ISGB 7978 Web Analytics

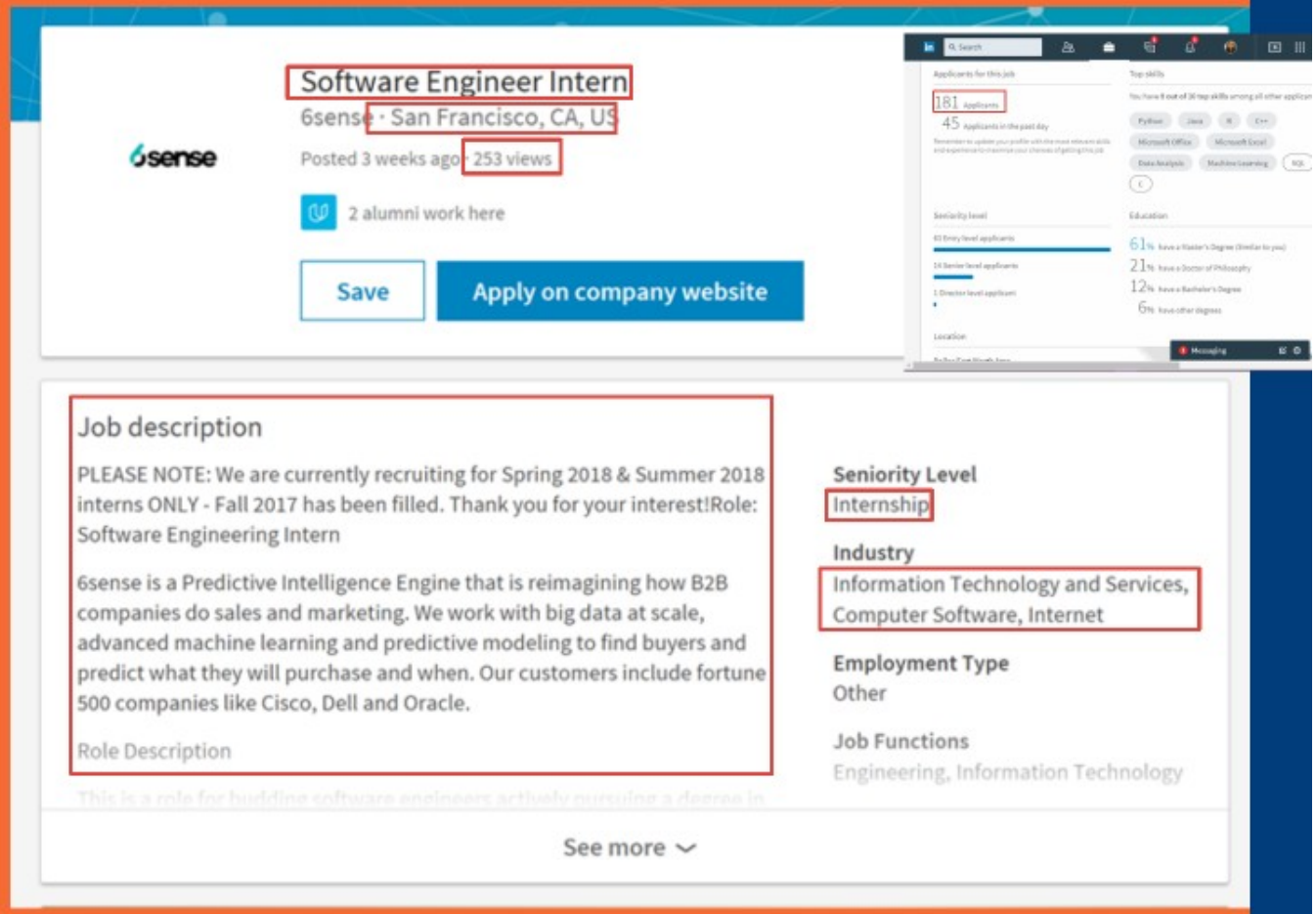
Group 7
Choudhury, Sumi
Gu, Jiahua
Huang, Feifei
Ji, Minxia

Mission

- Helping company to identify the import parts in Job Description to attract more candidates
- Prediction Model based on Skills & Qualification

Design

- Job Title
- Job Location
- # of Views
- # of Applicants
- Seniority Level
- Industry
- Job Description
- Skills & Qualification



Search

Applicants for this job

181 Applicants

45 Applicants in the past day

Remember to update your profile with the most relevant skills and experience to maximize your chances of getting this job

Top skills

You have 8 out of 10 top skills among all other applicants

Python | Java | R | C++ | Microsoft Office | Microsoft Excel | Data Analysis | Machine Learning | SQL | C

Seniority level

63 Entry level applicants

14 Senior level applicants

1 Director level applicant

Education

61% have a Master's Degree (Similar to you)

21% have a Doctor of Philosophy

12% have a Bachelor's Degree

6% have other degrees

Location

Messaging

Elements Console Sources Network Performance

Competitive intelligence about other applicants

Cancel

Styles Event Listeners DOM Breakpoints Properties

Console What's New

Highlights from the Chrome 60 update

New Audits panel, powered by Lighthouse

Find out whether your site qualifies as a Progressive Web App, measure the accessibility and performance of a page, and discover best practices.

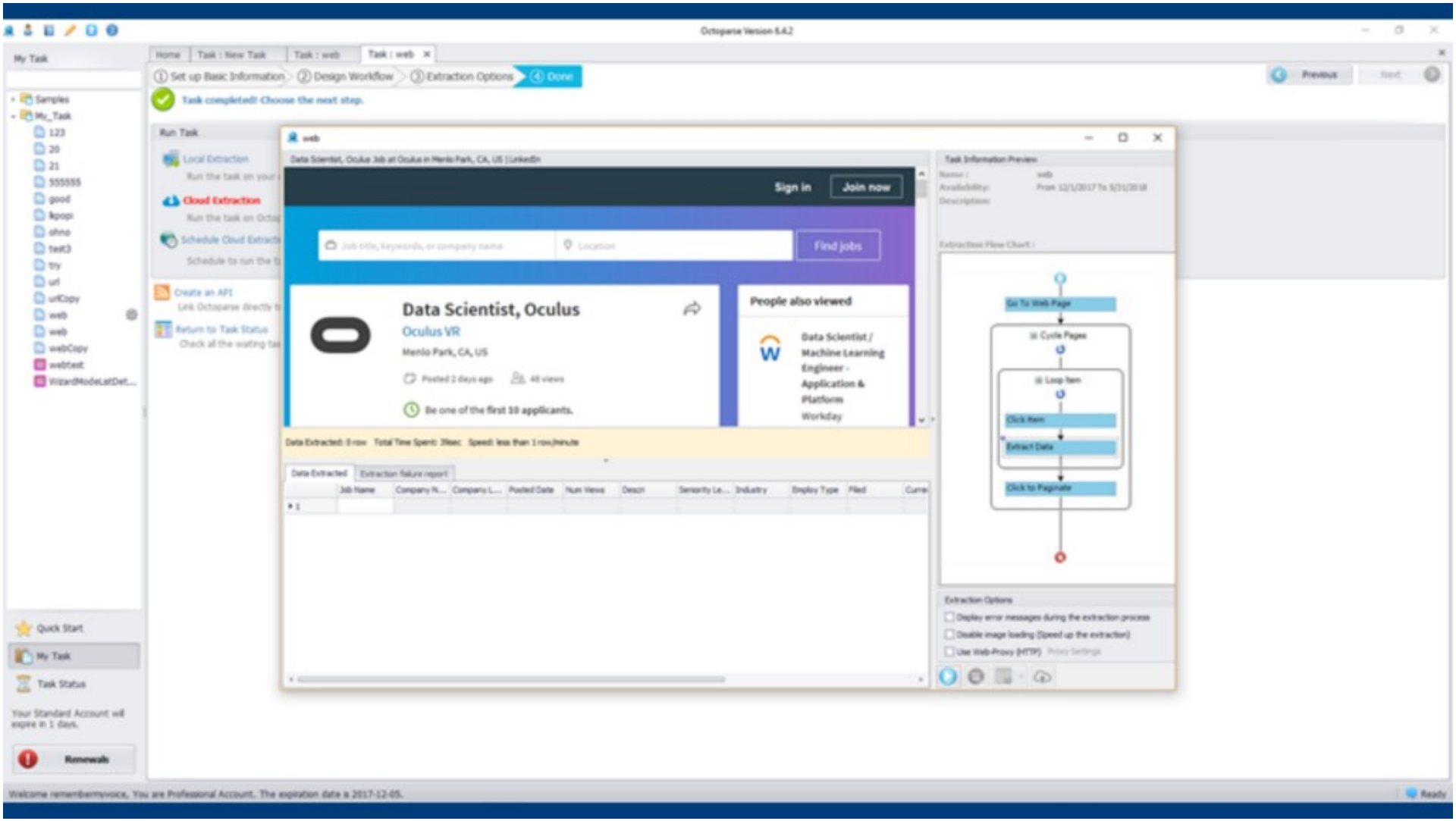
Third-party badges

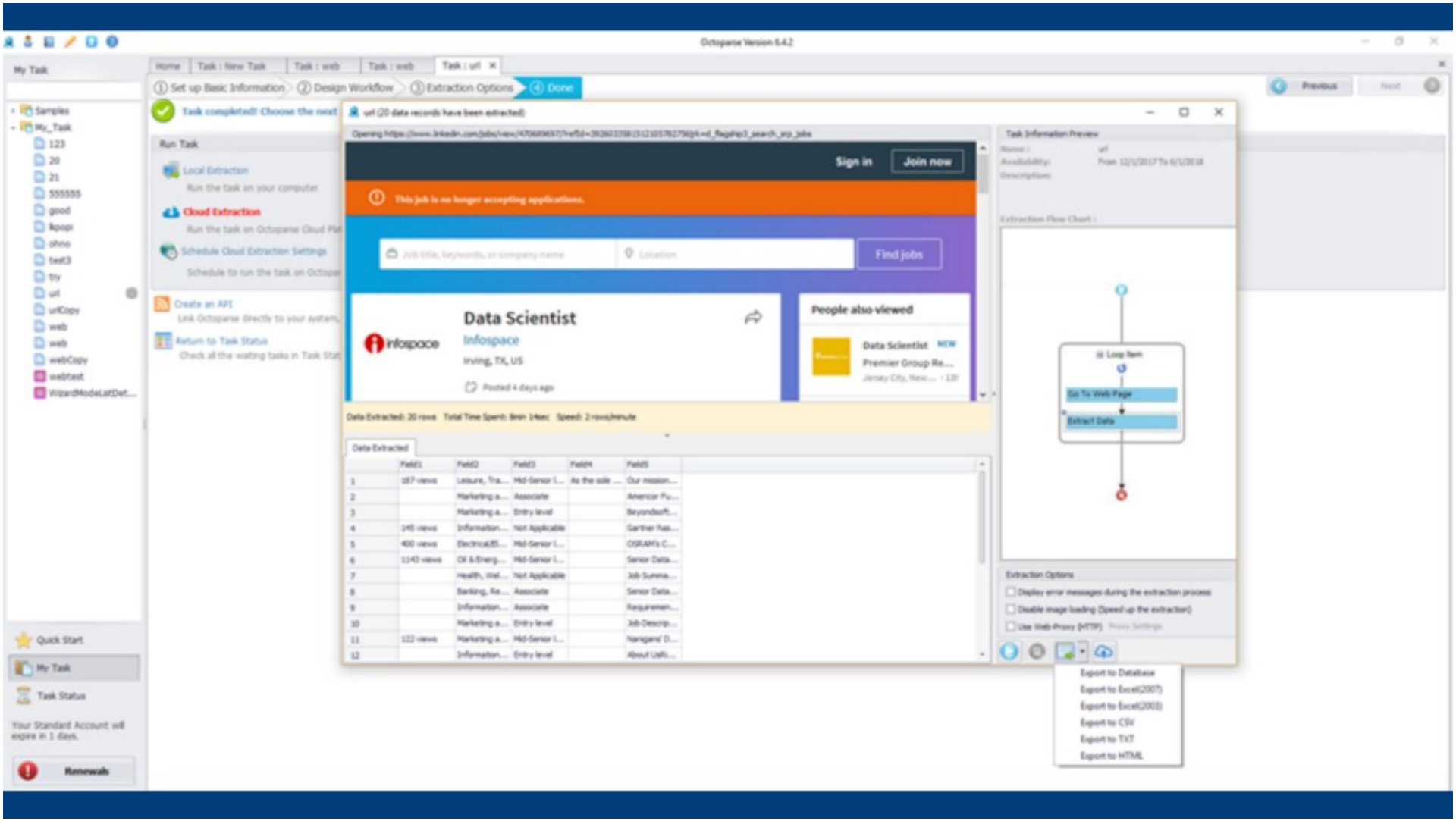
See what third-party entities are logging to the Console, making network requests, and causing work during performance

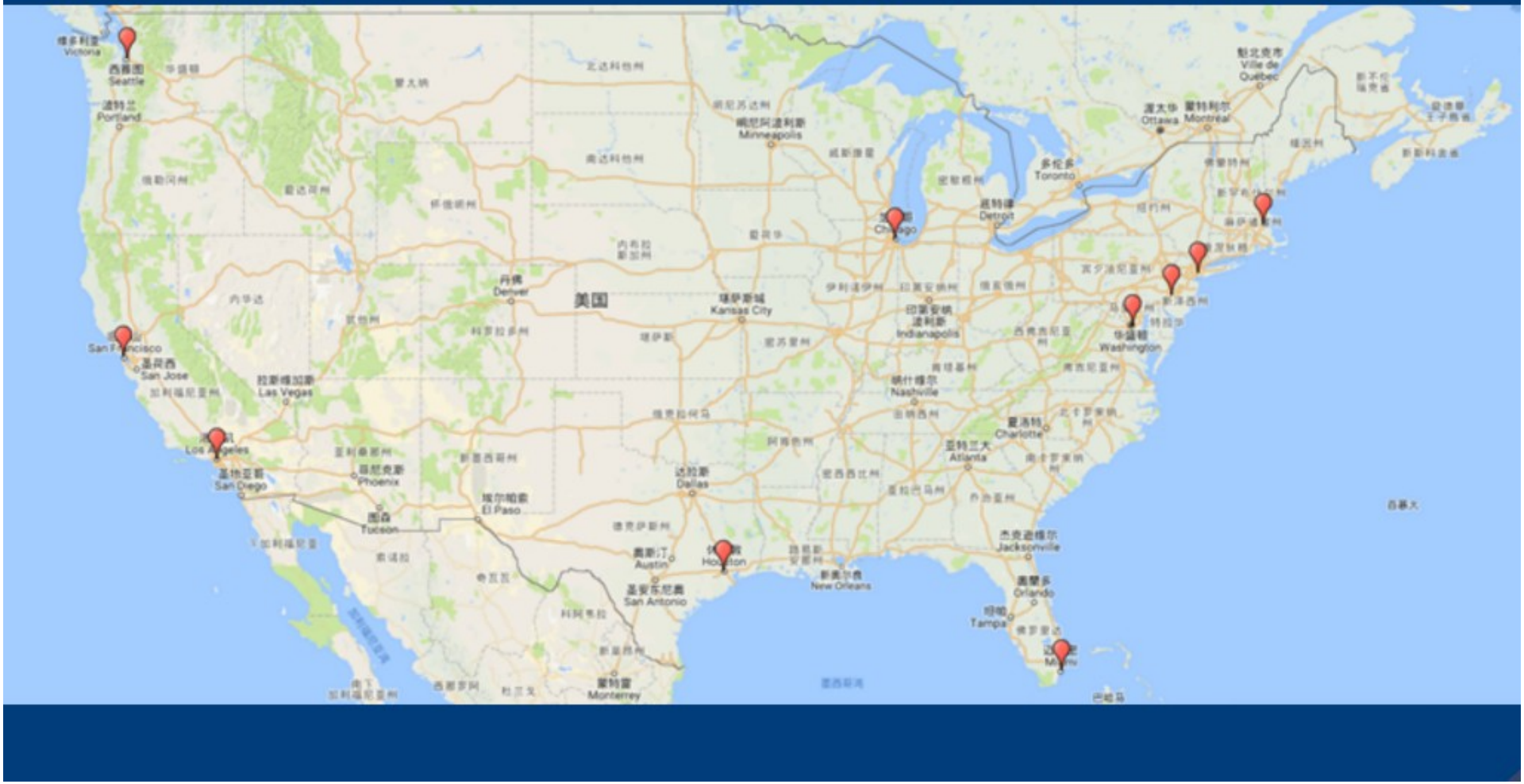
ISGB 7978 Web Analytics

Group 7
Choudhury, Sumi
Gu, Jiahua
Huang, Feifei
Ji, Minxia


Data Source & Collection
Octoparse
Happy Data Hunting
www.octoparse.com


Octoparse Version 6.4.2
My Task
Home
Task : New Task
Task : web
Task : web
Set up Basic Information
Design Workflow
Extraction Options
Done
Previous
Next
Task completed! Choose the next step.
Run Task
Local Extraction
Cloud Extraction
Create an API
Return to Task Status
Samples
My_Task
web
Sign in
Join now
Find jobs
Data Scientist, Oculus
Oculus VR
Menlo Park, CA, US
Be one of the first 10 applicants.
People also viewed
Data Scientist / Machine Learning Engineer - Application & Platform Workday
Data Extracted
Extraction failure report
Go To Web Page
Cycle Pages
Loop Item
Click Item
Extract Data
Click to Paginate
Extraction Options
Quick Start
My Task
Task Status
Renewals
Ready


Octoparse Version 6.4.2
My Task
Home
Task : New Task
Task : web
Task : web
Task : url
① Set up Basic Information
② Design Workflow
③ Extraction Options
④ Done
Previous
Next
Task completed! Choose the next
Run Task
Local Extraction
Run the task on your computer
Cloud Extraction
Schedule Cloud Extraction Settings
Create an API
Return to Task Status
Samples
My_Task
123
20
21
555555
good
kpopi
ohno
test3
try
url
urlCopy
web
web
webCopy
webtest
Quick Start
My Task
Task Status
Your Standard Account will expire in 1 days.
Renewals
url (20 data records have been extracted)
Sign in
Join now
This job is no longer accepting applications.
Find jobs
Data Scientist
Infospace
Irving, TX, US
Posted 4 days ago
People also viewed
Data Scientist NEW
Premier Group Re...
Data Extracted: 20 rows Total Time Spent: 8min 14sec Speed: 2 rows/minute
Data Extracted
Task Information Preview
Extraction Flow Chart :
Loop Item
Go To Web Page
Extract Data
Extraction Options
Display error messages during the extraction process
Disable image loading (Speed up the extraction)
Use Web Proxy (HTTP)
Export to Database
Export to Excel(2007)
Export to Excel(2003)
Export to CSV
Export to TXT
Export to HTML

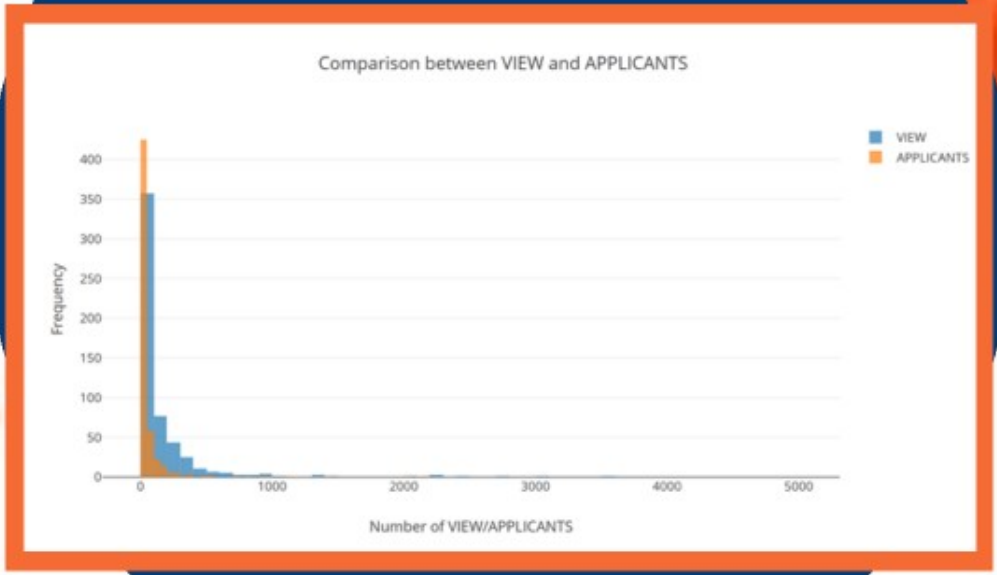

Victoria
Seattle
Portland
San Francisco
San Jose
Las Vegas
Los Angeles
San Diego
Phoenix
Tucson
El Paso
Denver
美国
Minneapolis
Kansas City
Chicago
Detroit
Toronto
Ottawa
Montreal
Ville de Québec
Indianapolis
Washington
Nashville
Charlotte
Atlanta
Dallas
Austin
Houston
San Antonio
New Orleans
Jacksonville
Orlando
Tampa
Miami
Monterrey

	A	B	C	D	E	F
767	129	29	Not Applicable	Information Technology and Services	Lead analytics consultin	Consultant / Principal Consultant / Consulting Partner – Banking-Insurance / Retail / TelecomPrimary res
768	1	362	Director	Management Consulting	Demonstrate technical b	As a Data Scientist, you will be: Providing essential insight and leverage on business problems through a
769	383	174	Mid-Senior level	Investment Banking	Experience in building m	Machine Learning Software Engineer will be a member of a growing team with responsibilities for design
770	58	73	Not Applicable	Financial Services	2+ years working as a da	The Moody's Analytics Content team has a great opportunity for a talented, self-motivated and responsibl
771	618	525	Executive	Computer Software	PhD or MS degree in Con	We are working with a multimedia content creator that is looking for a data scientist to join their researc
772	242	114	Not Applicable	Computer Software	If you have not written a	Work with internal client facing teams on broker/carrier analytics and reporting Perform data analysis b
773	286	58	Mid-Senior level	Financial Services Information Services Information Tec	Master's degree or highe	Senior Data Scientist - Machine Learning, Modeling, Ideation & Research Equifax empowers businesses
774	107	42	Entry level	Computer Software Information Technology and Services	Applying both technical	Opentext - The Information CompanyAs the Information Company, our mission at OpenText is to create so
775	319	78	Not Applicable	Construction Machinery	MS degree in Industrial E	Caterpillar Inc. is more than big, heavy equipment and much more than yellow iron. Our products have ev
776	305	91	Entry level	Information Technology and Services	Advanced degree (PhD pr	OpportunitySessionM is looking to add a talented Data Scientist to support the continued growth of our D
777	26	13	Entry level	Marketing and Advertising Information Technology and Serv	Mastery of either Python	We are looking to hire an entrepreneurial Data Scientist to join our team in Boston (Quincy). This person
778	63	14	Not Applicable	Consumer Goods Food & Beverages Retail	Basic QualificationsEduc	This job contributes to Starbucks success by guiding business decisions through utilizing data analysis a
779	131	39	Associate	Marketing and Advertising Information Technology and Services	You are awesome.You lov	Making projects happen: Coordinating and managing survey research projects on time and on budgets, w
780	200	73	Mid-Senior level	Computer Software Internet Information Technology and Service	PhD in CS, Stats, Mathem	TalentReach is a boutique "upstart" search service obsessed with finding the most talented individuals in
781	4	1	Associate	Real Estate	You enjoy exploring and	As a Principal Data Science Engineer, your job is to identify and solve challenging problems and productiz
782	18	2	Director	Information Technology and Services Computer Software Mecha	Minimum 5-8 years of ex	The Go To Market Director will be responsible for developing both short and long-term strategies for how
783	197	33	Not Applicable	Information Technology and Services Computer Software Compu	Bachelor's degree or equ	Work closely with engineers throughout the development cycle to reproduce the small and large customer
784	82	33	Entry level	Medical Devices Medical Practice Hospital & Health Care	Bachelors, MS or PhD in	As a Data Scientist you will be working on consulting side of our business. You will be responsible for ana
785	171	58	Entry level	Biotechnology Hospital & Health Care Pharmaceuticals	Level of education and ex	Be responsible for setting standards for core technical platforms and continual optimization and applica
786	3	0	Mid-Senior level	Internet	3 to 5 years experience b	We are looking for an experienced Sr. Data Scientist whos passionate about machine learning, model buil
787	70	7	Entry level	Electrical/Electronic Manufacturing Construction Automotive	Bachelor's degree in Com	We are looking for an experienced Data Engineer with an uncanny ability to integrate multiple heterogene
788	63	13	Entry level	Marketing and Advertising Internet Retail	2+ years of industry expe	Build complex statistical models that predict behavior and improve over time with additional data and le
789	163	34	Mid-Senior level	Computer Software	At Least 3 Years Of Exper	Passion for innovation & energetic company.Strong commitment & investment in our people's developmer
790	136	45	Entry level	Information Technology and Services Computer Software Compu	Experience in creating ar	Working in a very dynamic and exciting environment with top engineering and product talent the data sci
791	289	38	Not Applicable	Computer Software Information Technology and Services Interne	Master's degree in Math,	ABP is seeking an outstanding Data Scientist/Statistician who will help to build the next generation of crec
792	8	1	Associate	Information Technology and Services Staffing and Recruiting Fir	PhD in Machine Learning	We are looking for an Applied Scientist to join the Devices Demand Planning team for the entire device far
793	351	80	Director	Internet	PhD in CS or related field	Provide technical leadership, identify key business challenges and opportunities, and develop end-to-end
794	3	0	Associate	Computer Software Information Technology and Services	Building statistical mode	Work as a team to define and execute data science solutions that address client use cases and business r

ISGB 7978 Web Analytics

Group 7
Choudhury, Sumi
Gu, Jiahua
Huang, Feifei
Ji, Minxia

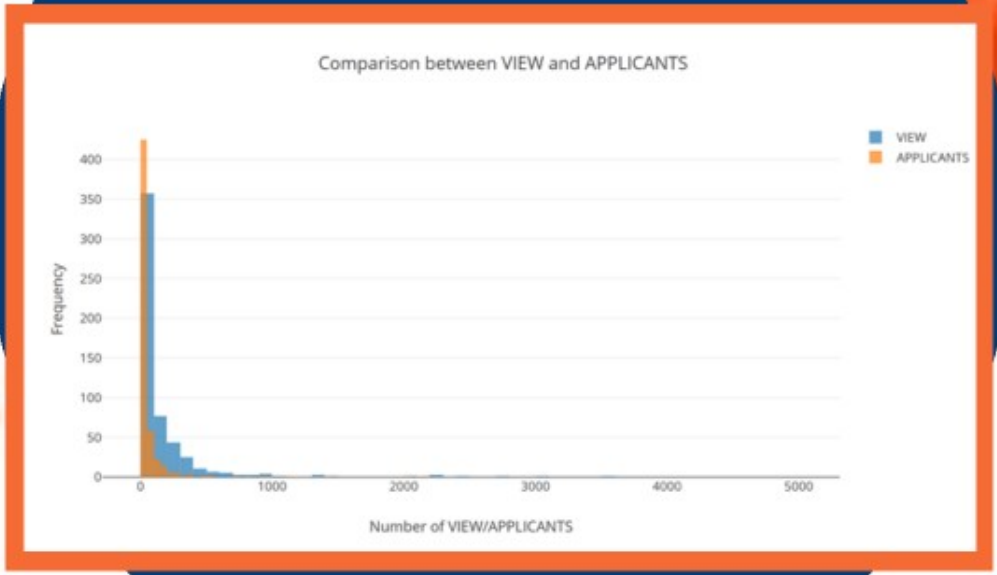
Visualization



Industry
Senority Level

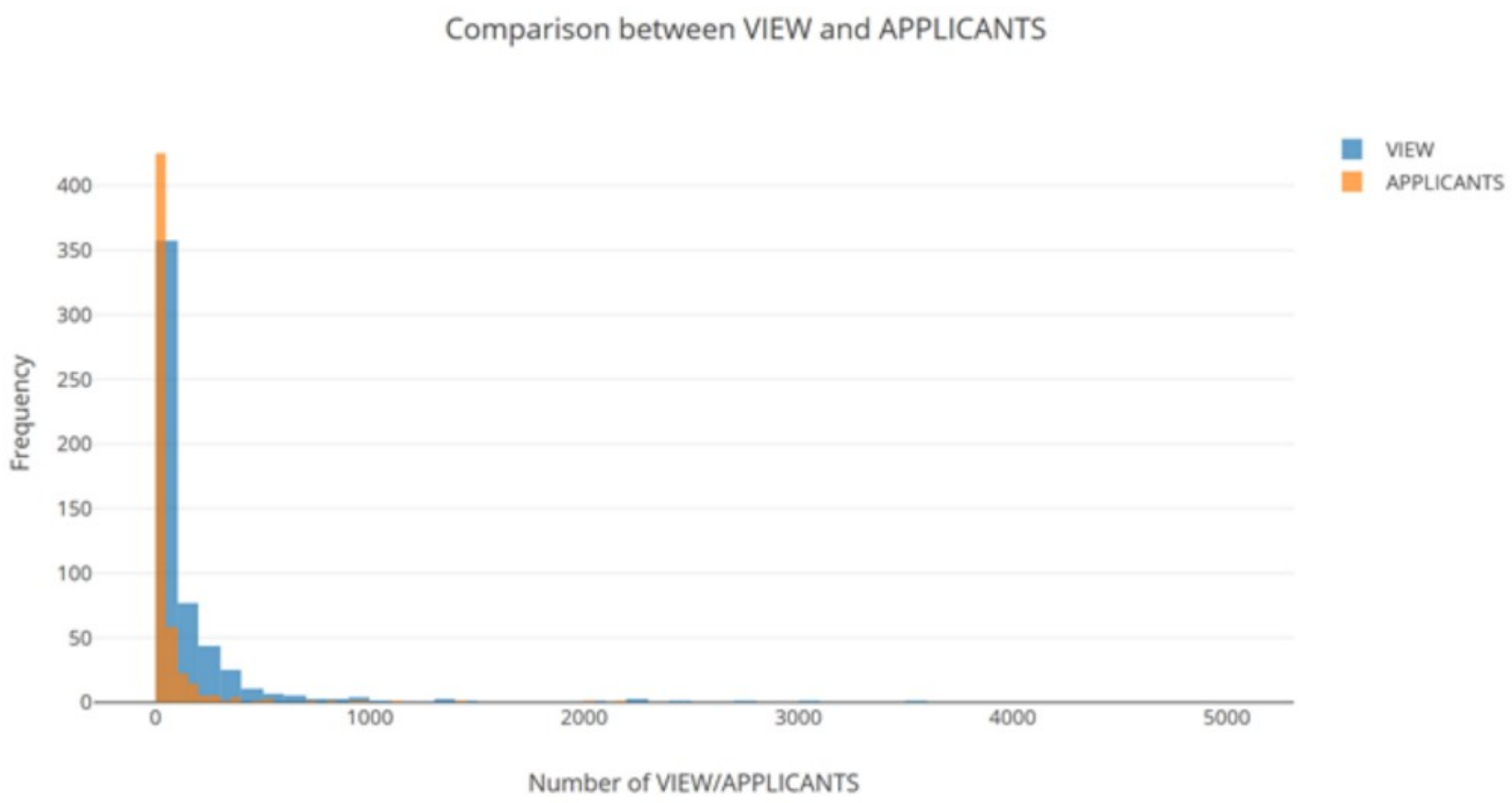
VS.
#Applicant
#Views

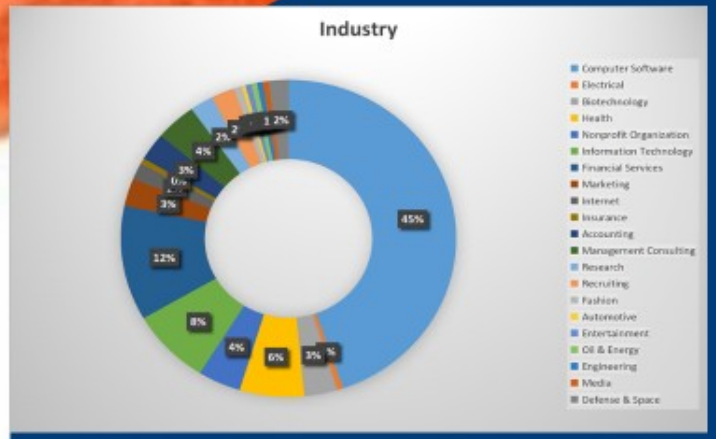
Visualization

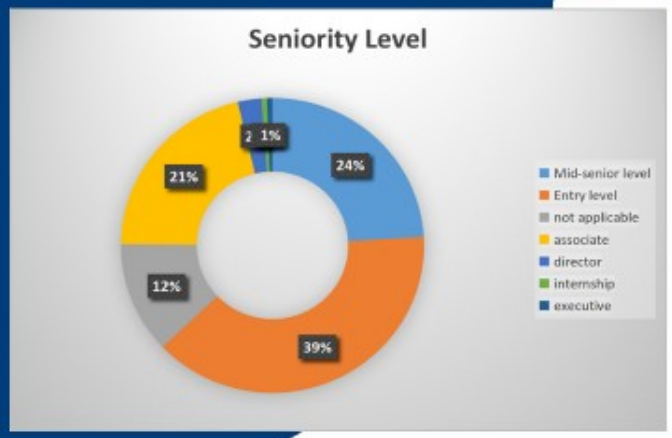


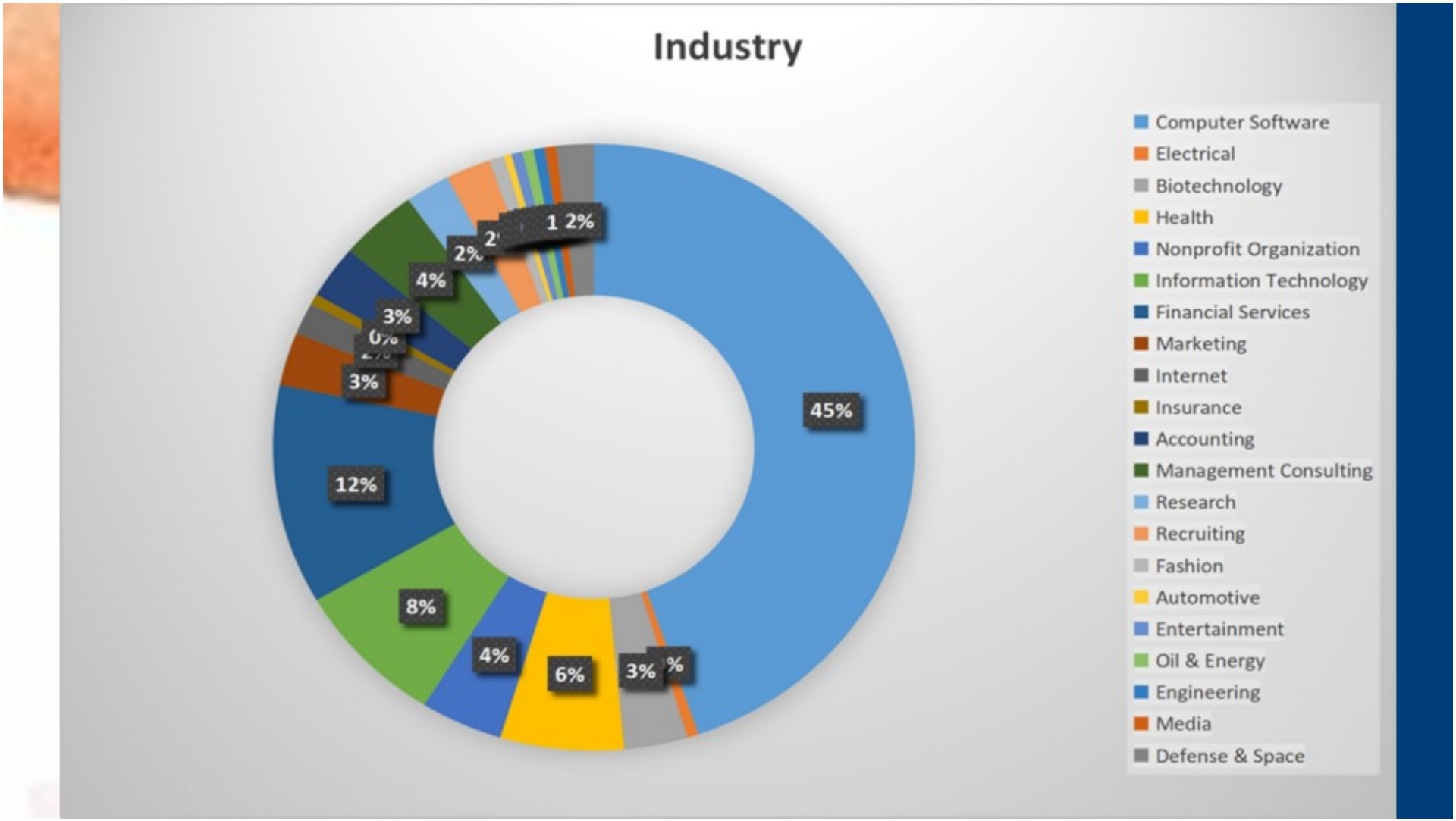
Industry
Senority Level

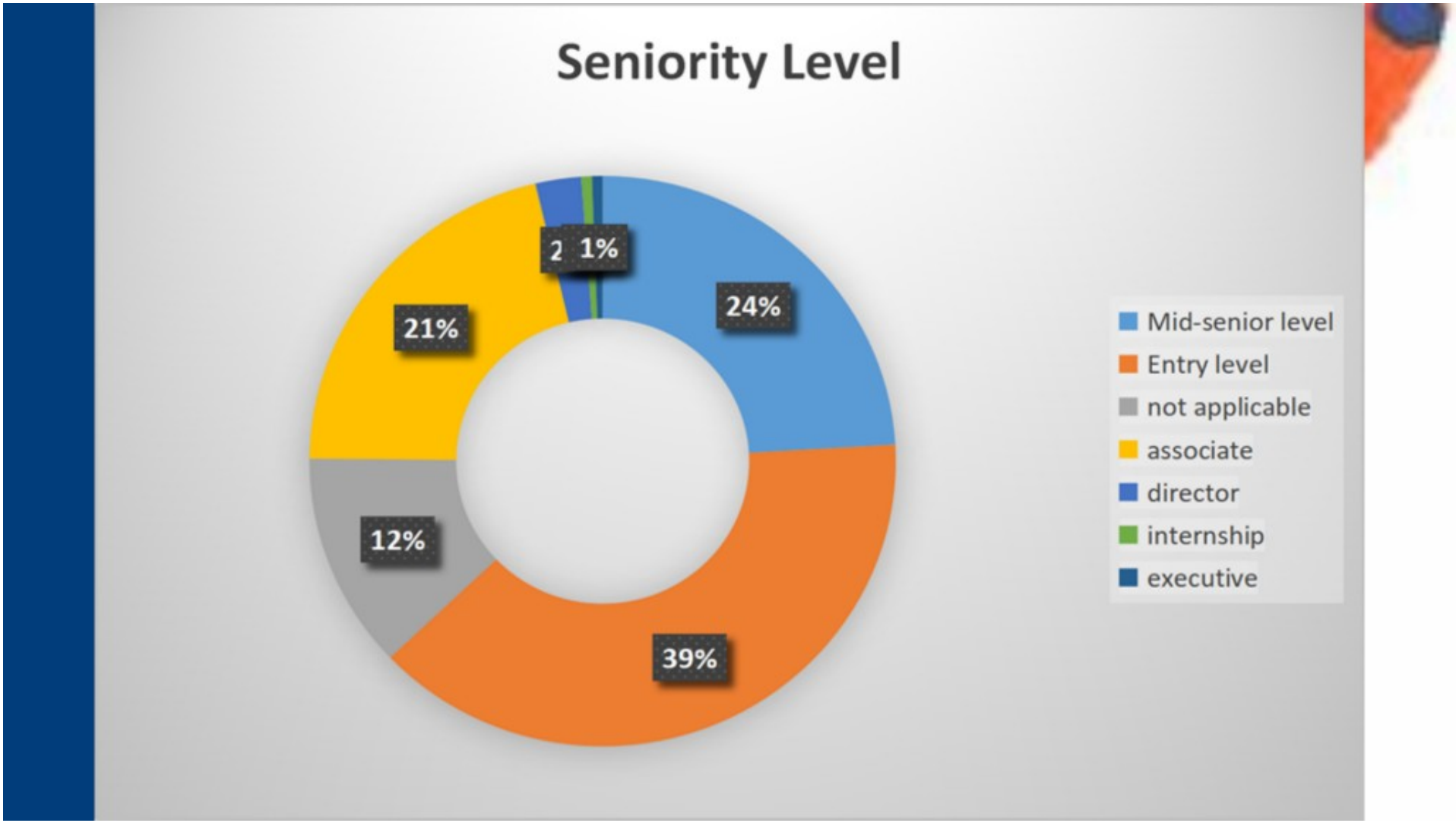
VS.
#Applicant
#Views

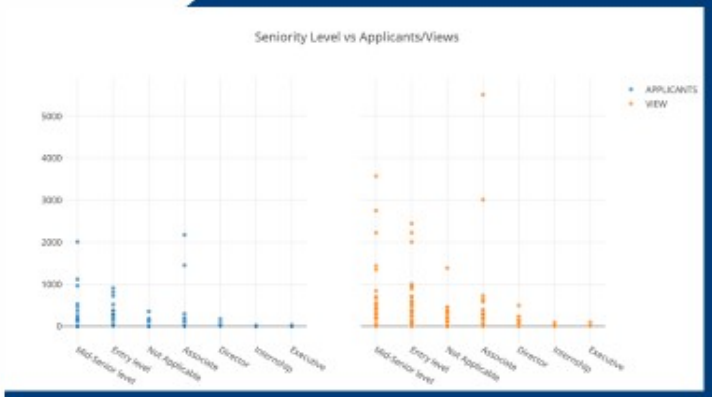

Comparison between VIEW and APPLICANTS
VIEW
APPLICANTS
Frequency
400
350
300
250
200
150
100
50
0
0
1000
2000
3000
4000
5000
Number of VIEW/APPLICANTS


Industry
45%
12%
8%
4%
6%
3%
3%
4%
2%
1 2%
Computer Software
Electrical
Biotechnology
Health
Nonprofit Organization
Information Technology
Financial Services
Marketing
Internet
Insurance
Accounting
Management Consulting
Research
Recruiting
Fashion
Automotive
Entertainment
Oil & Energy
Engineering
Media
Defense & Space

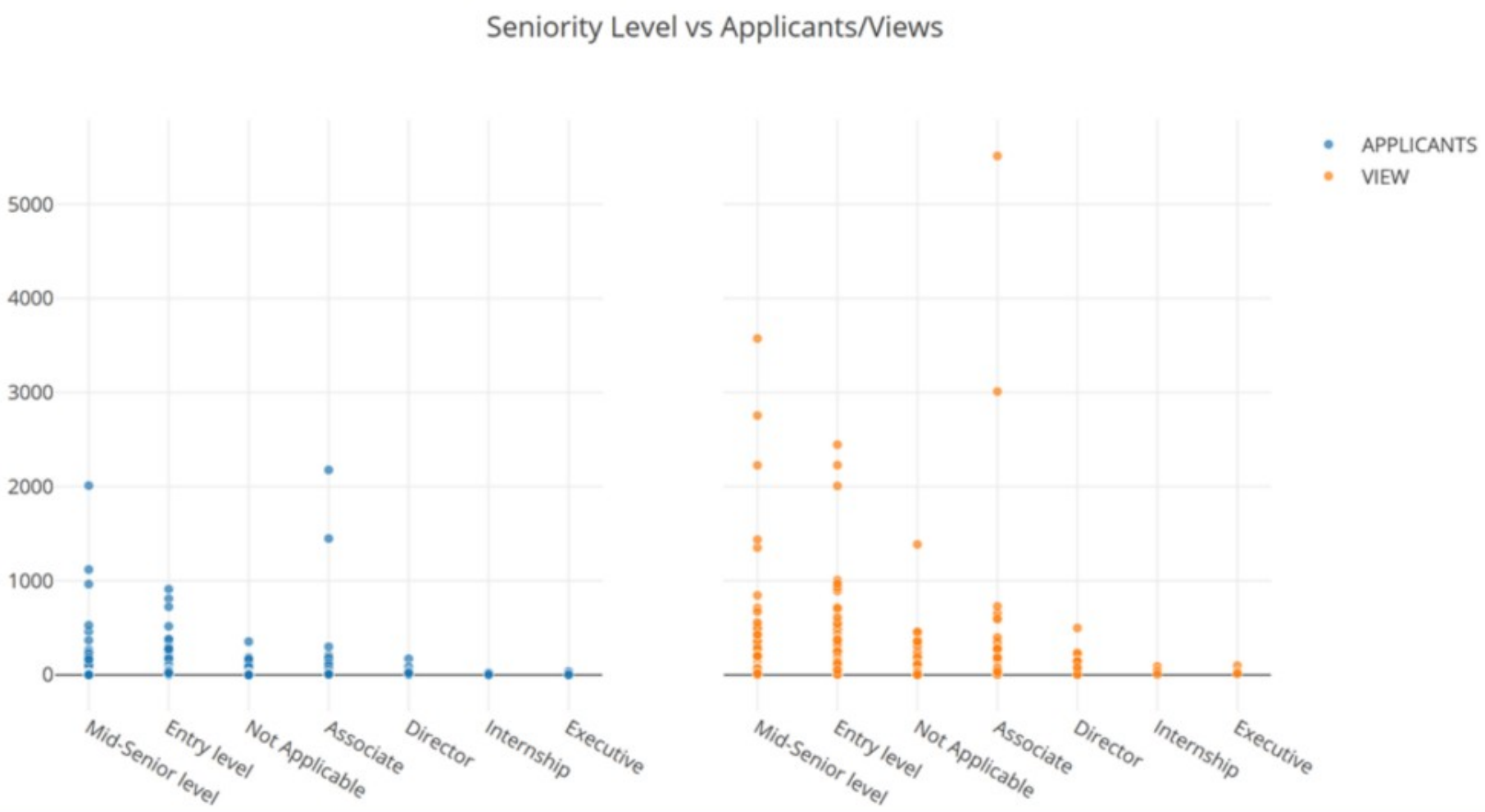

Seniority Level
24%
39%
12%
21%
2 1%
Mid-senior level
Entry level
not applicable
associate
director
internship
executive

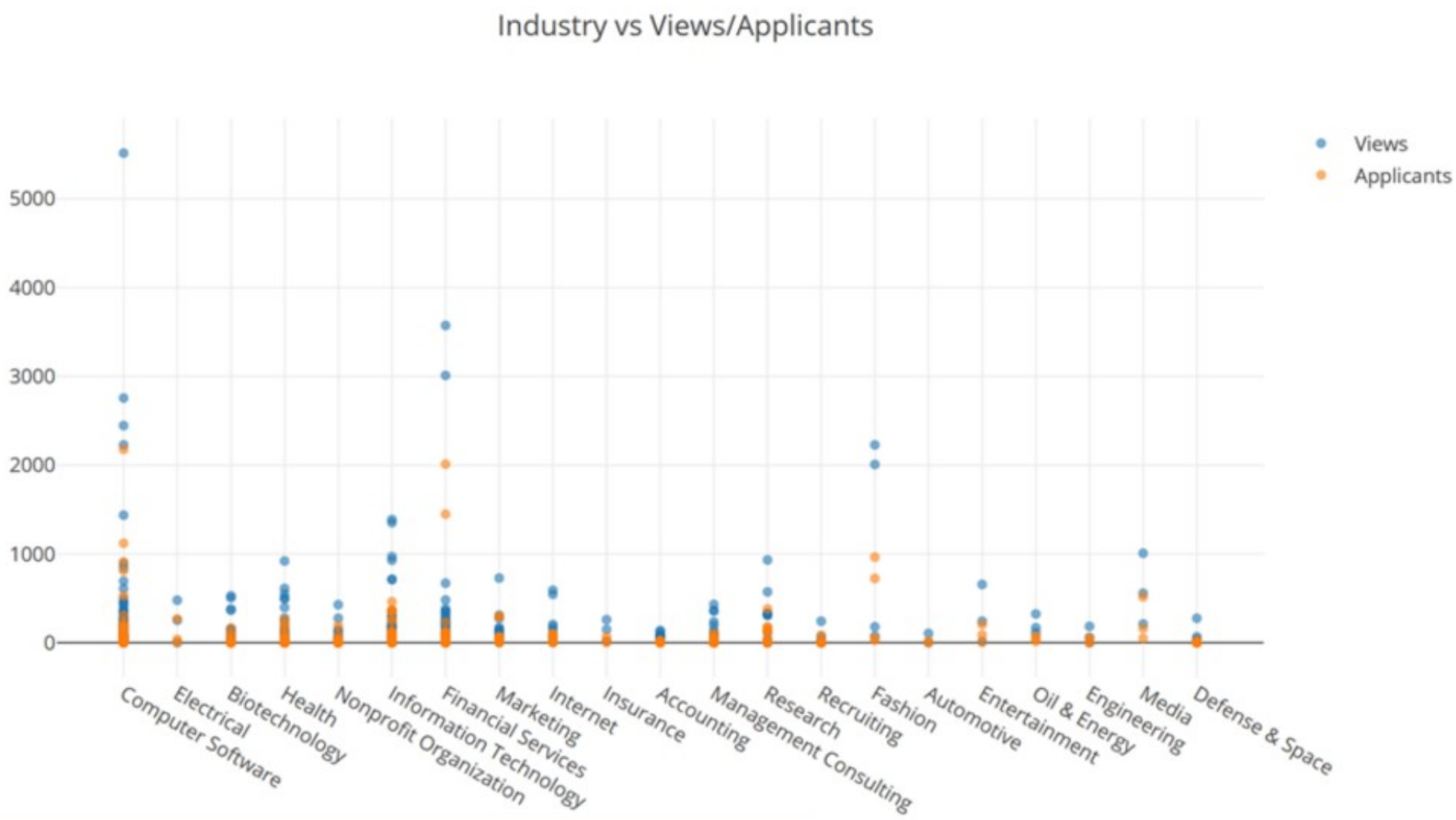

Industry
45%
12%
8%
4%
6%
3%
3%
3%
4%
2%
1 2%
Computer Software
Electrical
Biotechnology
Health
Nonprofit Organization
Information Technology
Financial Services
Marketing
Internet
Insurance
Accounting
Management Consulting
Research
Recruiting
Fashion
Automotive
Entertainment
Oil & Energy
Engineering
Media
Defense & Space

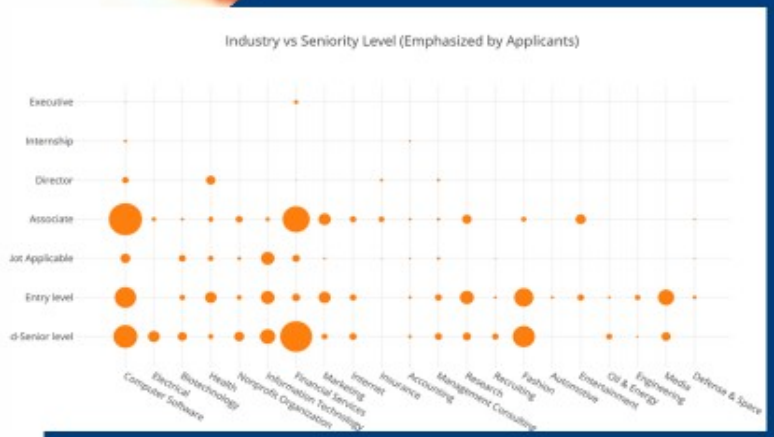

Seniority Level
24%
39%
12%
21%
2
1%
Mid-senior level
Entry level
not applicable
associate
director
internship
executive

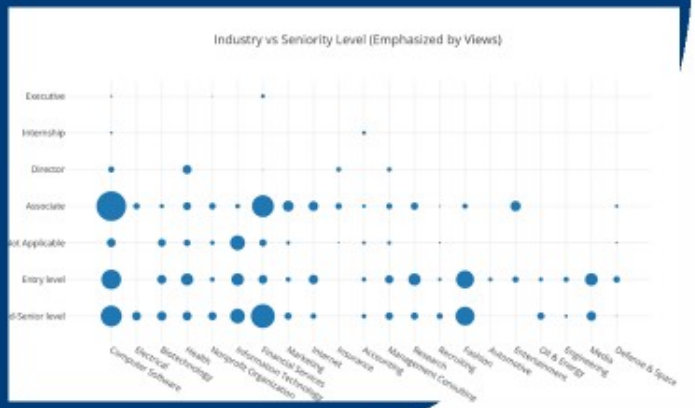

Seniority Level vs Applicants/Views
APPLICANTS
VIEW
5000
4000
3000
2000
1000
0
Mid-Senior level
Entry level
Not Applicable
Associate
Director
Internship
Executive

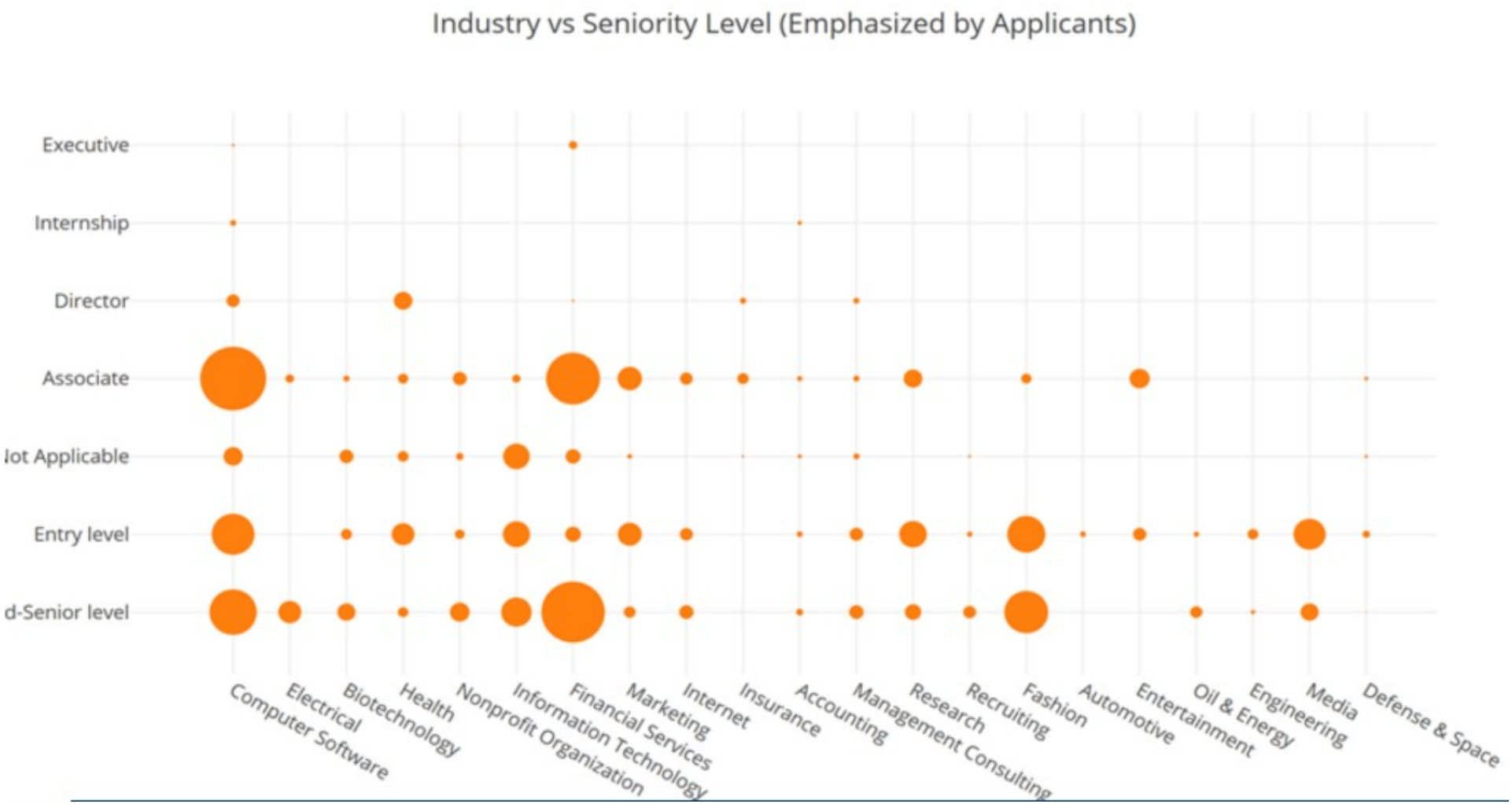

Industry vs Views/Applicants
Views
Applicants
5000
4000
3000
2000
1000
0
Computer Software
Electrical
Biotechnology
Health
Nonprofit Organization
Information Technology
Financial Services
Marketing
Internet
Insurance
Accounting
Management Consulting
Research
Recruiting
Fashion
Automotive
Entertainment
Oil & Energy
Engineering
Media
Defense & Space

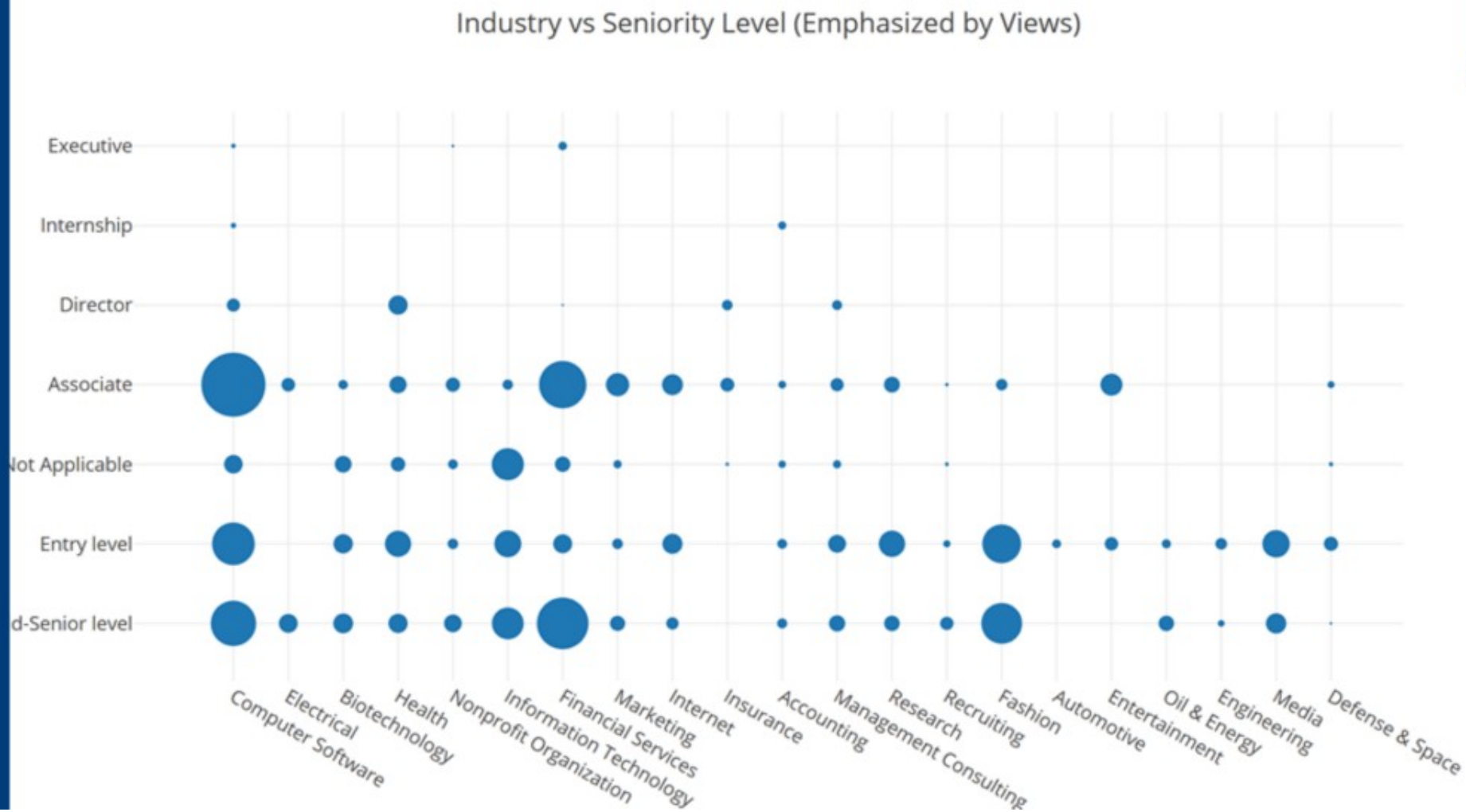

Seniority Level vs Applicants/Views
APPLICANTS
VIEW
5000
4000
3000
2000
1000
0
Mid-Senior level
Entry level
Not Applicable
Associate
Director
Internship
Executive
Mid-Senior level
Entry level
Not Applicable
Associate
Director
Internship
Executive


Industry vs Views/Applicants
Views
Applicants
5000
4000
3000
2000
1000
0
Computer Software
Electrical
Biotechnology
Health
Nonprofit Organization
Information Technology
Financial Services
Marketing
Internet
Insurance
Accounting
Management Consulting
Research
Recruiting
Fashion
Automotive
Entertainment
Oil & Energy
Engineering
Media
Defense & Space


Industry vs Seniority Level (Emphasized by Applicants)
Executive
Internship
Director
Associate
Not Applicable
Entry level
Mid-Senior level
Computer Software
Electrical
Biotechnology
Health
Nonprofit Organization
Information Technology
Financial Services
Marketing
Internet
Insurance
Accounting
Management Consulting
Research
Recruiting
Fashion
Automotive
Entertainment
Oil & Energy
Engineering
Media
Defense & Space


Industry vs Seniority Level (Emphasized by Views)
Executive
Internship
Director
Associate
Not Applicable
Entry level
Mid-Senior level
Computer Software
Electrical
Biotechnology
Health
Nonprofit Organization
Information Technology
Financial Services
Marketing
Internet
Insurance
Accounting
Management Consulting
Research
Recruiting
Fashion
Automotive
Entertainment
Oil & Energy
Engineering
Media
Defense & Space


Industry vs Seniority Level (Emphasized by Applicants)
Executive
Internship
Director
Associate
Not Applicable
Entry level
Mid-Senior level
Computer Software
Electrical
Biotechnology
Health
Nonprofit Organization
Information Technology
Financial Services
Marketing
Internet
Insurance
Accounting
Management Consulting
Research
Recruiting
Fashion
Automotive
Entertainment
Oil & Energy
Engineering
Media
Defense & Space


Industry vs Seniority Level (Emphasized by Views)
Executive
Internship
Director
Associate
Not Applicable
Entry level
Mid-Senior level
Computer Software
Electrical
Biotechnology
Health
Nonprofit Organization
Information Technology
Financial Services
Marketing
Internet
Insurance
Accounting
Management Consulting
Research
Recruiting
Fashion
Automotive
Entertainment
Oil & Energy
Engineering
Media
Defense & Space

Further
Ideas

ISGB 7978 Web Analytics

Group 7
Choudhury, Sumi
Gu, Jiahua
Huang, Feifei
Ji, Minxia

Algorithm

- **Feature: Skills & Qualifications**
- **Label: Conversion rate**

Conversion Rate =
Number of Applicants/Number of Views

The median of conversion rate is **0.2**, which means every **5** views has **1** applicants

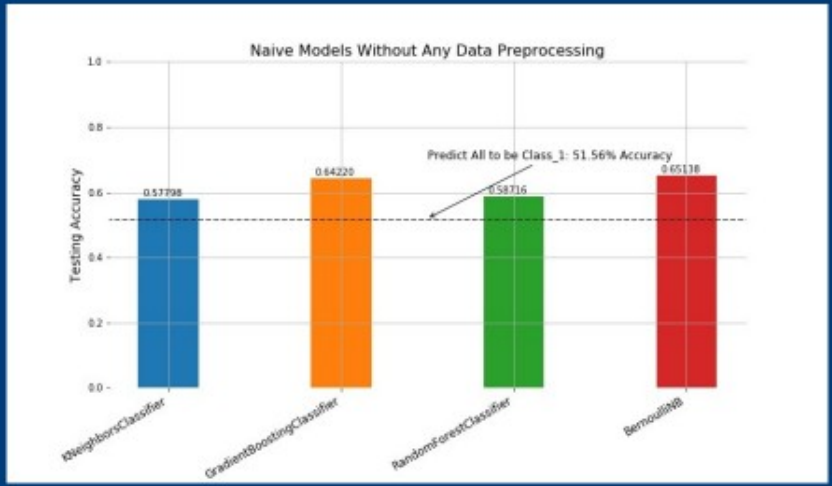
- label '1': conversion rate **> 0.2** (good)
- label '0': conversion rate **<= 0.2** (limited)

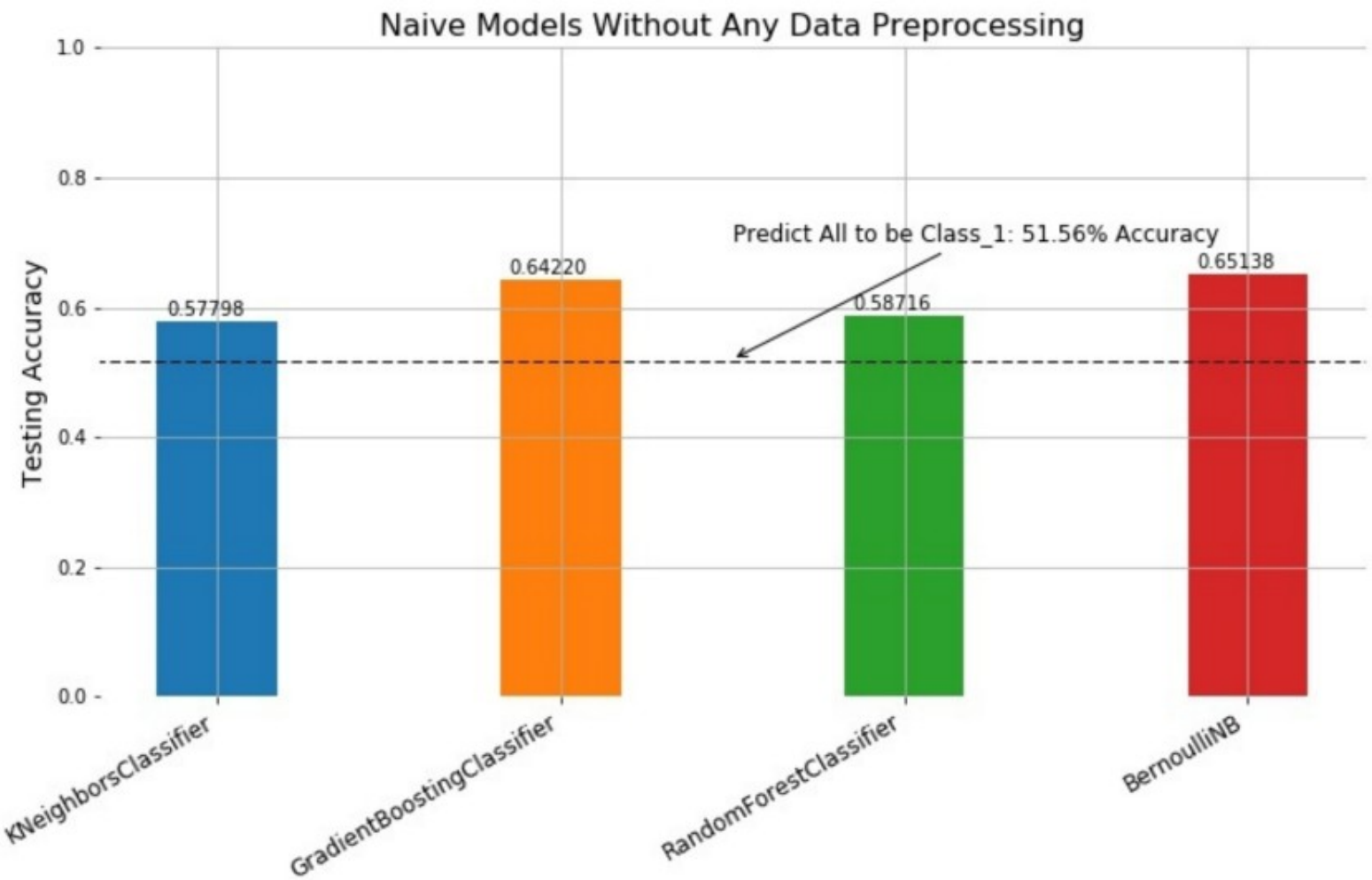
Naive Model

Data Cleaning & Transforming

Reduce Dimension & Classification

Naive Model


Naive Models Without Any Data Preprocessing
Testing Accuracy
1.0
0.8
0.6
0.4
0.2
0.0
Predict All to be Class_1: 51.56% Accuracy
0.57798
0.64220
0.58716
0.65138
KNeighborsClassifier
GradientBoostingClassifier
RandomForestClassifier
BernoulliNB


Naive Models Without Any Data Preprocessing
Testing Accuracy
1.0
0.8
0.6
0.4
0.2
0.0
Predict All to be Class_1: 51.56% Accuracy
0.57798
0.64220
0.58716
0.65138
KNeighborsClassifier
GradientBoostingClassifier
RandomForestClassifier
BernoulliNB

Data Cleaning & Transforming

- Removing Stopwords and Punctuations : remove commonly used words in the English language such as I, me, my, etc.
- Text Stemming: reduces words to their root form. For example, "moving", "moved" and "movement" to the root word, "move".
- Transformed each feature to TF*IDF

Reduce Dimension & Classification

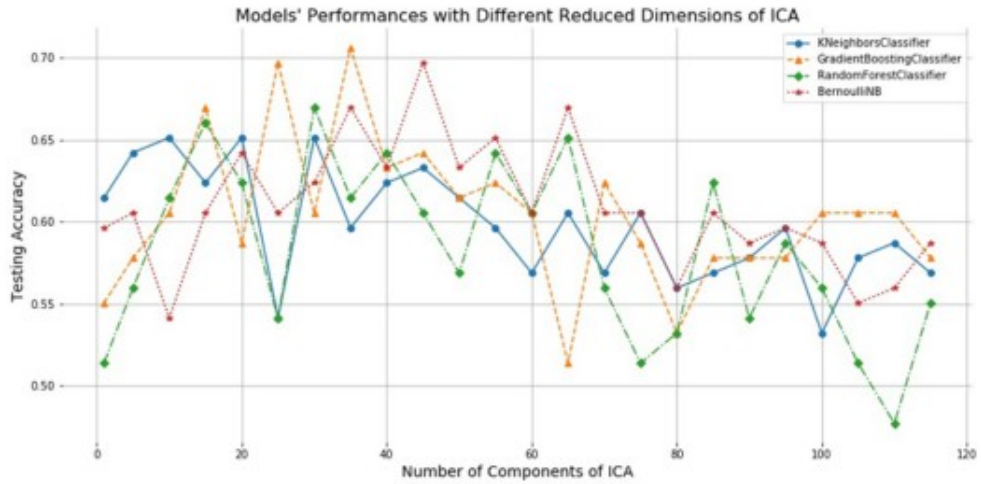
Reduce Dimension:

- ICA: Independent Component Analysis
- PCA: Principal Component Analysis
- KernelPCA: Kernel Principal Component Analysis

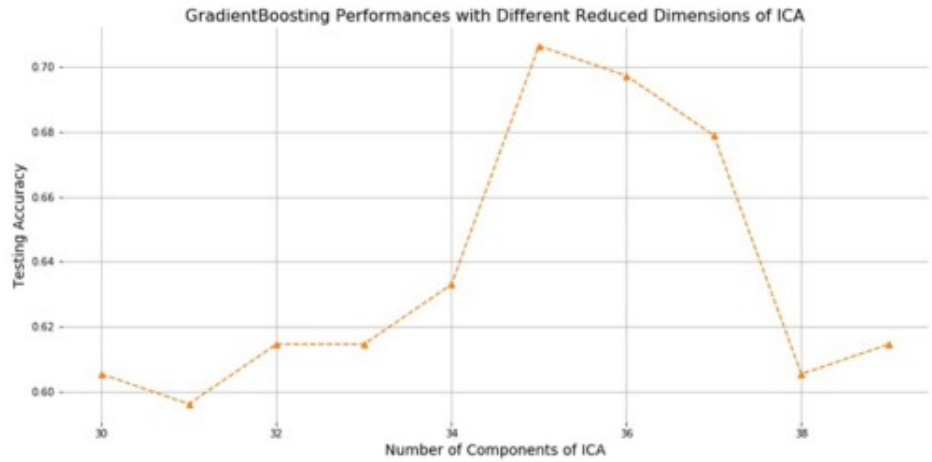
Classification:

- KNeighbors Classifier
- GradientBoosting Classifier
- RandomForest Classifier
- BernoulliNaiveBayes Classifier

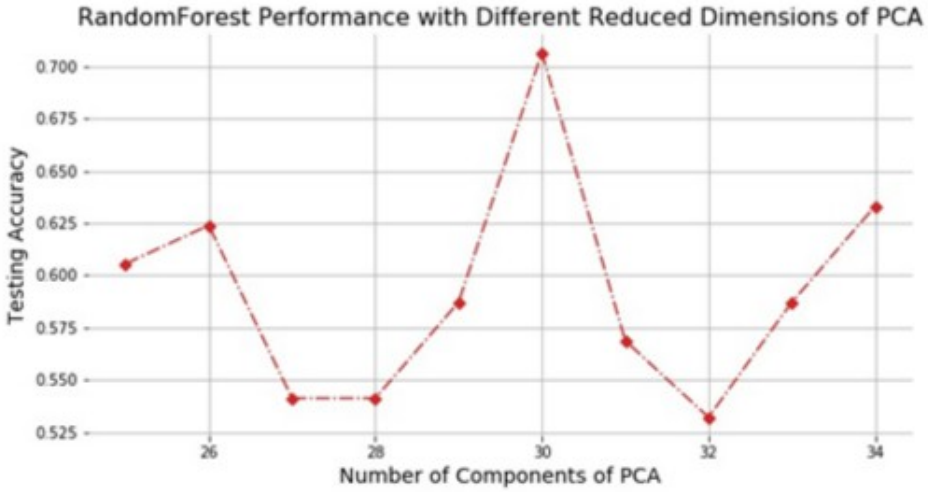
```python
def train_after_DR(learners,dr,X_train_tfidf,X_test_tfidf,y_train,y_test):
    """
    function to trian machine learning algorithms after dimension reduction method
    'learners': list of machine learning algorithms
    'dr': dimension reduction method
    'X_train_tfidf,X_test_tfidf,y_train,y_test': training data and testing data
    """
    result = {}
    # convert sparse matrix to ndarray
    X_train = X_train_tfidf.toarray()
    X_test = X_test_tfidf.toarray()
    # fit Dimension Reduction algorithm
    reduced_X_train = dr.fit_transform(X_train)
    reduced_X_test = dr.transform(X_test)
    # fit machine learning algorithms
    for learner in learners:
        learner.fit(reduced_X_train,y_train)
        pred = learner.predict(reduced_X_test)
        acc = np.mean(pred == y_test)
        result[learner.__class__.__name__] = acc

    return result
```

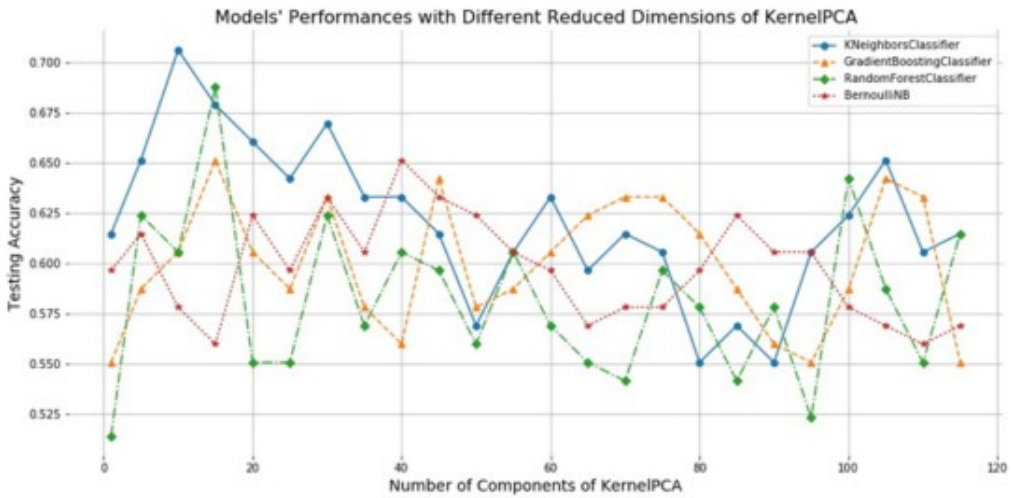


Gradient Boosting Classifier has the best performance based on first 35 ICs testing accuracy around 71%



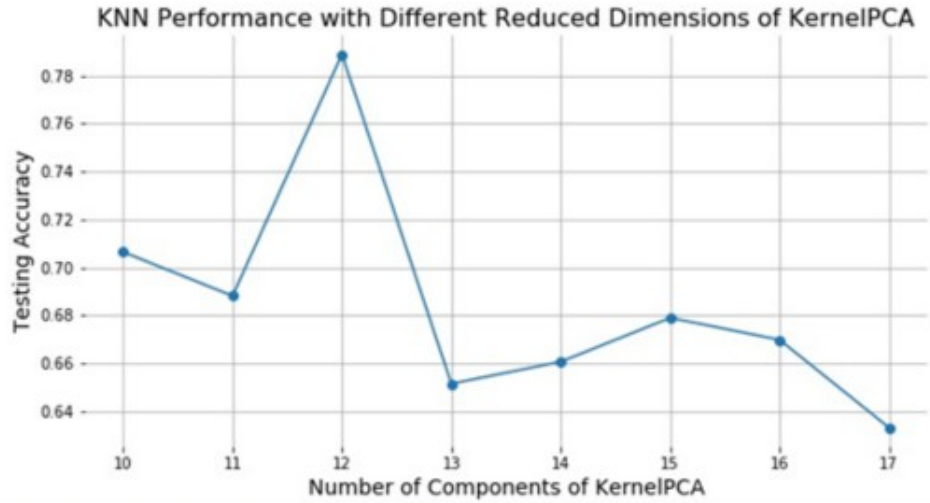
Random Forest Classifier has the best performance based on first 30 PCs: testing accuracy around 71%





Best! 79%
Data: first 12 PCs with Kernel PCA
Model: KNeighbors Classifier

KNeighborsClassifier has the best performance based on first 12 PCs:
testing accuracy around 79%



```
{'PC0': [u'experi', u'learn', u'machin', u'statist', u'data'],
 'PC1': [u'data', u'abil', u'year', u'model', u'busi'],
 'PC10': [u'knowledg', u'least', u'minimum', u'advanc', u'year'],
 'PC2': [u'year', u'learn', u'degre', u'analyt', u'experi'],
 'PC3': [u'data', u'year', u'least', u'experi', u'strong'],
 'PC4': [u'data', u'experi', u'viewpoint', u'plu', u'mathemat'],
 'PC5': [u'least', u'busi', u'abil', u'statist', u'quantit'],
 'PC6': [u'requir', u'knowledg', u'minimum', u'busi', u'model'],
 'PC7': [u'data', u'plu', u'least', u'big', u'languag'],
 'PC8': [u'languag', u'common', u'data', u'experi', u'program'],
 'PC9': [u'experi', u'softwar', u'requir', u'least', u'analyt']}
```

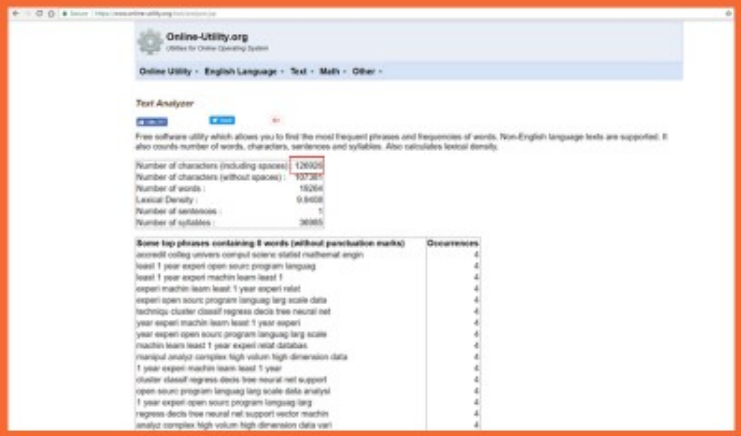
	precision	recall	f1-score	support
Unpopular_Post	0.84	0.70	0.76	53
Popular_Post	0.75	0.88	0.81	56
avg / total	0.80	0.79	0.79	109

Further Ideas

ISGB 7978 Web Analytics

Group 7
Choudhury, Sumi
Gu, Jiahua
Huang, Feifei
Ji, Minxia

Results



Online-Utility.org

Online Utility - English Language - Text - Math - Other -

Text Analyzer

Free software utility which allows you to find the most frequent phrases and frequencies of words. Non-English language texts are supported. It also counts number of words, characters, sentences and syllables. Also calculates lexical density.

Number of characters (including spaces) :	126926
Number of characters (without spaces) :	107381
Number of words :	18264
Lexical Density :	0.8458
Number of sentences :	1
Number of syllables :	36985

Some top phrases containing 8 words (without punctuation marks)	Occurrences
accredit colleg univers comput scienc statist mathemat engin	4
least 1 year experi open sourc program languag	4
least 1 year experi machin learn least 1	4
experi machin learn least 1 year experi relat	4
experi open sourc program languag larg scale data	4
techniqu cluster classif regress decis tree neural net	4
year experi machin learn least 1 year experi	4
year experi open sourc program languag larg scale	4
machin learn least 1 year experi relat databas	4
manipul analyz complex high volum high dimension data	4
1 year experi machin learn least 1 year	4
cluster classif regress decis tree neural net support	4
open sourc program languag larg scale data analysi	4
1 year experi open sourc program languag larg	4
regress decis tree neural net support vector machin	4
analyz complex high volum high dimension data vari	4

Secure | https://www.online-utility.org/text/analyzer.jsp

Online-Utility.org
Utilities for Online Operating System

Online Utility - **English Language** - **Text** - **Math** - **Other** -

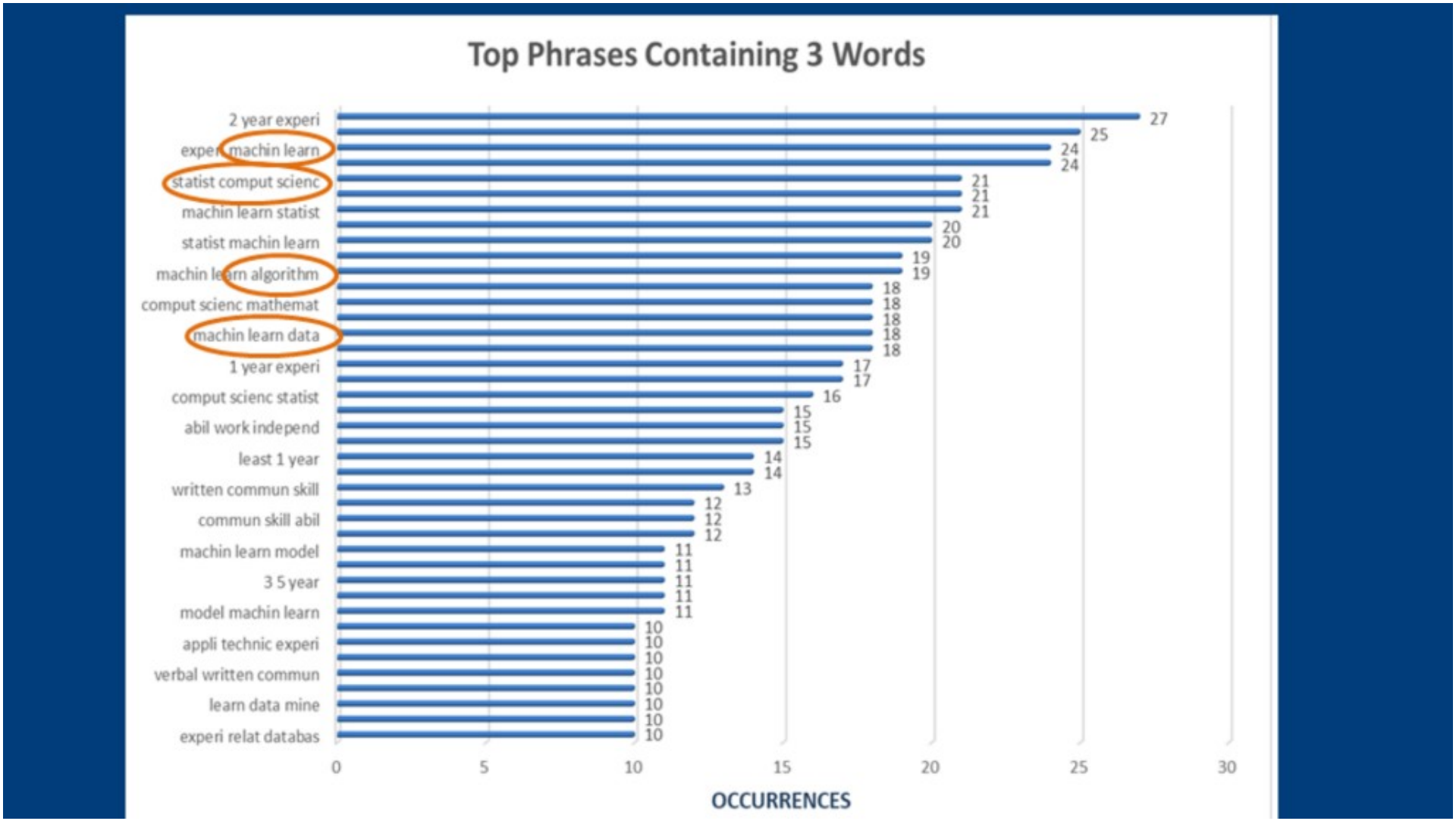
Text Analyzer

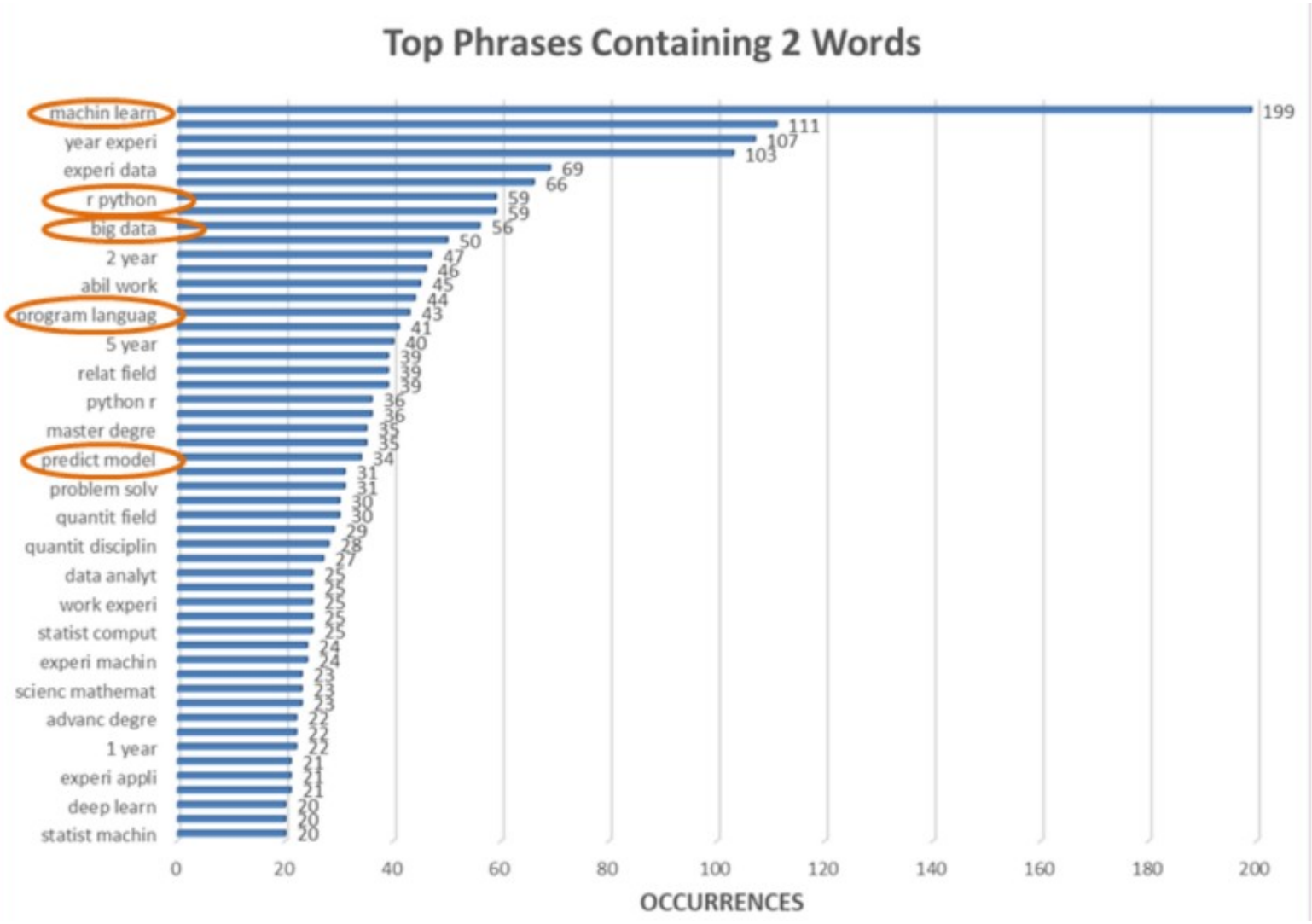
Like 210 Tweet G+

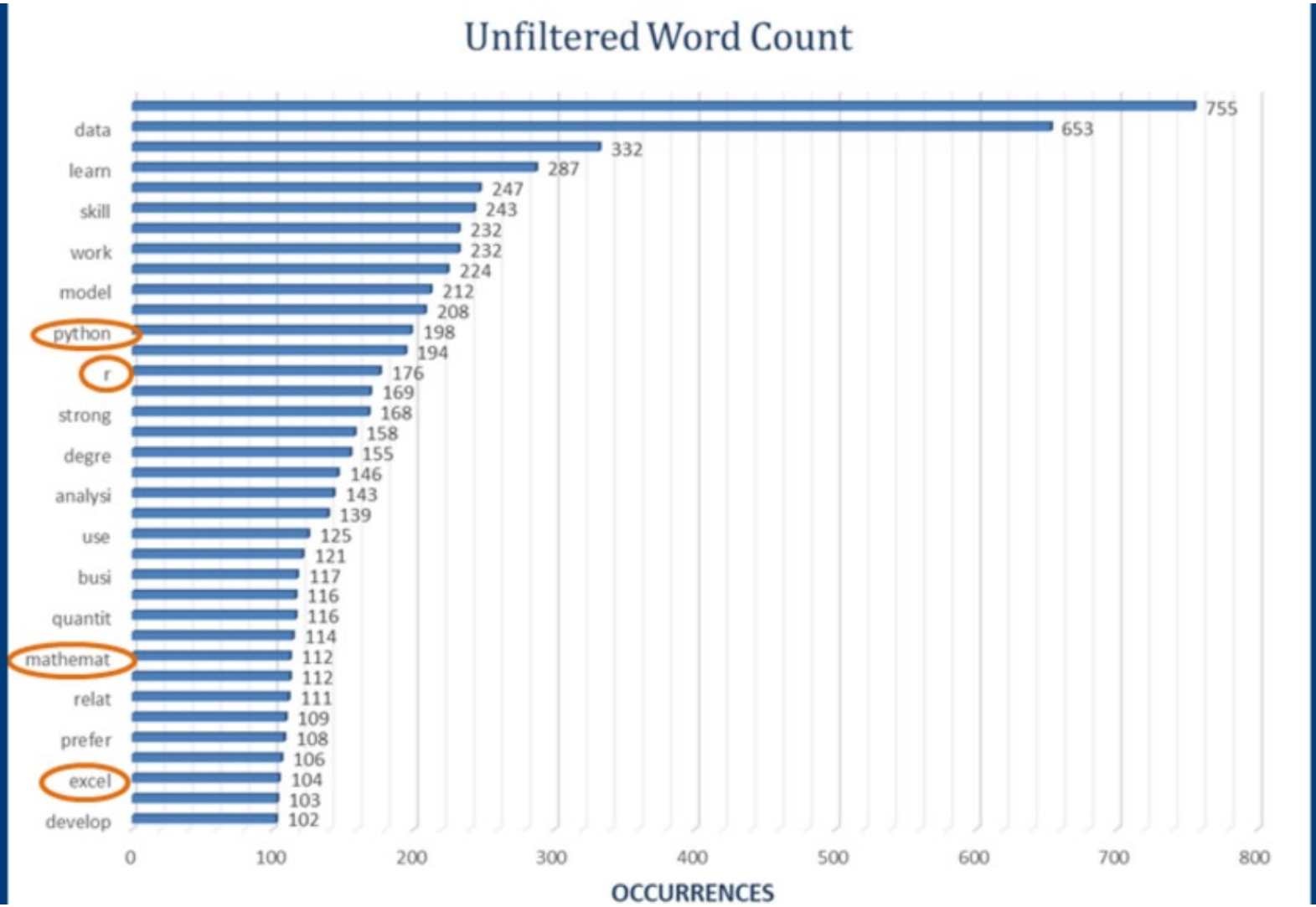
Free software utility which allows you to find the most frequent phrases and frequencies of words. Non-English language texts are supported. It also counts number of words, characters, sentences and syllables. Also calculates lexical density.

Number of characters (including spaces) :	126926
Number of characters (without spaces) :	107381
Number of words :	19264
Lexical Density :	9.9408
Number of sentences :	1
Number of syllables :	36985

Some top phrases containing 8 words (without punctuation marks)	**Occurrences**
accredit colleg univers comput scienc statist mathemat engin	4
least 1 year experi open sourc program languag	4
least 1 year experi machin learn least 1	4
experi machin learn least 1 year experi relat	4
experi open sourc program languag larg scale data	4
techniqu cluster classif regress decis tree neural net	4
year experi machin learn least 1 year experi	4
year experi open sourc program languag larg scale	4
machin learn least 1 year experi relat databas	4
manipul analyz complex high volum high dimension data	4
1 year experi machin learn least 1 year	4
cluster classif regress decis tree neural net support	4
open sourc program languag larg scale data analysi	4
1 year experi open sourc program languag larg	4
regress decis tree neural net support vector machin	4
analyz complex high volum high dimension data vari	4


Top Phrases Containing 3 Words
2 year experi
experi machin learn
statist comput scienc
machin learn statist
statist machin learn
machin learn algorithm
comput scienc mathemat
machin learn data
1 year experi
comput scienc statist
abil work independ
least 1 year
written commun skill
commun skill abil
machin learn model
3 5 year
model machin learn
appli technic experi
verbal written commun
learn data mine
experi relat databas
27
25
24
24
21
21
21
20
20
19
19
18
18
18
18
18
17
17
16
15
15
15
14
14
13
12
12
12
11
11
11
11
11
10
10
10
10
10
10
10
10
0
5
10
15
20
25
30
OCCURRENCES


Top Phrases Containing 2 Words
machin learn 199
year experi 111
experi data 107
r python 103
big data 69
2 year 66
abil work 59
program languag 59
5 year 56
relat field 50
python r 47
master degre 46
predict model 45
problem solv 44
quantit field 43
quantit disciplin 41
data analyt 40
work experi 39
statist comput 39
experi machin 39
scienc mathemat 36
advanc degre 36
1 year 35
experi appli 35
deep learn 34
statist machin 31
31
30
30
29
28
27
25
25
25
25
25
24
24
23
23
23
22
22
22
21
21
21
20
20
20
0
20
40
60
80
100
120
140
160
180
200
OCCURRENCES


Unfiltered Word Count
data
learn
skill
work
model
python
r
strong
degre
analysi
use
busi
quantit
mathemat
relat
prefer
excel
develop
755
653
332
287
247
243
232
232
224
212
208
198
194
176
169
168
158
155
146
143
139
125
121
117
116
116
114
112
112
111
109
108
106
104
103
102
0
100
200
300
400
500
600
700
800
OCCURRENCES

Secure | https://www.online-utility.org/text/analyzer.jsp

Online-Utility.org

Utilities for Online Operating System

Online Utility - English Language - Text - Math - Other -

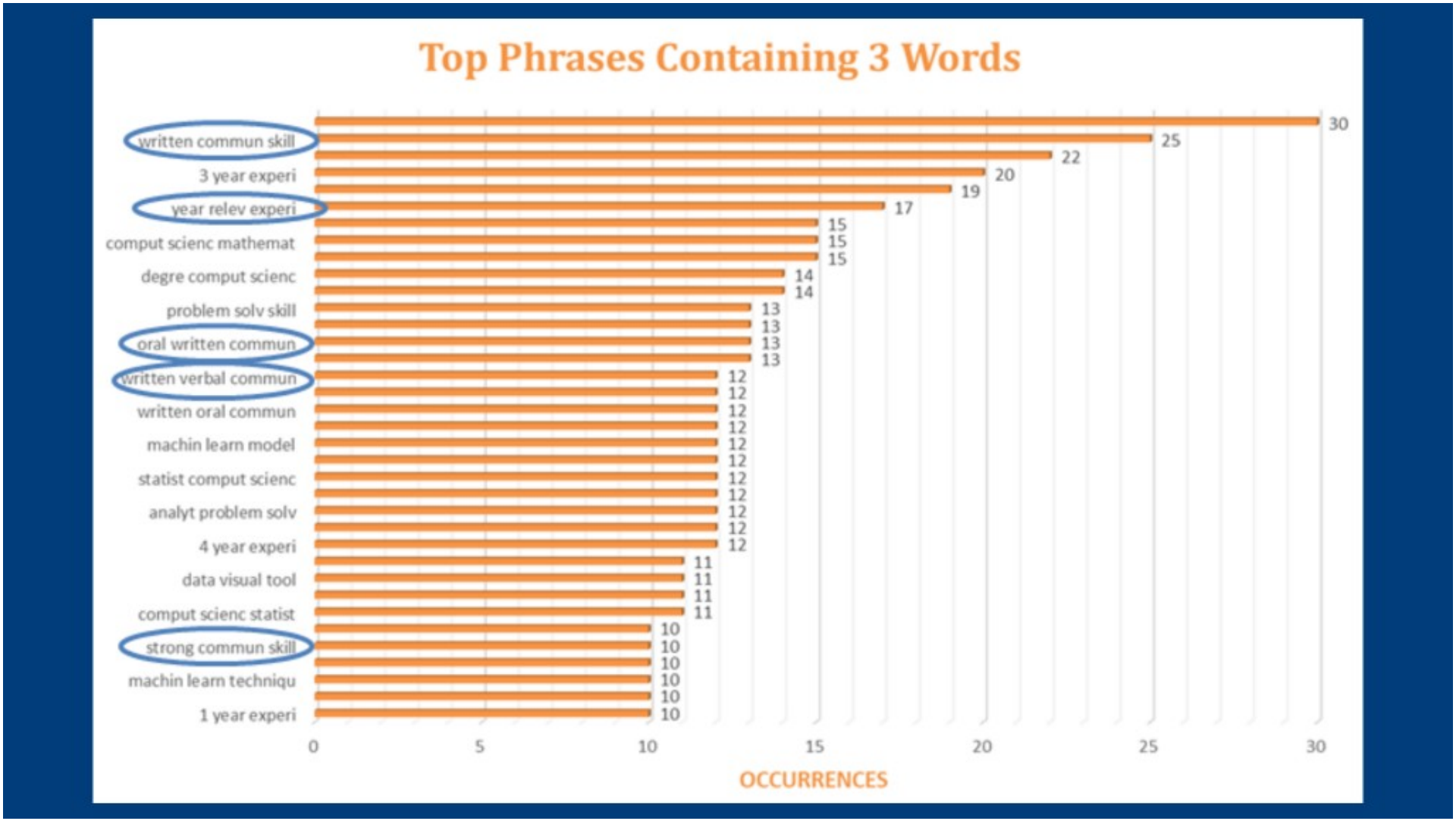
Text Analyzer

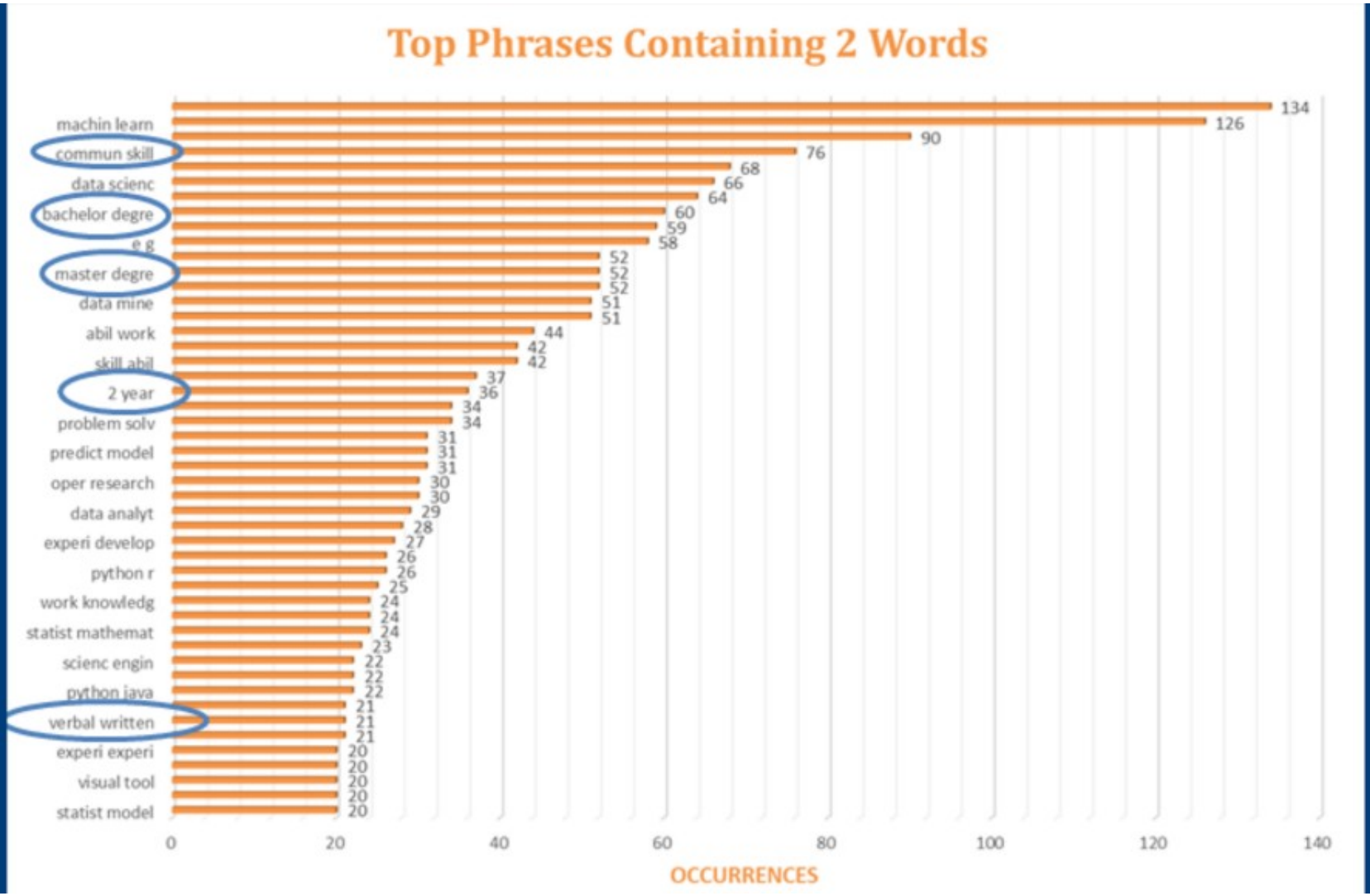
Like 217 Tweet G+

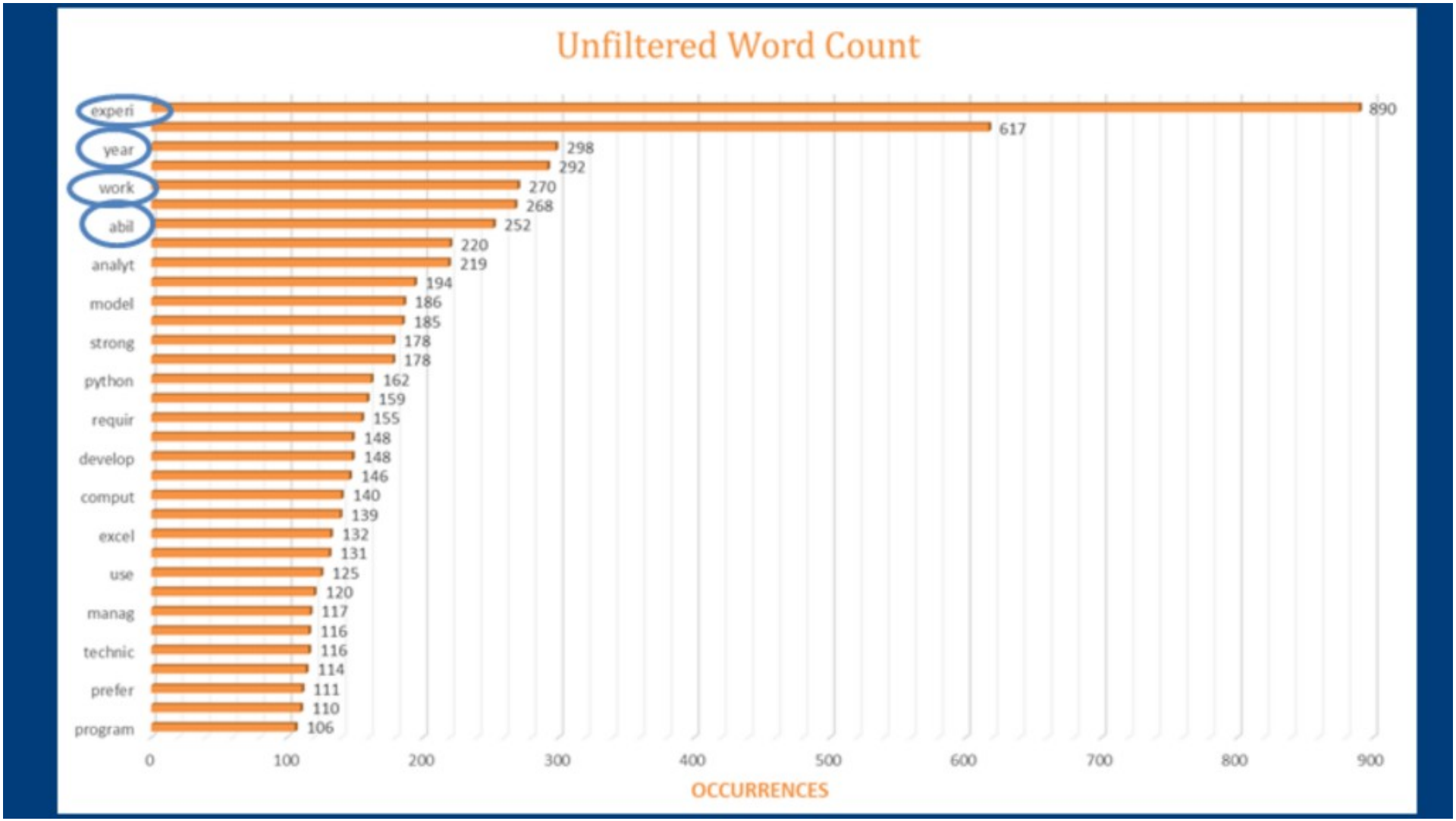
Free software utility which allows you to find the most frequent phrases and frequencies of words. Non-English language texts are supported. It also counts number of words, characters, sentences and syllables. Also calculates lexical density.

Number of characters (including spaces)	139158
Number of characters (without spaces) :	117856
Number of words :	21037
Lexical Density :	9.7067
Number of sentences :	1
Number of syllables :	41018

Some top phrases containing 8 words (without punctuation marks)	Occurrences
accredit colleg univers comput scienc statist mathemat engin	3
excel skill strong project manag skill abil priorit	3
knowledg data analyt busi intellig visual tool e	3
analysi skill abil clearli present complex inform simpl	3
stakehold experi data mine structur unstructur data sap	3
discern busi mean recommend action stakehold experi data	3
data mine data analysi skill abil clearli present	3
mine structur unstructur data sap hana sql etl	3
abil work matrix organ virtual team minim travel	3
commun skill strong knowledg data analyt busi intellig	3
sql etl data warehous etc busi environ larg	3
scienc master degre plu 5 year experi market	3
cluster regress analysi statist tool prefer r sa	3
busi mean recommend action stakehold experi data mine	3
degre plu 5 year experi market analyt excel	3
knowledg statist analys e g cluster regress analysi	3


Top Phrases Containing 3 Words
written commun skill
3 year experi
year relev experi
comput scienc mathemat
degre comput scienc
problem solv skill
oral written commun
written verbal commun
written oral commun
machin learn model
statist comput scienc
analyt problem solv
4 year experi
data visual tool
comput scienc statist
strong commun skill
machin learn techniqu
1 year experi
30
25
22
20
19
17
15
15
15
14
14
13
13
13
13
12
12
12
12
12
12
12
12
12
12
11
11
11
11
10
10
10
10
10
10
0
5
10
15
20
25
30
OCCURRENCES


Top Phrases Containing 2 Words
machin learn 134
commun skill 126
data scienc 90
bachelor degre 76
e g 68
master degre 66
data mine 64
abil work 60
skill abil 59
2 year 58
problem solv 52
predict model 52
oper research 52
data analyt 51
experi develop 51
python r 44
work knowledg 42
statist mathemat 42
scienc engin 37
python java 36
verbal written 34
experi experi 34
visual tool 31
statist model 31
31
30
30
29
28
27
26
26
25
24
24
24
23
22
22
22
21
21
21
20
20
20
20
20
0
20
40
60
80
100
120
140
OCCURRENCES


Unfiltered Word Count
experi
year
work
abil
analyt
model
strong
python
requir
develop
comput
excel
use
manag
technic
prefer
program
890
617
298
292
270
268
252
220
219
194
186
185
178
178
162
159
155
148
148
146
140
139
132
131
125
120
117
116
116
114
111
110
106
0
100
200
300
400
500
600
700
800
900
OCCURRENCES

Suggestion & Model

- Contain more specific & professional skills
- More consistent and pertinent to the Industry
- Try to avoid too general words, e.g. oral, written, ability, work
- Be gentle on the experience requirements
- Model to predict whether a job posting is popular ot not based on the skills and qualification description

ISGB 7978 Web Analytics

Group 7
Choudhury, Sumi
Gu, Jiahua
Huang, Feifei
Ji, Minxia

Further Ideas

- Try modeling with Deep Learning, such as CNN
- Try to inculde more data
- Through this research, we noticed some other related topics that we can dig into, such as the impact of Data Science thriving in the Fashion Industry

ISGB 7978 Web Analytics

Group 7
Choudhury, Sumi
Gu, Jiahua
Huang, Feifei
Ji, Minxia