

6. Benchmark

7. Visualize

Working on Jupyter Notebook

1. Download the BOILER PLATE code
2. Open Jupyter Notebook through the Anaconda navigator
3. In JN, open the downloaded folder
4. Go through the boiler plate code
5. Work on the data using Pandas and Matplotlib

What did we learn today?

1. Quick Course Updates
2. Benchmarking in jupyter notebook
3. Visualization using the data
4. Exporting the file

Thank You!

See you next Thursday!