Protecting Consumers from Counterfeit Medicines Using R & Google Maps

Sarah Eltinge
Johnson & Johnson Technology Advanced Analytics

The Task: Johnson & Johnson Global Brand Protection (GBP) works with local law enforcement agencies to identify counterfeit medicines worldwide. GBP asked our analytics team to perform an exploratory analysis with a novel data set, which contained information on potentially counterfeit products around the world. In particular, the GBP team wanted to identify possible hotspots of illicit trade from the location information in the dataset.

The Constraints:

- The location information in this dataset was free text, which was difficult to parse since address formats vary widely across the globe.
- The full dataset contained dozens of languages, but many language-detection tools failed because the location text was very brief.
- The final output of our analysis needed to be usable by non-technical business partners in a dynamic environment.
- The analysis needed to be flexible so that we could pull in additional geographic information, e.g. Johnson & Johnson manufacturing sites.
- For this exploratory analysis, we could not invest in expensive technology solutions such as commercial geographic information systems (GIS) software.

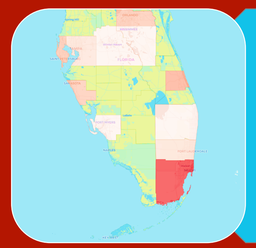
The Solutions:

Google Maps Geocoding API

- Transforms location data into latitude and longitude.
- Proved better than other geocoding solutions because it was cheap, handled many languages, and worked with various address formats.

`sf` (simple features) package in R

- Open-source package that handles most of the same tasks as commercial GIS software.
- Used to aggregate latitude and longitude data to relevant geographies, e.g. US counties.



Leaflet Maps

- Easy-to-use, interactive visualizations of geographic data.
- Can integrate with live data sources and processing scripts.

The Result: In a short period of time, we used a variety of data science tools to build a low-cost, dynamic visualization of location information for our GBP partners. This tool will be used to identify hotspots of illicit activity, and it may protect consumers around the globe from harmful counterfeit pharmaceuticals and medical devices.

Acknowledgements: Advanced Analytics co-op Adam Yormark did extensive research on commercially available geocoders for this project.