



NSSE Sentiment Analysis: A cluster-based approach of survey respondents

Afshin Karimi
Rohit Murarka
Su Swarat

Office of Assessment and Institutional Effectiveness
CSU Fullerton

2018 CAIR Conference

NSSE Open Ended Question

Students were asked in the 2016 NSSE to respond to the following question in an open text box (444 responded with valid responses):

"If you have any additional comments or feedback that you'd like to share on the quality of your educational experience, please type them below."

Overview

- Automatic Sentiment Analysis on each response (assign polarity/confidence)
- Cluster Analysis (groups of 'similar' NSSE respondents)
 - Look up institutional data on respondents
 - K-Means method to create clusters
 - Label and describe the created clusters
- Text Analysis (create a vector of words, etc.) – set up for visualization

Use Tableau dashboards to explore common sentiments for each cluster

Sentiment Analysis

- Sentiment analysis uses Linguistics, AI, Data Mining concepts to extract opinion and tone from text
- Basic tasks:
 - Determine polarity (positive/negative/neutral)
 - Determine subjectivity/objectivity
- For this project we use a software tool (RapidMiner) to do Sentiment Analysis (will not/cannot re-invent the wheel)

Clustering

- Exploratory method to find hidden patterns or groupings in data (not a predictive method)
- Process of dividing the dataset into groups (clusters) such that:
 - Members in each cluster are as similar to each other as possible
 - Clusters are as dissimilar from one another as possible
- Defining the similarity criteria is key

K-Means Clustering

- Non-hierarchical (partitive) method
- Divide the data in pre-specified number of clusters (K) such that clusters are as homogeneous as possible
- How to determine best K?

Similarity Criteria

- Campus GPA (ranges – 5 values)
- First Generation to Degree (binary)
- Student Type (3 groups: native freshman, native sr., transfer sr.)
- Sex (binary)
- UR Status (binary)

K-Means Clustering Model (RapidMiner)

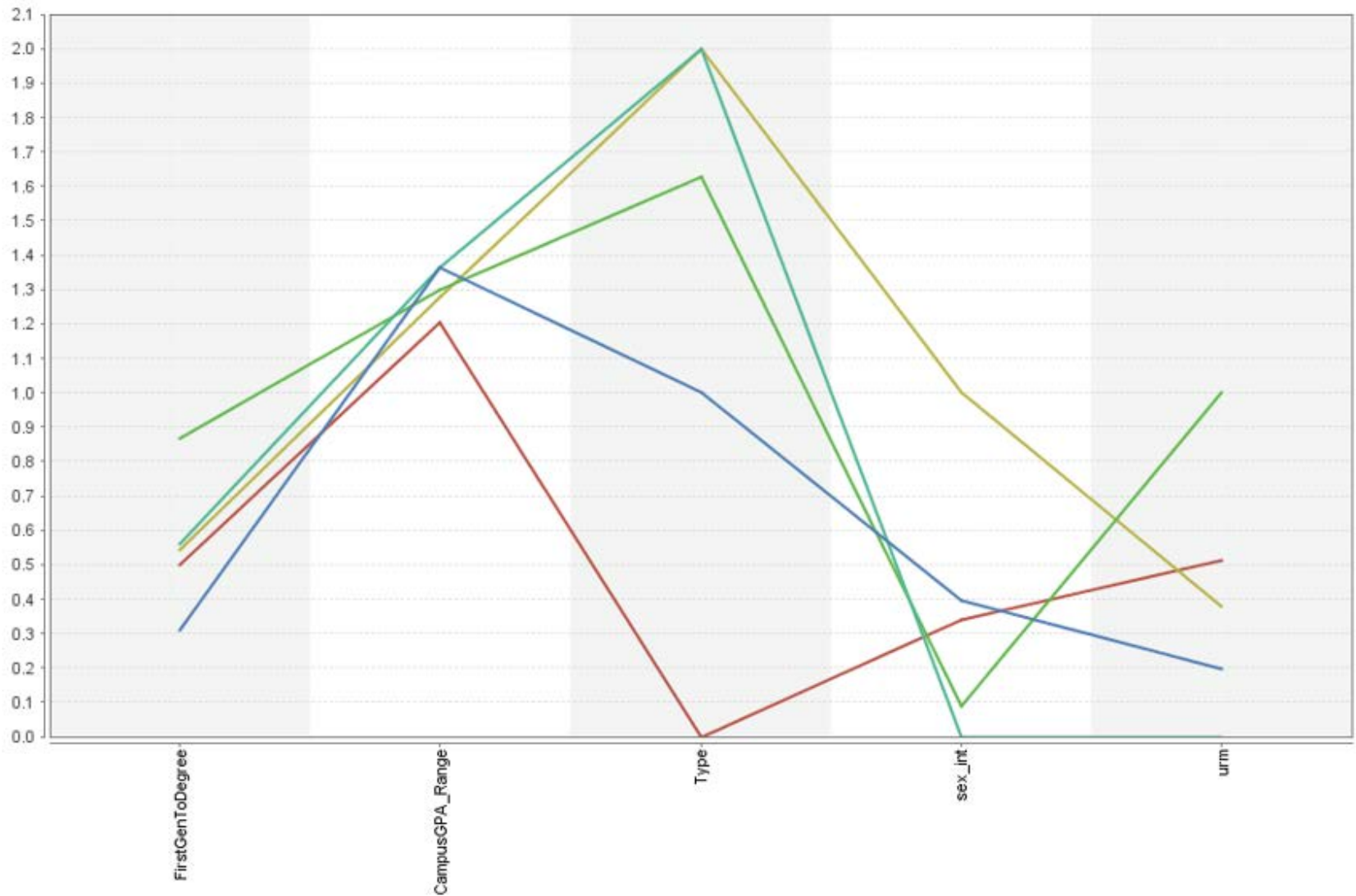
The screenshot displays the RapidMiner interface with a workflow for K-Means Clustering. The workflow consists of the following operators: Retrieve temp5, Select Attributes, Set Role, Clustering, and Cluster Model Visualizer. The Clustering operator is highlighted in orange, indicating it is the active operator. The parameters for the Clustering (k-Means) operator are shown on the right side of the interface:

- add cluster attribute
- add as label
- remove unlabeled
- k: 5
- max runs: 10
- determine good start values
- measure L: Numeric...
- numerical me: Euclidean

Below the parameters, there is a link to [Hide advanced parameters](#) and a warning icon with the text [Change compatibility \(9.0.001\)](#). At the bottom of the parameters panel, there is a **Help** section for **k-Means** with a sub-section for **Concurrency**.

Clusters Centroid Plot

Cluster 0 (81) Cluster 1 (100) Cluster 2 (91) Cluster 3 (98) Cluster 4 (74)



Clusters' Labels

Cluster 0: Native seniors with top campus GPA

Cluster 1: White & Asian female transfer seniors

Cluster 2: Hispanic female, first generation-to-degree seniors

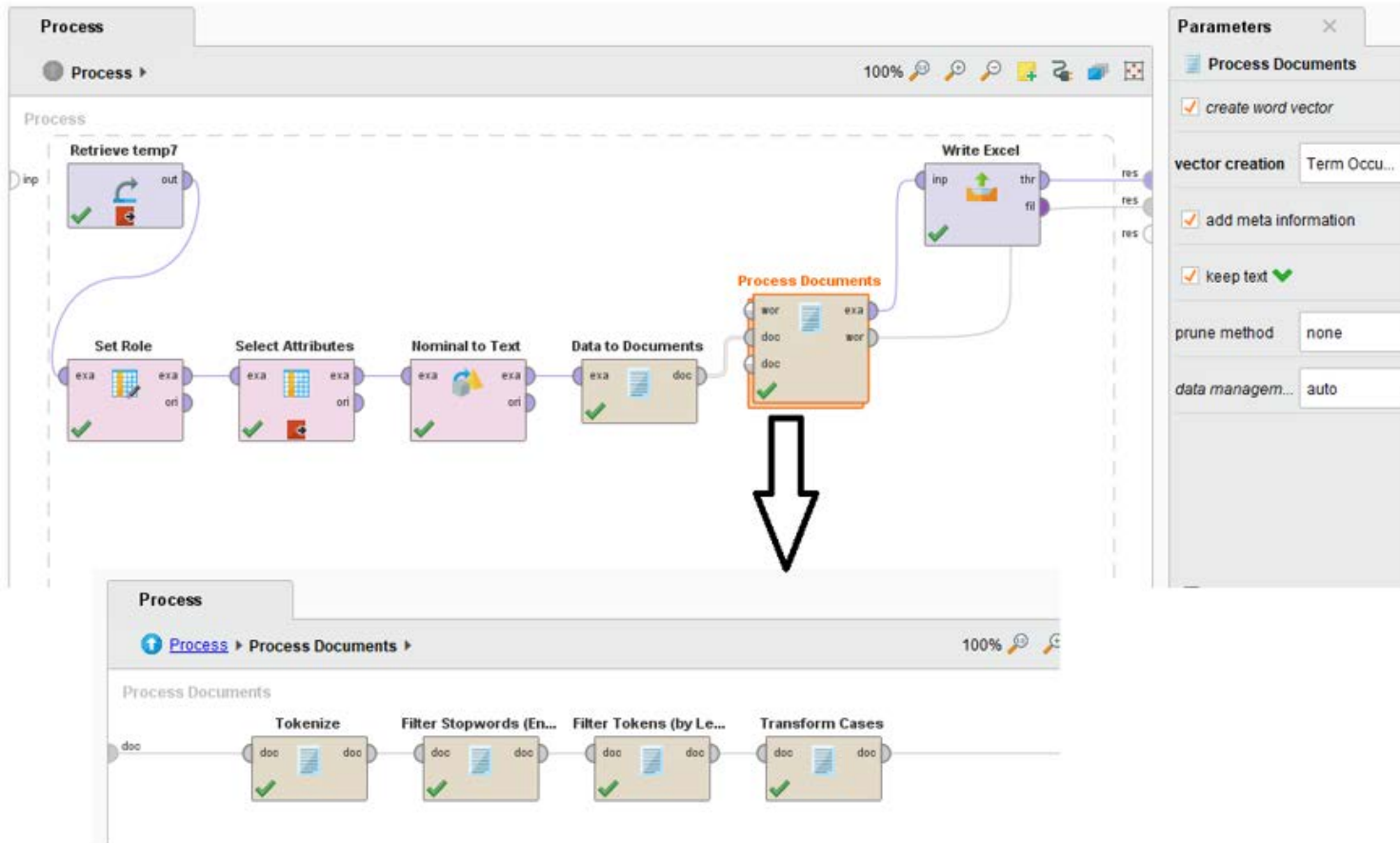
Cluster 3: Male transfer students

Cluster 4: Freshmen

Text Analysis

- Create vector of words for each response (in order to explore sentiments in Tableau in both response and word levels)
- Process includes:
 - Tokenizing
 - Filtering out English stop words
 - Filtering tokens by length
 - Convert all words to either lower case or upper case

Text Processing Steps (RapidMiner)



Text Analysis Output

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Student ID	Response	abhorrent	abilities	ability	able	abominab	abroad	absent	absolute	absolutely	abundanc	academia	academic a
2	#####	supported faculty education department feel alway make appointment email respond time	.0	.0	.0	1.0	.0	.0	.0	.0	.0	.0	.0	.0
3	#####	given chance gone csuf chosen colleges chose csuf family money reason make choices	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0
4	#####	chicana studies faculty department helped student flourish major helped critically think br	.0	.0	.0	1.0	.0	.0	.0	.0	.0	.0	.0	.0
5	#####	support help transfer students transition university	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0
6	#####	fullerton helped achieve educational goals professors awesome involved making student r	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0
7	#####	faculty energetic assist students help remember students money huge problem statistics s	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0
8	#####	state department deplorable chairs easels fall apart covered decades dirt dust dark concret	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0
9	#####	think student know rights notified rules advisor keep working academic year make student	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	1.0
10	#####	college student leisure time entertain example hangout friends watch movies make procr	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0
11	#####	tuition fees international students reduced help reduce financial burden international stuc	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0
12	#####	instructors strikes make scared please support instructors students reduced tuition textbo	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0
13	#####	thankful attending csuf	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0
14	#####	think focus undocumented students provide resources participate take advantage resource	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0
15	#####	suggest studying space facility midterm final weeks difficult find good spot studying parkin	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0
16	#####	advisors international students bigger office knowledgeable advisors definitely process pa	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0
17	#####	answers fairly accurate birth year survey recommend dividing parts reduce number total qu	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0
18	#####	reason choose school commute school closer live	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0
19	#####	bring octa pass free huge help students money pass difficult ride campus morning practice	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0
20	#####	teach explain	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0
21	#####	parking residence halls greatly inhibited experience state fullerton expensive difficult find	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0
22	#####	liked core business classes mihaylo professors knowledgeable helpful picking concentratic	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0
23	#####	thing striked transferring csuf cypress community college quality professors professors tak	.0	.0	1.0	.0	.0	.0	.0	.0	.0	.0	.0	2.0
24	#####	thank csuf	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0
25	#####	senior ready graduate months useful help interaction student community part time studen	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0
26	#####	unversity great place beautiful teachers accents make difficult understand books expensiv	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0
27	#####	practical learning example group proiets help students imorove teamwork find curriculum	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0

Data Format Suitable for Visualization

Student ID	Cluster #	Level	Sex	Ethnicity	Major	College	CampusGPA	Response	Word	polarity	polarity con
8.01E+08	cluster_1	SR	WOMEN	Multi	MARKT	CBE	2.36	There should be more support to help transfer students transition to university.	support	positive	0.50
8.01E+08	cluster_1	SR	WOMEN	Multi	MARKT	CBE	2.36	There should be more support to help transfer students transition to university.	help	positive	0.50
8.01E+08	cluster_1	SR	WOMEN	Multi	MARKT	CBE	2.36	There should be more support to help transfer students transition to university.	transfer	positive	0.50
8.01E+08	cluster_1	SR	WOMEN	Multi	MARKT	CBE	2.36	There should be more support to help transfer students transition to university.	students	positive	0.50
8.01E+08	cluster_1	SR	WOMEN	Multi	MARKT	CBE	2.36	There should be more support to help transfer students transition to university.	transition	positive	0.50
8.01E+08	cluster_1	SR	WOMEN	Multi	MARKT	CBE	2.36	There should be more support to help transfer students transition to university.	university	positive	0.50
8.01E+08	cluster_2	SR	WOMEN	Hispanic	SOC	H&SS	2.35	CSU Fullerton has helped me achieve my educational goals. The professors are awes	fullerton	positive	0.98
8.01E+08	cluster_2	SR	WOMEN	Hispanic	SOC	H&SS	2.35	CSU Fullerton has helped me achieve my educational goals. The professors are awes	helped	positive	0.98
8.01E+08	cluster_2	SR	WOMEN	Hispanic	SOC	H&SS	2.35	CSU Fullerton has helped me achieve my educational goals. The professors are awes	achieve	positive	0.98
8.01E+08	cluster_2	SR	WOMEN	Hispanic	SOC	H&SS	2.35	CSU Fullerton has helped me achieve my educational goals. The professors are awes	educational	positive	0.98
8.01E+08	cluster_2	SR	WOMEN	Hispanic	SOC	H&SS	2.35	CSU Fullerton has helped me achieve my educational goals. The professors are awes	goals	positive	0.98
8.01E+08	cluster_2	SR	WOMEN	Hispanic	SOC	H&SS	2.35	CSU Fullerton has helped me achieve my educational goals. The professors are awes	professors	positive	0.98
8.01E+08	cluster_2	SR	WOMEN	Hispanic	SOC	H&SS	2.35	CSU Fullerton has helped me achieve my educational goals. The professors are awes	awesome	positive	0.98
8.01E+08	cluster_2	SR	WOMEN	Hispanic	SOC	H&SS	2.35	CSU Fullerton has helped me achieve my educational goals. The professors are awes	involved	positive	0.98
8.01E+08	cluster_2	SR	WOMEN	Hispanic	SOC	H&SS	2.35	CSU Fullerton has helped me achieve my educational goals. The professors are awes	making	positive	0.98
8.01E+08	cluster_2	SR	WOMEN	Hispanic	SOC	H&SS	2.35	CSU Fullerton has helped me achieve my educational goals. The professors are awes	student	positive	0.98
8.01E+08	cluster_2	SR	WOMEN	Hispanic	SOC	H&SS	2.35	CSU Fullerton has helped me achieve my educational goals. The professors are awes	reaches	positive	0.98
8.01E+08	cluster_2	SR	WOMEN	Hispanic	SOC	H&SS	2.35	CSU Fullerton has helped me achieve my educational goals. The professors are awes	full	positive	0.98
8.01E+08	cluster_2	SR	WOMEN	Hispanic	SOC	H&SS	2.35	CSU Fullerton has helped me achieve my educational goals. The professors are awes	potential	positive	0.98
8.01E+08	cluster_2	SR	WOMEN	Hispanic	SOC	H&SS	2.35	CSU Fullerton has helped me achieve my educational goals. The professors are awes	change	positive	0.98
8.01E+08	cluster_2	SR	WOMEN	Hispanic	SOC	H&SS	2.35	CSU Fullerton has helped me achieve my educational goals. The professors are awes	education	positive	0.98
8.01E+08	cluster_2	SR	WOMEN	Hispanic	SOC	H&SS	2.35	CSU Fullerton has helped me achieve my educational goals. The professors are awes	facility	positive	0.98
8.01E+08	cluster_2	SR	WOMEN	Hispanic	SOC	H&SS	2.35	CSU Fullerton has helped me achieve my educational goals. The professors are awes	return	positive	0.98
8.01E+08	cluster_2	SR	WOMEN	Hispanic	SOC	H&SS	2.35	CSU Fullerton has helped me achieve my educational goals. The professors are awes	heart	positive	0.98
8.01E+08	cluster_2	SR	WOMEN	Hispanic	SOC	H&SS	2.35	CSU Fullerton has helped me achieve my educational goals. The professors are awes	beat	positive	0.98

