Since my goal is trying to break into marketing/advertising analytics, I've been trying to think of datasets that would support this kind of analysis. Public datasets such as sales data for customer segmentation analysis are probably hard to come by since most of that information is private and proprietary.

Then I remembered Yelp has a massive public dataset on its users, businesses, and reviews.

A couple ideas I’ve come up with so far:

* Can we predict what businesses users will like based on their past reviews?
* How much influence do Yelp Elite users have on businesses?
* Market trends in businesses over time – Where are the hotspots? Can we predict upcoming hotspots? What makes a hotspot/location prime for businesses?
* Market over-saturation/under-saturation for different types of businesses