Problem Statement

Transactional data generated in our Inventory Management System (OLTP) is not being utilized in the business decision-making process on a regular and timely basis, due to the lack of an efficient and effective data pipeline that supports OLAP.

The main opportunity is to be able to evaluate the success of each product in the different marketplaces, and take this valuable input into the decision-making process of what products and/or sales platforms to keep or obsolete every season.

Context

S Corp. is a social mission-oriented company that unites woven-goods from developing countries with wholesale and consumer marketplaces. One of their competitive advantages is the ability to adapt, which includes high innovation in product development and being able to test-market it.

As much as this ability has grown overtime, it still has enormous potential for improvement in terms of how much transactional data is being used in the decision-making process, as well as how fast and efficiently.

Criteria for Success

Having a reliable Data Pipeline and user-friendly Dashboard that allow the organization to obtain accurate data and timely generate valuable information that can be translated into actionable insights, hence improving the decision-making process.

Solution Scope

Solution proposed includes the following key elements:

- Definition of key metrics
- Design of Pipeline and ETL Process
- Implementation of Warehouse and Dashboard

Proposed pipeline integrates technology currently being used by the organization with new components into a reliable pipeline. This includes the following:

- Shopify: Main sales platform, currently used for E-Commerce and as POS.

- Fishbowl (FBL): Inventory-control software in place (transactional system), that includes multiple business modules, such as: Sales Orders, Purchase Orders, Customers, Vendors, etc.

  It is comprised by:

  - MySQL Database (server part). **This will be the transactional source to derive from.**
  - Clients (software used in the workstations)

  All sales transactions are processed on FBL, either directly or indirectly. This includes:

  - Shopify's sales (approx. 90% of the total transactions), which are synced into FBL approx. every hour, through a plugin offered and developed by FBL
  - National Retailers' sales, which are manually entered into FBL.
  - Sales from any of the multiple external platforms (approx. 7-8) where S. Corp products are offered. These orders are either

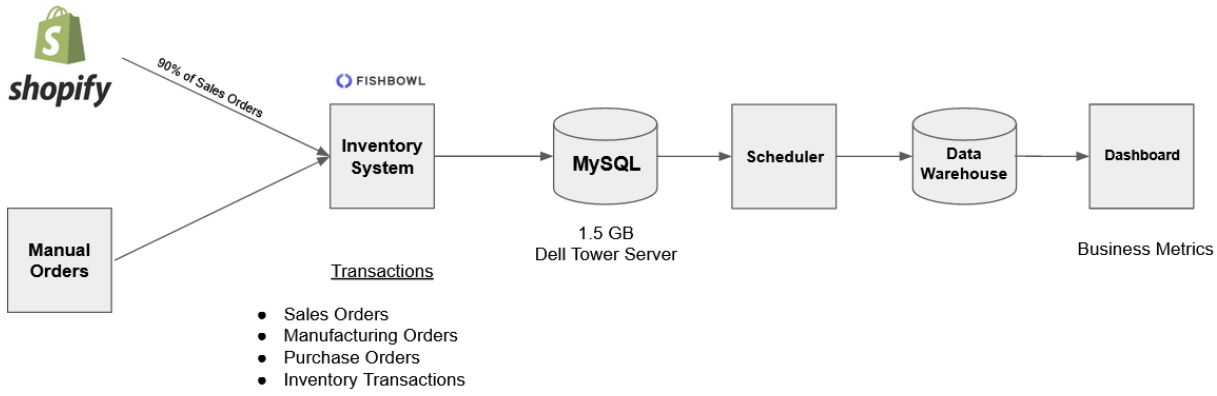
manually entered into FBL or Shopify (which eventually syncs with FBL)

- Data Workflow Scheduler: This tool will allow us to author, schedule and monitor the pipeline workflows. Application TBD (potentially Apache Airflow).

- Data Warehouse: This dedicated database will allow us to store historical data and support the BI activities, without impacting the transactional MySQL database used by FBL. Application TBD (potentially Azure).

- Analytics Dashboard or Reporting Tool: To be defined if for this specific scenario, a dashboard or a reporting tool is more convenient (Power BI?)

Proposed Pipeline



Since S. Corp currently is a medium organization, proposed architecture is Kimball's Architecture (over Inmon's), as this bottom-up approach requires less resources (time, human capital, overall investment, etc.) to implement and maintain.

Deliverables

GitHub repo that includes:

- Slide Deck
- Readme markdown file with descriptions of the project and its working