**Project Overview – Road Database and Analysis**

Hi, I’m Richard Anstice. The goal of this project is to demonstrate my skills in SQL and Power BI.

I have created a database and then used SQL and Power BI to analyse the data.

All the code and data in this project is mine from scratch.

Previously, I completed an introduction to SQL at University (Oracle SQL), including a basic database design and build. Then, I learned more advance SQL (SQL Lite) and Power BI via a Dataquest online course. This project is a refresh of my Uni project, to show my new skills.

The database is a fictional roading database for a city council to identify its roads, it’s people, and to manage budgets and expenses for projects and contracts for works on those roads.

The data has been created by me, although the roads in the database are real roads in West Auckland NZ (the coordinates work in Google Maps).

I have chosen to use Microsoft SQL Server because it is free, it has the server-client paradigm (unlike SQLite), and it works well with Power BI.

This project:

* Converts my university project from Oracle to SQL server,
* Adds separate payment table to allow analysis of payments by aggregate, including by date,
* Adds more advanced select statements, from my online learning,
* Adds reporting in a Power BI report and a Dashboard.

**Overview of files in this Github Project**

**First part:** This introduction, plus the ERD for this database

**Second part:** The SQL Server Queries to create and populate tables, originally written for Oracle

.csv of payment records, and script to bulk-load from csv

.bak file of the functioning SQL Server database

**Third Part:** SQL Query file with all Select queries, numbered 1 to 14

Word Document reporting on those queries 1 to 14, explaining the business need for the query, and results including manual graphs from Excel where relevant.

**Fourth Part:** Power BI Report and Dashboard based on database

Excel file with all raw data, so a user can play with the report without having to access the database.

**Overview of Database**

The ERD (drawn in Visual Paradigm) is attached below, and as a separate image.

**Key features in the tables are:**

**Locations**: Points of Latitude and longitude from Google Maps that can be used to identify the start and end of a road.

**Roads**: Roads can be ‘inside’ another road – reflecting that part of a physical road can have special status as a memorial or town-centre.

**Projects**: One or more roads can be linked to one or more projects. Key people and dates are listed

**Contracts**: Each contract is linked to one Project, although a project can have more than one contracts attached. Budgets are listed at the contract level (i.e. a Project budget must be aggregated as the sum of contract budgets). Key people and dates are listed.

**Descriptions**: There are separate tables providing more detailed descriptions for contracts and projects. They are in separate tables just so that the large text strings are out of the way until they are needed.

**Contractors**: A contract must have one contractor. The contractor table identifies the name, key contact and contact details for contractors.

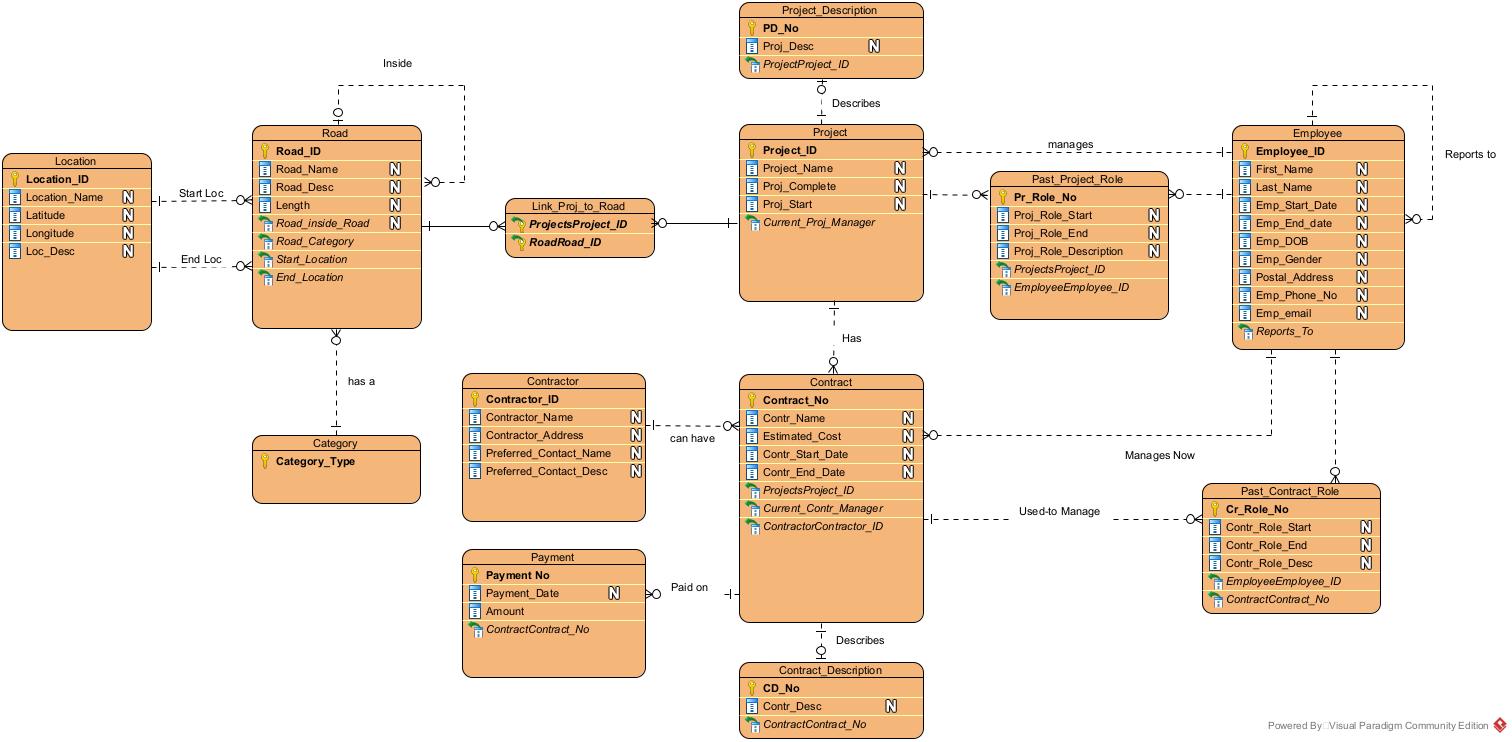
**Payments**: A key upgrade from my University project was to add a separate payment table. Each payment is linked to a contractor, with a payment date. Therefore, the actual spend on a contract is the aggregate of the payments made for the relevant contract.

**Employees**: There is an employee list. It also shows the employee’s direct report. Employee data is used in Project and Contract tables to show the current manager, and also may be used to show past roles for that employee.

**Past roles**: The original assignment asked for a sequence of listings for past-roles for employees in contracts or projects. In this project, I have limited this to one employee – there is a question mark about the probity of that employees actions paying funds out, while he was in past roles.

**Lessons Learned and Future opportunities**

* It is a quirk of Oracle SQL to need a lot of individual INSERT statements. After I had adapted my Oracle code to SQL Server, an experienced DBA explained to me that other forms of SQL have more efficient commands to add large amounts of data.
* Database would work better with a standard date table – added in Power BI reporting. E.g. would have allowed calculations for whole financial years, and accounting for months where there are no payments in that month.

****

**ERD for Fictional Road Database**