Data Analyst – Technical Test Answers

Jafar Shodiq

1. What are your strategies to enable self-service analytics in the company?
2. **Introduction:**
3. What is self-service business intelligence (SSBI)?

Self-service business intelligence is an approach that empowers business users to perform data analytics such as filter, sort, analyse, and visualize on their own without the needs for BI background or related fields.

1. What are the benefits of SSBI?

* Business users will be able to do their own ad-hoc analysis, meaning that BI and IT teams can focus more on higher-value priorities that require more technical skills.
* Accelerates business processes and decision-making.
* Creates a fully data-driven culture in the organization.

1. **Strategies:**
2. Overview

My strategies would involve using specific application/software for the data input process so that the values are all in the uniform format. The data values will then be stored in a database server. Lastly, Tableau will be able to read the server database as a data source which will then able to continuously monitor the updates in the database.

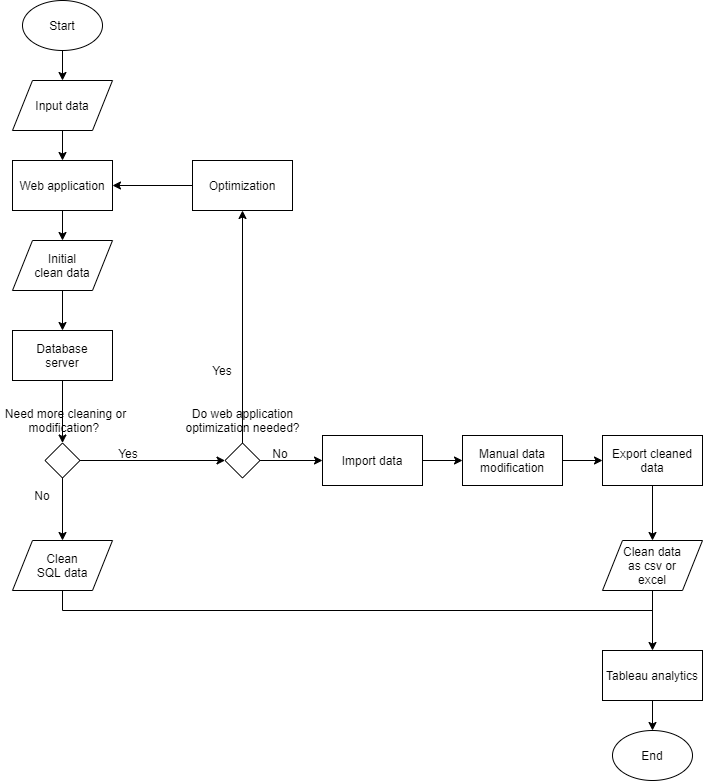
1. Technology/tools that could be used

* Web application (with APIs)
* Python libraries
* Database server
* Tableau

1. Why Tableau?

Tableau is known for taking any kind of data from almost any system, and turning it into actionable insights with speed and ease. Tableau makes exploring data more intuitively and easily with only dragging and dropping the components. Tableau also supports different devices such as mobile phone or tablet for displaying dashboards. Moreover, Tableau community has more than one million of active, diverse, creative, and supportive members worldwide.

1. Flowchart



Description:

* *Input data:* useful information such as names, location, dates, etc
* *Web application:* from my own experience and knowledge, web application can be built with backend and frontend development. The former requires API, which can be developed using Python Flask, and database management while the latter JavaScript. This web application includes component like textboxes, dropdowns, buttons, etc.
* *Initial clean data:* data sent to the database after hitting the ‘submit’ button.
* *Database server:* Server to store data, e.g. using MySQL.
* *Clean SQL data:* data if the initial clean data does not need further cleaning or modification.
* *Optimization:* Optimize or modify the web application if further cleaning or modification needed and crucial for base data input.

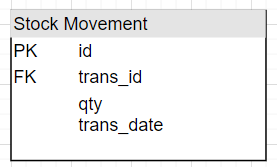
For example, there is a column addition in database server. An additional input component will be needed in the web application.

* *Import data:* data from server will be imported if there is no need to optimize or modify the web application.
* *Manual data modification:* Data cleaning using tools like Pandas or NumPy.
* *Export cleaned data:* Export the data that had been cleaned or modified to csv or excel file.
* *Clean data as csv or excel:* data after manual cleaning and modification.
* *Tableau analytics:* Data either from SQL or as csv or excel file will then be analysed and visualized using tools in Tableau.

1. What will be our role as data analysts?

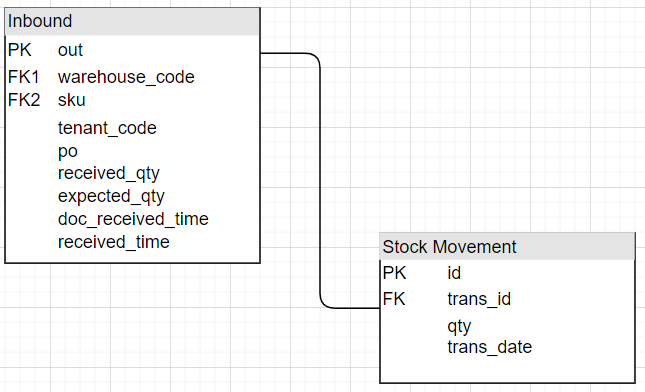
If self-service business intelligence is going to be implemented, as a data analyst, while we can help business users analyzing and visualizing using Tableau, we can also do database maintenance and produce cleaned data when needed.

1. Please create a data warehouse schema and explain your steps.
2. Stock Movement



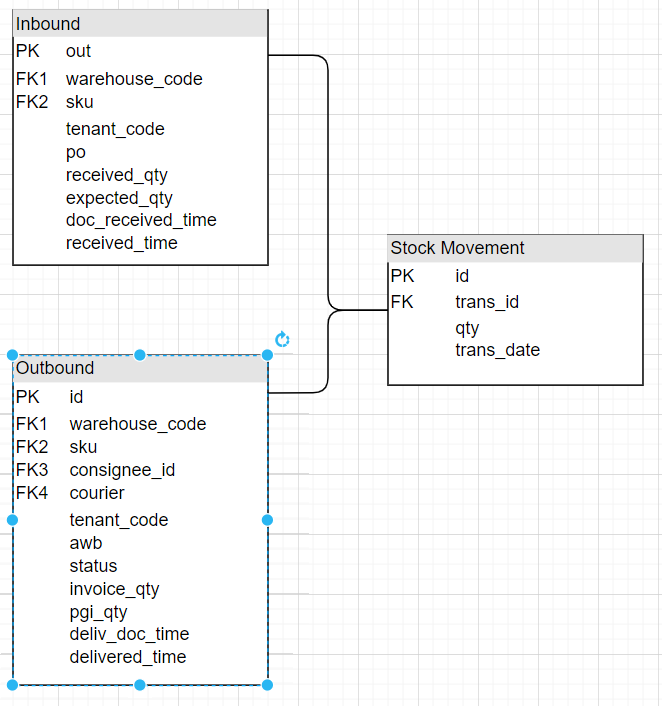
Stock Movement table has attributes: id, trans\_id, qty, and trans\_date with id as the primary key and trans\_id as a foreign key.

1. Inbound



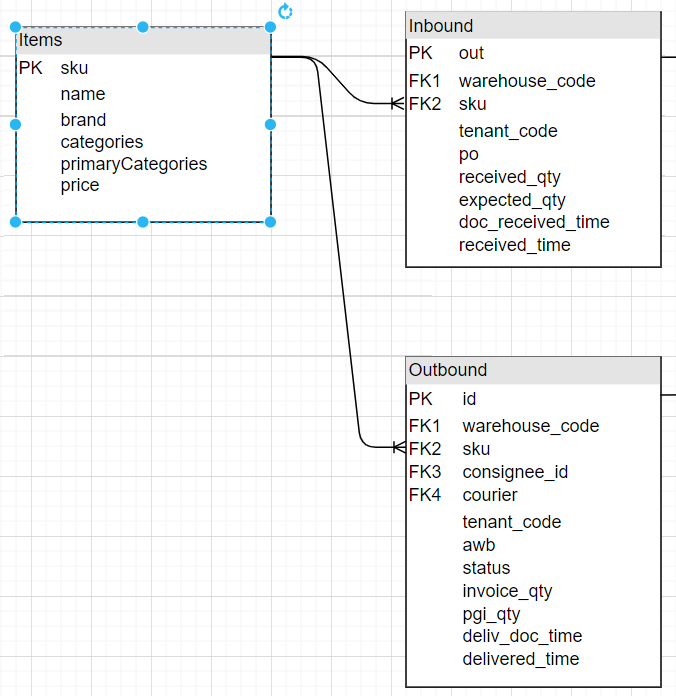
Inbound table has attributes: out, warehouse\_code, sku, tenant\_code, po, received\_qty, expected\_qty, doc\_received\_time, and received\_time with out as the primary key and warehouse\_code and sku as foreign keys. Attribute out from Inbound table has one-to-one relationship with attribute trans\_id from Stock Movement table.

1. Outbound



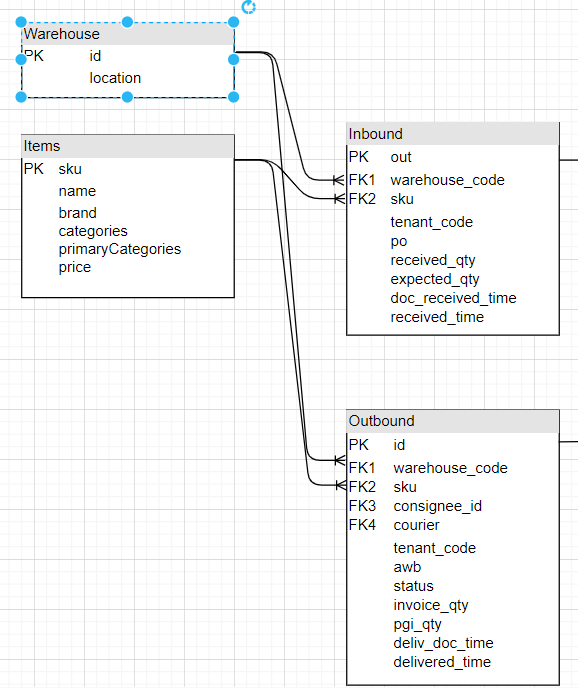
Outbound table has attributes: id, warehouse\_code, sku, consignee\_id, courier, tenant\_code, awb, status, invoice\_qty, pgi\_qty, deliv\_doc\_time, and delivered\_time with id as the primary key and warehouse\_code, sku, consignee\_id, and courier as foreign keys. Attribute id from Outbound table has one-to-one relationship with attribute trans\_id from Stock Movement table.

1. Items



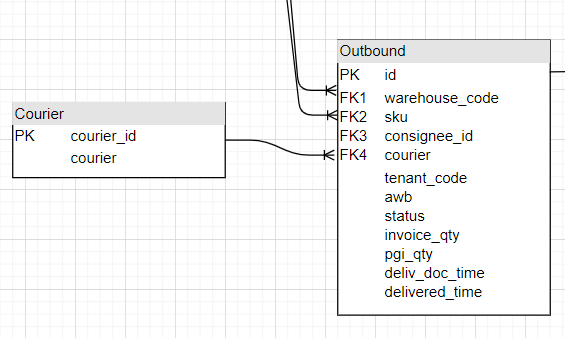
Items table has attributes: sku, name, brand, categories, primaryCategories, and price with sku as the primary key. Attribute sku has one-to-many relationship with attribute sku from both Inbound and Outbound table.

1. Warehouse



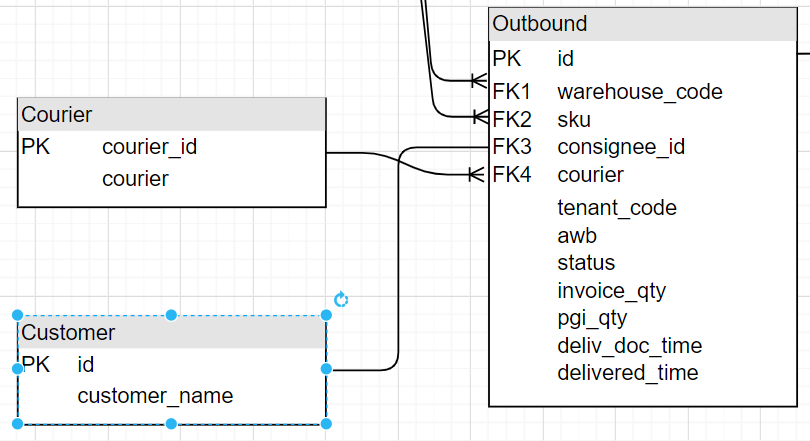
Warehouse table has attributes: id and location with id as the primary key. Attribute id has one-to-many relationship with attribute warehouse\_code from both Inbound and Outbound table.

1. Courier



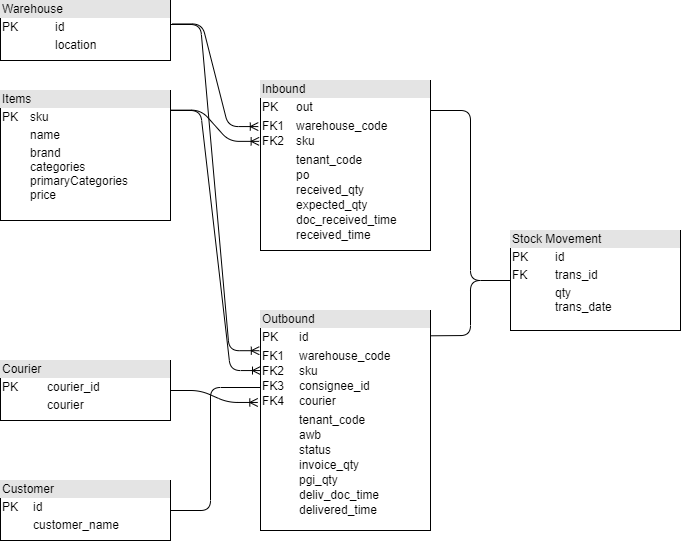
Courier table has attributes: courier\_id and courier with courier\_id as the primary key. Attribute courier\_id has one-to-many relationship with attribute courier from Outbound table.

1. Customer



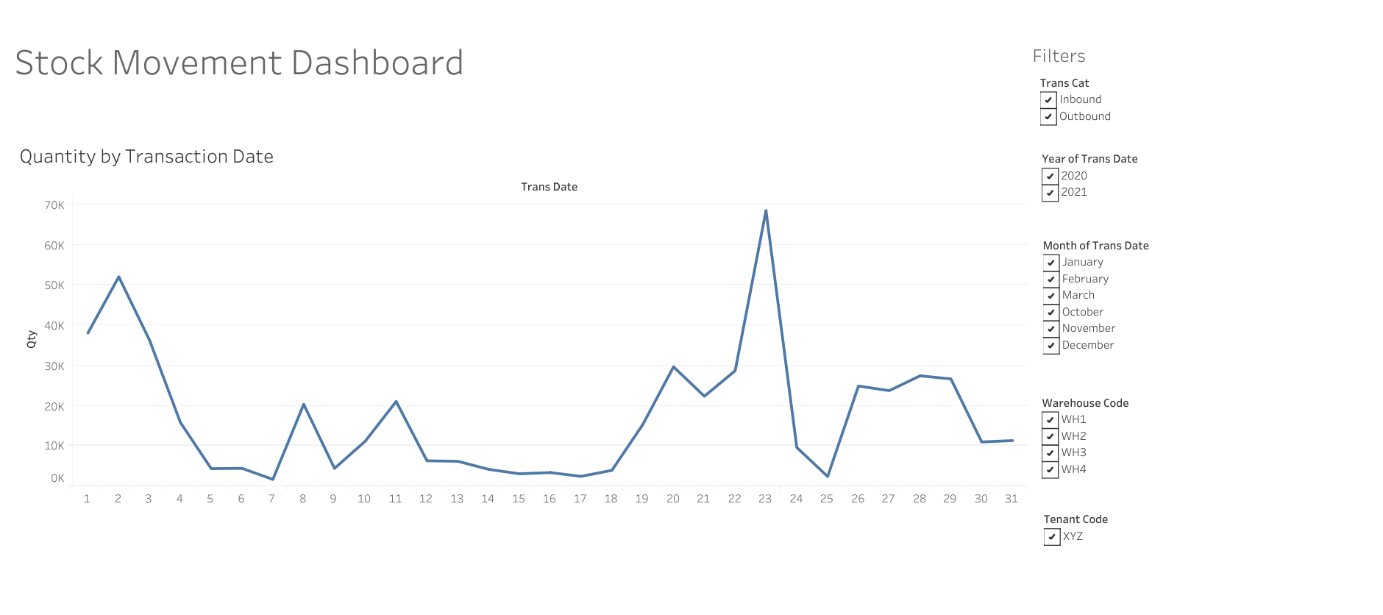
Customer table has attributes: id and customer\_name with id as the primary key. Attribute id has one-to-one relationship with attribute consignee\_id from Outbound table.

Final diagram

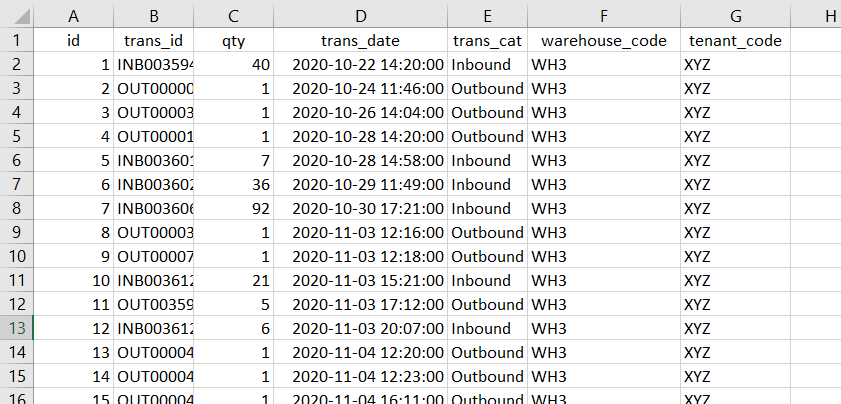


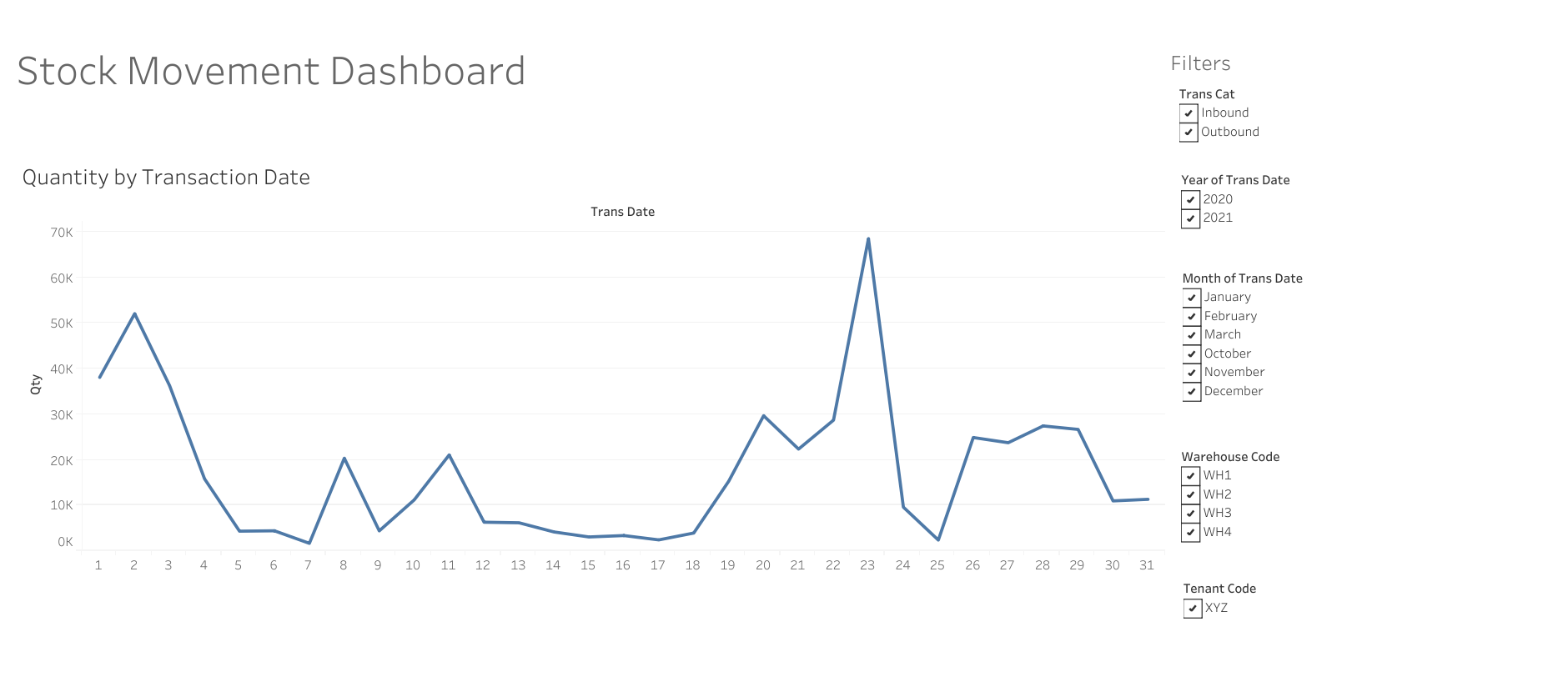
1. Please create dashboard mock-ups for these use cases. (At minimum you can provide stock movement dashboard)

The dashboard can be found on this link: <https://public.tableau.com/app/profile/jafar.shodiq/viz/MIG-assessment-dashboard/stock-movement-dashboard>

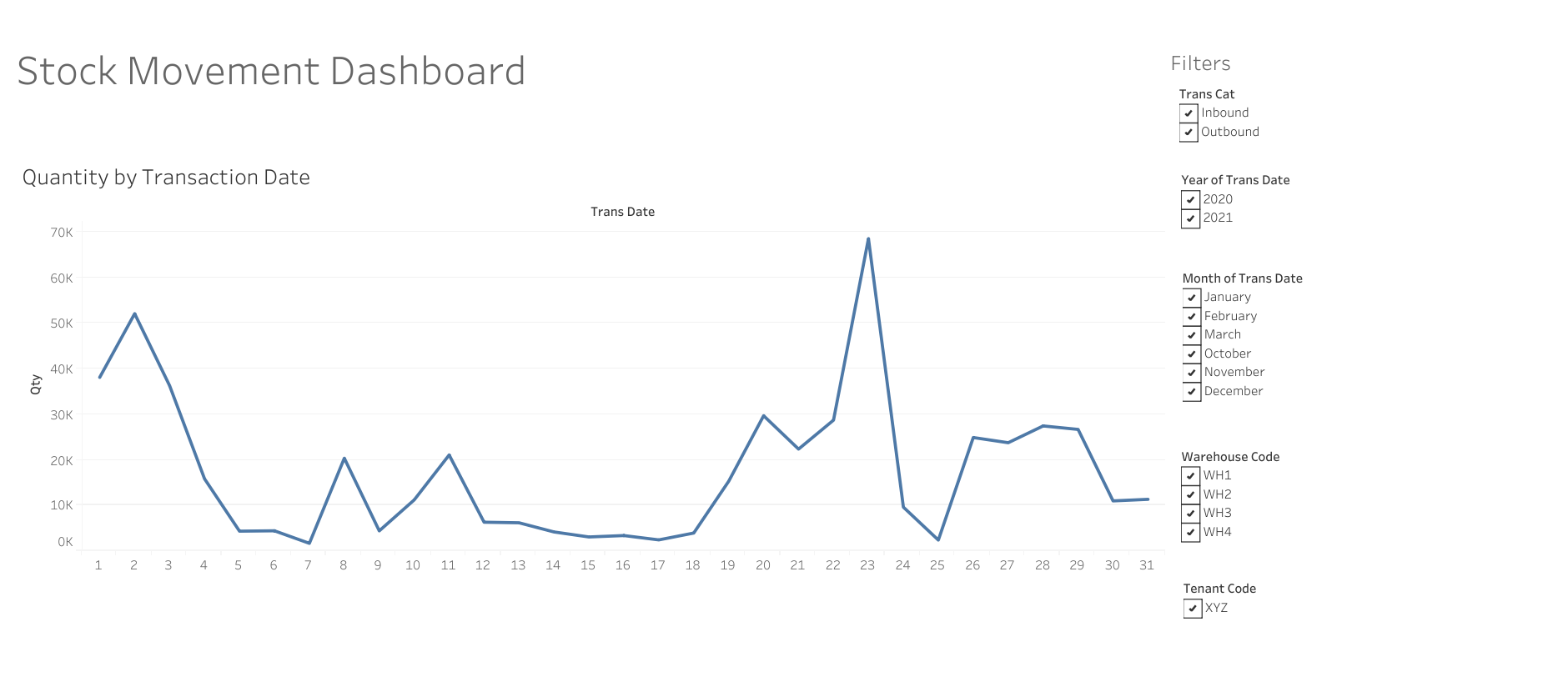


Before I loaded the excel file into Tableau, I did some modification on the Stock Movement table so that it contains additional information such as transaction category, warehouse code, and tenant code. Transaction category is based on the trans\_id column, while warehouse\_code and tenant\_code are based on Inbound and Outbound tables.





This dashboard takes the sum of the qty column as the y-axis and the day from trans\_date column as the x-axis .



It can do filters on the transaction categories (inbound or outbound), year, month, warehouse code, and tenant code.

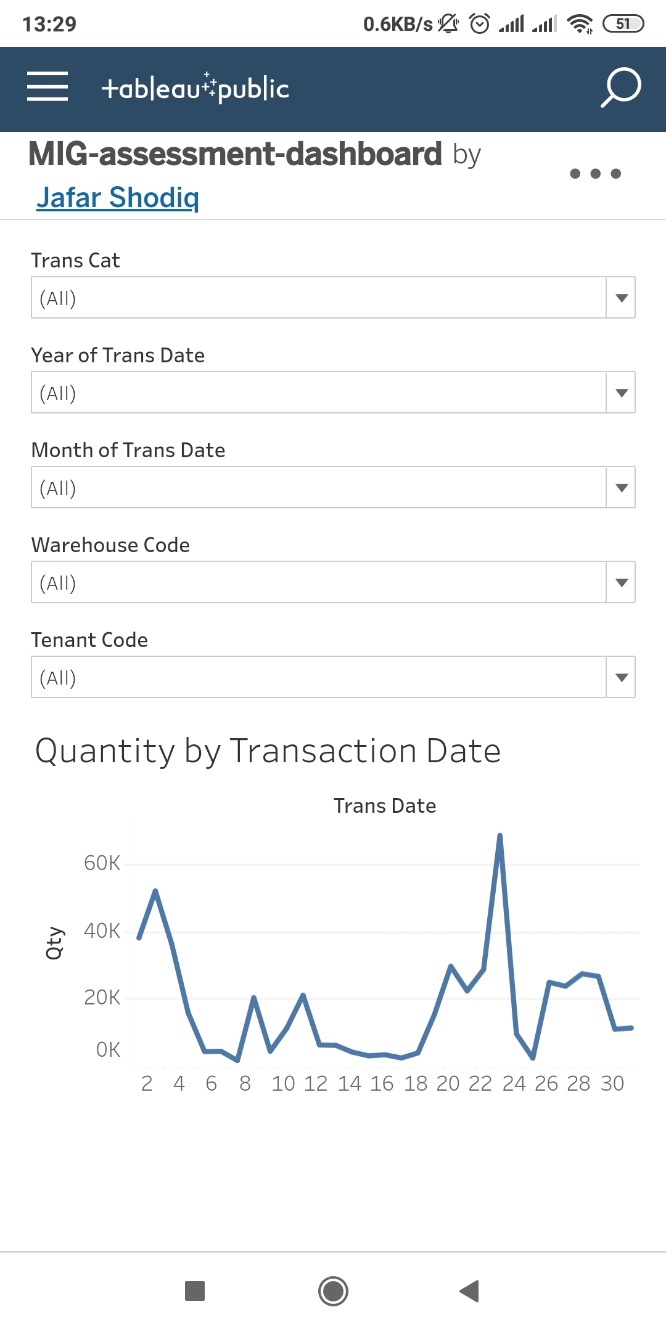
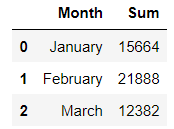
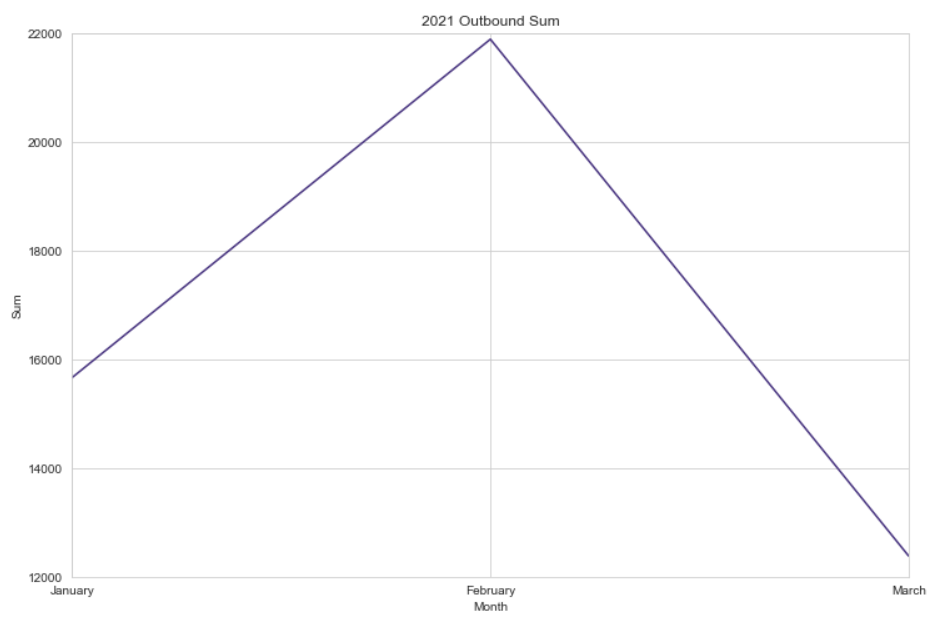


Tableau also supports dashboard view on different devices like mobile phone.

1. Give us top 3–8 interesting insights you could find from the datasets.
2. Outbound in 2021 from January – March

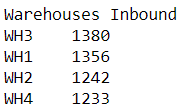


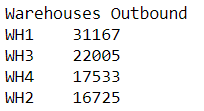


There’s an increase of 39.73% from January to February and decrease of 43.43% from February to March.

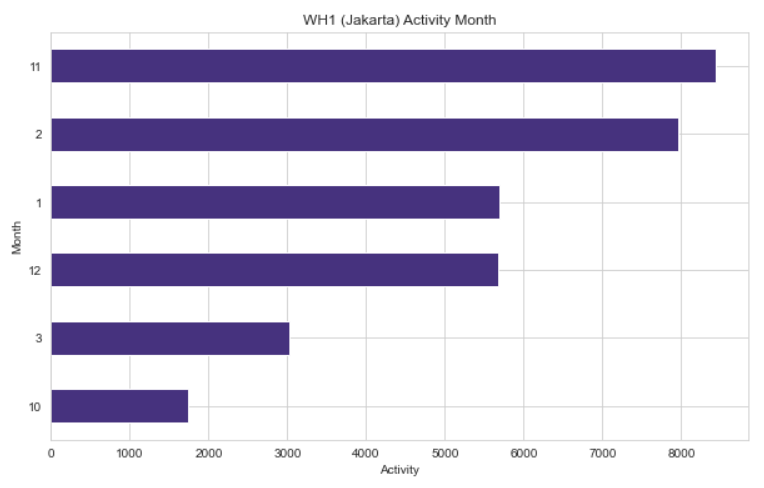
1. Busiest warehouse

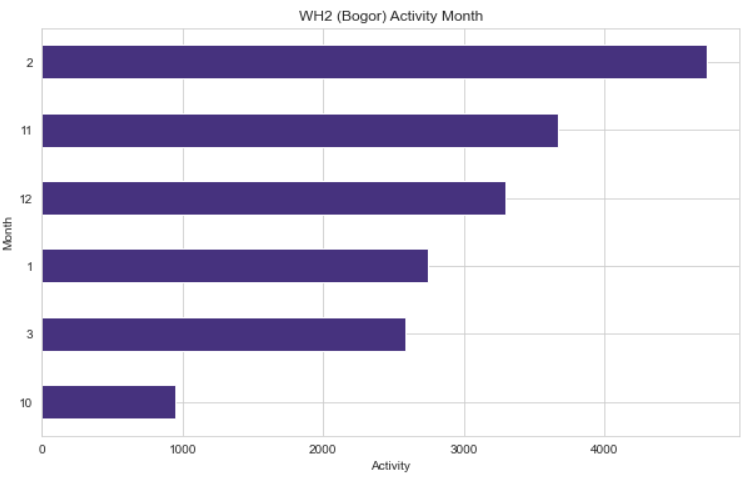
Activities (inbound and outbound) in each warehouse.

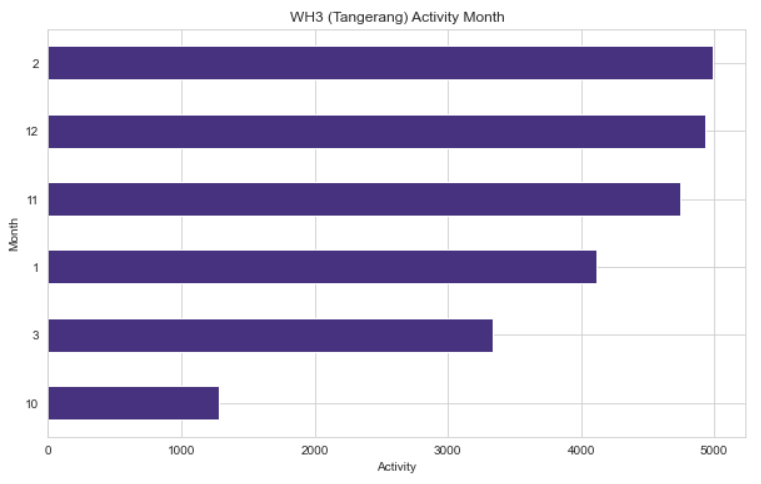


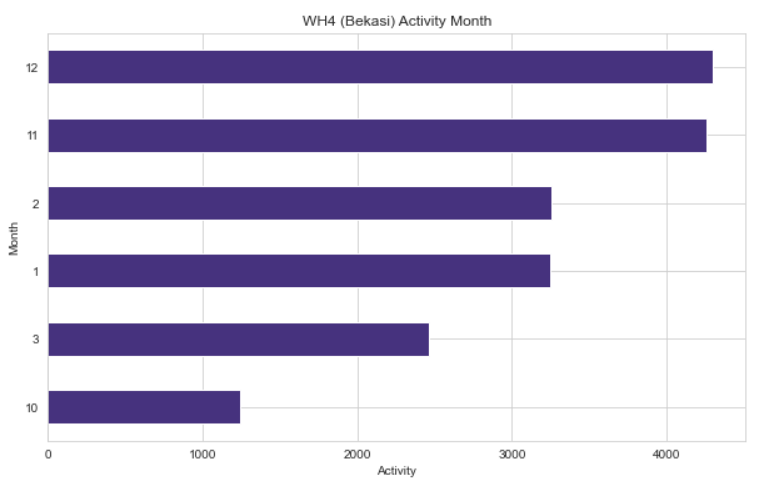


Looking at what month each warehouse has the most activity.









1. Categories in Items table

Table Items has 1546 rows. In categories column, values were being input on top of each other with multiple comma splits inside. In result, there are 782 unique values in column categories.



What words are the most frequent can be found within this wall of text with a wordcloud. The bigger the word, the more frequent that word appear in categories column.



Turned out that Electronic, Computer, Accessories, Audio, and TV are the top 5.

1. Top 5 metrics to define the success of this product, what, why and how would you choose?

Links and references:

Why choose Tableau - <https://www.tableau.com/why-tableau>

Brief Tableau tutorial - <https://www.youtube.com/watch?v=jEgVto5QME8>

MySQL with Tableau - <https://help.tableau.com/current/pro/desktop/en-us/examples_mysql.htm>

Draw.io - <https://app.diagrams.net/>

My dashboard assessment - <https://public.tableau.com/app/profile/jafar.shodiq/viz/MIG-assessment-dashboard/stock-movement-dashboard>