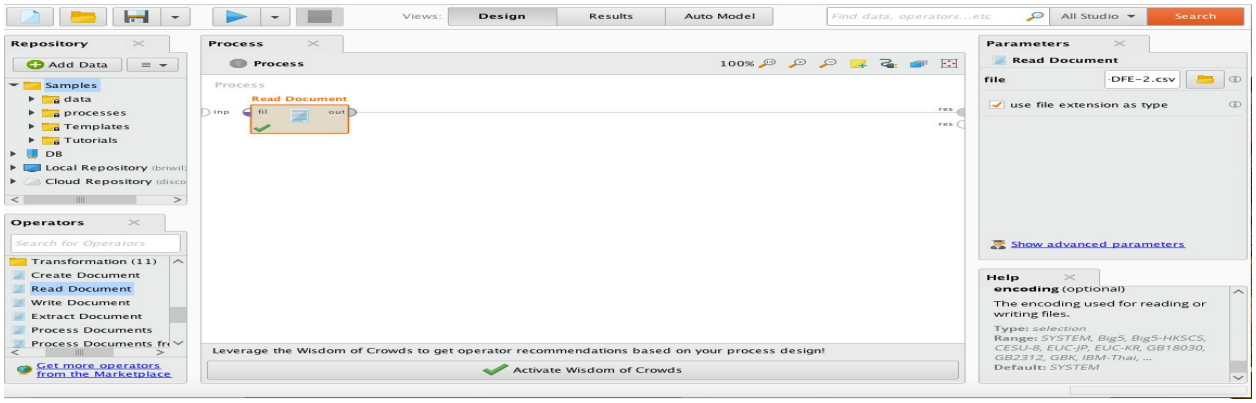
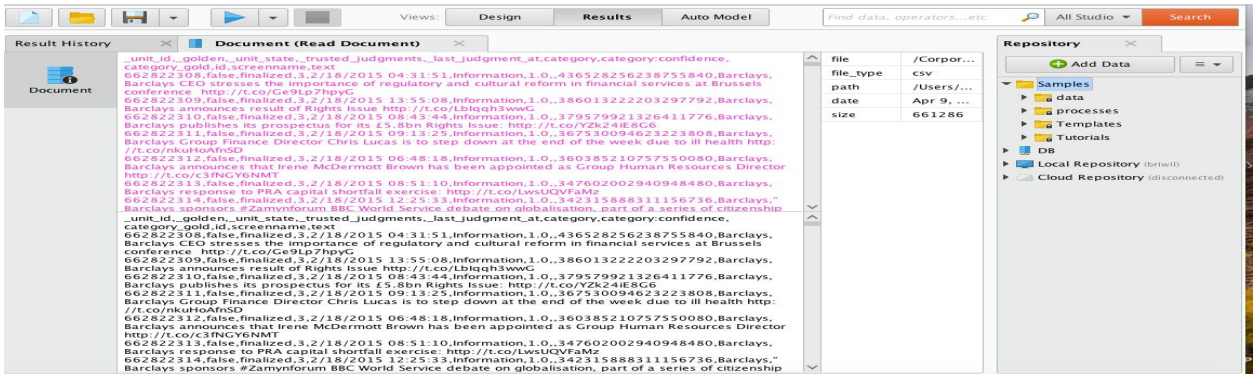
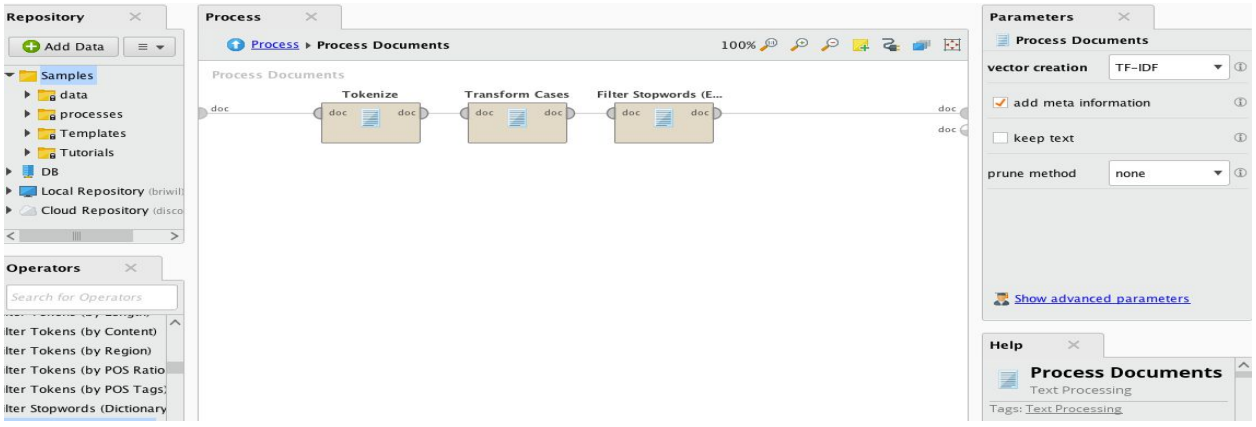
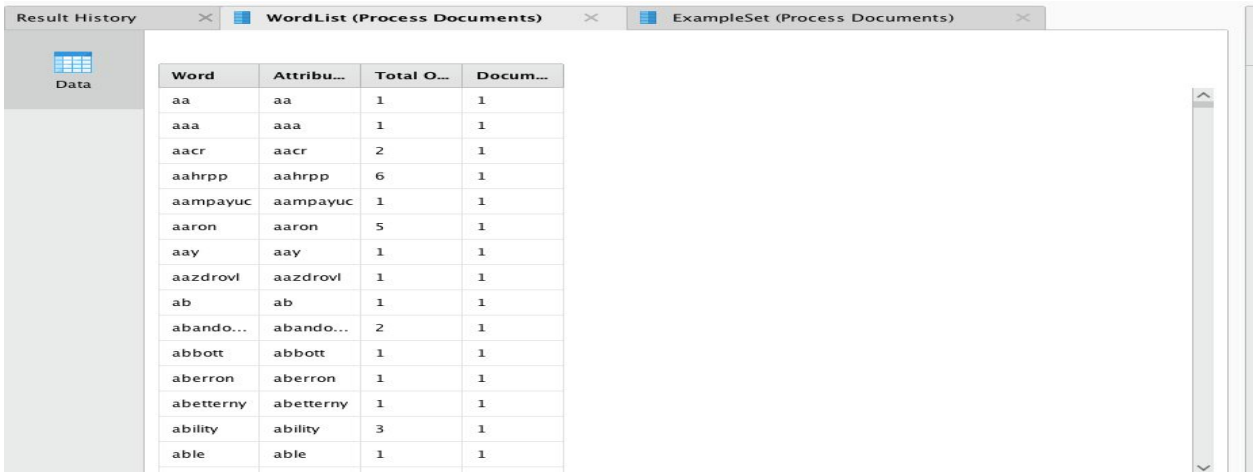
After connecting the Read Document out node to the res node, I then selected the Corporate-messaging-DFE-2.csv file. From here I connected the Read Document operator to the output to import the text, and return to the design window.





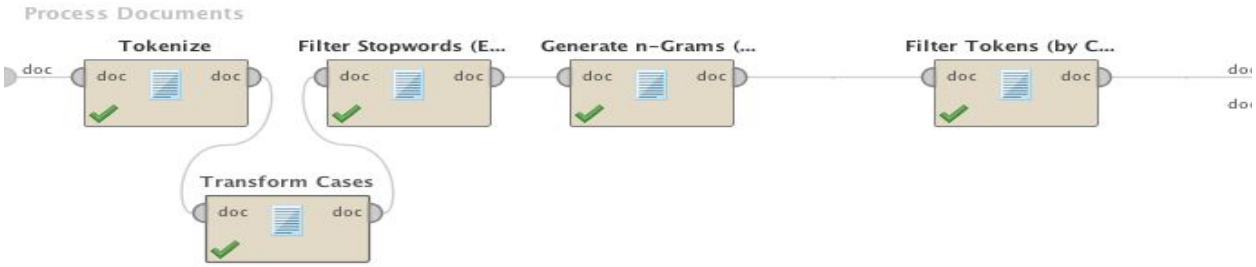
With the Read Document, the Process Documents, Transform Cases, Filter Stopwords(English), and the Tokenize operator connected to the doc node I ran the process, which displays Word, Attribute Name, Total Occurrences & Document Occurrences as they appear in the corporate-messaging-DFE-2 document.





Word	Attribu...	Total O...	Docum...
aa	aa	1	1
aaa	aaa	1	1
aacr	aacr	2	1
aahrpp	aahrpp	6	1
aampayuc	aampayuc	1	1
aaron	aaron	5	1
aay	aay	1	1
aazdrovl	aazdrovl	1	1
ab	ab	1	1
abando...	abando...	2	1
abbott	abbott	1	1
aberron	aberron	1	1
abetterny	abetterny	1	1
ability	ability	3	1
able	able	1	1

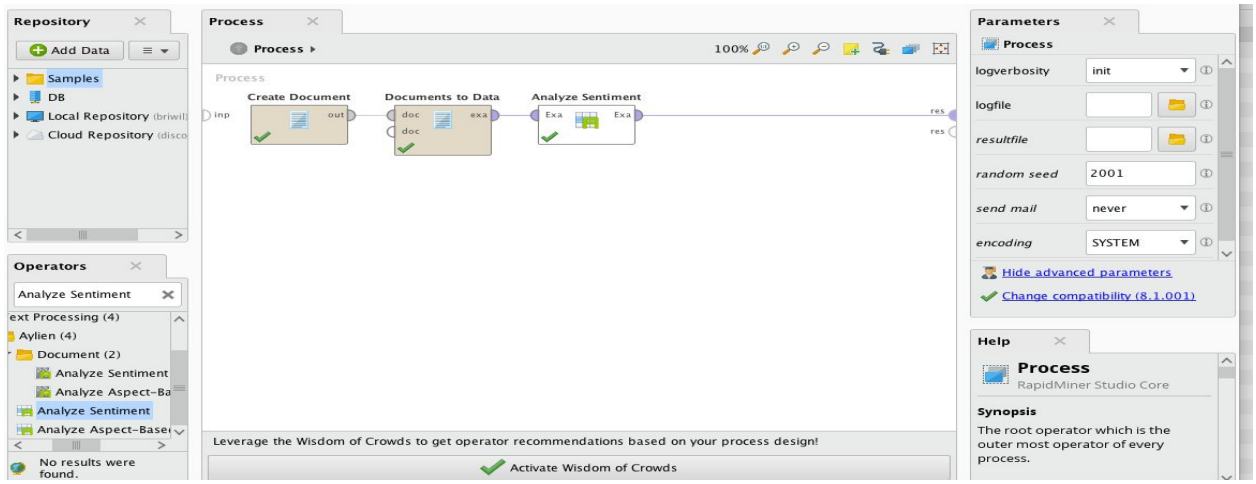
I added the Generate n-Grams (Terms) operator to the sub process within the Process Documents operator. I added Filter Tokens (by Content) and set the parameters to conditions = contains and string = _ (underscore). We can see the name Aaron Gurwitz appears five times in one document.

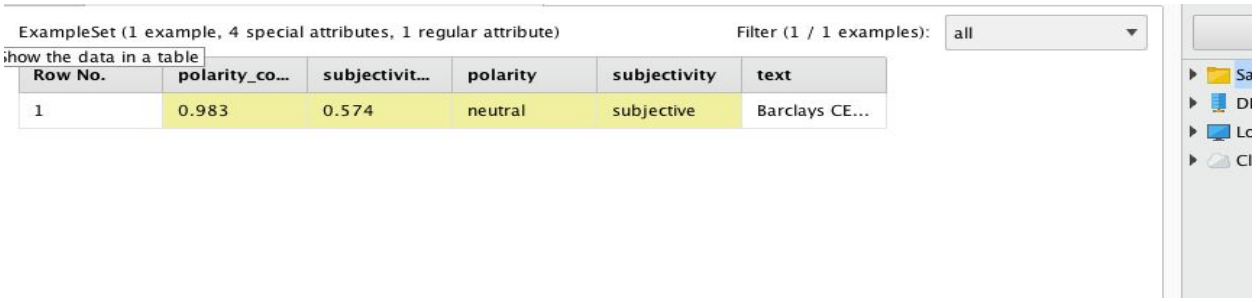


Word	Attribu...	Total O...	Docum...
aa_false	aa_false	1	1
aaa_pro...	aaa_pro...	1	1
aacr_look	aacr_look	2	1
aahrpp...	aahrpp...	4	1
aahrpp...	aahrpp...	2	1
aampay...	aampay...	1	1
aaron_g...	aaron_g...	5	1
aay_o	aay_o	1	1
aazdrov...	aazdrov...	1	1
ab_false	ab_false	1	1

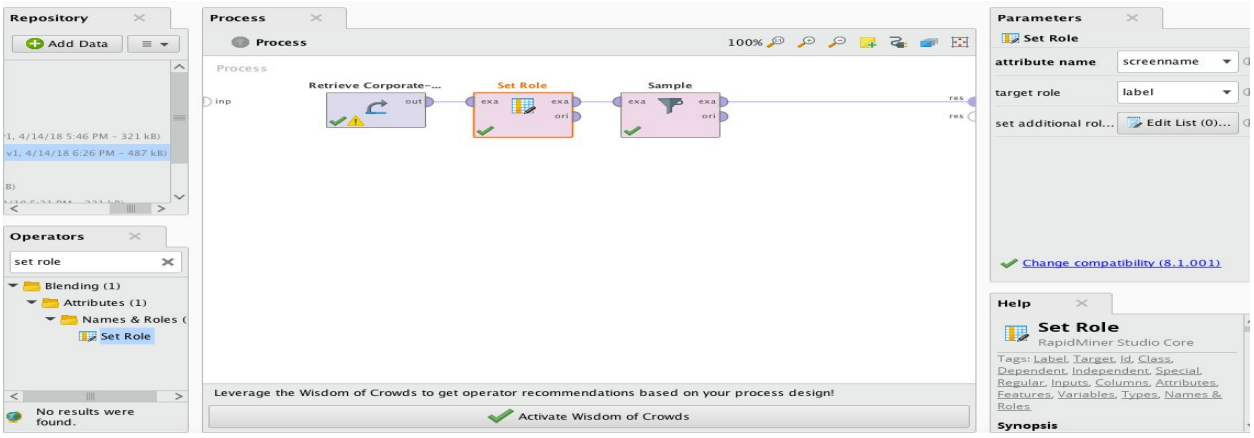
aaron_gurwitz

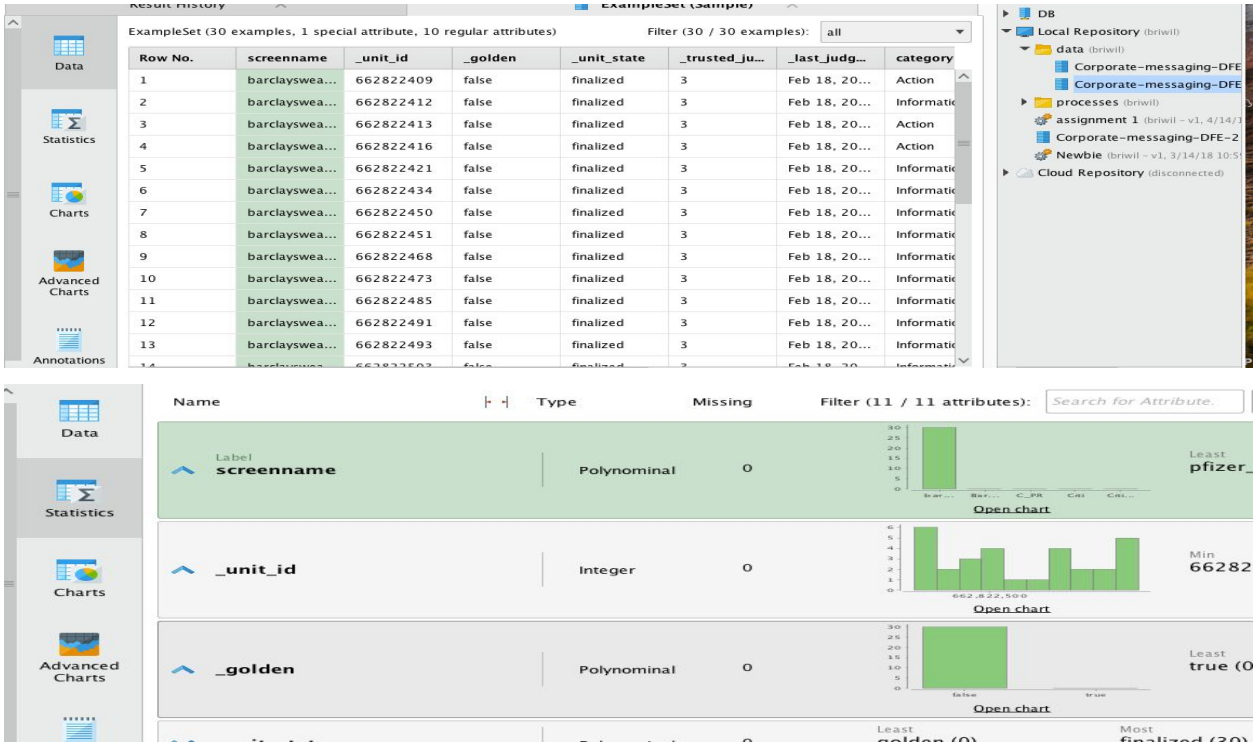
I created a new process and then created a new document and entered the tweet text file into the Parameter window. I selected the Documents to Data operator, and set text attribute to = text. I the setup Aylien by going to Connections -> Manage Connections. Add a connection for Aylien. Then I added the Analyze Sentiment operator.



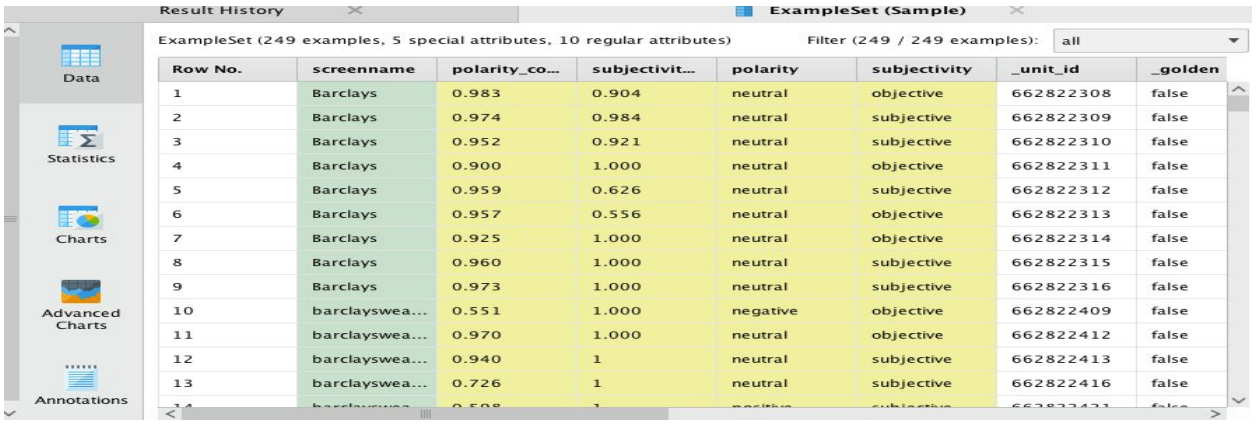


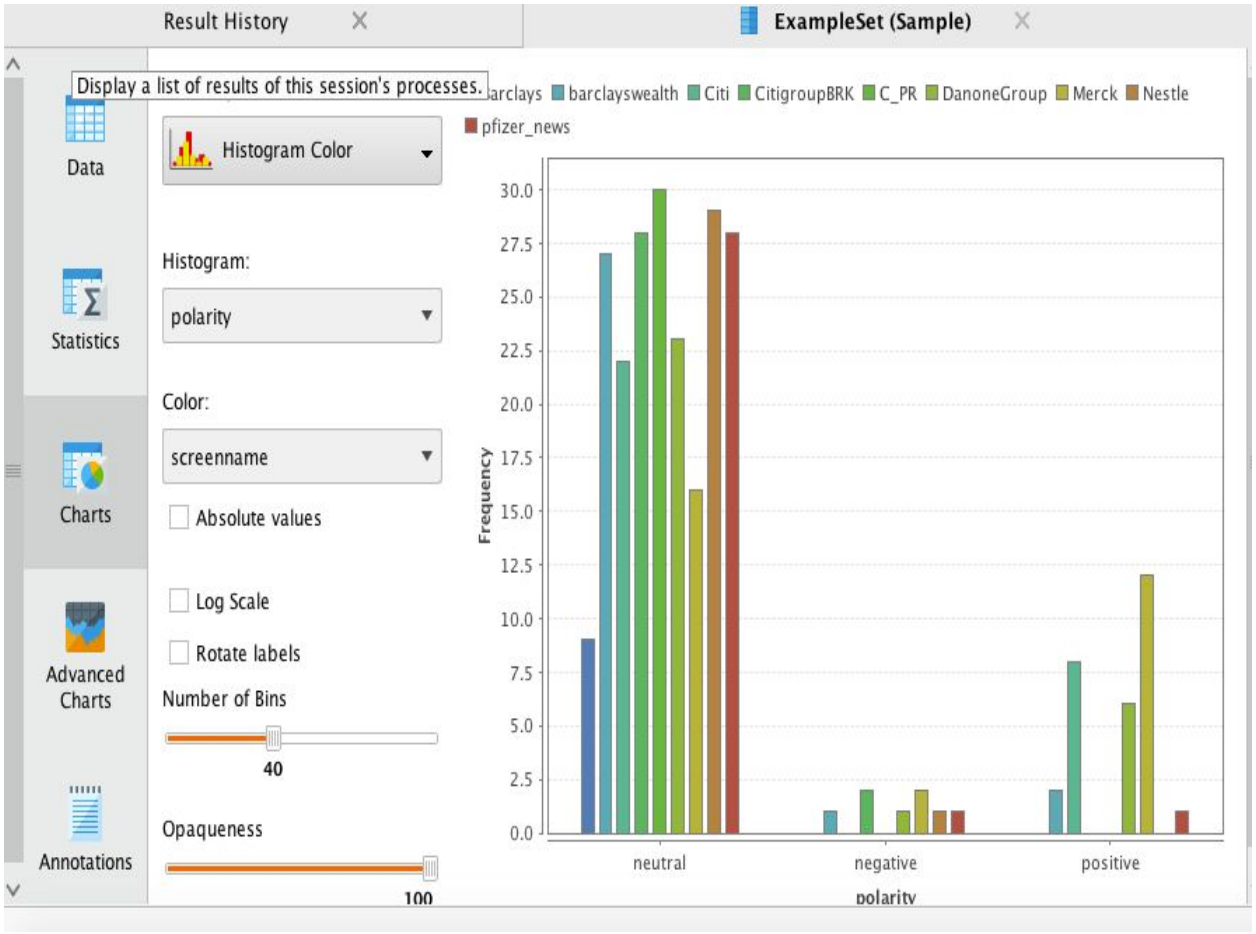
I then imported the corporate-messaging-DFE-2 file and connected it to a Set Role operator, and a Sample operator.





After connecting to the Analyze Sentiment operator I ran the process. I then made a Histogram Color chart with polarity. You can see that Merck had the most positive posts, though nearly ties with Citi in negative posts.




Result History
ExampleSet (Sample)
Display a list of results of this session's processes.
Data
Statistics
Charts
Advanced Charts
Annotations
Histogram Color
Histogram:
polarity
Color:
screenname
Absolute values
Log Scale
Rotate labels
Number of Bins
40
Opaqueness
100
arclays
barclayswealth
Citi
CitigroupBRK
C_PR
DanoneGroup
Merck
Nestle
pfizer_news
Frequency
30.0
27.5
25.0
22.5
20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5
0.0
neutral
negative
positive
polarity