DataFest 2025 Team Vitamin E One-Page Write Up

Mukto Akash, Jasmine Sun, Adrian Tang, Irene Chung

Research Question: *What are the different trends and market shifts between legal services in the Phoenix market that have been classified as such by Savills, and those that are left unclassified? What are some market opportunities that Savills may have missed due to unclassified data?*

Starting with an objective of informing Savills of trends within the commercial real estate market, our team noticed that a large number of lease transactions (172208 out of 194685) left the internal industry unlabelled. A quick word cloud of the company names of these "unclassified" transactions showed that many of these transactions involved companies with "Law" in their name. Since legal services form a bulk of Savills' clientele, we decided to investigate the unclassified companies that provide such services.

  Filtering the unclassified transactions for company names containing "Law", "Attorney", "Legal" and "Pllc" as substrings, we found 524 entries that fit our criteria. This is substantial as there are only 2109 entries that have been labelled as "Legal Services" in the dataset. So we decided to create a subset of the data by combining all transactions labelled "Legal Services", with those from the unclassified transactions that fit the bill. We then wrote a Python script to get the geolocations (latitudes, longitudes) for these entries by making calls to the Google Maps API.

  Once we identified the major markets for legal services, we decided to focus only on Phoenix as this had the largest proportion of unclassified versus classified transactions. We then tracked the time evolution of the geospatial data. Over the years, legal companies in the Phoenix market have spread out more to cities other than Phoenix. These cities in the Phoenix market have high median income, high population, and decreasing crime rates which make them appealing for potential clients. This could indicate that the industry is expanding into different cities and that there is an emerging market for legal services.

  However, classified legal services are mostly clustered in Northern Phoenix and Scottsdale. On the other hand, unclassified entries seem to be spreading into nearby cities like Goodyear, Glendale, Tempe, and Mesa. In fact, judging by recent trends, legal services hoping to step into the Phoenix market may wish to open their location in one of these latter cities. Most of these newer locations also have a smaller footprint in terms of leased square footage. These facts would be completely missed if Savills only advises their clients based on transactions that were labelled as "Legal Services".

  Finally, a disproportionate amount of the classified companies lease higher quality ('A') locations with very few leasing other quality ('O') buildings. A chi-squared (χ^2) test shows overwhelming evidence (p-value $\approx$ 0) that labelling the industry for the transaction depends on the quality of the location leased. Other factors, such as the leased square footage, also tend to influence whether the transaction would be labelled. Further investigations showed that Phoenix is not a unique market exhibiting these features, neither are these issues limited to the Legal Services industry. Savills mostly labelling transactions with larger footprint and higher building quality leads them to miss clients' changing preferences and indication that companies may be downsizing their office space.