David Schwartzman

12/07/2024

Assignment 6 Report

The indexes I created are designed to optimize query performance by targeting specific query patterns.

1. **Index on account_type and balance:** This index improves performance for range queries (for example: WHERE account_type = 'savings' AND balance > 20000) by organizing rows based on account_type and sorting balance within each type. Without this index, the database performs a full table scan, which is much slower for large datasets.
2. **Index on branch_name and balance:** This index optimizes point queries targeting specific branches and balances (for example, WHERE branch_name = 'Perryridge' AND balance = 5000). The index enables quick lookups by matching both branch_name and balance directly. Without it, the database scans the entire table to find matching rows, which increases execution time.
3. **Index on branch_name and account_type:** This index is beneficial for queries filtering by branch and account type (e.g., WHERE branch_name = 'Downtown' AND account_type = 'savings'). It organizes the data to quickly locate rows with the given branch and type combination. Without the index, the database must evaluate each row individually, significantly increasing execution time.

As we can see from the table below, every query, without any exceptions, was faster when we used the appropriate index for that specific query.

Query Type	Description	Dataset Size	Index Type	Microseconds
Point Query 1	Finds the number of accounts that are in the “Downtown” branch and are savings an account.	50,000	With index on branch name and account type	16,771
			Without index	32,619
Point Query 2	Finds the number of accounts that are in the “Perryridge” branch and have a balance of exactly 5000$	50,000	With index on branch name and balance	187
			Without index	33,335
Range Query 1	Finds the number of accounts that are a savings account and have a balance greater than 20,000	50,000	With indexes on account type and balance	14,192
			Without index	21,611
Range Query 2	Finds the number of counts that are a savings account and have a balance between 5000$ and 20000$	50,000	With index on account type and balance	19,656
			Without index	26,026
Point Query 1	Finds the number of accounts that are in the “Downtown” branch and	100,000	With index	36,103

	are savings an account.			
			Without index	196,449
Point query 2	Finds the number of accounts that are in the “Perryridge” branch and have a balance of exactly 5000$	100,000	With Index	175
		100,000	Without index	41,488
Range query 1	Finds the number of accounts that are a savings account and have a balance greater than 20,000	100,000	With index	24,849
		100,000	Without index	37,262
Range query 2	Finds the number of counts that are a savings account and have a balance between 5000$ and 20000$	100,000	With index	38,045
			Without index	48,727
Point query 1	Finds the number of accounts that are in the “Downtown” branch and are savings an account.	150,000	With index	54,229
			Without index	76,501

Point query 2	Finds the number of accounts that are in the “Perryridge” branch and have a balance of exactly 5000$	150,000	With index	3,397
			Without index	65,616
Range query 1	Finds the number of accounts that are a savings account and have a balance greater than 20,000	150,000	With index	36,306
			Without index	55,348
Range query 2	Finds the number of counts that are a savings account and have a balance between 5000$ and 20000$	150,000	With index	57,280
			Without index	72,954