PRD: Restaurant Data ETL

Product Engineering Project 1
by Nick, Justin, Chi

Change history:
- Document created 1/22/24

Overview:

As businesses constantly are collecting and storing data, anomalies may arise with the expansive dataset. To help combat this, our project will take company data about customers and orders, conduct a thorough examination, and remove anomalies, duplicates and redundancies. By doing so, we can provide the customer with a cleaned and transformed database which displays and makes use of originally unclean or erroneous data.

Objectives:
- Convert all data into dataframes
- Work within dataframes to remove null, duplicate and outlier values
- Merge files together
- Design a database schema
- Transport cleaned data from files into tables in the database

Success Metrics:
- A working database setup with a schema with proper constraints
- Data is correctly inserted into each table
- No duplicates, null values or outliers exist in database

Messaging:
- The deliverable of this product will be a database with cleaned, optimized and ready-to-use data.

Timeline and Release Planning:
- Release Date 1/25/24

Personas:
- Restaurant Owner
- Restaurant Manager

Scenarios:
- The restaurant owner seeks a user-friendly database to view behavior of customers, menu orders and analytics about marketing to drive sales and operations.
- The restaurant manager seeks to view day-to-day operations, inventory, and types of orders to make data-driven decisions.

Features In:
[M] denotes minimum viable experience requirement for Restaurant Data ETL
- Providing detailed information about customer preference, patterns and orders or menu items that might be popular to help target a group of people.
- Ability to evaluate trends or frequency of menu items ordered to help customize menu items or to remove or add items.
- Ability to see real time sales or orders to help make data-driven decisions.
- [M] Ability to view clean, organized data to ensure accurate information to help make decisions.

Features Out:
- Backup functionality

Designs:
- ER Diagram
- Presentation Slides

Open Issues:
- None at this moment, currently experimenting with the dataframes to test for duplicates and such as well as looking into setting up the database.