

Industry	Chemicals
Use Case Title	Pipeline Price Benchmarking
About the Customer	The Client is a multinational chemical company based in North America and The Netherlands. The Client is one of the largest licensors of polyethylene and polypropylene technologies. It also produces ethylene, propylene, polyolefins, and oxyfuels.
Business Problem	In the region of Texas, many vendors provide transportation services of different materials. The Client is one of the many vendors. The objective of the project is to benchmark Client's service / transportation charges against its competitors' charges.
Solution	<p>Stats</p> <ul style="list-style-type: none"> A total of ~700 PDF documents worth of data are available. Data is available at Texas' Govt. website. <p>Deployment</p> <ul style="list-style-type: none"> Power BI tool on private machines. <p>Approach</p> <ul style="list-style-type: none"> All vendor data is available publicly on US Govt. website. Scrape the website for PDF documents. Different vendors have differently structured documents i.e., the data isn't entered in a standardized format. This makes automating the process of extraction difficult. The required pricing information is present in tables in the respective PDF documents. The PDF documents are subjected to both manual as well automated extraction. The automated extraction uses Python, OpenCV and NLP. The manual extraction involves reading through all the ~700 documents manually and entering it into an excel sheet manually. The pricing information contains the source address, destination address, and the corresponding pricing information for each vendor. It is important to standardize the price in Cents / Pound / Mile. For this, the distance between the source and the destination is required. Google Distance Matrix API and Google Geocoding API are used to calculate the road-distance and crow-flies distance between the source and the destination. Once the data is finalized, Power BI and Python-Dash are leveraged to create a visual platform to navigate the data. <p>Implementation</p> <ul style="list-style-type: none"> Manual extraction of required information took around one month.
Outcome	<ul style="list-style-type: none"> One-stop excel database of collated pricing data. Power BI and Dash visualizations for easy navigation. Different statistical numbers and representations for lateral investigations. Different hierarchical grouping of data based on user discretion. Visual platform to browse pricing information. 90%-time reduction for identifying customers. Improved Profit-plan estimations.