Project Planning and Risk Analysis

Project Scope

The project helps to overcome potential constraints associated in establishing the business obtained from the numerical metrics and supports the concept of entrepreneurship by harnessing costs, miscellaneous expenditure to resource efficiency aided by an efficient management and analysis system to effectively impact and enhance Canada's growth in the economy drastically affected by Covid-19 that would preferably specify a competitive platform for emerging business start-ups over the next few years.

Entrepreneurship is seen as one of the driving factors behind creativity, profitability, employment opportunities, and economic progress in Canada. Countries with higher level of entrepreneurial activity are often economically prosperous. Investors have had a tremendous impact on Canadian civilization, and today, an increasing number of Canadians from across all kinds of backgrounds are becoming or are contemplating becoming entrepreneurs. Canadian entrepreneurs are well-liked in their neighborhoods and in the media, and at a time when many celebrities are viewed with distrust, they are arising as our new mentors, motivating youths to consider entrepreneurship as a prospective career.

Furthermore, knowledge understanding regarding management's output will assist to better forecast future objectives and provide the groundwork for their achievement. Market competitions between firms may be evaluated in comparison using several indicators. As a result, it promotes small-scale start-ups to take the first step and presents strategies for competing with large scale commercial companies. As a result, this will instill motivation in the minds of aspiring tech entrepreneurs, allowing them to profit from inferring analyses for the successful expansion of their firms.

The team’s key contributions towards the solution’s outcome will potentially employ advanced database concepts such as query optimization, normalization on large scale employee records to derive individual performance efficiency based on query given parameters where the queried results will be generated as a report via email using a frontend UI that can be viewed by managers for further evaluation. The SQL query execution plans can be viewed and analyzed using an external software ApexSQL by importing the query file into the software tool. Originally, the team will effectively develop a custom dataset with curated related set of tables with python script to produce large-scale data. The database fields will be created in the database server (MS SQL Server). The queried results will be generated in the shell terminal as a performance report to managers with a frontend user interface.

Possible Solutions and Design Alternatives

The act of examining SQL queries and selecting the most efficient execution mechanism is known as SQL query optimization which is the significant advance database concept involved in retrieving employee record. It is usually a trial-and-error procedure in which many queries are examined to discover which one provides the greatest performance while still delivering the desired data. Although query optimizers are occasionally included into database management systems, external, third-party solutions are frequently regarded to provide better performance results. For each query, a typical query optimizer will create one or more query plans, each of which provides a method for running the query. Each query plan's performance (i.e., execution time) is measured, and the most efficient query plan is chosen and utilized to execute the query.

The database tables can be imported to perform query optimization which can be implemented using various techniques by employing several software tools to serve as possible solutions is as follows;

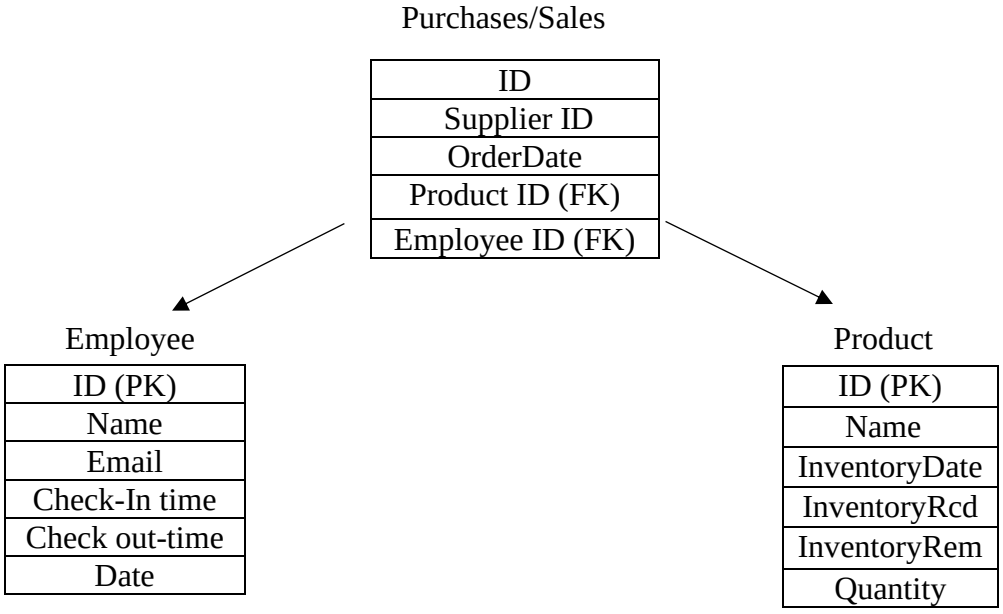
SolarWinds Database Performance Analyzer (DPA) is a database management tool designed for monitoring, analyzing, and improving SQL query performance. It analyses your SQL database instances and, using an Answer Time Analysis approach that focuses on the time between a query request and the associated database response, examines wait kinds and events to locate database bottlenecks. It also enables to view database instances, wait times, query advice, CPU (with warning and critical alerts), memory, disk and sessions.

SQL Query Tuner uses database profiling to rapidly and simply identify SQL queries that cause poor database performance and displays a graphical representation of wait time analysis. It will continuously profile an entire data source over a specified time period.

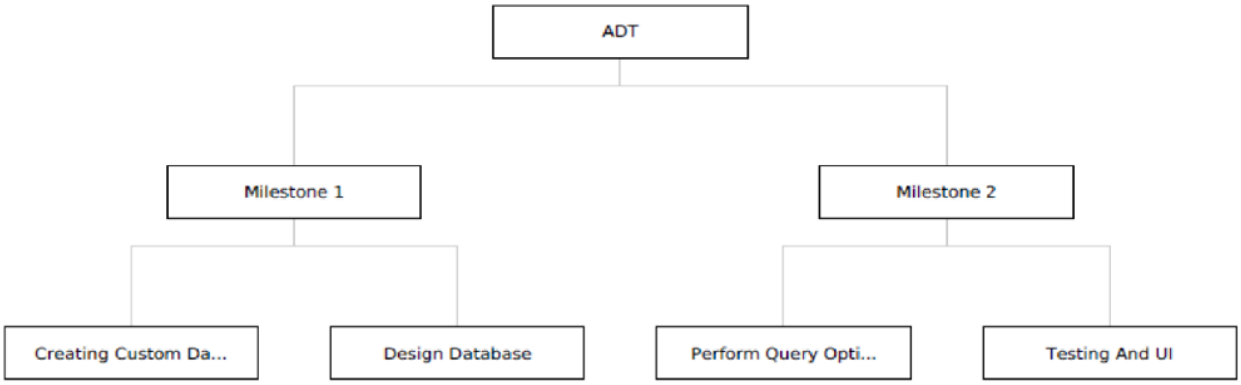
SQL Server Database Engine Tuning Advisor (DTA) can be used for debugging a specific issue query's performance, adjust a large group of queries across one or more databases, do an exploratory what-if examination of prospective physical design modifications, and handle storage space.

The programming language that the project will involve is Python. Since the project includes analysis as one of the key components, it provides support to data visualization and statistical computation of data. Also, it has fewer syntactical keywords, increased code readability, faster processing speed and facilitates automation of tasks. The database management choice will be relational DB - MS SQL Server as the concept of query optimization can be better demonstrated with set of relational tables and structured data. In addition, ODBC can be made between SQL server and Python using Pyodbc that can further be explored for an insightful analytical study.

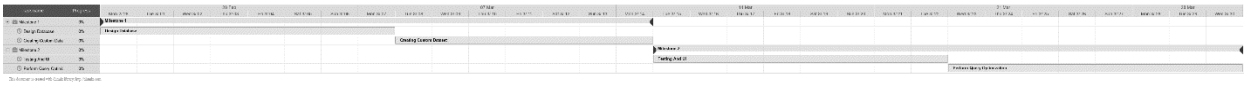
Initial Technical Design



Work Breakdown Structure (WBS)



Gantt Chart



Link: For reference, https://drive.google.com/file/d/16GSnWcbCr6Lg-xuinez2v--r1yC1P0nN/view?usp=sharing

Software Development Methodology

The software development methodology that is used to develop this project is waterfall model. The waterfall model is a software development methodology that emphasises a linear flow from the start of a project to its completion. The waterfall model is the most appropriate for this project since each step of the project necessitates the completion of the phase before i.e. (each phase is dependent on the completion of its previous phase).

Individual Contribution

JIRA is used to keep track of all of these individual contributions. Also, JIRA will be able to see all of the reports about specific team members as well.

Team Details

Time Frame	Team Member	Duration	Start Date	End Date
Custom Dataset	Vishal Jayaraman	5 days	02/28/2022	03/07/2022
Database Design	Eswaran Badrinarayanan Venkateswaran	5 days	03/08/2022	03/14/2022
Query optimization	Hani Pankajkumar Bhavsar	5 days	03/15/2022	03/21/2022
UI and Testing	Hani Pankajkumar Bhavsar	5 days	03/22/2022	03/28/2022

Risk Identification

These are some of the possible and potential risks that were discovered at the original proposal stage of the project and are listed below.

- It is possible that the custom data that was produced for the project will not cover all of the edge scenarios.
- Throughout the database design process, it is essential that the data be normalized. This helps the database in adhering to the criteria of ACID properties. A significant risk exists that the database will not comply to the ACID features if the normalization process is not performed properly.
- Because the major purpose of the project is to create query optimization, the volume of data generated is enormous, as is the number of participants. There is a high possibility that poor indexing is performed, which may have an adverse effect on performance. This might result in queries that are inefficient.

Risk Mitigation

These are the solutions that will be used during the project's development in order to mitigate the risks that have been identified.

- The majority of the edge cases are discovered during the testing phase of the project's development. Finding the edge cases requires both creativity and critical thinking. End-to-end testing is also used in order to identify edge situations.
- The database should ensure that it has achieved the maximum level of normalization possible for the database in interest. This will prevent the occurrence of data redundancy from occurring. It will also ensure that the data dependencies are logical. In the end, this will guarantee that the database complies with the ACID properties of the database.
- Poor indexing is defined as any SQL Server table setup in which performance suffers as a result of an excessive number, an incorrect number, or a lack of indexes. Incorrect indexing techniques include creating an index on a column that does not allow for simpler data manipulation, or creating an index on many columns that, instead of speeding up queries, slows them down. A way to solve this problem is to index the right columns or choose the right columns to index, which will help you make an ideal clustered index.

GitHub Repository Link

https://github.com/EswarVenkat2070/ADT-Project-Query-Optimization-of-Employee-Performance-in-SCM.git