Predictors of Teams' High Performance

CIND 820 Final Project Presentation

By Ibrahim Ibrahim

Dec 2020

Ryerson University

Overview

Ryerson University

Research question

what is my intent and why does it matter?

- Organizations heavily invest to hear from their employees and get very little in return
- Predictive analytics and advances in AI can complement our ability to deeply listen and give us super listening powers
- My research question is: What are the predictors of team's high performance? Can we figure out what makes some teams succeed and make this broadly available to help ALL teams succeed?

Ryerson University

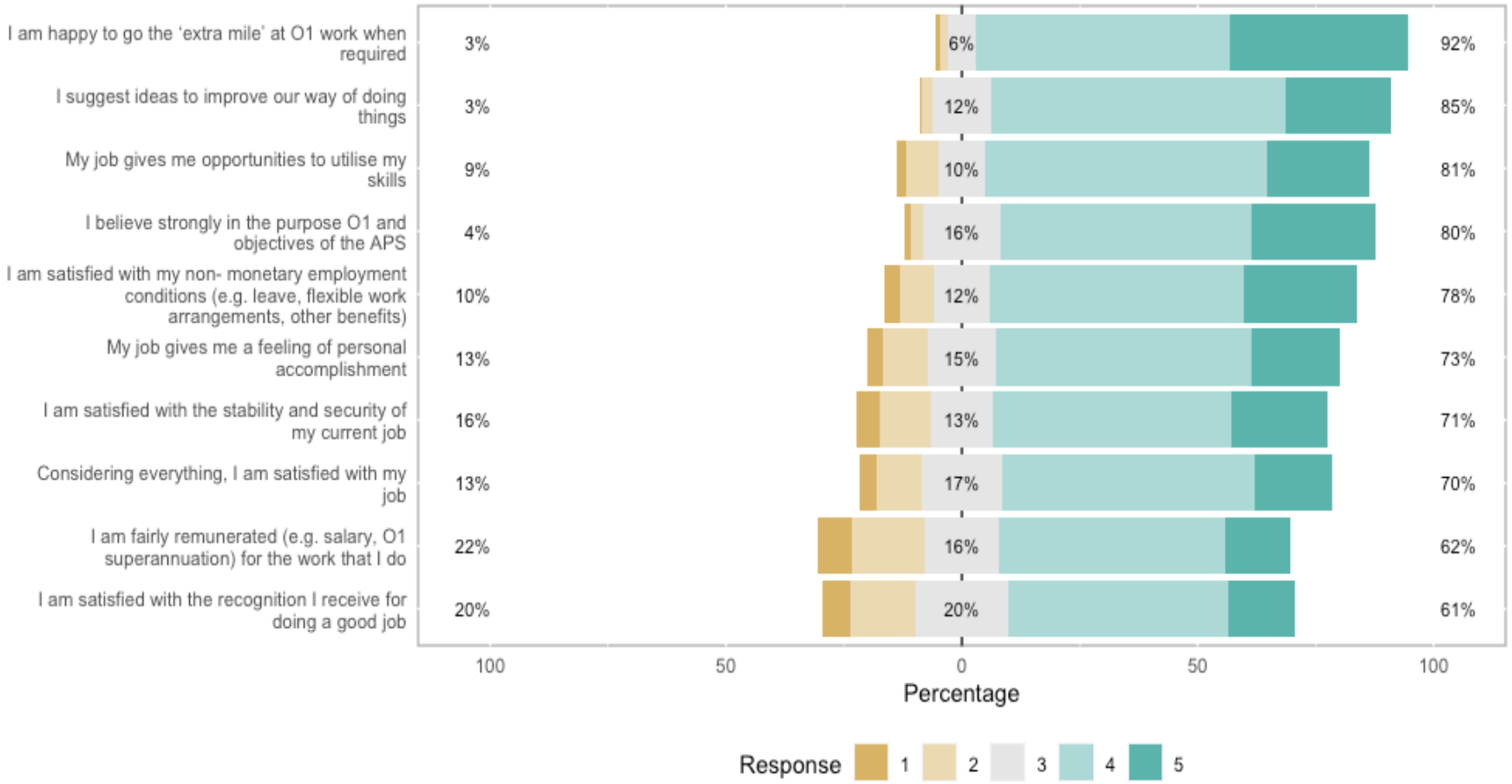
Data set Overview

- **2018 census (training set)**: https://data.gov.au/data/dataset/aps_employee_census_2018. The 2018 census shows individual answers to 301 questions by 103,137 respondents who answered the survey in 2018.

- **2019 census (test set)**: https://data.gov.au/data/dataset/aps_employee_census_2019. The 2019 census shows individual answers to 380 questions by 104,471 respondents who answered the survey in 2019.

Data set

Questions making up 1 independent variable: Job engagement



Ryerson University

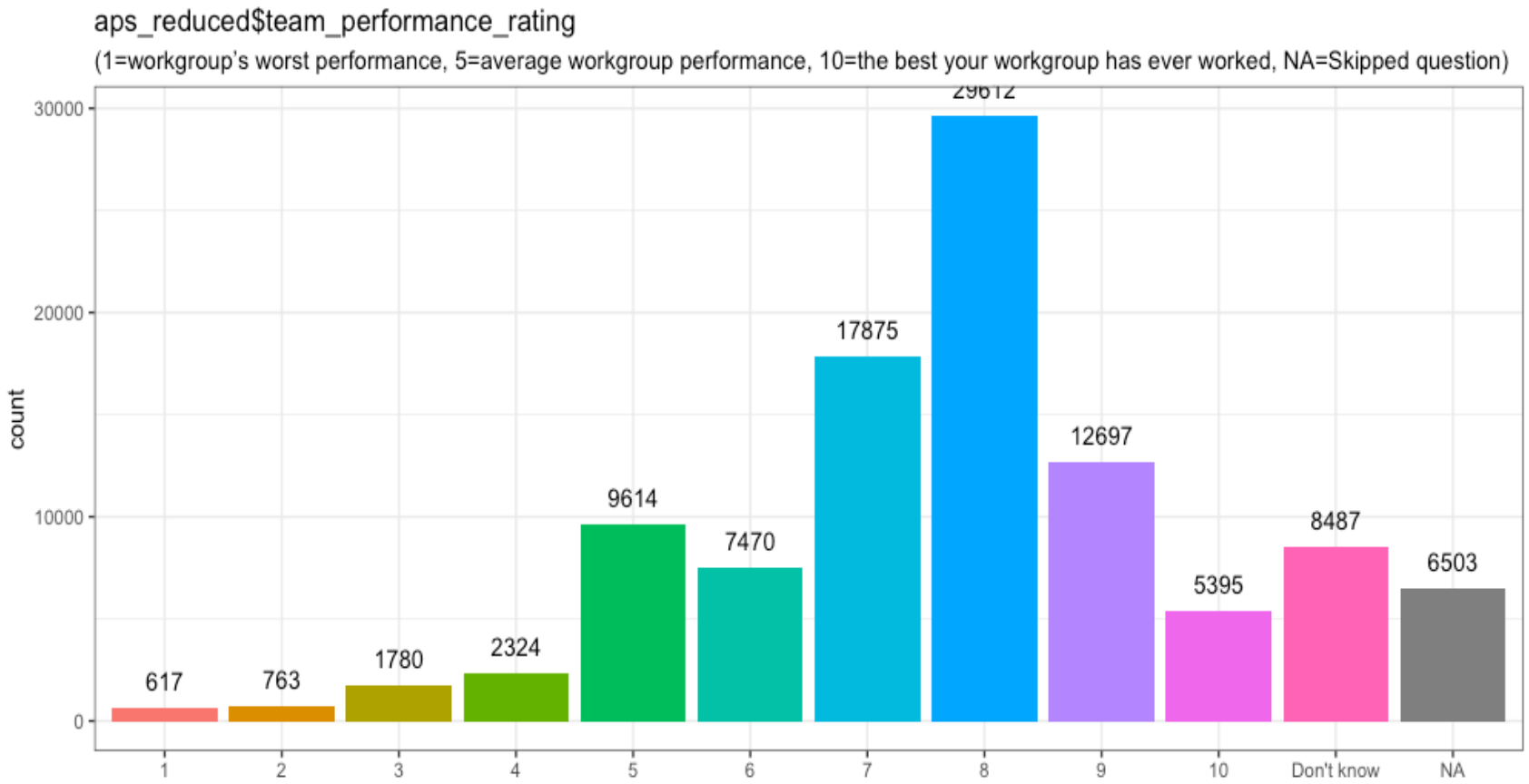
Data set

11 independent variables and 1 dependent variable

Category of questions to be included (11 factors under study + dependent var)	# of questions
1- job engagement	10
2- team engagement	4
3- supervisor engagement	11
4- senior manager engagement	12
5- agency engagement	17
6- leadership engagement	7
7- wellbeing	13
8- team performance support	4
9- risk culture	5
10- innovation	5
11- values	3
dependent variable (team performance rating assessment)	1
total number of questions to be included (variables)	92

Data set

The dependent variable



Ryerson University

Data set

The dependent variable has no strong correlations and the least number of moderate correlations vs. other variables

	Job eng.	Team eng.	Supervisor eng.	Senior manager eng	agency eng.	Team performance support	Risk culture	innovation	Leadership eng.	wellbeing	values	Team performanc rating
Job engagement	1.00	0.54	0.57	0.58	0.75	0.55	0.50	0.63	0.55	0.71	0.46	0.40
Agency engagement	0.75	0.53	0.52	0.63	1.00	0.62	0.65	0.70	0.71	0.74	0.53	0.45
wellbeing	0.71	0.57	0.64	0.57	0.74	0.65	0.60	0.67	0.58	1.00	0.56	0.45
innovation	0.63	0.45	0.56	0.56	0.70	0.58	0.65	1.00	0.57	0.67	0.47	0.42
Senior manage engagement	0.58	0.44	0.49	1.00	0.63	0.48	0.50	0.56	0.71	0.57	0.48	0.36
Supervisor engagement	0.57	0.59	1.00	0.49	0.52	0.52	0.41	0.56	0.38	0.64	0.52	0.40
Team performance support	0.55	0.51	0.52	0.48	0.62	1.00	0.55	0.58	0.49	0.65	0.48	0.53
Leadership engagement	0.55	0.38	0.38	0.71	0.71	0.49	0.58	0.57	1.00	0.58	0.48	0.34
Team engagement	0.54	1.00	0.59	0.44	0.53	0.51	0.40	0.45	0.38	0.57	0.53	0.46
Risk culture	0.50	0.40	0.41	0.50	0.65	0.55	1.00	0.65	0.58	0.60	0.45	0.38
values	0.46	0.53	0.52	0.48	0.53	0.48	0.45	0.47	0.48	0.56	1.00	0.39
Team performance rating	0.40	0.46	0.40	0.36	0.45	0.53	0.38	0.42	0.34	0.45	0.39	1.00

.40-.59 moderate

.60-.79 strong

Ryerson University

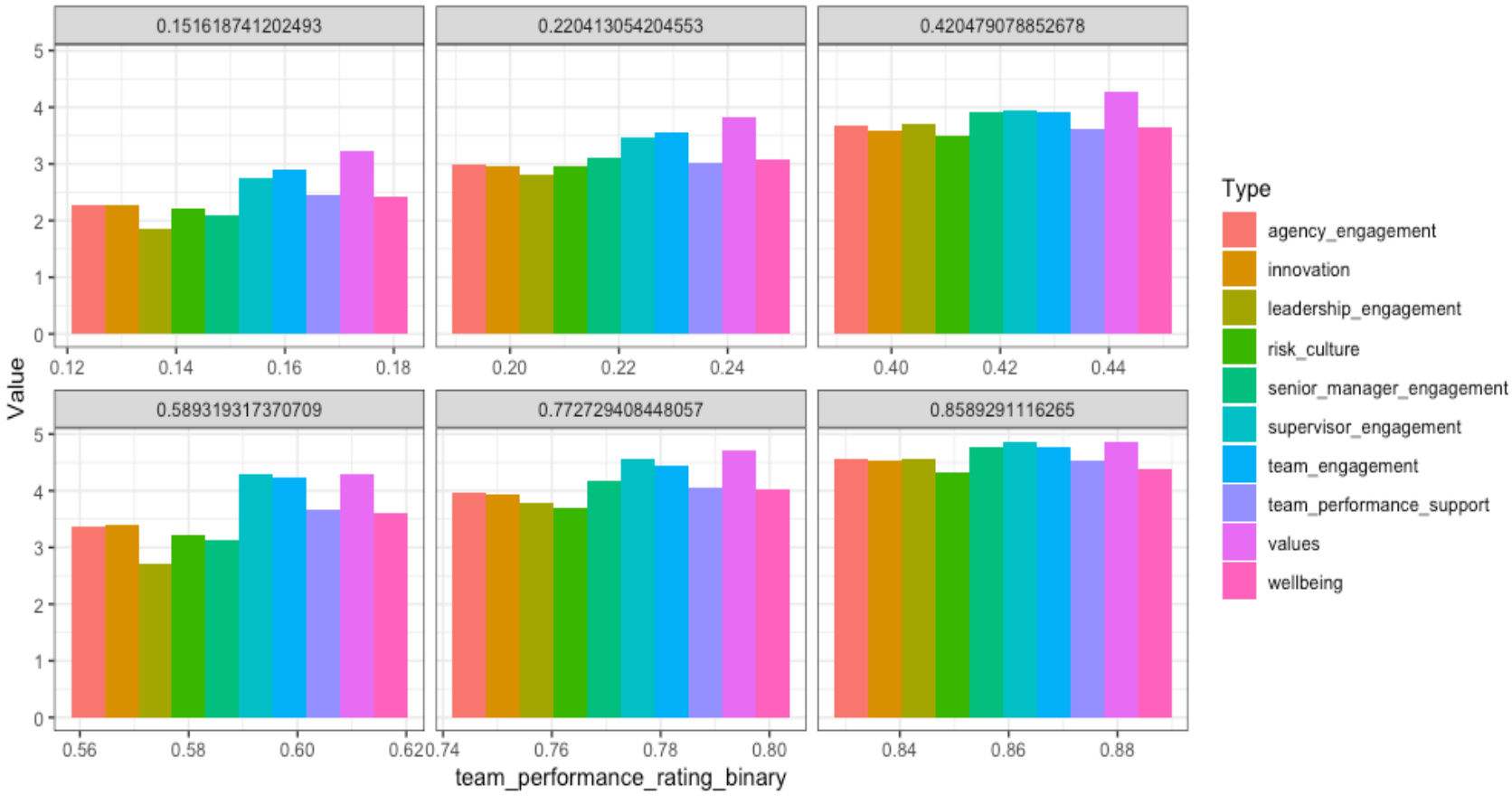
Data set K-means cluster analysis

6 Team performance clusters

1- Low

2- Low

3- Medium



4- Medium

5- High

6- High

Ryerson University

Data set

20% of participants from the High and Low performing clusters did not answer the dependent variable question in line with their cluster

cluster #	1 - Low	2- Low	3- Medium	4- Medium	5- High	6- High
Dependent Variable Mean	0.15	0.22	0.42	0.59	0.77	0.85
0 #observations	4,219	11,060	11,274	6,329	4,837	1,399
% of cluster	**85%**	**78%**	58%	41%	**23%**	**14%**
1 #observations	754	3,127	8,180	9,082	16,446	8,518
% of cluster	**15%**	**22%**	42%	59%	**77%**	**86%**
total #observations	4,973	14,187	19,454	15,411	21,283	9,917
% total	6%	16%	23%	18%	25%	12%

Ryerson University

Predictive models' Performance Metrics

	Logistic regression	Decision tree	Ordinal logistic regression		
Accuracy			Low	Medium	High
2018 training set and 2019 test set	69.7%	69.4%	66.2%	66.2%	66.2%
2018 data 70% training - 30% test	69.7%	68.9%	63.9%	63.9%	63.9%
2018 data 10-fold cross validation	69.2%	68.6%	63.7%	63.7%	63.7%
Sensitivity / Recall / true +ve rate					
2018 training set and 2019 test set	69.8%	68.4%	18.6%	47.8%	82.3%
2018 data 70% training - 30% test	69.5%	68.6%	17.8%	46.6%	82.0%
2018 data 10-fold cross validation	69.4%	68.6%	17.0%	46.7%	81.6%
Specificity / true -ve rate					
2018 training set and 2019 test set	69.6%	70.7%	99.3%	78.5%	53.6%
2018 data 70% training - 30% test	69.9%	69.1%	99.3%	76.8%	53.5%
2018 data 10-fold cross validation	68.9%	68.6%	99.3%	76.2%	53.5%
Precision					
2018 training set and 2019 test set	75.7%	76.0%	59.4%	56.8%	70.7%
2018 data 70% training - 30% test	73.0%	72.2%	62.8%	57.2%	67.3%
2018 data 10-fold cross validation	72.5%	72.0%	63.4%	56.3%	67.4%
F1 score					
2018 training set and 2019 test set	72.6%	72.0%	28.3%	51.9%	76.1%
2018 data 70% training - 30% test	71.2%	70.4%	27.7%	51.4%	73.9%
2018 data 10-fold cross validation	70.9%	70.3%	26.8%	51.1%	73.8%

Predictive model 1 : Binomial Logistic Regression

5 independent variables are significant predictors of teams' high performance

Coefficients:	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	7.67605	0.07685	-99.885	< 2e-16 ***
team_performance_support	0.90322	0.01595	56.619	< 2e-16 ***
team_engagement	0.57836	0.01622	35.652	< 2e-16 ***
values	0.20671	0.01611	12.831	< 2e-16 ***
agency_engagement	0.18524	0.02332	7.944	1.96e-15 ***
innovation	0.16930	0.01717	9.862	< 2e-16 ***
job_engagement	-0.06464	0.02172	-2.976	0.002920 **
supervisor_engagement	0.04255	0.01525	2.790	0.005272 **
senior_manager_engagement	0.07220	0.01387	5.206	1.93e-07 ***
risk_culture	0.05582	0.01646	3.390	0.000698 ***
leadership_engagement	-0.07569	0.01511	-5.010	5.44e-07 ***
wellbeing	-0.03461	0.02350	-1.473	0.140825

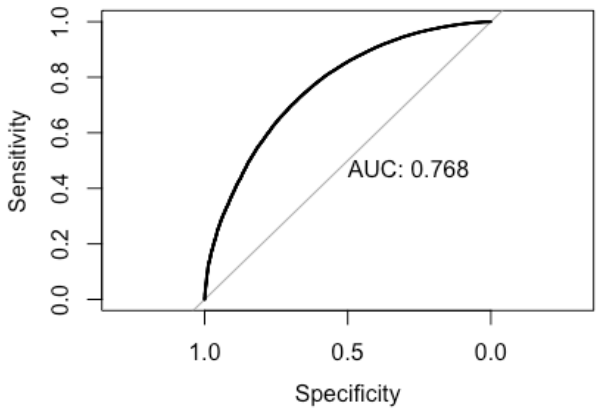
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Null deviance: 117573 on 85224 degrees of freedom

Residual deviance: 97985 on 85213 degrees of freedom

AIC: 98009

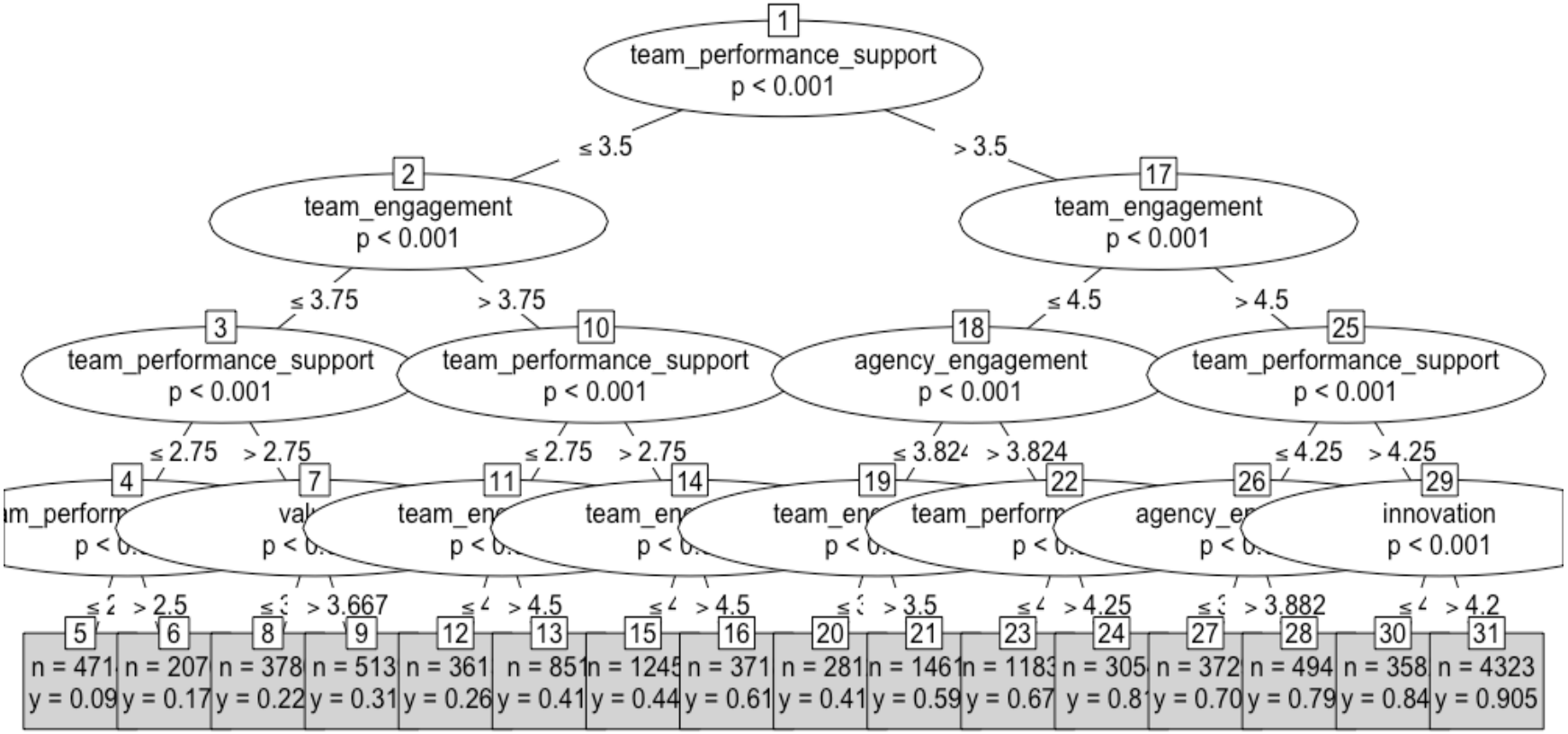
Number of Fisher Scoring iterations: 4



Ryerson University

Predictive model 2 : Decision Tree

The top predictors are in line with Logistic Regression



Ryerson University

Predictive model 3: Ordinal Logistic Regression

The top predictors are in line with Logistic Regression

	Value	Std. Error	p value
team_performance_support_f2	0.839858010	0.13140423	0.000
team_performance_support_f3	1.776660670	0.13232253	0.000
team_performance_support_f4	2.505495553	0.13275177	0.000
team_performance_support_f5	3.149926054	0.13757789	0.000
team_engagement_f3	0.623653342	0.13763417	0.000
team_engagement_f4	1.215077198	0.13722008	0.000
team_engagement_f5	1.774232920	0.13837937	0.000
values_f4	1.145285128	0.28748704	0.003
values_f5	1.334526559	0.28785668	0.000
agency_engagement_f4	0.382823379	0.12008885	0.002
agency_engagement_f5	0.603785250	0.12725730	0.000
innovation_f4	0.459841664	0.09503297	0.000
innovation_f5	0.538268330	0.10113800	0.000
risk_culture_f4	0.277387240	0.07329326	0.000
risk_culture_f5	0.397013802	0.08725755	0.000
wellbeing_f5	0.367403603	0.18221050	0.006

Ryerson University

Conclusion

what have I learnt and what do I hope for?

We can accurately predict teams' high performance year over year with the following as the most significant predictors:

1. Team performance support (availability of tools, resources, skills, knowledge, work processes and high team standards)
2. Team engagement (team refers to people that the participant works with on daily basis)
3. Agency engagement (agency refers to the immediate organization the participant belongs to)
4. Values (team, supervisor and senior management acting in accordance to organizational values)
5. Innovation (support + recognition for innovation and new ways of working and acceptance that failure is part of innovation)

- Organizations can benefit immediately and achieve higher team performance by focussing on the top 5 predictors.

- Medium and large organizations, with the scale and needed resources, would benefit immensely from building their own high-performance teams predictive models using a combination of survey data and linking it to actual team performance data.