Vape Scrape Spec

1. **Goal:** Identify vape shops and hookah bars. From the grant (some paraphrasing): Among researchers, there is disagreement about how to define a vape shop. Some studies require that vape shops do not sell any conventional tobacco products and others require 50% or more of their product line be for vaping. Our goal is to identify stores that retailers or customers identify as vape shops on social media (Yelp) and Google.
2. **Time frame:** Every six months. Ideally, first scrape in December 2018
3. **Geography:** Six states of interest: CA, GA, MA, MN, OK, WA
4. **Scrape sources:** Yelp & Google
5. **Needed data per record:** store name, street address, city, zip, phone number if possible, type of store if possible (vape only vs. vape + other conventional tobacco products vs. hookah lounge), source, store type category “hit”? (see below)
6. **Update previous scrapes by:**
   1. Eliminate false hits that are primarily tobacco stores, such as the chains Cigarettes 4 Less and Cigarettes Cheaper.
   2. Possibly identify whether store sells vape products only, as opposed to vape AND other tobacco products (cigarettes, cigars, cigarillos, or chew, etc.)
   3. Previous scrape results seemed inconsistent and not clear in terms when a record was included.
   4. For Yelp scrape, up to n=3 store type categories returned, as least one of which was “Vape shops”.
   5. For Google scrape, between n=28 to n=39 Store Types reported per store (varied by draw), most of the time one of the store types was “Vaporizer store”, but not always.
   6. In last scrape, there were n=85 records that did not have store type + Vape Shop in Yelp or Google file.
7. We do not have a comparison “gold standard” list for vape shops. In CA vape shops are required to have a California State tobacco retailer license, so we can check to see whether or not a vape shop from the scrape appears on the CA licensed tobacco retailer list (“LTR”). The CA LTR list does not include store type and is known to have great variation in name and address entries when compared to other lists.