**PROJECT PROPOSAL**

**PROJECT OVERVIEW:**

The main focus of this project:

* To design a data warehouse model for the business model described in this project.
* Identify available open source database to use for the project.
* Create a ETL plan to extract, load and transform this data from the database into data warehouse schema.
* Finally load this sample data into the data warehouse schema and create some visualization dashboard to represent analytics.

**BUSINESS DESCRIPTION:**

PROBLEM STATEMENT:

The main business model I am targeting via this project is the manufacturing. Primarily in manufacturing business the multiple entities of the business are tied in and interlinked i.e. starting from employee records to purchase orders and online transactions are very important for the success of the business.

In this business of manufacturing of cycles by Adventure Works marketing/sales analytics division would like to closely analyze product specific data. Data of materials, billing and sale per category is available in the database and there is no existing model to identify the needed dimensions and schema to load them into data warehouse.

SOLUTION:

We aim at designing a data warehouse model with the available database of all online transactions. Down the lane this will be used for various analytics activities to significantly improve any business-related issues as well as understand market trend worldwide and how the business and transactions fluctuate as a function of different entities.

**DATABASE FINALIZE:**

In this project I plan to use the online open source Adventure Works Database available from Microsoft.

Database: <https://github.com/Microsoft/sql-server-samples/releases/tag/adventureworks>

Regarding the Database - all the data generated and stored is in regards to Adventure works bicycle manufacturing company. Database consists of many tables (<https://technet.microsoft.com/en-us/library/ms124438(v=sql.100).aspx>) relating to Customer vendor and employee details, Product specifics, transactions and ordering specifics, billings details etc.

**INFORMATION NEEDS:**

* Needed data from the database would be of Customer related tables, Product tables, Sales Tables, Purchase, vendor, and store tables.
* This information will be extracted, filtered and loaded into schema and used for visualization.
* The research and analytics process down the lane can use the data to understand business trends and possibly correlate different factors contributing towards sales.