BidBetter

Leveraging Visual Insights for Smarter Tender Bidding

SMU SINGAPORE MANAGEMENT UNIVERSITY

Background

Issues & Problems

While the database of historical tenders is extensive, it is not organized or visualized in a manner that allows for patterns or insights to be easily derived without having some knowledge of data analytics methods and applications.

Therefore, our R Shiny application seeks to address this.

Motivation

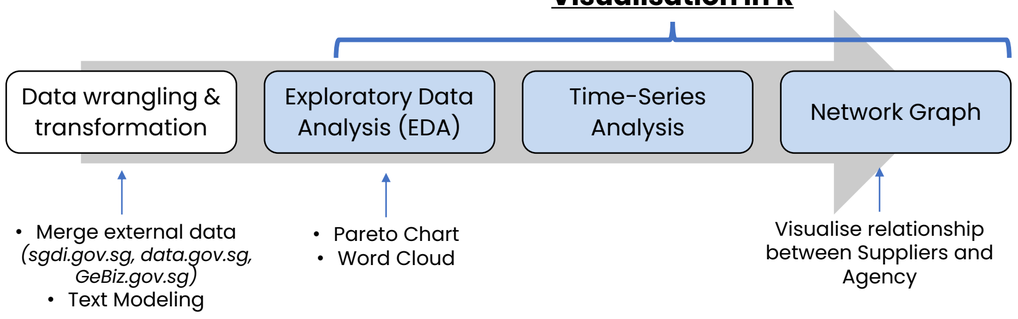
This project aims to provide the public with an interactive app to better understand trends in government procurement and observe changes in the new norm. It also allows interested bidders with time-series forecast on government demand in the various categories.

Data Preparation

- Enriched procurement information on GeBiz (e-procurement portal) with Ministry information and companies' Industry Code.
- Bids are categorized into products or services using Text Mining and Tokenization
- Bids below SGD 6,000 are determined as small-valued item and excluded from analysis.

Approach

Visualisation in R



STEPS

1 Data Wrangling

The following data were combined through the process flow outlined below to enable categorization and complete the intended visualizations:

(i) Government directory to map Agencies to their respective Ministries
(ii) GeBiz on past government procurement tenders and results
(iii) ACRA company information to map suppliers to their business type

Government directory – sgdi.gov.sg

Ministry	Ministry Abbrev

a. Joined **government directory** with **tender details** by `agency`

Tender details table – GeBiz.gov.sg

agency	supplier_ name	tender_no.	tender_ description	award_date	tender_ detail_status	awarded_ amt

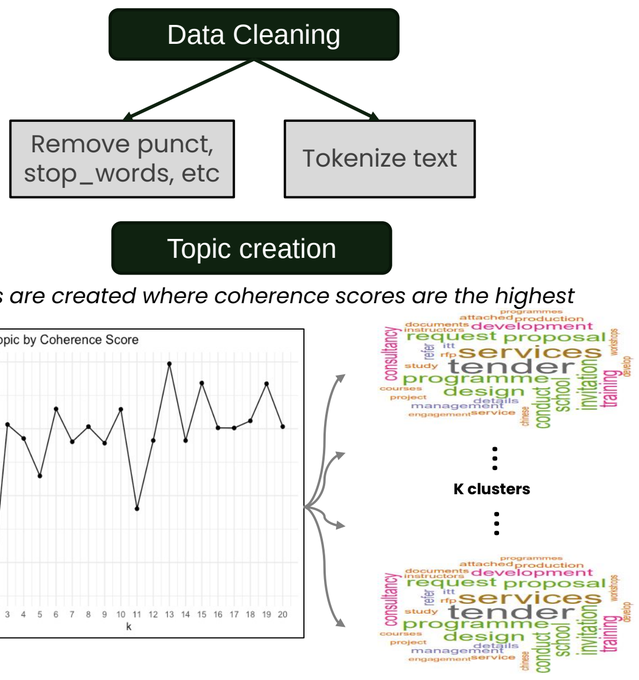
b. Joined **ACRA** with **tender details** by `supplier_name`

ACRA – data.gov.sg

uen	postal_code	primary_ ssic_code	paid_up_ capital1_ordinary

2 Text Modeling

A key challenge we faced was that there is no categorized grouping for these tenders. Our team then investigated the possibility of using tender descriptions to categorize each tender into products and services. We used **TextMineR** to process the tender descriptions, which enabled us to categorize tenders automatically.



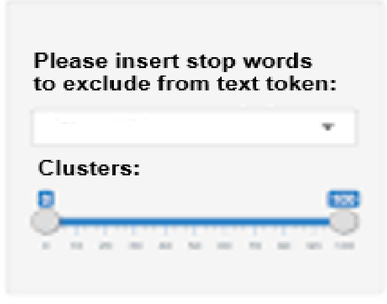
3 Visualization in R Shiny

Now with a usable data set, The R Shiny app was built using more than 35 packages, with the key ones shown below, to visualize useful insights on government procurement tenders. The app was then deployed to Shinyapps.io for public access.

Future Improvements

Interactivity with TextMineR

- Further leverage on TextMineR for user interactivity
- Users may:
  - insert own stop words to customize and finetune the text tokenization factor
  - Browse through the number of clusters using a scroll bar to determine which group of k-cluster brings about the most value for each use case



Application Design

Descriptive Analysis

The app provides users with different visualisation and descriptive analytics of GeBiz data through **Sankey diagram, Pareto Chart , Word Cloud and Network Analysis.** These visualisation capitalize on **focus topics** tagged to individual bids, identified through text analysis.

Application 1: Flow of Tender Value to Product Category

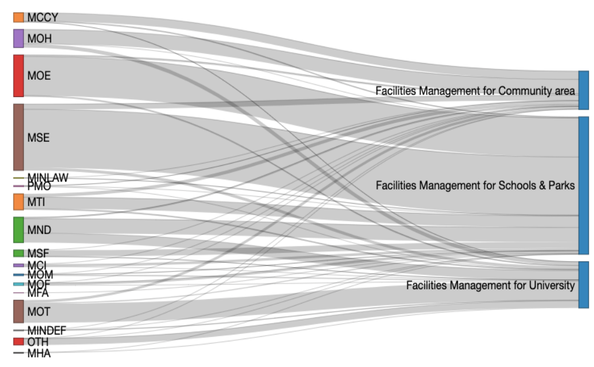


Chart 1: Sankey Diagram for procurement in product categories related to facilities management

The Sankey Diagram quantifies the value of bids awarded by each Ministry per Product Category.

This application provides users with an overview of the key product categories that Agencies under each Ministry is interested in.

Chart 1 (left) provides an example of bids related to facilities management categories (for community area, schools & parks, university) in the post-Covid period defined by the user as 1 April 2021 to 31 Mar 2022. From the Sankey plot, users can derive key insights; MOE / MSE / MND are the key stakeholders in such bids, and most of the investment in this area are dedicated to schools and parks.

Application 2: Procurement Spending Proportion

The Pareto Chart and Word Cloud allows users to derive more detailed information on specific categories and ministries. This builds on the insights obtained from the earlier Sankey plot, enriching the narrative for selected Ministry-Category pair. In addition, the word cloud allows users a deeper understanding on recurring themes.

Chart 2 (right) develops on our earlier example on facility-related categories; narrowing the dataset to procurements by agencies under Ministry of Sustainability and Environment. The Pareto Chart provides a simple visualisation that 40.5% (~SGD43.9M) of total expenditure across the same time period is dedicated to schools & parks.

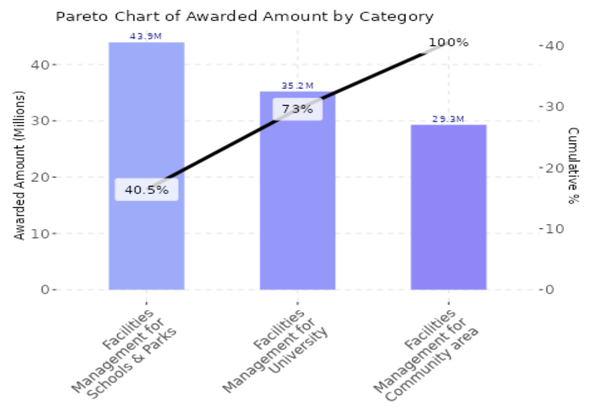


Chart 2: Pareto Chart of total awarded value in SGD'Ms for procurements under Ministry of Sustainability and Environment

Application 3: Top Tender Product Categories

Chart 3: Word cloud for procurement by agencies under MSE for facilities management of schools and parks

The word cloud provides users with more context when users narrow their inputs to specific ministry-category relation.

For example, Chart 3 (left) narrows the dataset to bids grouped under Facilities Management for Schools & Parks, procured by agencies under MSE. Recurring words such as "food, centre, insurance, refuse" enrich users' understanding of possible recurring theme(s). The Pareto Chart and Word Cloud is supported by a table with information on actual bid information provided by the procuring agency.

Application 4: Ministry & Supplier Relationship

The network analysis graph visualizes the relationship, through representing each as a singular node, between each Ministry (blue square node), its Agencies (red triangle node), and their suppliers (green circle node).

Agencies have reasons to choose different suppliers (e.g. for risk management) however there can also be compelling reasons to choose the same reliable suppliers for cost savings and efficiencies. This network analysis visual easily identifies

Agencies within the same Ministry using the same suppliers through the edges (lines) that connect them as seen in Chart 4 (right). Suppliers who are used by multiple agencies could represent strong and reliable suppliers that other agencies can use as well.

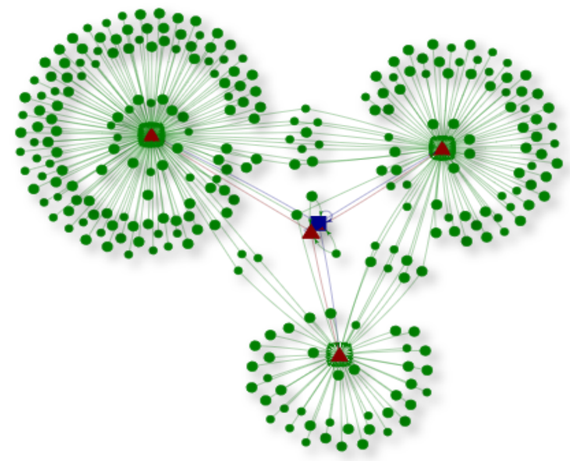


Chart 4: Network Analysis of MSE relationship with their suppliers in 2018, 2019 and 2021

Application 5: Analysis of Variance

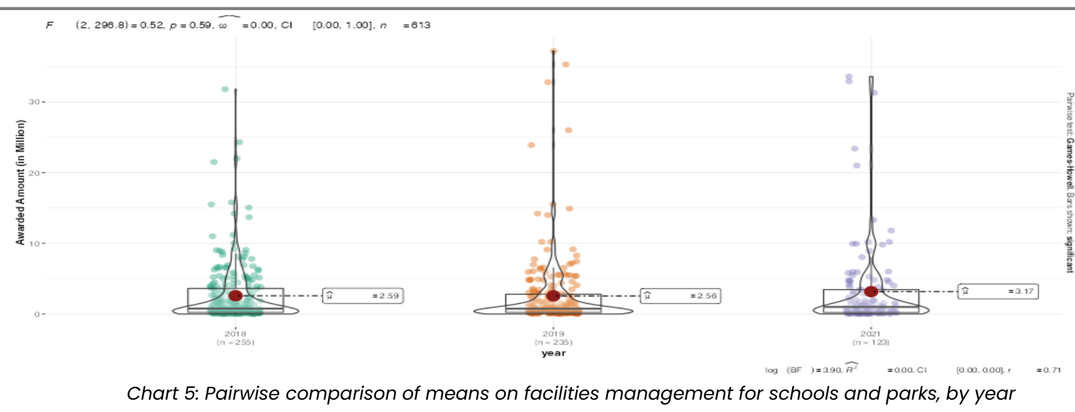


Chart 5: Pairwise comparison of means on facilities management for schools and parks, by year

The Violin Plot in this app provides pairwise-comparison with one-way ANOVA analysis, allowing users to narrow the spectrum of analysis across years. This analysis is highly relevant in today's setting, as it can be used to compare changes in procurement pre- and post- Covid.

To illustrate based on earlier examples, Chart 5 (above) compares government procurement related to Facilities Management for Schools & Park, across all ministries and agencies, filtered for outlier bids higher than SGD100M. As the data for 2022 is incomplete, the exercise compares procurement means between years 2018, 2019, 2021. Results of the analysis show that increased spending in 2021 in this category is statistically significant, possibly due to increased focus on hygiene and safe distancing measures.