

Chart-to-Text

A large-scale benchmark for chart summarisation

Student: Leong Ko Rixie Tiffany Supervisor: Dr Joty Shafiq Rayhan

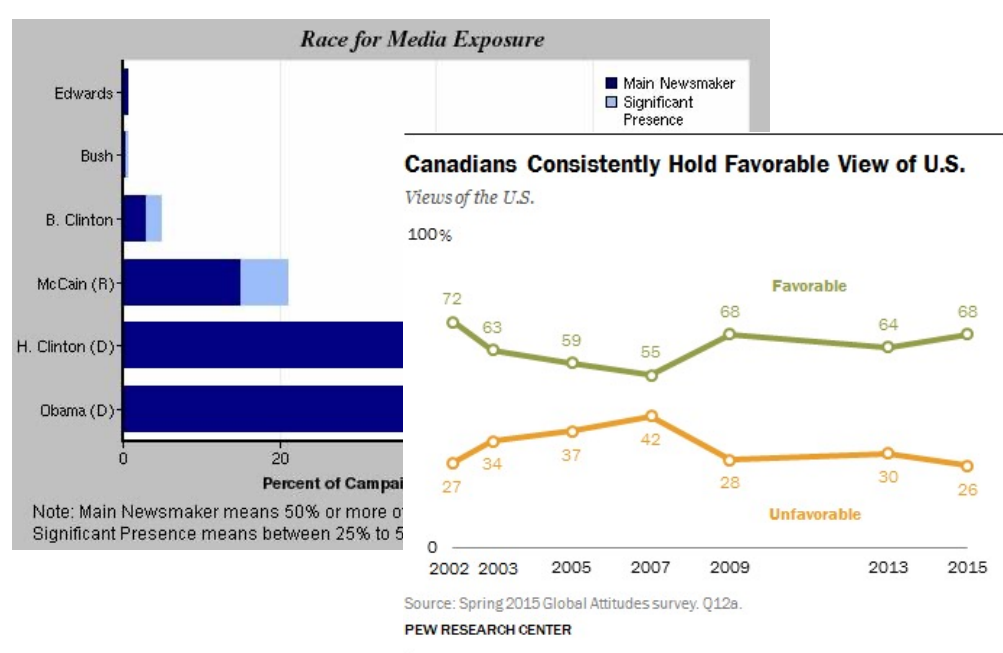
Project Motivation:

Charts are ubiquitous in our data-driven world. They are used for exploring data and communicating insights. However, charts are often difficult for readers to interpret. Automatically-generated chart summaries can act as a useful tool for readers to more easily and more accurately interpret charts.

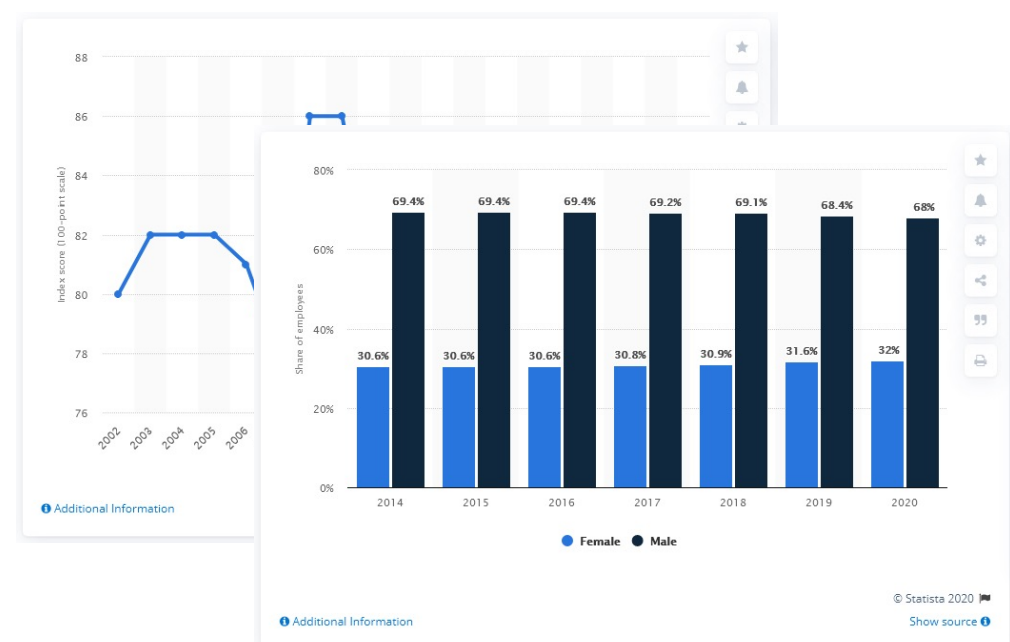
Project Objectives:

- To construct a large-scale dataset for training models to generate chart summaries
- To introduce state-of-the-art machine learning models as baseline models

Dataset Sources:

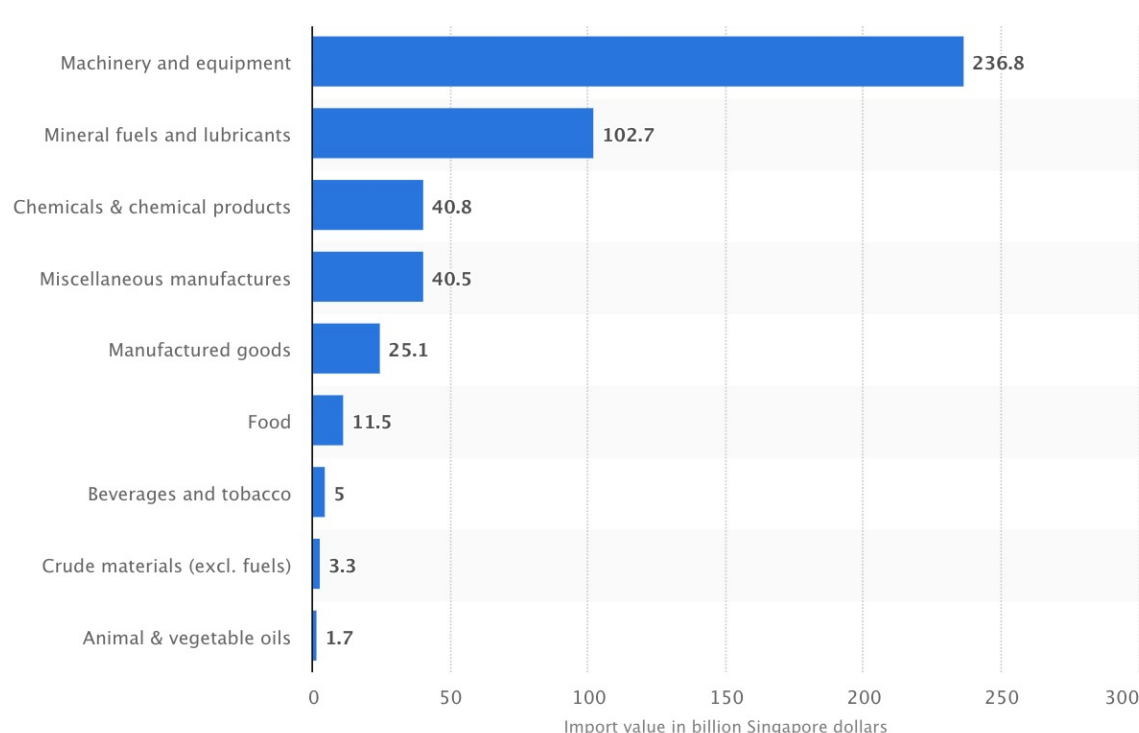


Pew Research Center



Statista

Chart Summarisation Task:



Actual Summary:

In 2019, Singapore imported approximately 236.8 billion Singapore dollars worth of machinery and equipment, making it the country's largest import commodity by value. This was followed by the import of mineral fuels and lubricants, valued at 102.7 billion Singapore dollars.

Model-generated Summary (T5):

Machinery and equipment was the most valuable commodity for Singapore in 2019, with an import value of 236.8 billion Singapore dollars. Mineral fuels and lubricants were the second most valuable commodity for Singapore, with an import value of 102.7 billion Singapore dollars.