

# Scribe Consulting Project

Michael Gooch



## Scribe scores sales leads for B2B SAAS companies

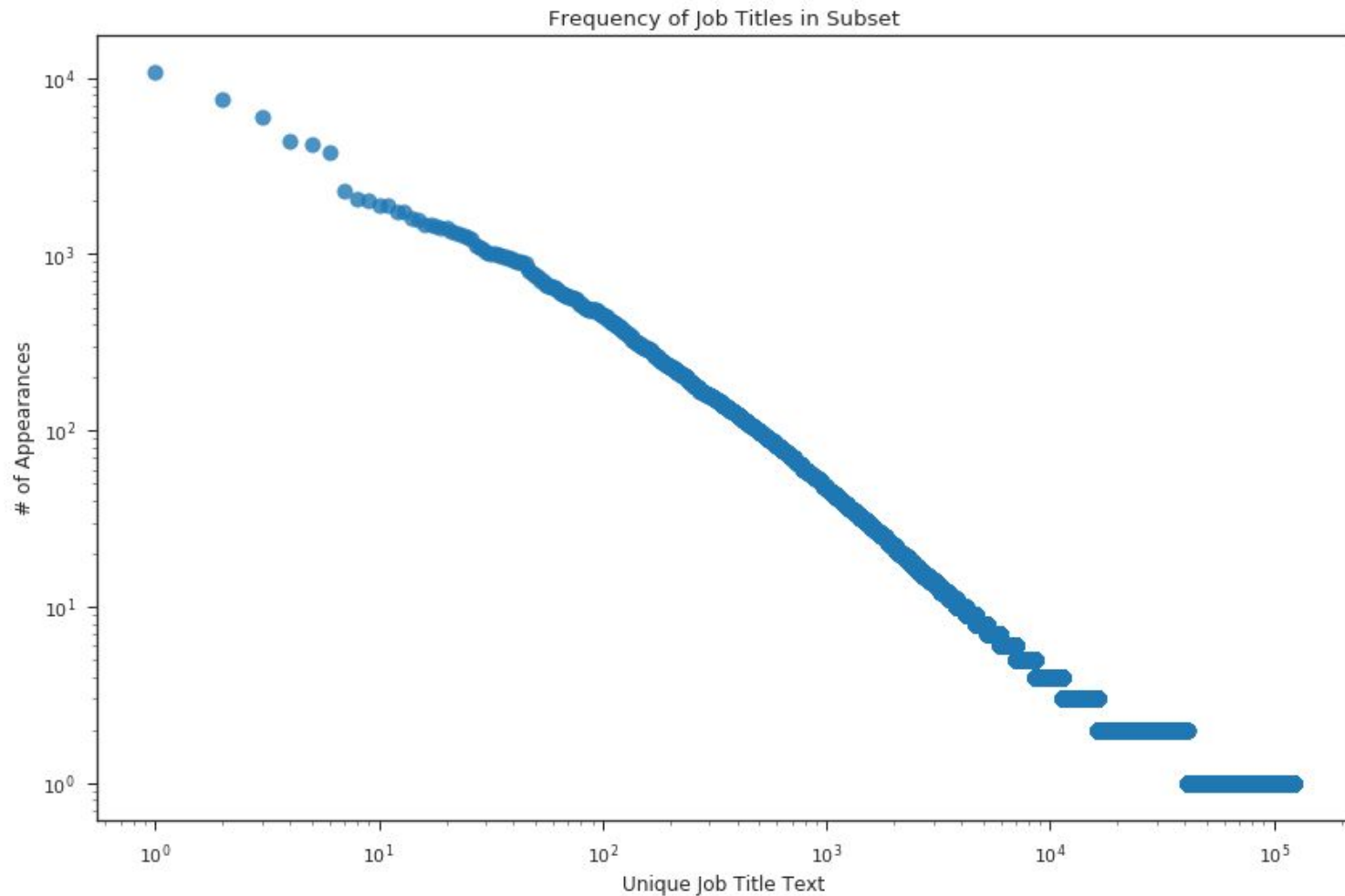
Lead ID	Name	Email	Company	Company size	Title	Category
1	Bob	bob@...	Google	72,053	Accountant	?
2	John	jko@...	JohnK inc	15	CEO	?
3	Seth	sbg@...	Amazon	>350,000	Security Guard	?
4	Shana	shan@....	13th Precinct	80	Police Chief	?
5	Siobhan	sbmcoy@..	Furniture R US	8000	Journeyman Carpenter	?



# Scribe scores sales leads for B2B SAAS companies

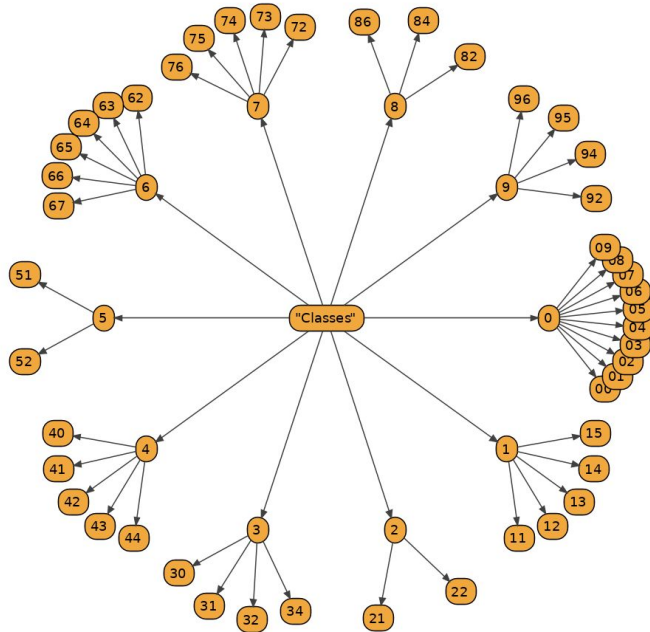
Lead ID	Name	Email	Company	Company size	Title	Category
1	Bob	bob@...	Google	72,053	Accountant	?
2	John	jko@...	JohnK inc	15	CEO	?
3	Seth	sbg@...	Amazon	>350,000	Security Guard	?
4	Shana	shan@....	13th Precinct	80	Police Chief	?
5	Siobhan	sbmcoy@..	Furniture R US	8000	Journeyman Carpenter	?

**Few Titles → high freq; Many titles → low freq**



# Job Classes available from external source

## Canadian Job National Occupational Classification



0011 attorney general  
0111 accounting department manager  
1111 bank branch accountant  
1511 clerk, mail distribution  
2111 elementary particle physicist  
2211 analytical chemistry technician  
3111 cardiac surgeon  
4011 chairman/woman, food sciences department - university  
5111 computer search librarian  
5254 yoga instructor  
6211 sales clerk supervisor  
7441 chain link fence erector  
8211 foreman/woman, logging  
9411 agglomerating attendant, iron ore pellets  
9521 aileron fitter - aircraft assembly  
9611 aluminum sheet separator

# Use Bag-of-Words with a classifier

“Police Officer” →

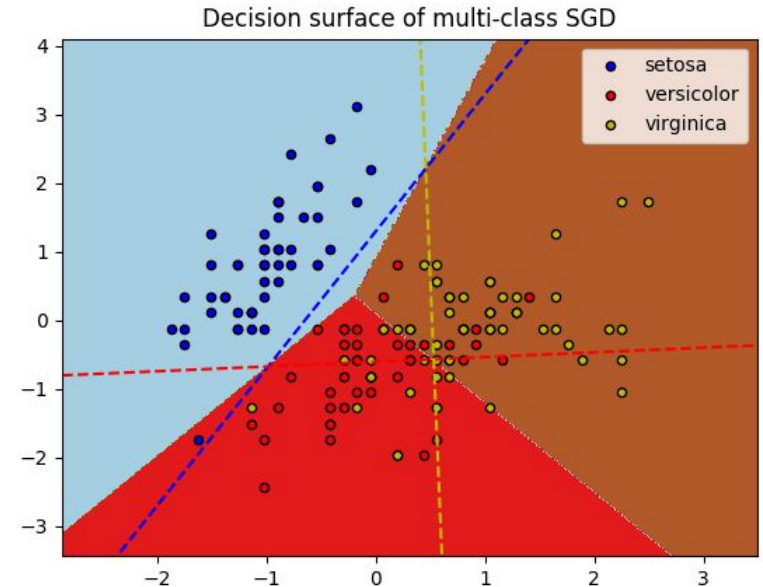
“police”	“chief”	“officer”	“doctor”	...
1	0	1	0	...

Stochastic Gradient Descent classifier  
N one-vs-all classifiers

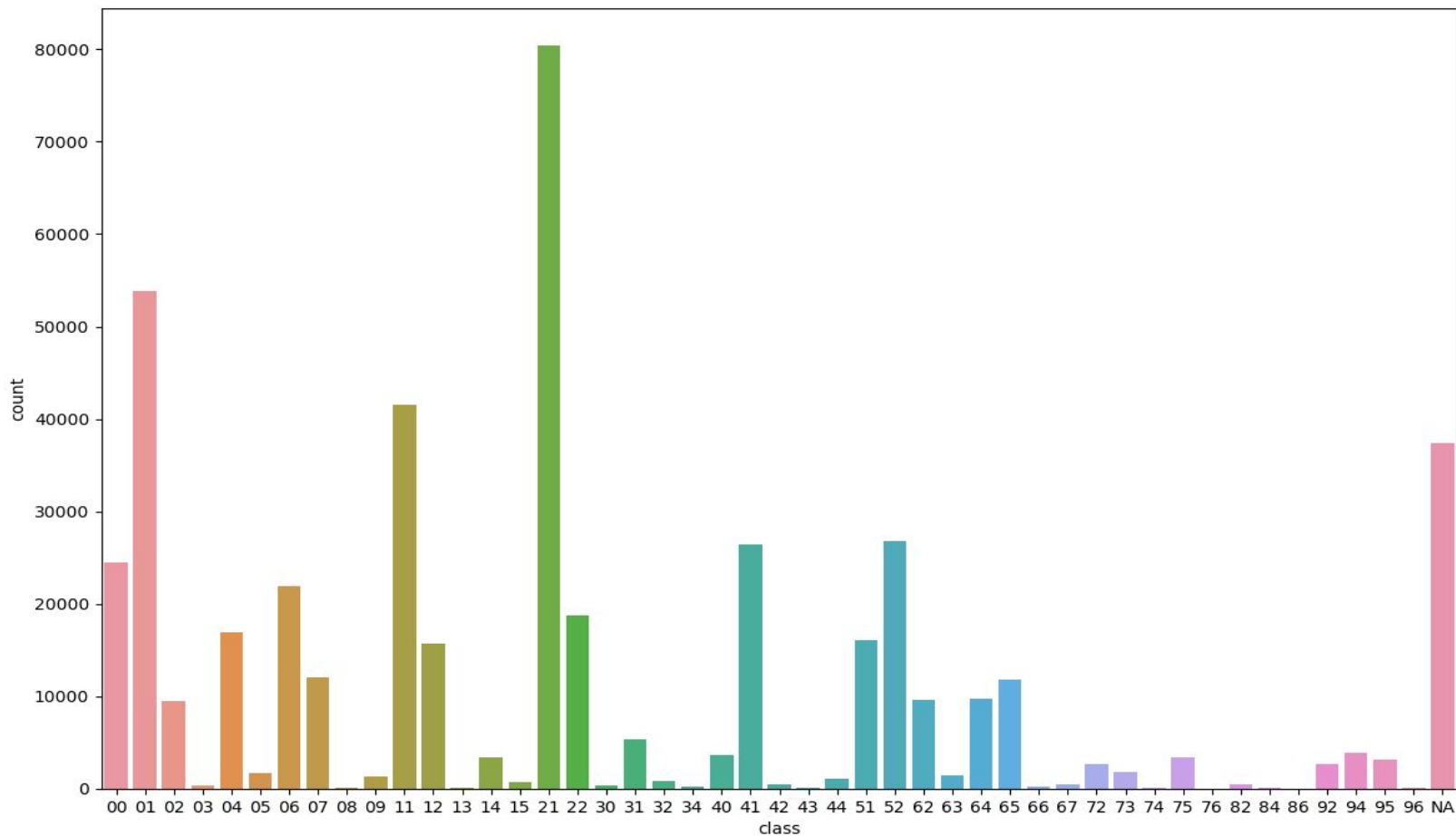
4

Occupations in education, law and social, community and government services

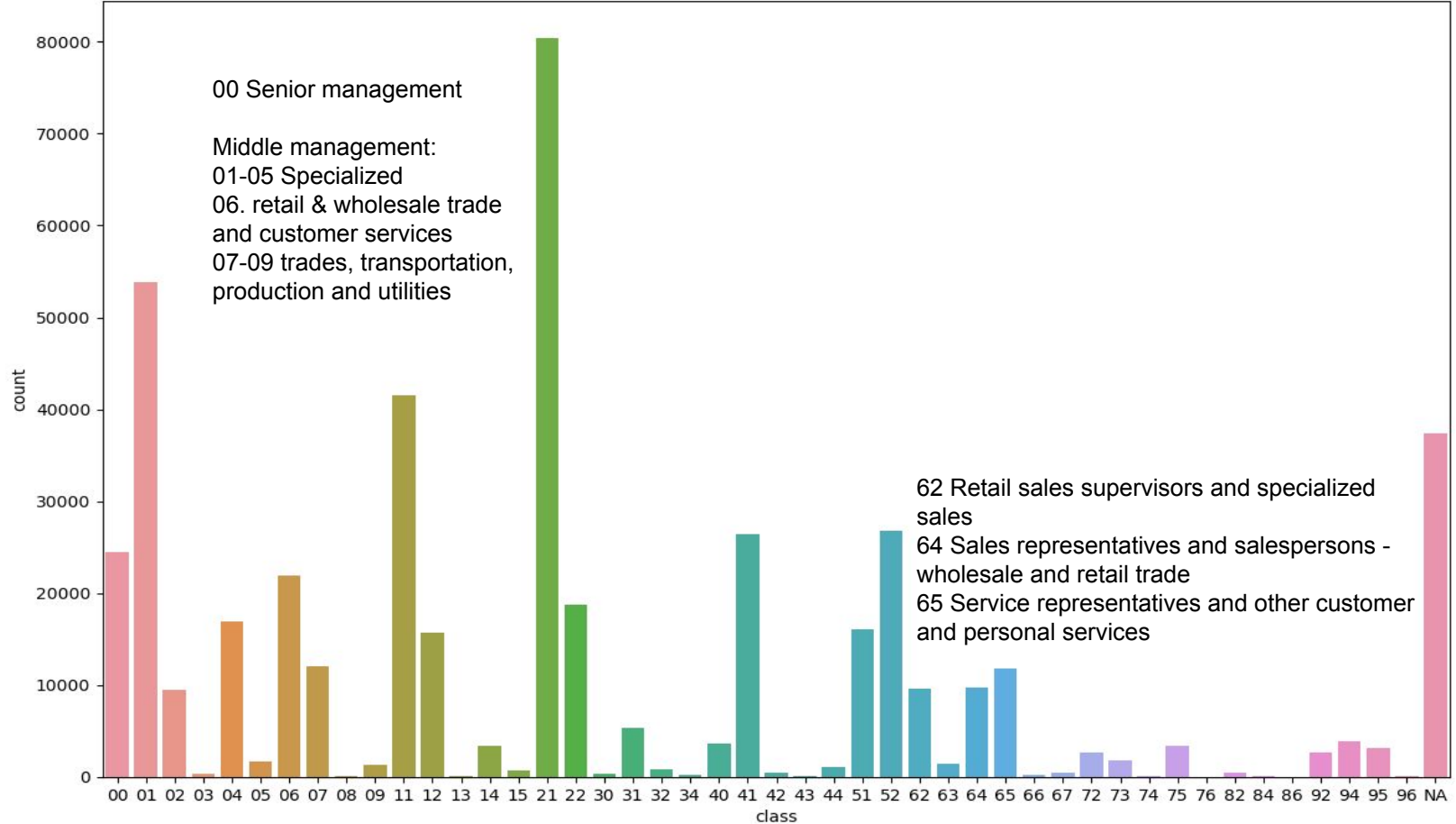
43 & 431 Occupations in front-line public protection services



## Applying classifier to scribe's records nets believable results

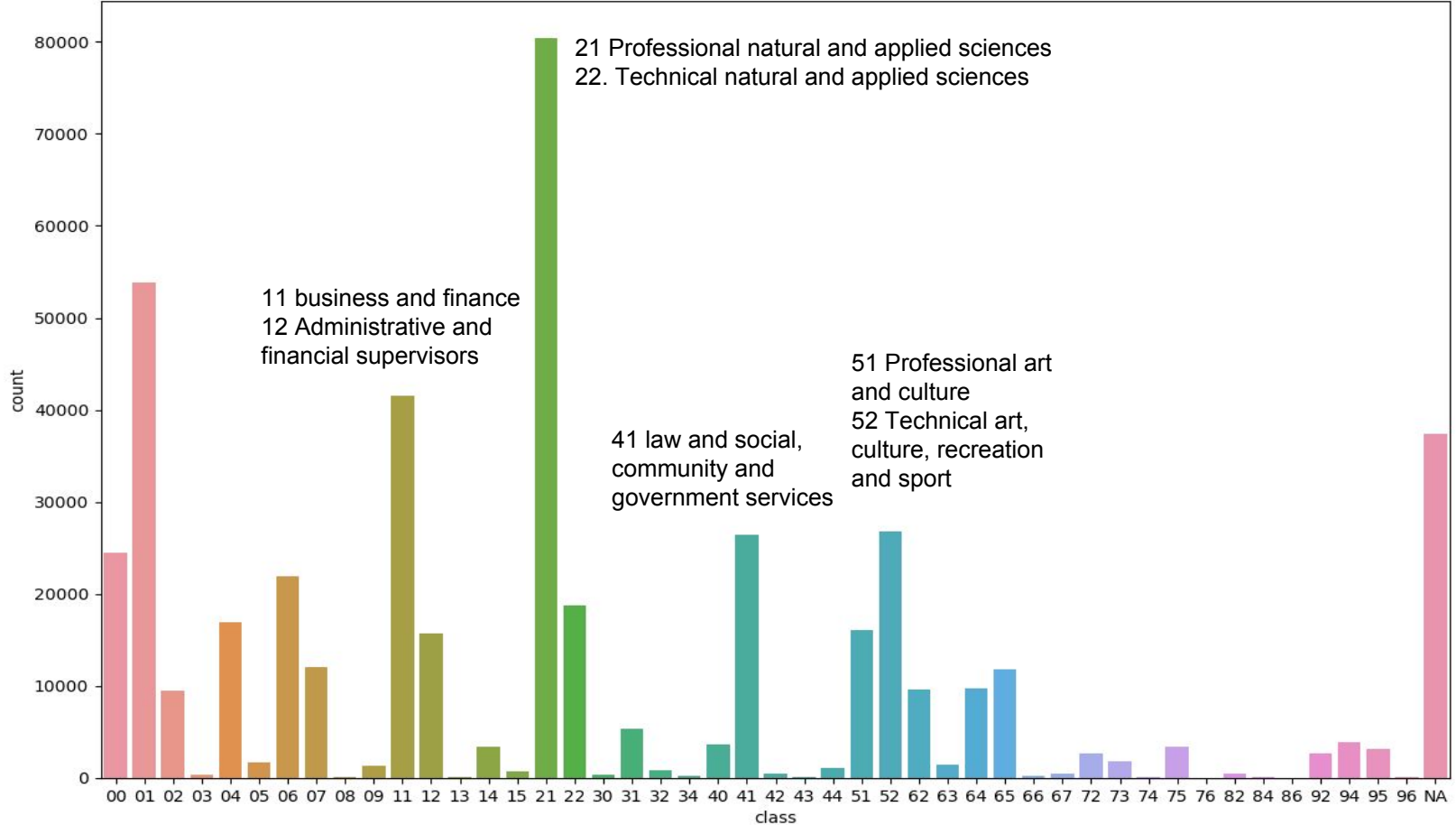


# Applying classifier to scribe's records nets believable results





# Applying classifier to scribe's records nets believable results



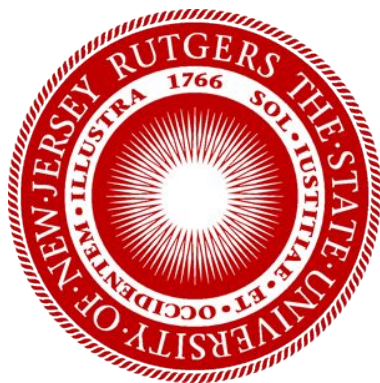
# Classification produces results useful for scoring!

Lead ID	Title	1	2	3	desc
1	Accountant	1	1	1	Auditors, accountants and investment professionals
2	Chief Executive Officer	0	0	1	Legislators and senior management
3	Security Guard	6	5	4	Security guards and related security service occupations
4	Police Chief	0	4	3	Managers in public protection services
5	Journeyman Carpenter	7	2	7	Carpenters and cabinetmakers



# Michael Gooch

BS Rutgers 2009



Dual Major Biochem & Biotech (bioinformatics)

PhD UConn Health 2017 Biomedical Science  
Genetics & Genome Sciences Dept.

Studied differences in patterns of mutation between  
hypermutating colon tumors and mutationally stable  
tumors using TCGA database.

