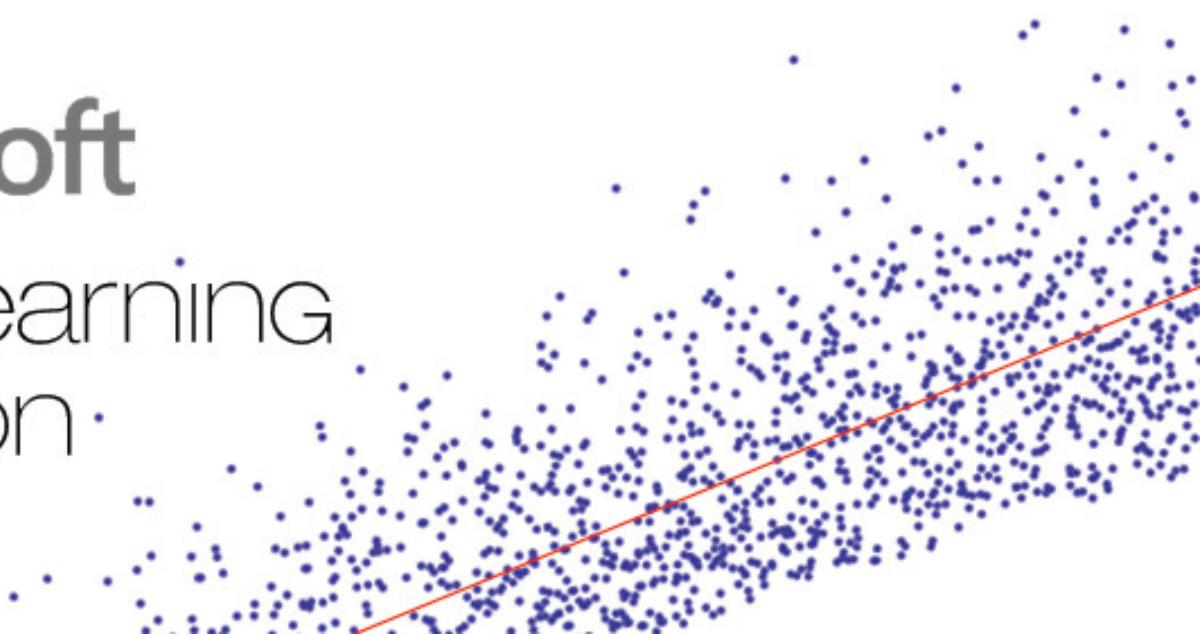


DataSense presents...



Machine Learning COMPETITION

Predicting salaries



.II DataSense - Make Data Make Sense

Educating and Inspiring the Next Generation of Data Scientists



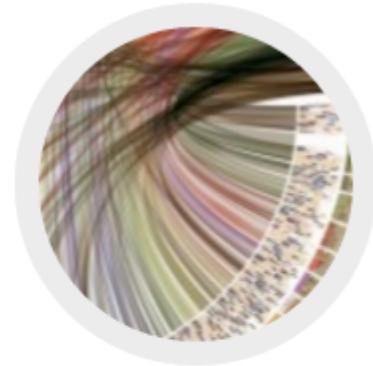
LEARN

Monthly Seminar Series



DO

Data Science Competition



WORK

Data Science Job Board

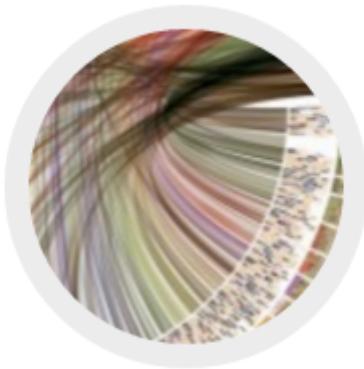
Are you interested in Data Science?

Sign up at MakeDataSense.ca to become a member (**free!**)

In fact, use the membership fee you would have spent for other clubs to buy yourself a cup of coffee – think of it as our treat!

.IDataSense - Make Data Make Sense

Educating and Inspiring the Next Generation of Data Scientists



WORK

Data Science Job Board

.IDataSense

Home What We Do Why Data Science About Us BECOME A MEMBER

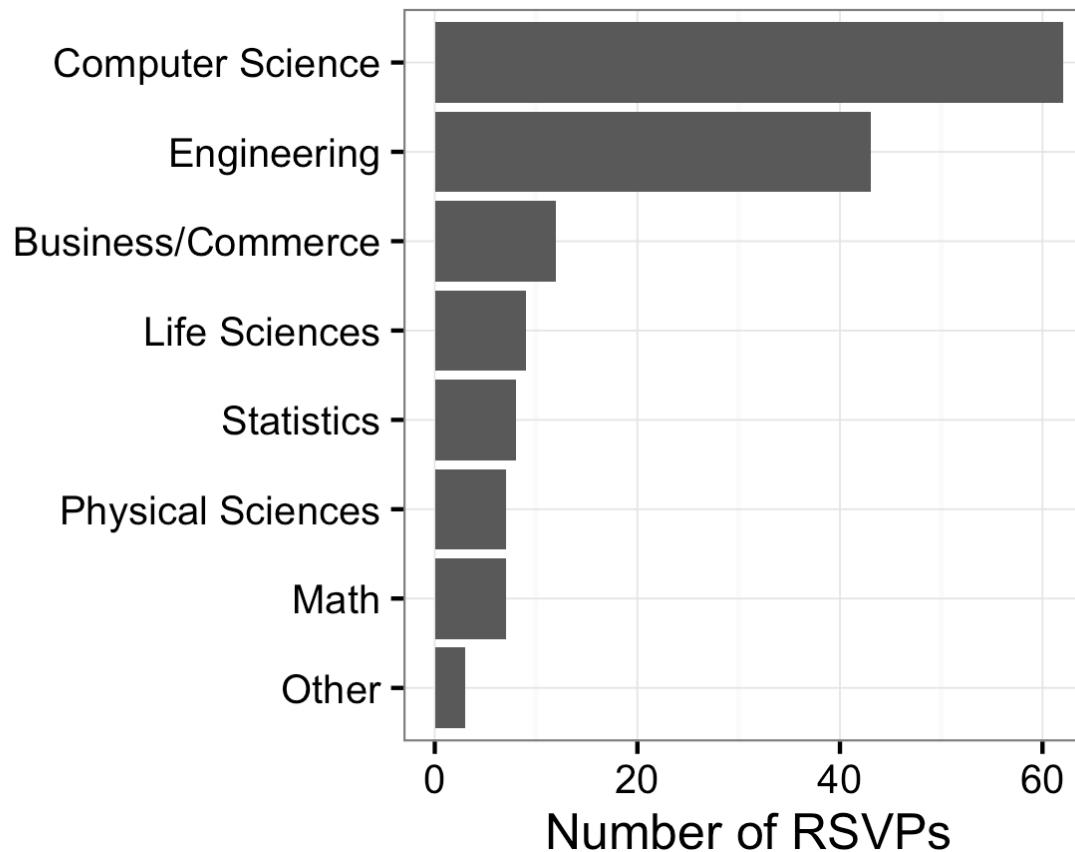
Job Listings Job Seeker Profiles

Job Listings RSS Reset

	Undergraduate Machine Learning Internship Empirical Results	1080 Mainland Street, Suite 225-7, Vancouver, BC, Canada V6B 2T4	Job Listings 2 days ago
	Server Scalability Engineer 3vGeomatics	202-2339 Columbia Street, Vancouver, BC	Job Listings 6 days ago
	Applied Researcher- Machine Learning 1QB Information Technologies	458-550 Burrard Street, Vancouver BC V6C 2B5	Job Listings 1 week ago
	Machine Learning Scientist D-Wave Systems	3033 Beta Ave, Burnaby	Job Listings 2 weeks ago
	Sr. Big data engineer INFOX Consulting	Richmond BC	Job Listings 3 weeks ago
	Data Scientist INFOX Consulting	Downtown Vancouver	Job Listings 3 weeks ago



Breakdown of degree levels





Introduction

Statistics Canada produces a national Census. You will use a subset of the 2006 Census data to better understand what factors lead to higher wages in Canada.



The Data

Training data ~51,000 responses

S18	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
1	CASEID	cma	pr	hhsiz	agegrp	marst	sex	citizen	fol	kol	attsch	cip	hdgree	hrswrk	locstud	mtnno	mfs	powst	cow	naics	nochrd	noc	morethan60kyr	
2	29	Ottawa_Gatineau	Quebec	3	25	Legally_marr	female	Canada_birt	French	Both_English	Not_Attende	Humanities	9	40	Ontario	No_non_ffi	Humanities	_Worked_in_Paid_worker	Public_admit	Professional_Occupations			FALSE	
3	36	Vancouver	BritishColum	7	35	Legally_marr	male	Other	English	English	Not_Attende	Business_Ma	12	40	Southeast_asia_Indo_Iranian	Commerce	_Worked_in_Paid_worker	Finance_and_Administrativ	Financial_se			FALSE		
4	39	other	Quebec	1	55	Single	male	Canada_birt	French	French	Not_Attende	Architecture	9	32	Quebec	No_non_ffi	Engineering	_Worked_in_Self_employ	Professional_Professionals	Occupations		FALSE		
5	57	Montreal	Quebec	2	30	Single	female	Canada_birt	French	French	Not_Attende	Education	9	40	Quebec	No_non_ffi	Educational	_Worked_in_Paid_worker	Educational_Professionals	Teachers_and_teachers		FALSE		
6	81	Edmonton	Alberta	3	50	Legally_marr	female	Canada_birt	French	Both_English	Not_Attende	Education	10	40	Ontario	No_non_ffi	Educational	_Worked_in_Paid_worker	Educational_Administrativ	Financial_se		TRUE		
7	84	Ottawa_Gatineau	Ontario	4	45	Legally_marr	male	Canada_birt	English	Both_English	Not_Attende	Social_Behav	9	50	Ontario	No_non_ffi	Social_scien	_Worked_in_Paid_worker	Other_servic	Professionals_Occupations		FALSE		
8	100	other	BritishColum	3	40	Legally_marr	male	Canada_birt	English	English	Not_Attende	Social_Behav	9	50	Ontario	No_non_ffi	Social_scien	_Worked_in_Paid_worker	Public_admi	Skilled_sales_Occupations		TRUE		
9	101	other	Ontario	2	35	Legally_marr	female	Canada_birt	English	English	Not_Attende	Business_Ma	9	35	Ontario	No_non_ffi	Social_scien	_Worked_in_Paid_worker	Health_care_Professionals	Occupations		FALSE		
0	102	Toronto	Ontario	4	55	Legally_marr	male	Canada_birt	English	English	Not_Attende	Business_Ma	12	0	Ontario	No_non_ffi	Commerce	_Worked_in_Paid_worker	Wholesale_t	Managers_Other mana		TRUE		
1	103	Toronto	Ontario	2	35	Single	male	Canada_birt	English	Both_English	Not_Attende	Architecture	12	50	Ontario	Spanish	Engineering	_Worked_in_Paid_worker	Manufacturi	Managers_Other mana		TRUE		
2	127	Toronto	Ontario	3	60	Legally_marr	male	Canada_birt	English	English	Not_Attende	Physical_Life	13	98	Ontario	No_non_ffi	Mathematic	_Worked_in_Paid_worker	Professional_Professionals	Occupations		TRUE		
3	217	Toronto	Ontario	2	25	Single	male	Other	English	Both_English	Not_Attende	Architecture	12	40	Ontario	No_non_ffi	Engineering	_Worked_in_Paid_worker	Professional_Professionals	Occupations		FALSE		
4	218	Brantford_Guelph_E	Ontario	4	50	Legally_marr	male	Canada_birt	English	English	Not_Attende	Education	10	40	Ontario	No_non_ffi	Educational	_Worked_in_Paid_worker	Educational_Professionals	Teachers_and_teachers		TRUE		
5	226	Toronto	Ontari	7	30	Legally_marr	female	Canada_birt	French	French	Not_Attende	Physical_Life	12	40	Ontario	No non ffi	Agricultural	_Worked_in_Paid_worker	Information_Professionals	Occupations		FAISF		

Testing data ~12,000 responses

A1	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	
1	CASEID	cma	pr	hhsiz	agegrp	marst	sex	Formula Bar	citizen	fol	kol	attsch	cip	hdgree	hrswrk	locstud	mtnno	mfs	powst	cow	naics	nochrd	noc
2	67	Montreal	Quebec	1	25	Single	female	Canada_nati	French	Both_English	Not_Attende	Humanities	9	50	Quebec	Other_Europe	Humanities	_Worked_in_Paid_worker	Professional_Professionals	Occupations_in_art_culture			
3	193	Vancouver	BritishColum	3	40	Legally_marr	male	Canada_nati	English	English	Not_Attende	Architecture	12	40	Eastern_Asia	Chinese_lang_Engineering	_Worked_in_Paid_worker	Professional_Professionals	Occupations_in_natural_and				
4	286	other	Ontario	4	45	Legally_marr	male	Canada_birt	English	English	Not_Attende	Business_Ma	9	66	Ontario	No non ffi	Commerce	_Worked_in_Paid_worker	Manufacturi	Managers_Other management_occu			
5	307	Oshawa	Ontario	2	35	Legally_marr	male	Canada_birt	English	English	Not_Attende	Humanities	12	38	Ontario	No non ffi	Humanities	_Worked_in_Paid_worker	Health_care_Professionals	Occupations_in_social_scien			
6	312	Montreal	Quebec	3	50	Legally_marr	female	Canada_birt	French	French	Not_Attende	Social_Behav	9	35	Quebec	No non ffi	Social_scien	_Worked_in_Paid_worker	Public_admi	Professional_Occupations_in_social_scien			
7	348	other	BritishColum	3	30	Legally_marr	female	Canada_birt	English	English	Not_Attende	Humanities	9	40	Manitoba	No non ffi	Humanities	_Worked_in_Paid_worker	Other_servic	Professionals_Occupations_in_social_scien			
8	429	other	BritishColum	2	75	Single	female	Canada_birt	English	English	Not_Attende	Education	9	48	Alberta	No non ffi	Educational	_Worked_in_Paid_worker	Educational_Professionals	Teachers_and_professors			
9	454	Toronto	Ontario	3	55	Legally_marr	male	Canada_nati	English	English	Not_Attende	Architecture	9	50	Ontario	Indo_Iranian	Engineering	_Worked_in_Paid_worker	Manufacturi	Skilled_crafts_other_trades_occu			
10	476	Sherbrooke_Quebec	Quebec	3	50	Divorced	female	Canada_birt	French	Both_English	Not_Attende	Education	12	0	Quebec	No non ffi	Educational	_Worked_in_Paid_worker	Educational_Professionals	Teachers_and_professors			
11	477	Ottawa_Gati	Ontario	2	50	Legally_marr	female	Canada_nati	English	English	Not_Attende	Business_Ma	9	0	Quebec	No non ffi	Commerce	_Worked_in_Paid_worker	Health_care_Semi_profs	Occupations_in_social_scien			
12	481	Ottawa_Gati	Ontario	3	40	Single	female	Canada_nati	English	English	Not_Attende	Business_Ma	12	40	United_State	No non ffi	Commerce	_Worked_in_Paid_worker	Wholesale_t	Managers_Other management_occu			
13	520	Toronto	Ontario	5	40	Legally_marr	female	Canada_birt	English	English	Not_Attende	Visual_Perfo	9	50	Ontario	No non ffi	Fine_and_ap	_Worked_in_Paid_worker	Retail_trade	Managers_Senior_management_occu			
14	534	Ottawa_Gati	Ontario	3	50	Legally_marr	male	Canada_birt	English	Both_English	Not_Attende	Social_Behav	9	38	Ontario	Italian	Social_scien	_Worked_in_Paid_worker	Public_admi	Professionals_Occupations_in_social_scien			
15	666	London	Ontario	4	50	Legally_marr	male	Canada_birt	English	Both_English	Not_Attende	Social_Behav	9	48	Ontario	No non ffi	Social_scien	_Worked_in_Paid_worker	Information_Managers	Other_management_occu			
16	670	Toronto	Ontario	4	35	Legally_marr	male	Canada_nati	English	English	Not_Attende	Mathematics	9	40	Southeast_asia_Indo_Iranian	Mathematics	_Worked_in_Paid_worker	Professional_Professionals	Occupations_in_natural_and				



Your Objective

Tell me what I should do to make more money when I graduate.

- **Build a machine learning model that predicts whether or not a household earns more than 60k/year**
- **Use this model to figure out what someone should do to maximize their income**



Guiding questions

- Does a Masters degree make me earn more money?
- Does living in Vancouver make me earn more money?
- Do rich people commute or drive to work?
- Do men make more money than women?

Submission details

- Due November 27, 2016 @ midnight online – email hello@makedatasense.ca with “Census competition – [team name]”
- Submit two things:
 - A ~500 word/5 figures writeup (HTML or PDF only) of your findings telling me what I should do to make money
 - A CSV of your predictions

CASEID	morethan60
67	TRUE
193	FALSE
286	FALSE
307	FALSE
312	FALSE
348	FALSE

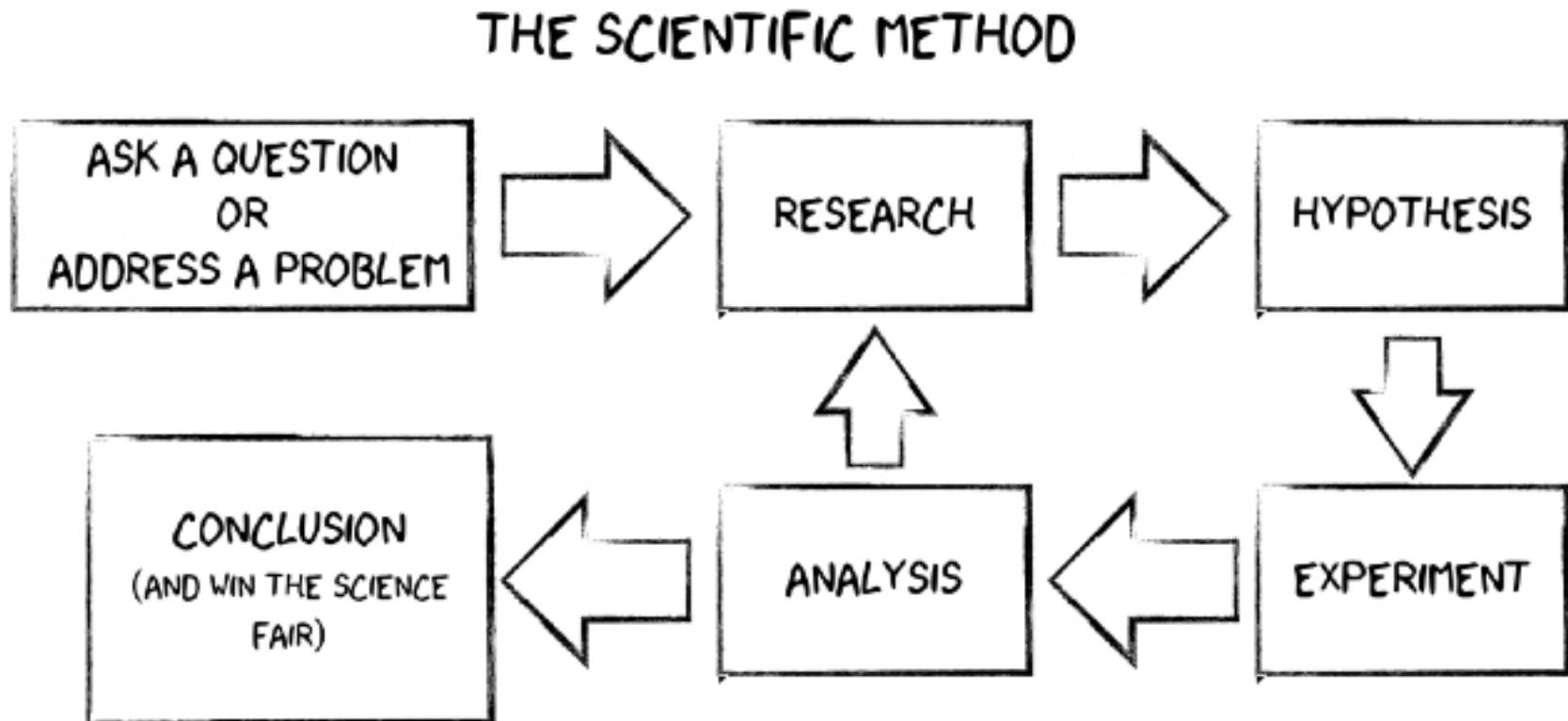


How you're judged

- Round 1 - Your model's accuracy
 - Top 5 teams with highest accuracy invited to present
- Round 2 - How convincing your recommendations are
 - 5 minute presentation + 5 minute Q&A with judges

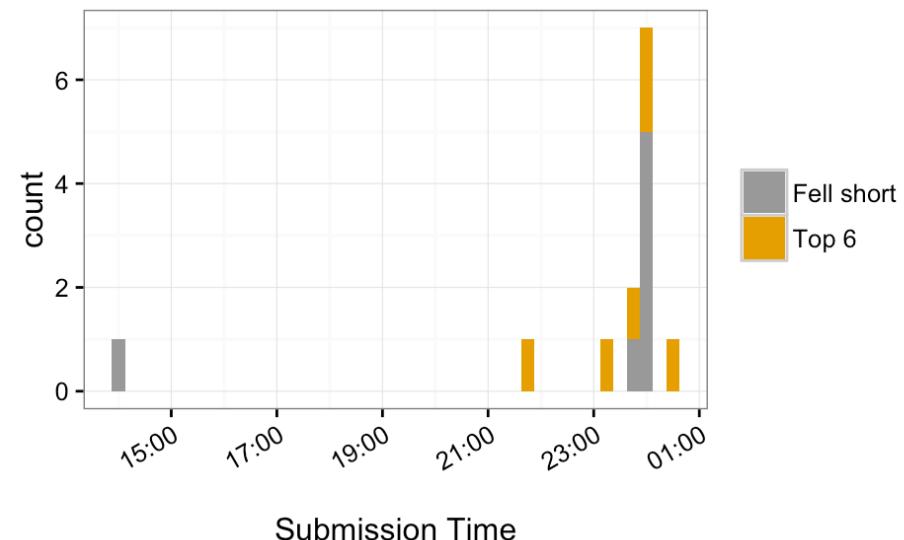
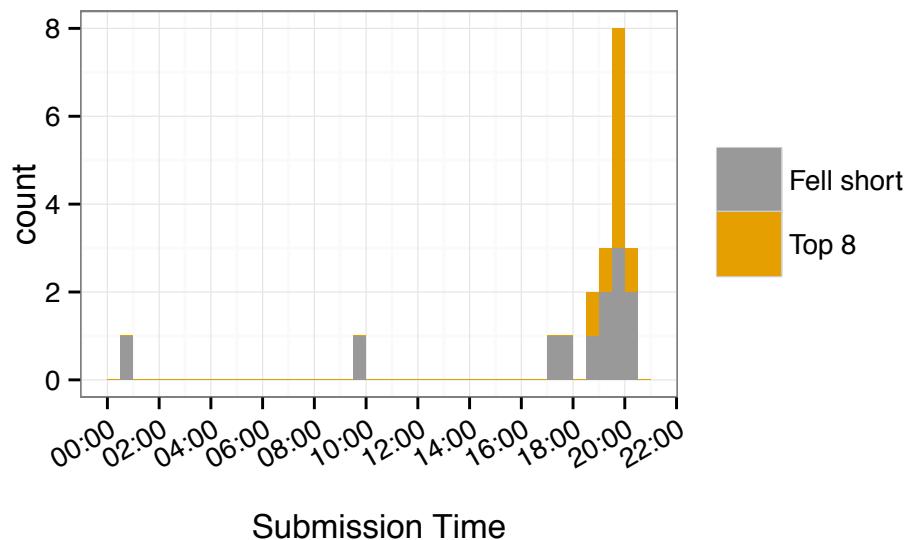


The (data) scientific method



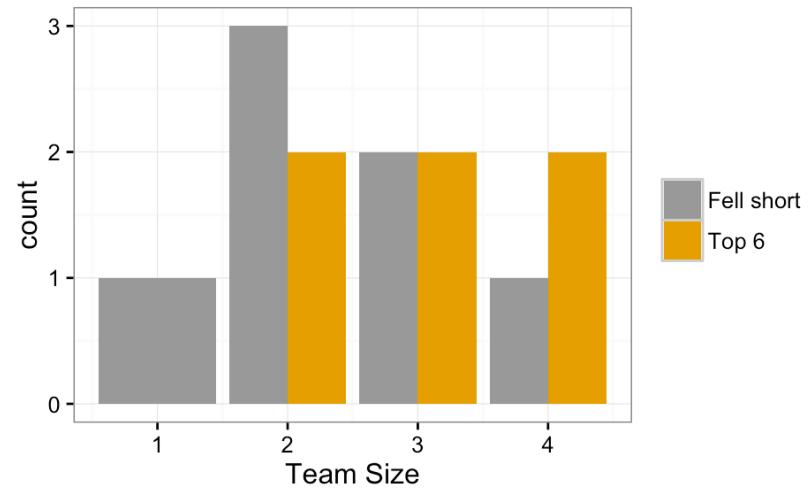
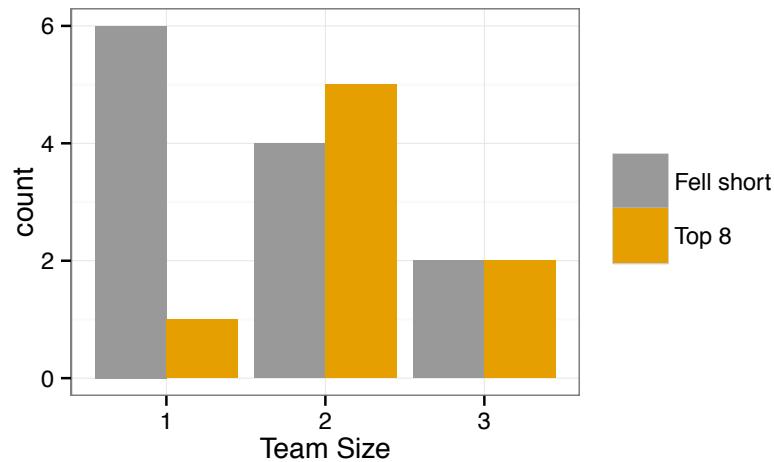


Who won our previous competitions?





Who won our previous competitions?



makedatasense.ca/microsoft

Access these set of slides and
download the data!