

MET CS 689 Designing and Implementing a Data Warehouse - Term Project Update 3

Hicran Arnold

April 9, 2023

Part 1: Are you working on your own or with a partner? If with a partner provide their name. If on your own, just state that this is the case.

I am working alone in this project

Part 2: Determine the project scope

- In a short paragraph, describe the topic you wish to explore --an update if any
- Update the five business questions that your data warehouse will answer.

I have added the project objective part

Project Description The topic I wish to explore for my data warehouse project is the management of Parts Unlimited's EV parts business. This topic was inspired by "The Unicorn Project," which describes the challenges and opportunities of digital transformation in a large organization. Specifically, I plan to focus on storing and analyzing data related to charging stations, EV products, customer purchases, and geographic location. Parts Unlimited already sells EV parts, and my goal is to improve the organization's data management, reporting, and analysis capabilities related to this business.

Project Objective The primary objective of this project is to create a concept data warehouse project using a slowly changing dimension approach. I will identify the business requirements by formulating specific business questions that the data warehouse will help answer. Based on these requirements, I will develop a consolidated ERD schema for the data warehouse.

I will focus on storing and analyzing data related to EV charging stations, product information, customer purchases, and geographic location. To accomplish this, I

will use a slowly changing dimension approach to ensure that data is properly tracked over time. Specifically, I will use SCD Type 2 to track changes to the customer's location, SCD Type 3 to track changes to the EV product information, and SCD Type 1 to track changes to the charging station information.

Finally, I will use Tableau and Postgres SQL to create reports that will help the business make data-driven decisions based on the data stored in the data warehouse. By using a consolidated ERD schema and slowly changing dimensions, I will ensure that the data warehouse accurately reflects the business needs and can be used to provide valuable insights to the organization.

Business Questions:

Business Question 1

Is there a correlation between the number of EV charging stations in a particular area, the number of EV cars registered in that area, and the time period in which they were registered? And if so, how can we use this information to optimize our expansion strategy and better serve our customers over time?

Parts Unlimited Business Development team is interested in using this information to plan and prioritize their expansion strategy for EV parts and charging station installations. The marketing team can use this information to tailor their marketing campaigns and promotions to specific regions.

Business Question 2

What is the current range of EV parts manufacturers and their product offerings that Parts Unlimited is working with, and how can this information be leveraged to optimize their product mix and pricing strategy for increased revenue in the growing EV market? Additionally, how has this range of manufacturers and their product offerings evolved over time, and what trends can be identified for future business planning?

Purchasing team is interested in this question, they are responsible for sourcing and procuring products for Parts Unlimited, including EV parts from different manufacturers. By understanding the current range of EV parts manufacturers and their product offerings, the Procurement department can make informed decisions about which manufacturers to work with, what products to stock, and how to optimize their product mix and pricing strategy for increased revenue in the growing EV market.

Business Question Q3

How does the popularity of different EV models and plug types vary by geographic region, and how can Parts Unlimited use this information to target their marketing and sales efforts?

The marketing and sales department in Parts Unlimited can use the information to target their efforts more effectively. For example, if a certain geographic region shows a higher preference for a particular EV model or plug type, the marketing and sales team can focus their promotional activities and campaigns in that region to increase sales. They can also use this information to tailor their messaging and product offerings to better meet the needs and preferences of customers in each region.

Business Question Q4

What is the relationship between the location and price of existing EV charging stations over time, and how can this information be used to determine the feasibility of adding new charging stations in the vicinity of Parts Unlimited's stores in partnership with companies like Tesla? Additionally, how can this information be leveraged to increase revenue and customer convenience?

Parts Unlimited is considering providing EV charging stations as a new service. Therefore the Research and Development team can determine if there is a relationship between price and location. This information can provide insights into the cost and demand for EV charging stations in different locations and inform the decision on where to install new stations.

Business Question Q5

What are the most popular types of EV charging plugs for different model years, and how have these trends changed over time? How can Parts Unlimited leverage this information to ensure that they stock the appropriate parts for EV charging stations and stay ahead of industry trends?

The Supply Chain department at Parts Unlimited would need this information to know the most popular types of EV charging plugs for different model years, as well as how these trends have changed over time, in order to ensure that they stock the appropriate parts. By analyzing this information, they can optimize their inventory management, pricing strategy, and product mix to meet the changing demand and stay ahead of industry trends.


Part 3: Data Sources

Provide two data sources you will be using, for each data source list the number or columns and rows that are in each data source. Provide a header and first 5 rows from each source.

- What is the URL or location of the data?
- What information does this data provide that will help answer one or more of the above questions?
- Do you see any issues in the data that will require transformation.

Dataset Name	Columns	Rows	URL
1-Product Info - EV West	5	50	EV West

The data set provides information about EV parts, manufacturer, weight and price information. The data is not available in the PDF version and requires a Python script to extract and format the information.




Elcon 3.3kW UF CAN Bus Charger with EVCC
The UFC3300 is an excellent low cost solution to h ...

Model : UFC33-CAN
Manufacturer : [Elcon](#)
Weight : 19.00

Price: **\$1,199.00**

[ADD TO CART](#) [Details](#)




Elcon PFC1500 Charger
This item has been discontinued by the manufacture ...

Model : ELC1500
Manufacturer : [Elcon](#)
Weight : 15.00

Price: **\$569.00**

[ADD TO CART](#) [Details](#)



Manzanita Micro PFC 20-XM Charger
This product is no longer available through EV Wes ...

Model : MM PFC20-XM
Manufacturer : [Manzanita](#)
Weight : 16.00

Price: **\$2,390.00**

[ADD TO CART](#) [Details](#)

Dataset Name	Columns	Rows	URL
2-Alternative Fueling Station Locations	65	56,800	Data.gov

Provides information about the current list of charging stations and their locations. There are a lot of missing values that need to be cleaned and lat and long location info needs to be identical with the EV Car Population dataset. Price information is available in a text format and needs to be formatted so it is a number.

```

import pandas as pd
import warnings
warnings.filterwarnings('ignore')

ev_stat= pd.read_csv('data/alt_fuel_stations.csv')
ev_stat.columns.to_list()

['Fuel Type Code', 'Station Name', 'Street Address', 'Intersection Directions', 'City', 'Sta

```

Dataset Name	Columns	Rows	URL
3- EV Population Data	17	121,978	Data.gov

This dataset shows EV cars that are currently registered through Washington State Department of Licensing (DOL). Provides information about EV cars, car types, and registered locations. Location fields have to match with the format of the EV Charging dataset. There are some missing fields. Currently, I do not see any other issues related to this dataset.

Dataset Name	Columns	Rows	URL
4- EV Car Dataset	14	103	Kaggle

Provides information about EV car and models and charging plug types. This dataset is clean, currently I do not see any issues.

Index	EV Car Dataset Column Names	Index	EV Car Dataset Column Names
1	Brand	9	PowerTrain
2	Model	10	PlugType
3	AccelSec	11	BodyStyle
4	TopSpeed_KmH	12	Segment
5	Range_Km	13	Seats
6	Efficiency_WhKm	14	PriceEuro
7	FastCharge_KmH		
8	RapidCharge		