

WhereSIOUS - San Diego

Where Should I Open Up Shop?



Jeffrey Silverman, Insight DS Fellow
with Maksim Pecherskiy, San Diego CDO
for the City of San Diego, CA

Where Should I Open Up Shop?



Where Should I Open Up Shop?



- *We want to predict how successful a business will be in a given location*

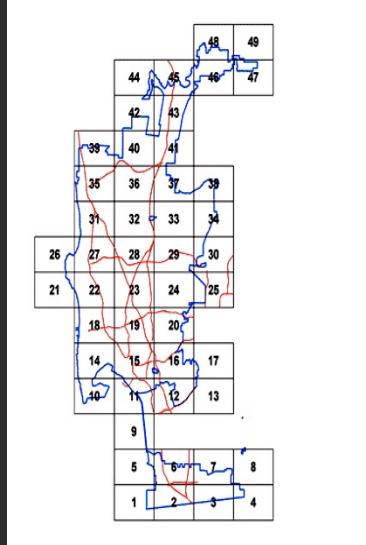
Where Should I Open Up Shop?



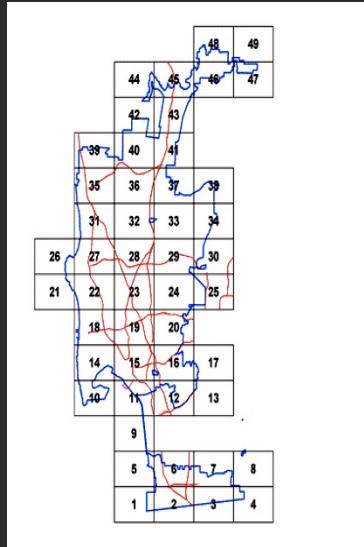
- *We want to predict how successful a business will be in a given location*
- *We want to predict the best place for a business to open*

Things that could affect the success of a business at a given location (i.e., features)

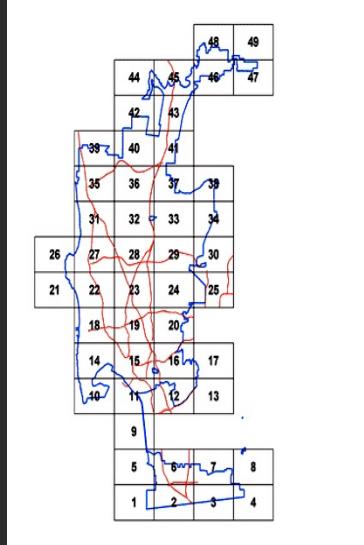
Things that could affect the success of a business at a given location (i.e., features)



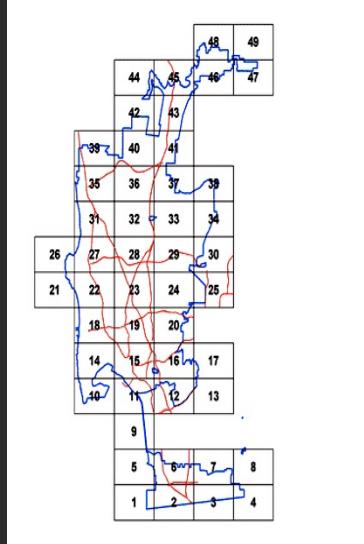
Things that could affect the success of a business at a given location (i.e., features)



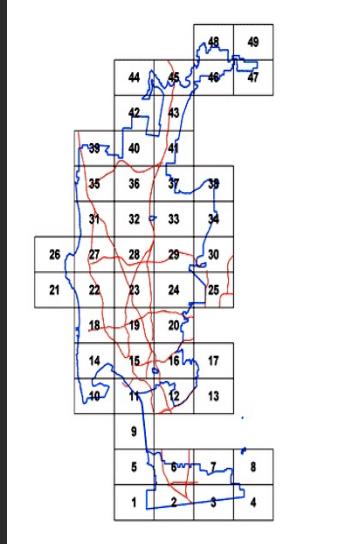
Things that could affect the success of a business at a given location (i.e., features)



Things that could affect the success of a business at a given location (i.e., features)



Things that could affect the success of a business at a given location (i.e., features)



Measures of success of a business at a given location (i.e., labels)

- Profits (per employee?)
- Change in profits over time
- Change in # of employees over time
- Yelp reviews
- Other?

Measures of success of a business at a given location (i.e., labels)

- Profits (per employee?)
- Change in profits over time
- Change in # of employees over time
- Yelp reviews
- Other?

How well have similar businesses in a similar location performed?

Data Sets

- City treasurer's business listings
- Census data (population, median income, etc.)
- Parking meter locations & transactions
- Bike lanes
- Walk score
- Waze
- City zoning & land use

Deliverables – Basic

- A model that predicts a “success score/probability” for a given business location in San Diego

Deliverables – Basic

- A model that predicts a “success score/probability” for a given business location in San Diego
- A blog post with the City of San Diego describing the process and results

Deliverables – Basic

- A model that predicts a “success score/probability” for a given business location in San Diego
- A blog post with the City of San Diego describing the process and results
- A simple web interface where you enter an address and it returns a “success score/probability”, e.g.,
[here](#)

Deliverables – Advanced (i.e., Stretch Goals)

- Multiple success scores/probabilities for different categories (e.g., foot traffic, distance from competition, etc.)

Deliverables – Advanced (i.e., Stretch Goals)

- Multiple success scores/probabilities for different categories (e.g., foot traffic, distance from competition, etc.)
- Users can weight the importance of different categories

Deliverables – Advanced (i.e., Stretch Goals)

- Multiple success scores/probabilities for different categories (e.g., foot traffic, distance from competition, etc.)
- Users can weight the importance of different categories
- Different types of businesses can give different scores

Deliverables – Advanced (i.e., Stretch Goals)

- Multiple success scores/probabilities for different categories (e.g., foot traffic, distance from competition, etc.)
- Users can weight the importance of different categories
- Different types of businesses can give different scores
- Given a type of business, calculate the best (or top 10?) locations

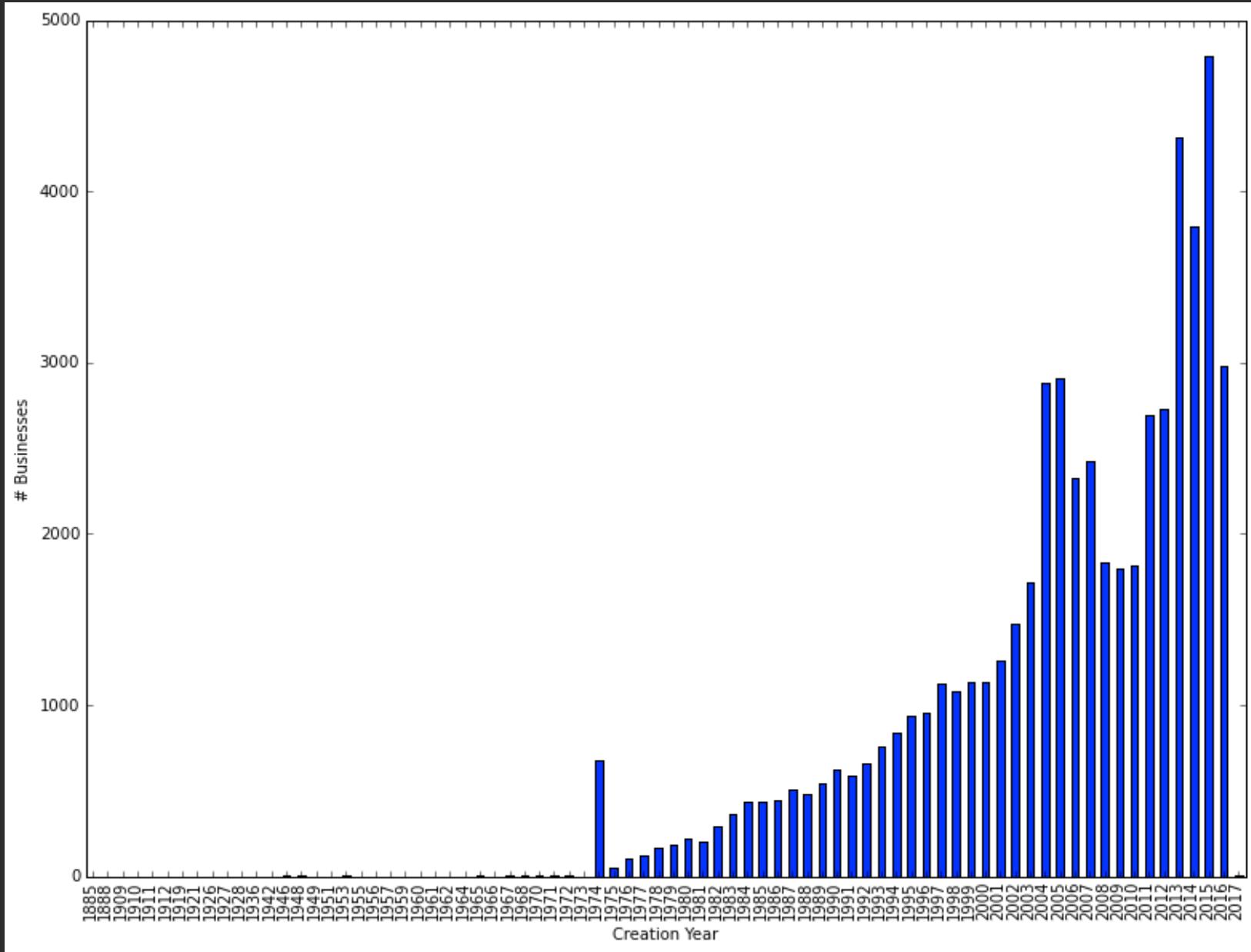
Deliverables – Advanced (i.e., Stretch Goals)

- Multiple success scores/probabilities for different categories (e.g., foot traffic, distance from competition, etc.)
- Users can weight the importance of different categories
- Different types of businesses can give different scores
- Given a type of business, calculate the best (or top 10?) locations
- The web app allows all of the above possible variations

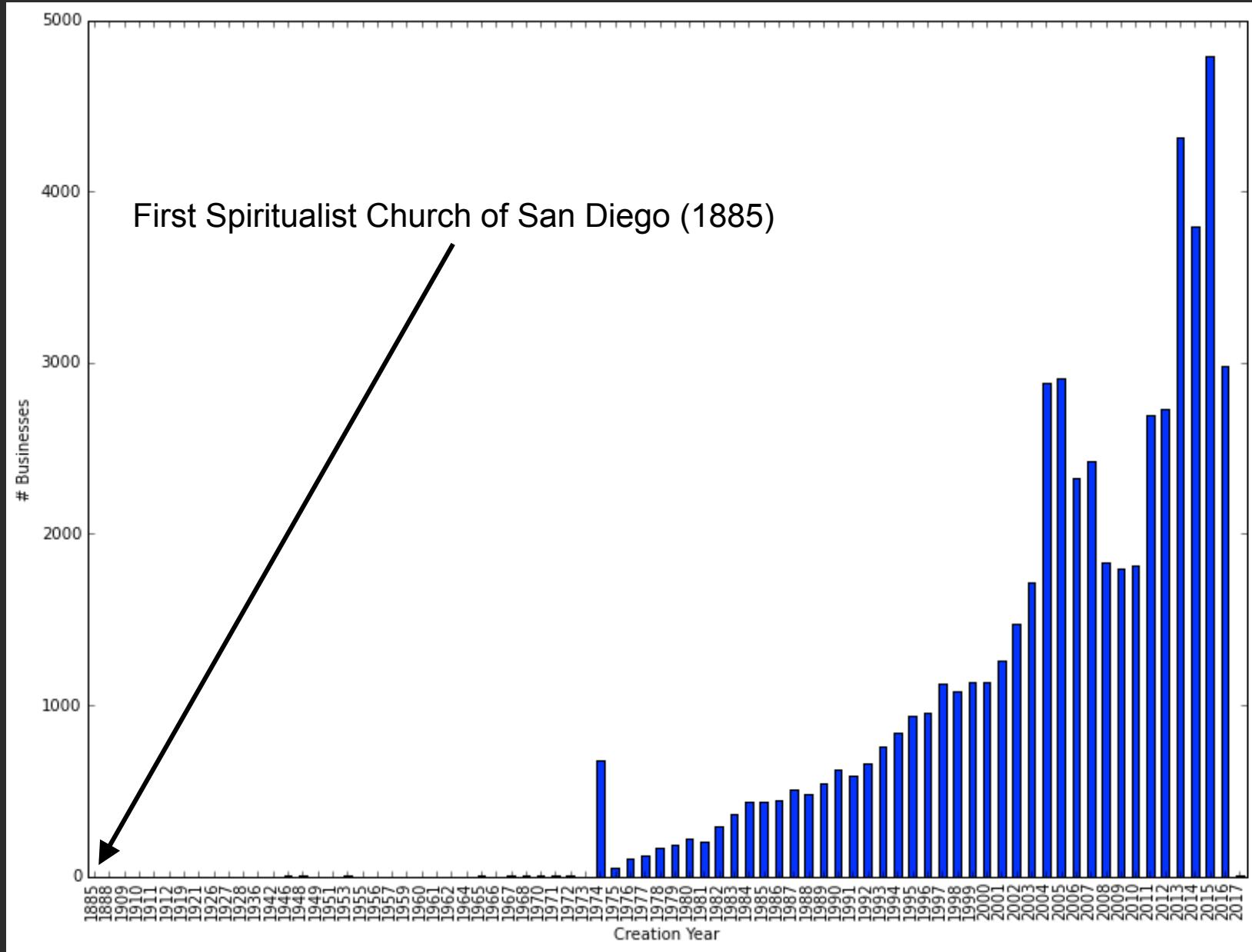
Possible Algorithms for the Model

- Linear Regression
- Something more complicated!?

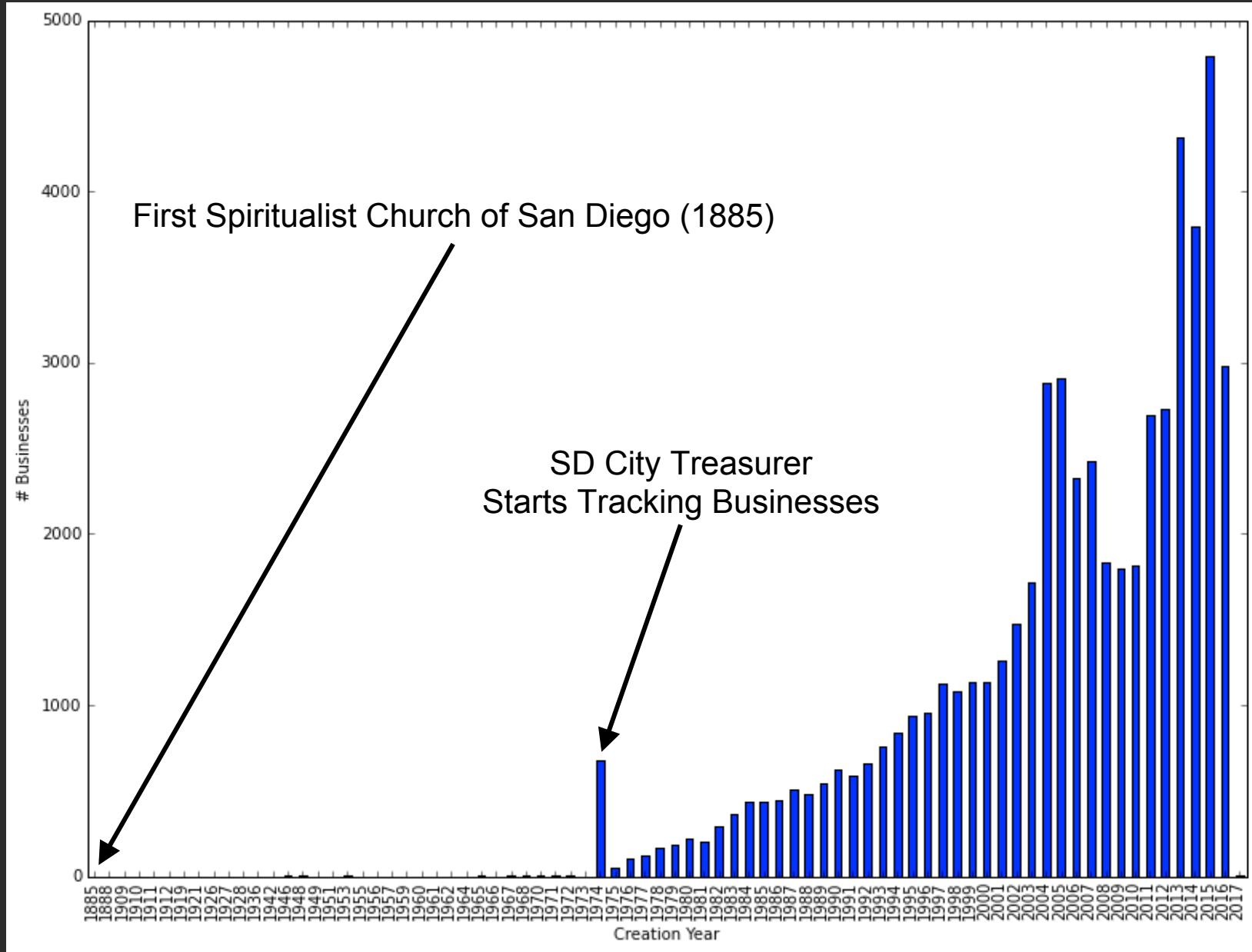
Business Info – Creation Date



Business Info – Creation Date



Business Info – Creation Date



Business Info – NAICS (North American Industry Classification System)

Business Info – NAICS (North American Industry Classification System)



	# Stores	NAICS Code	NAICS Description
0	31	722	FOOD SERVICES & DRINKING PLACES
1	27	722213	SNACK & NONALCOHOLIC BEVERAGE BARS
2	6	7224	DRINKING PLACES (ALCOHOLIC BEVERAGES)
3	4	4452	SPECIALTY FOOD STORES
4	3	445299	OTHER MISCELLANEOUS STORE RETAILERS
5	3	4539	ALL OTHER SPECIALTY FOOD STORES
6	2	7223	SPECIAL FOOD SERVICES
7	1	45299	OTHER SPECIALTY FOOD STORES
8	1	44	ALL OTHER GENERAL MERCHANDISE STORES
9	1	7221	GENERAL MERCHANDISE STORES
10	1	7222	FOOD & BEVERAGE STORES
11	1	445	MISCELLANEOUS STORE RETAILERS
12	1	44531	LIMITED-SERVICE EATING PLACES
13	1	452	CAFETERIAS
14	1	453	FULL-SERVICE RESTAURANTS
15	1	5614	RETAIL TRADE
16	1	44529	BEER, WINE & LIQUOR STORES
17	1	722212	BUSINESS SUPPORT SERVICES

Business Info – NAICS (North American Industry Classification System)



	# Stores	NAICS Code	NAICS Description
0	31	722	FOOD SERVICES & DRINKING PLACES
1	27	722213	SNACK & NONALCOHOLIC BEVERAGE BARS
2	6	7224	DRINKING PLACES (ALCOHOLIC BEVERAGES)
3	4	4452	SPECIALTY FOOD STORES
4	3	445299	OTHER MISCELLANEOUS STORE RETAILERS
5	3	4539	ALL OTHER SPECIALTY FOOD STORES
6	2	7223	SPECIALTY FOOD SERVICES
7	1	45299	OTHER SPECIALTY FOOD STORES
8	1	4451	NONSTORE GENERAL MERCHANDISE STORES
9	1	7221	GENERAL MERCHANDISE STORES
10	1	7220	FOOD & BEVERAGE STORES
11	1	445	MISCELLANEOUS STORE RETAILERS
12	1	44531	LIMITED-SERVICE EATING PLACES
13	1	452	CAFETERIAS
14	1	453	FULL-SERVICE RESTAURANTS
15	1	5614	RETAIL TRADE
16	1	44529	BEER, WINE & LIQUOR STORES
17	1	722212	BUSINESS SUPPORT SERVICES

4 “simple” codes

79% are Code 72

13% are Code 44

The End...
...for now