



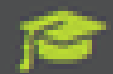
THE
CAPSTONE PROJECT

AN OBJECTIVE LIGHT INTO SUBJECTIVE OPINIONS

BY ZAFRIN MINOTI
DATA SCIENTIST, GALVANIZE
AUGUST 8, 2018



PROFS



SCHOOL



RATE



BLOG



Karen Unger

Professor in the Physical Education department
at [Rio Hondo College](#), Whittier, CA

[ARE YOU KAREN?](#)

OVERALL
QUALITY

4.5

WOULD TAKE
AGAIN

N/A

LEVEL OF
DIFFICULTY

1.1

Tags for this Professor

See how other students describe this professor.

CARING (2)

SKIP CLASS? YOU WON'T PASS. (2)

EXTRA CREDIT (1)

PARTICIPATION MATTERS (1)

CHOOSE YOUR TAGS

REVIEWS ARE NOT PERFECT

11/12/2016



AVERAGE

3.0

OVERALL QUALITY

4.0

LEVEL OF DIFFICULTY

STAT231

For Credit: Yes

Attendance: Not Mandatory

Textbook Used: Yes

Would Take Again: Yes

Grade Received: N/A

Enjoys her work, makes the tutorials fun. Approachable, and understandable.



1 person found this useful



6 people did not find this useful

[report this rating](#)

REVIEWS ARE NOT PERFECT

11/04/2015



AWFUL

1.0

OVERALL QUALITY

5.0

LEVEL OF DIFFICULTY

STAT230

TOUGH GRADER

For Credit: Yes

Attendance: Mandatory

Textbook Used: No

Would Take Again: N/A

Grade Received: A+

terrible old whatever



11 people found this useful



2 people did not find this useful

[report this rating](#)

REVIEWS ARE NOT PERFECT

04/05/2010



GOOD

4.0

OVERALL QUALITY

4.0

LEVEL OF DIFFICULTY

MATH249

For Credit: N/A

Attendance: N/A

Textbook Used: Yes

Would Take Again: N/A

Grade Received: N/A

A nice person and expects a lot from his students. The closest mathematician in mannerism to G.H. Hardy in Waterloo. The midterm was very fair, but the assignments were challenging. Expect to spend a lot of time on them, and learn a lot if you start early.

 0 people found this useful

 0 people did not find this useful

[report this rating](#)

REVIEW BIAS

Gender bias

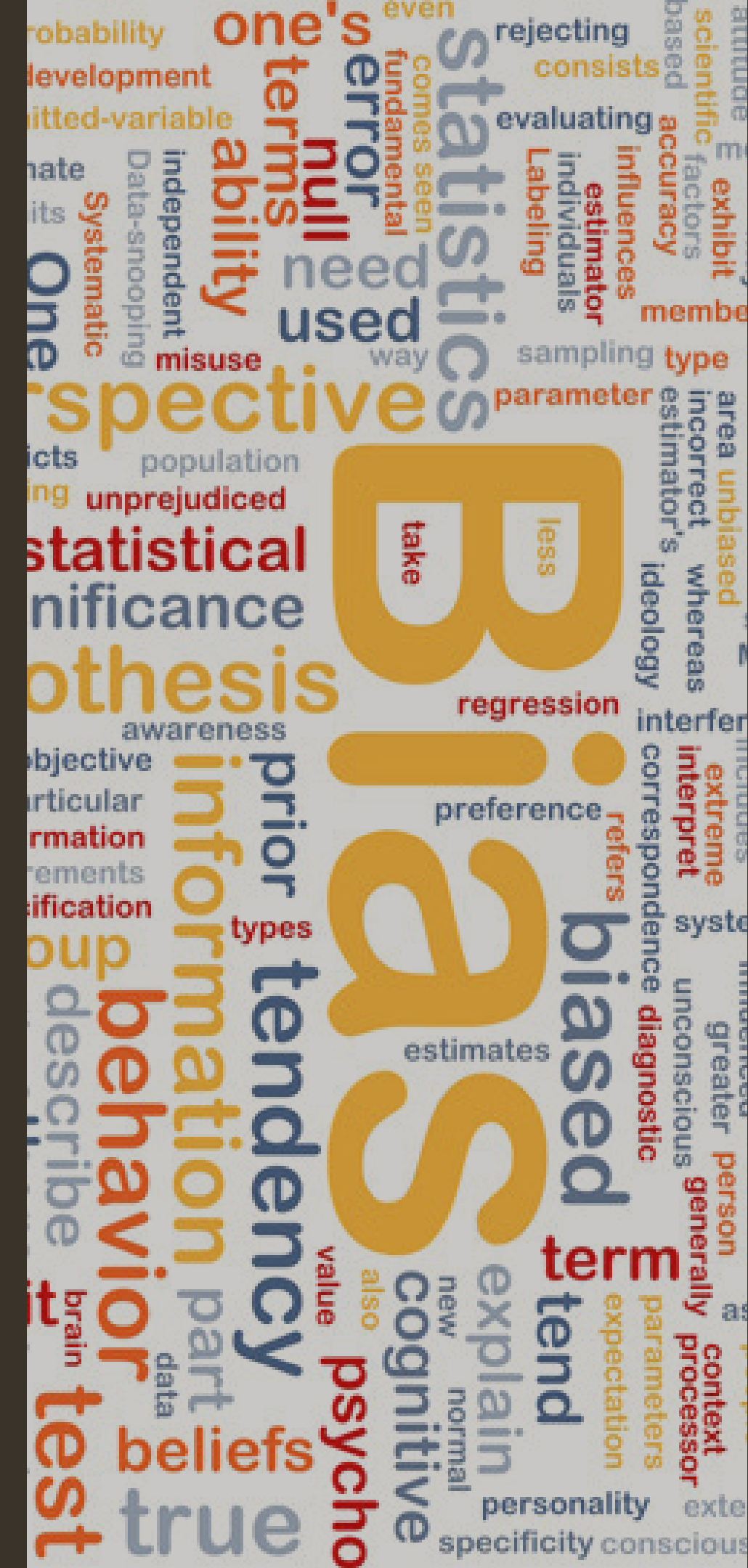
Subject matter bias

Personal bias

Reviewer's ability for constructive criticism

Halo

Horn



OBJECTIVE

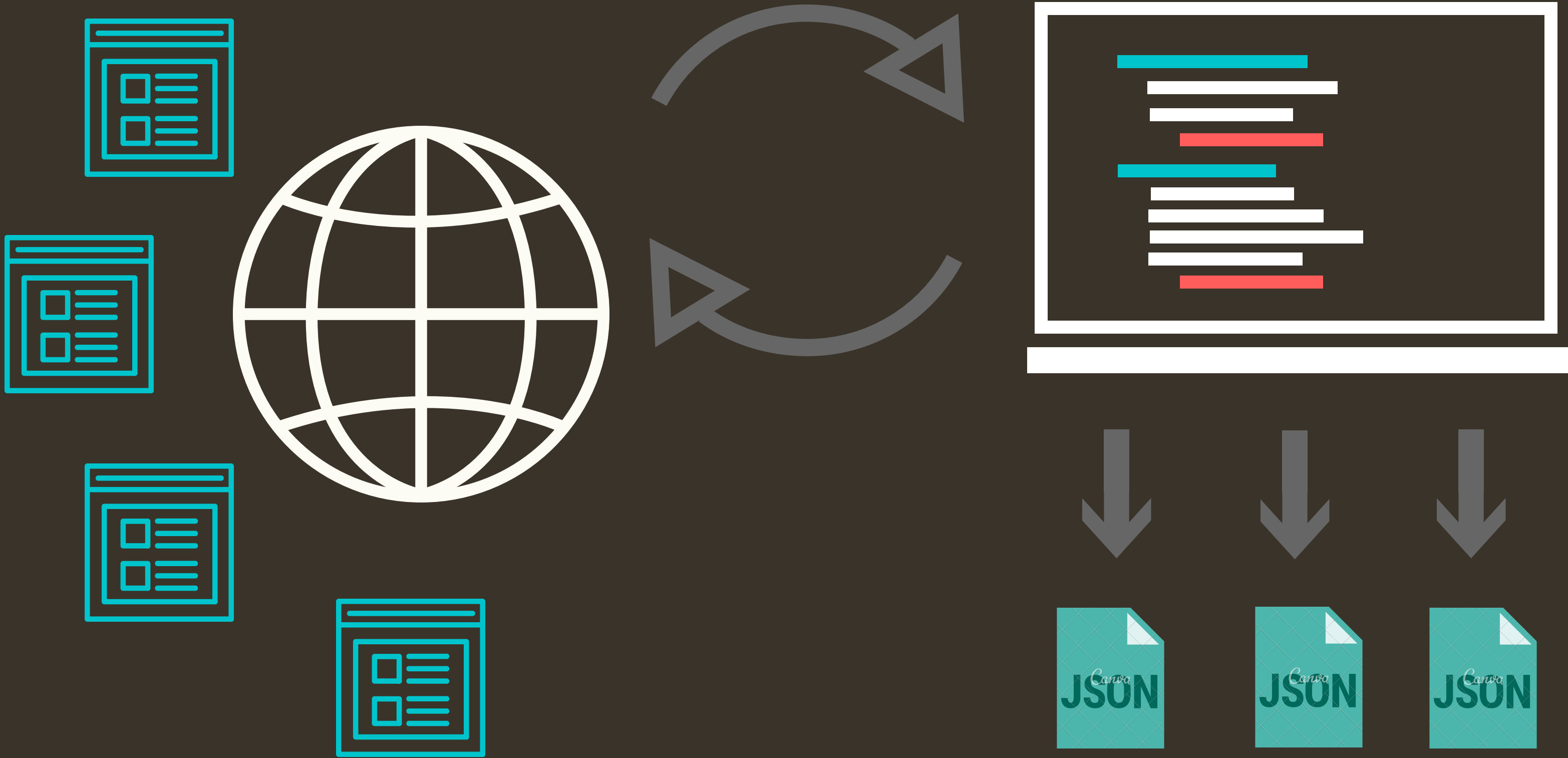
A scoring system based on the reviews that is focusd solely on professors' teaching skills.

HOW?

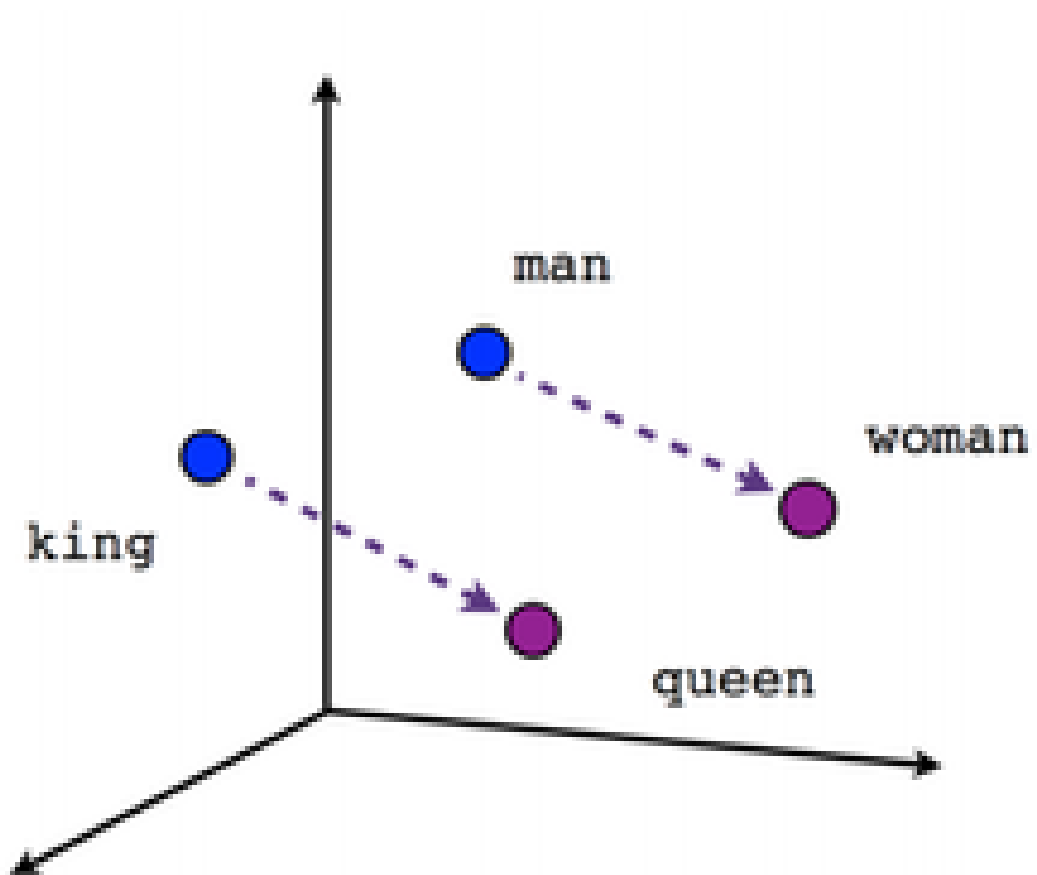


STEPS

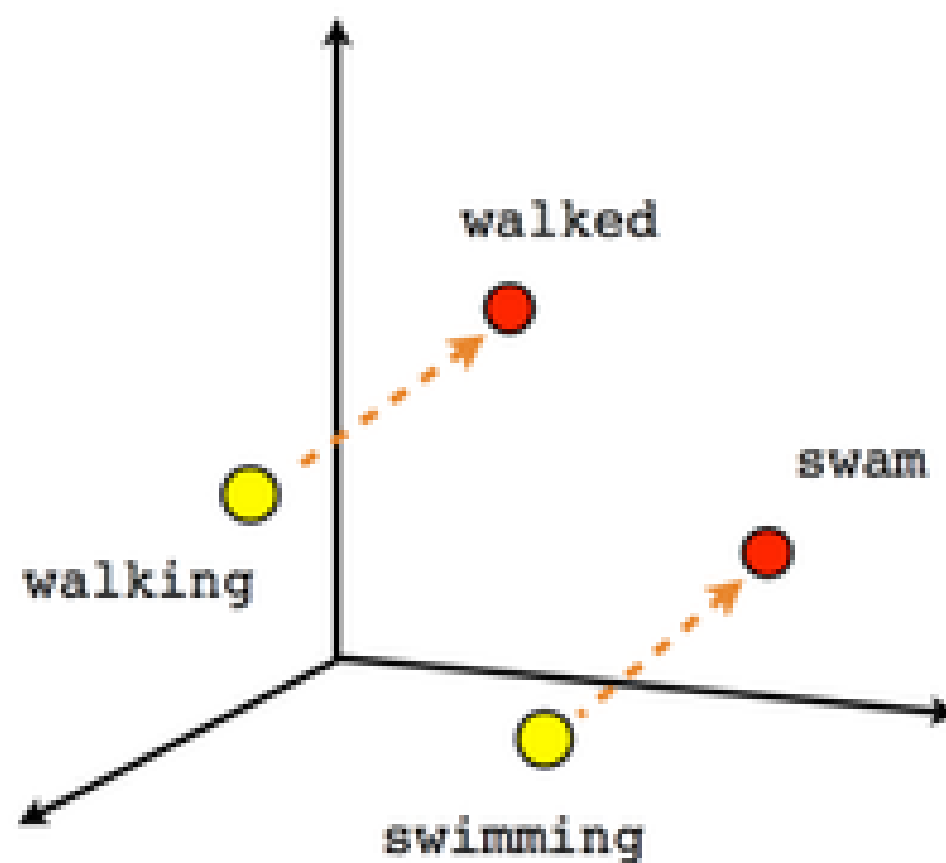
GATHERING DATA



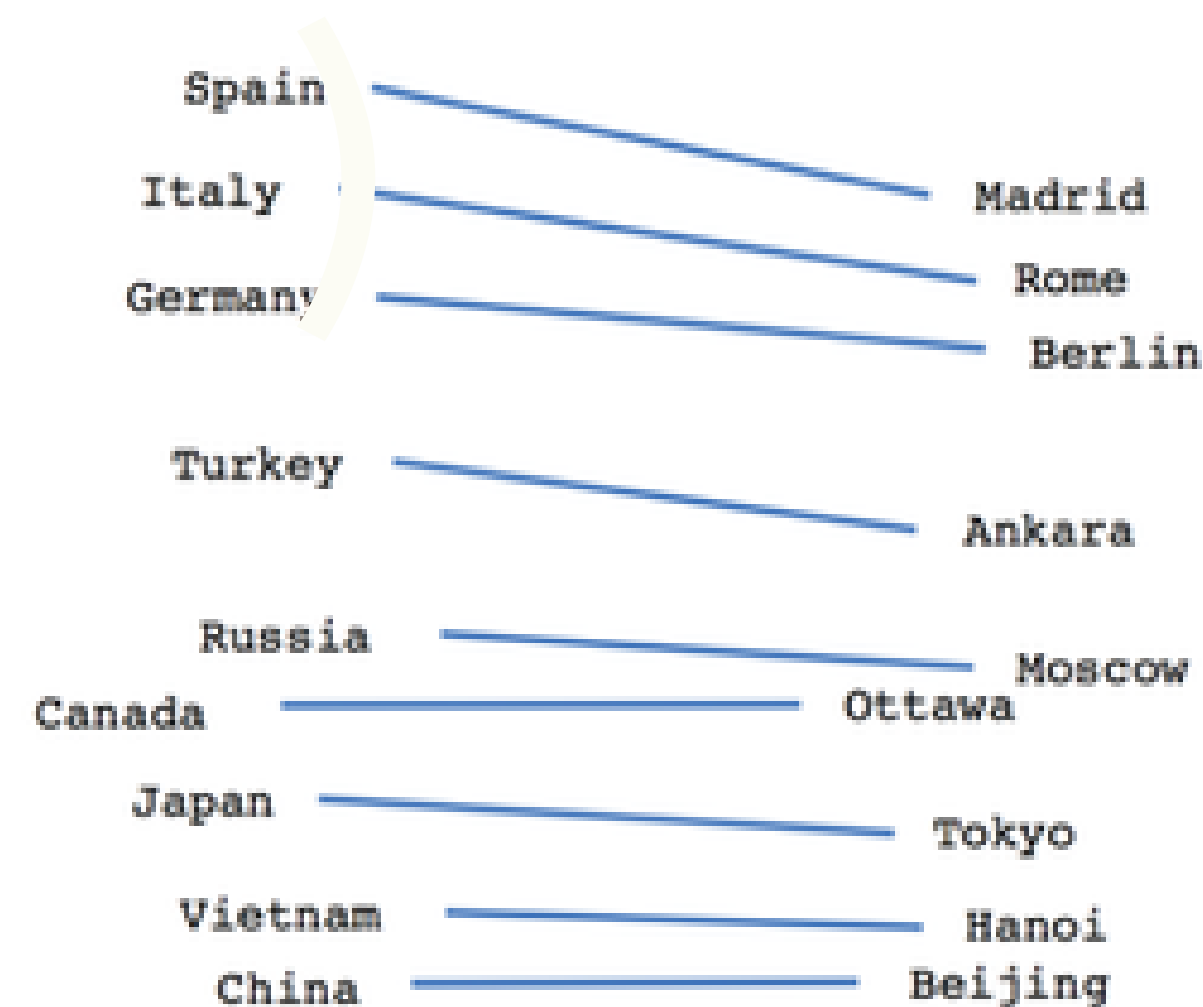
WORD2VEC



Male-Female



Verb tense



Country-Capital

PREPROCESSING

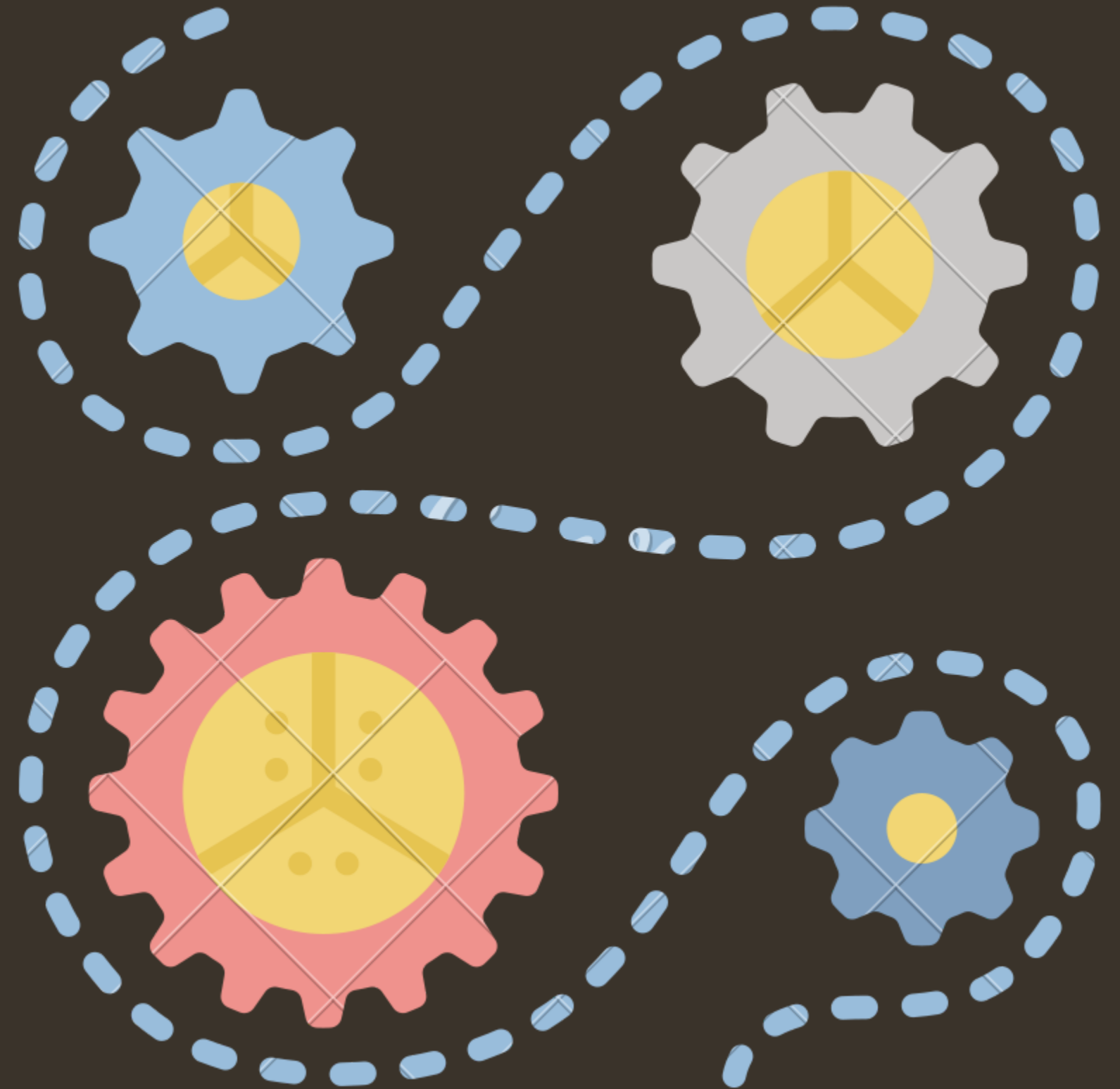
Tokenize

Lemmatize

phrase modeling

Lemmatizing and phrase modeling work best before feeding the data into word2vec model. Removing stopwords actually affect the performance of the model negatively.

Phrase modeling requires ignoring stop words and some additional keywords.



WORDS REPRESENTING “TEACH”

In [67]:

```
model.wv.most_similar('learn')
```

Out[67]:

```
[('teach', 0.9923907518386841),  
( 'expect', 0.985884428024292),  
( 'understand', 0.9851656556129456),  
( 'much', 0.9845415353775024),  
( 'want', 0.9839208126068115),  
( 'anything', 0.9837123155593872),  
( 'how', 0.9827748537063599),  
( 'talk', 0.9820513725280762),  
( 'well', 0.9812719225883484),  
( 'need', 0.9809737205505371)]
```

In [68]:

```
model.wv.most_similar('teach')
```

Out[68]:

```
[('much', 0.9935899972915649),  
( 'expect', 0.9931437373161316),  
( 'understand', 0.9925994873046875),  
( 'learn', 0.9923907518386841),  
( 'how', 0.9915162920951843),  
( 'can', 0.9911298751831055),  
( 'anything', 0.9898964762687683),  
( 'well', 0.9897662401199341),  
( 'like', 0.987104594707489),  
( 'tell', 0.9854578375816345)]
```

11/12/2016



AVERAGE

3.0

OVERALL QUALITY

4.0

LEVEL OF DIFFICULTY

STAT231

For Credit: Yes

Attendance: Not Mandatory

Textbook Used: Yes

Would Take Again: Yes

Grade Received: N/A

Enjoys her work, makes the tutorials fun. Approachable, and understandable.

👍 1 person found this useful

👎 6 people did not find this useful

[report this rating](#)

11/04/2015



AWFUL

1.0

OVERALL QUALITY

5.0

LEVEL OF DIFFICULTY

STAT230

For Credit: Yes

Attendance: Mandatory

Textbook Used: No

Would Take Again: N/A

Grade Received: A+

TOUGH GRADER

terrible old whatever



11 people found this useful



2 people did not find this useful

[report this rating](#)

04/05/2010



GOOD

4.0

OVERALL QUALITY

4.0

LEVEL OF DIFFICULTY

MATH249

For Credit: N/A

Attendance: N/A

Textbook Used: Yes

Would Take Again: N/A

Grade Received: N/A

A nice person and expects a lot from his students. The closest mathematician in mannerism to G.H. Hardy in Waterloo. The midterm was very fair, but the assignments were challenging. Expect to spend a lot of time on them, and learn a lot if you start early.



0 people found this useful



0 people did not find this useful

[report this rating](#)

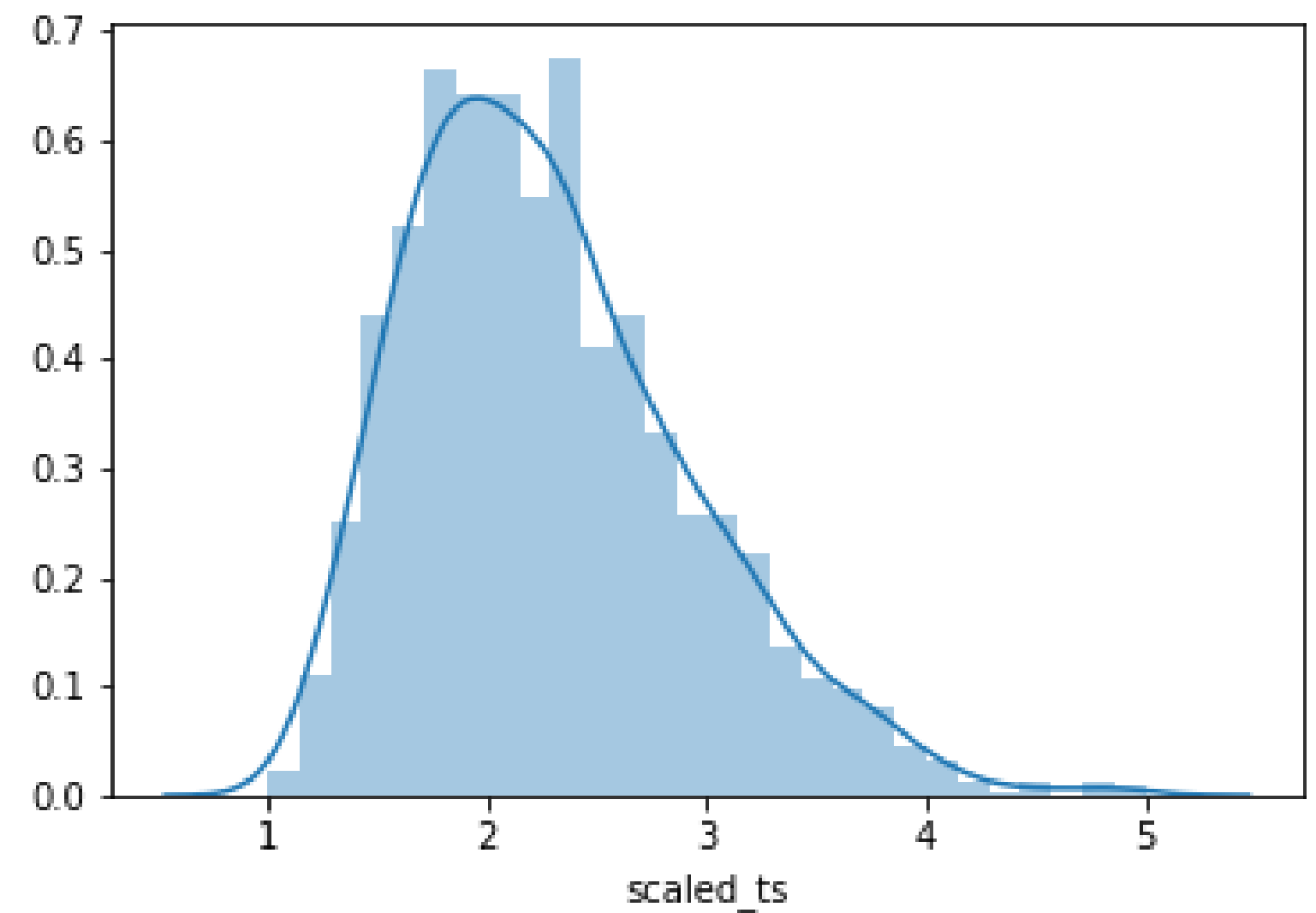
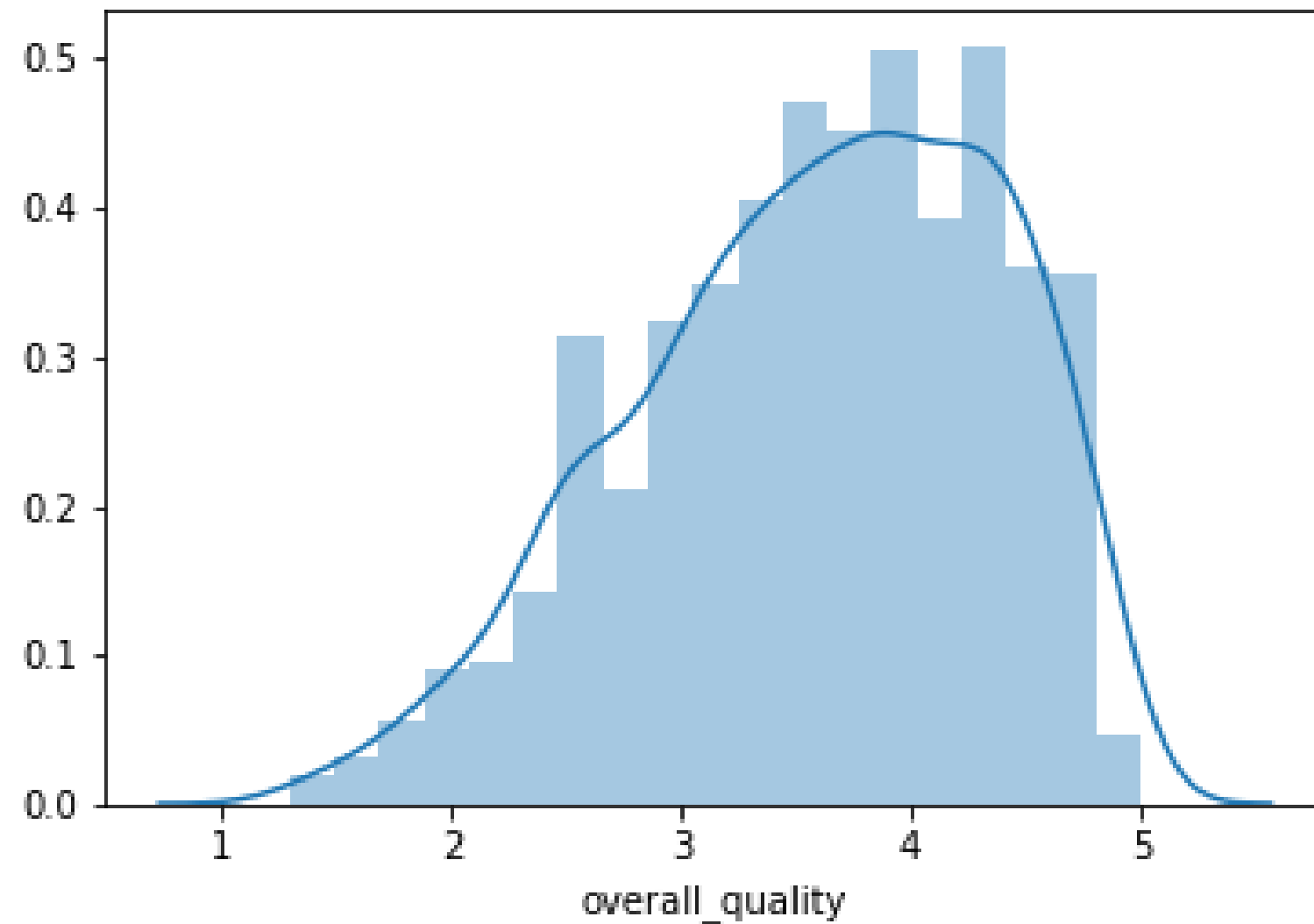


$$TS(P) = \frac{\sum (TAS(r) \times S(r))}{N(r)}$$



RESULTS

SUBJECTIVE

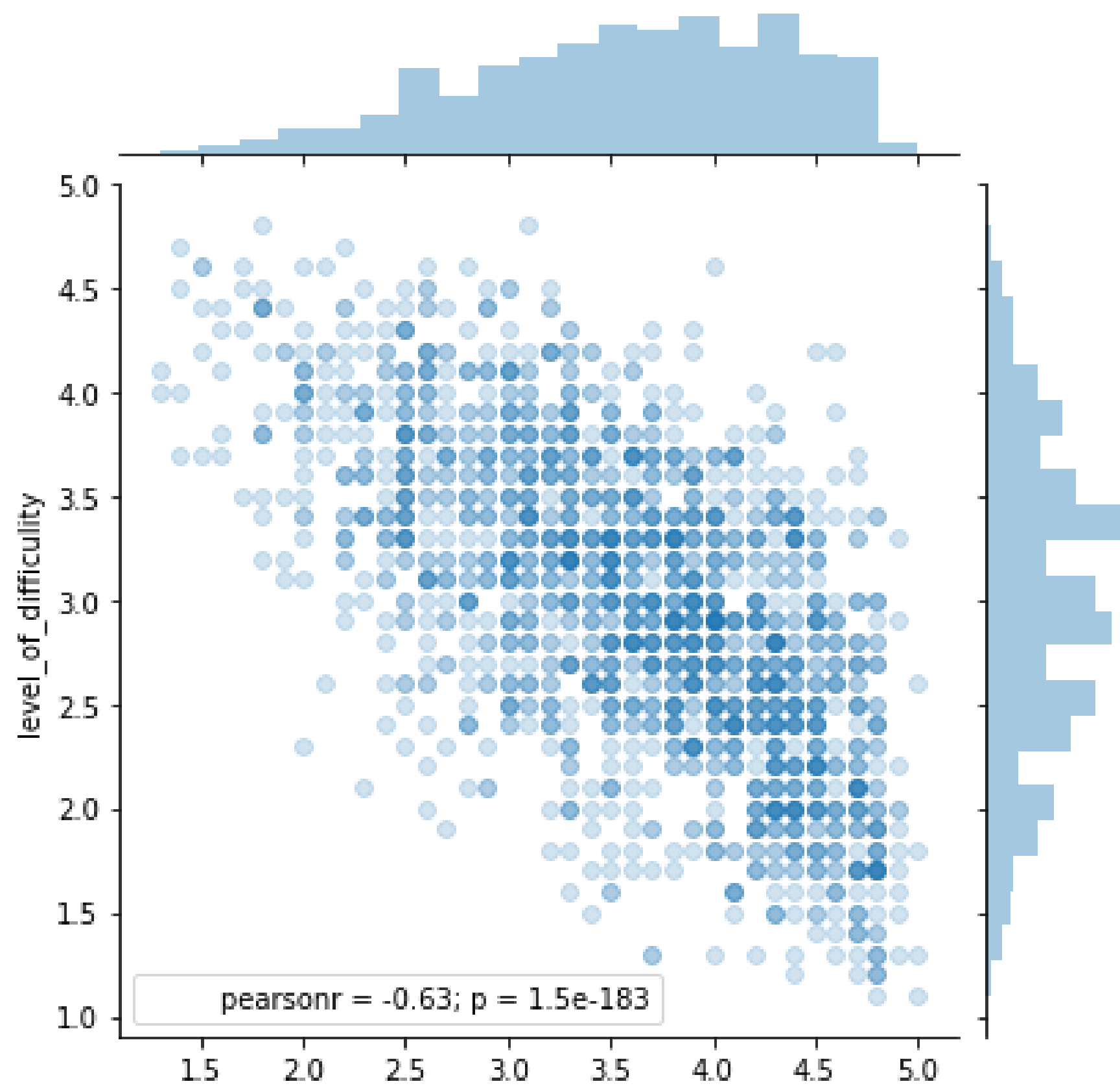


OBJECTIVE

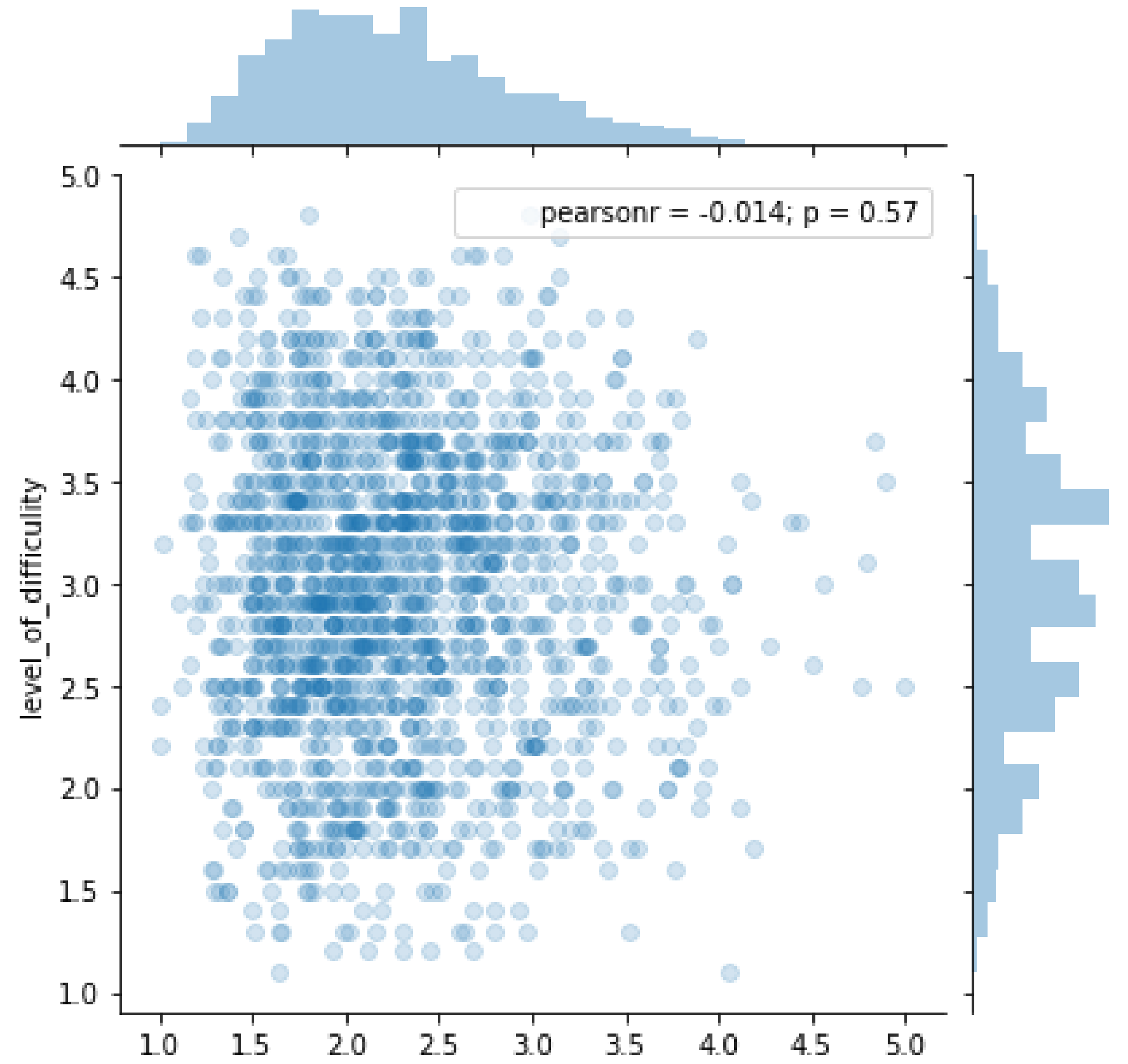




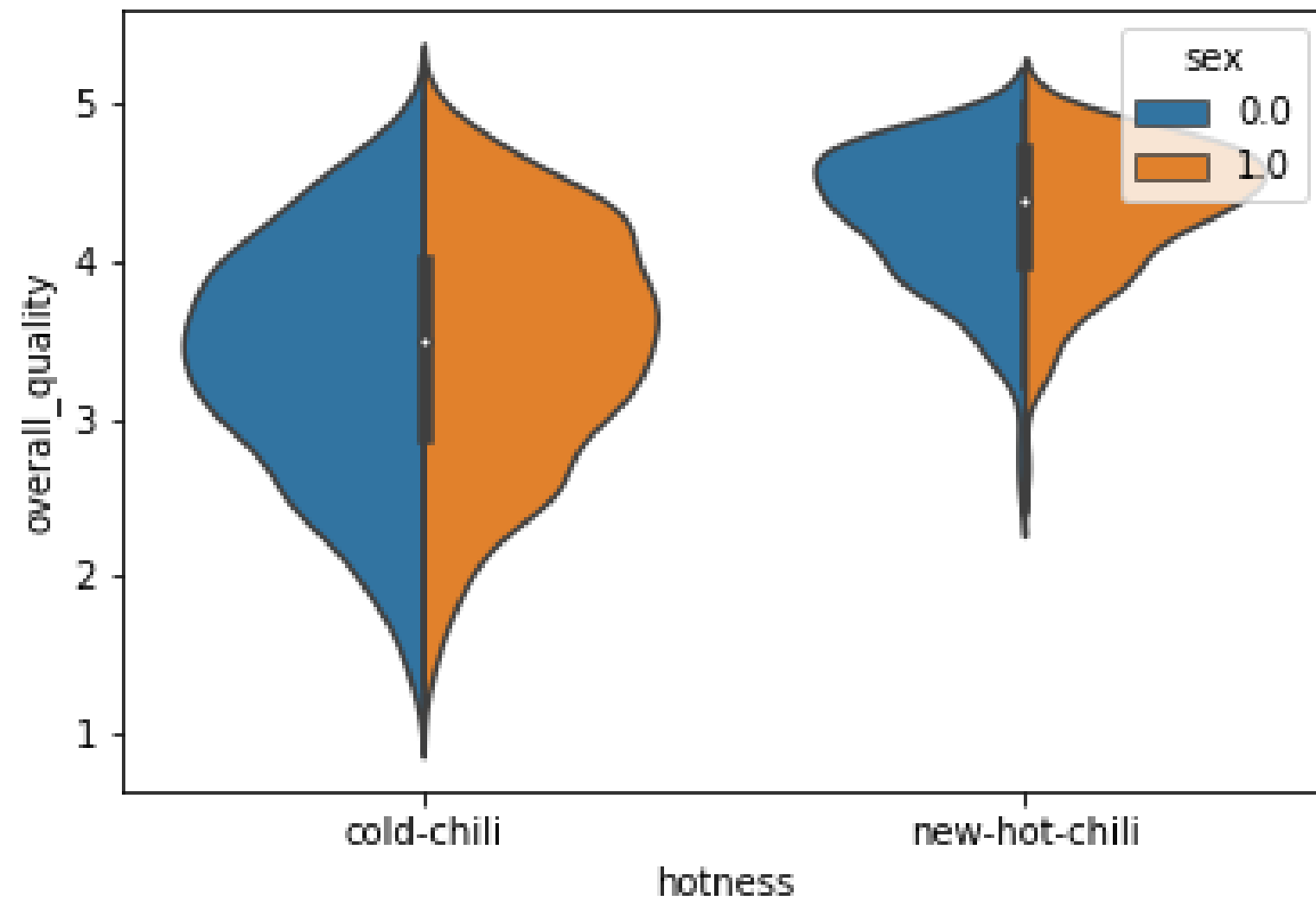
SUBJECTIVE



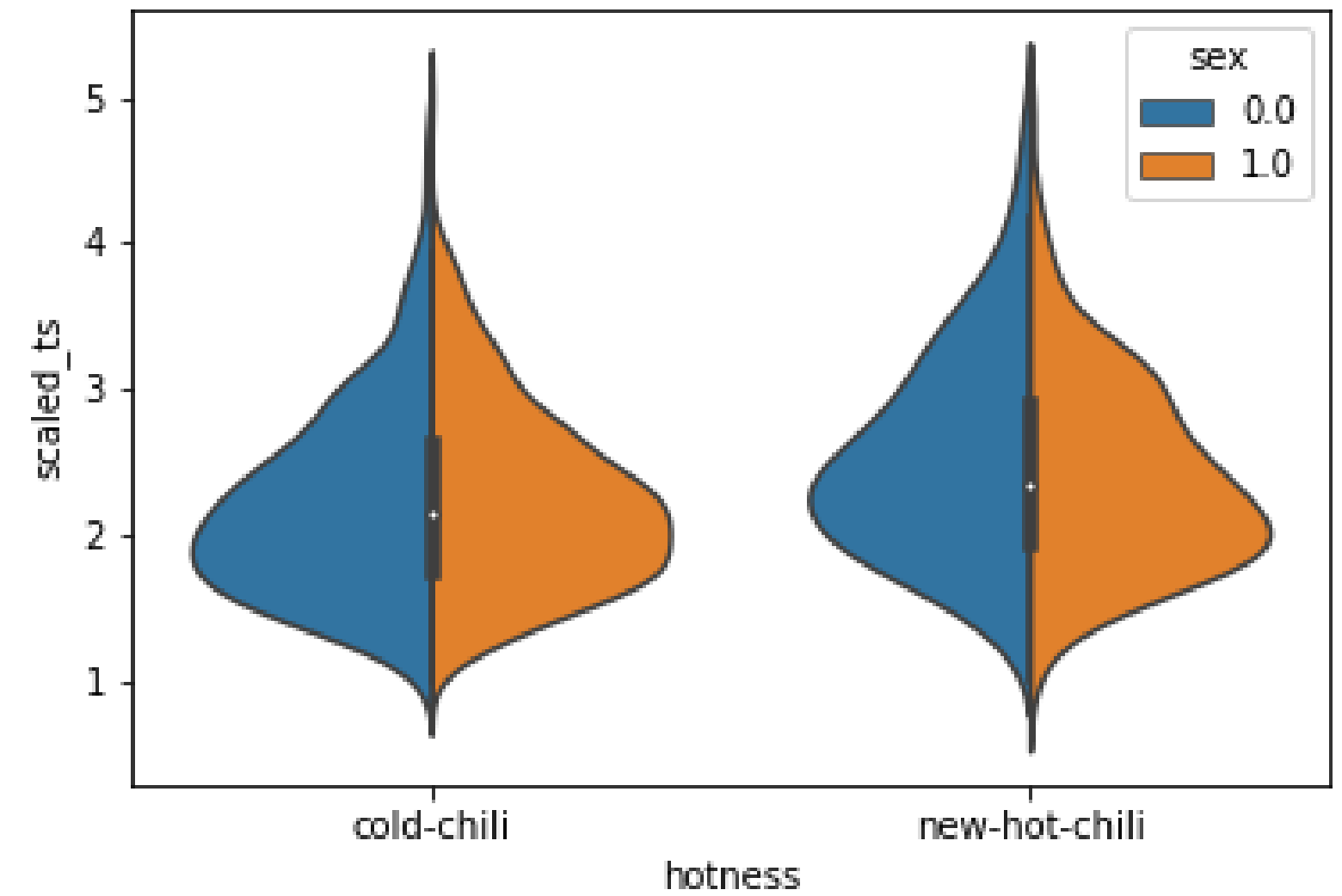
OBJECTIVE



SUBJECTIVE



OBJECTIVE



SECOND OBJECTIVE

What are the characteristics associated with high-value professors?

WHAT MAKES UP A HIGH- VALUE- PROF?

HIGH VALUE PROFESSORS

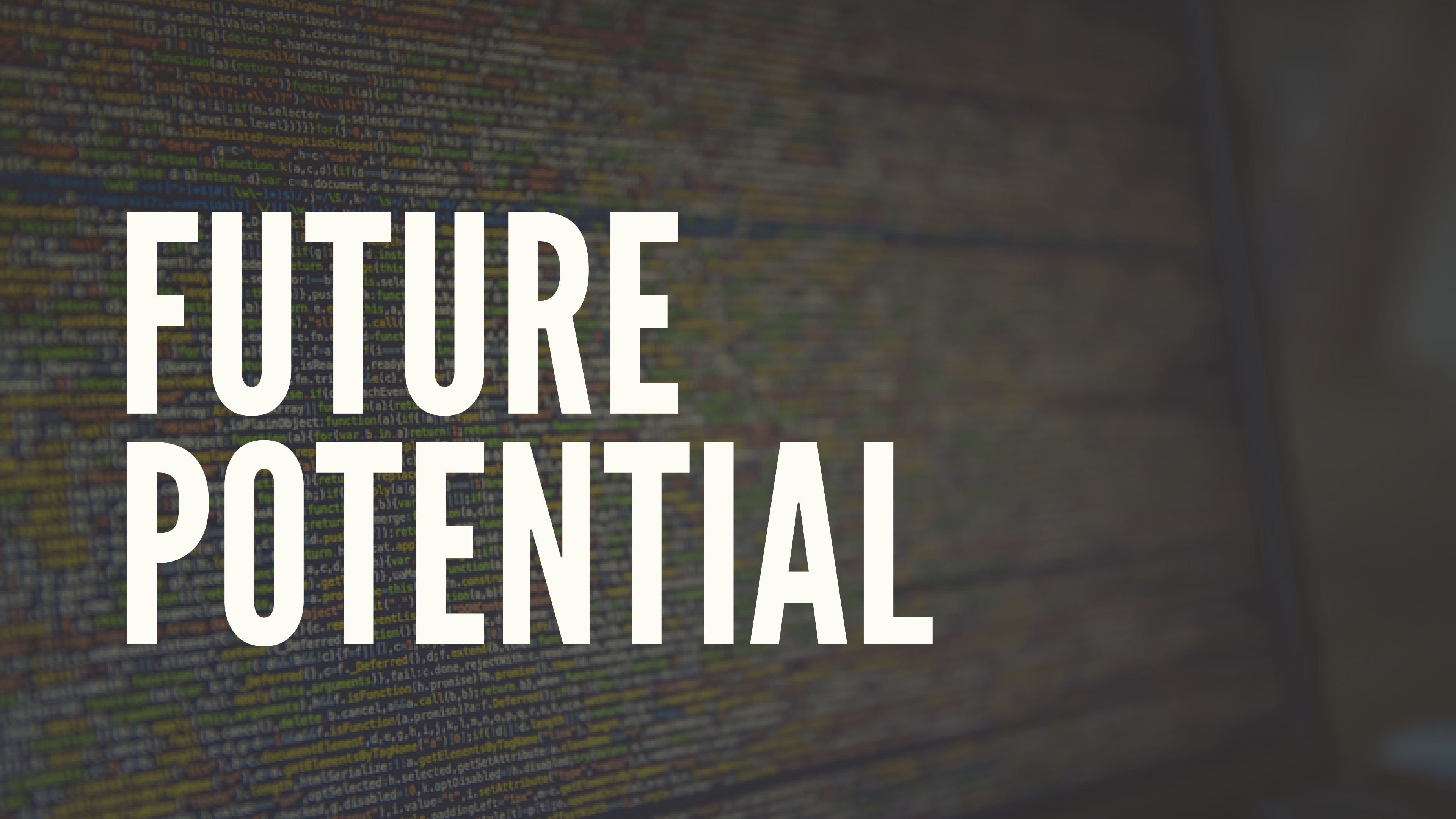
good 281.2291108155739
great 269.99325015107365
easy 268.21085424241386
hard 255.76953876222788
that 252.77575936018366
helpful 243.94068342390324
nice 239.4876566845607
which 236.98214285714286
sure 233.38078643578646
bad 230.79081196581197
other 226.06792929292934
difficult 225.49121295371293
all 221.78015873015872
clear 220.6218364968365
much 218.48780663780664
final 215.64200937950943
more 215.39166666666665
willing 213.35457393483708
little 212.79126984126984
awesome 209.2418831168831

good 342.26197963382987
easy 339.53680437652554
great 315.442085116466
hard 314.3917287623827
that 309.0327801425626
bad 297.9803429626639
which 293.53984626484623
nice 291.05268356125737
boring 288.14585548052855
helpful 284.95764790764787
interesting 270.67433602392765
all 262.68968253968256
other 260.50318570318575
little 254.13412698412696
much 249.84285714285716
sure 238.4742424242424
clear 237.75357142857138
more 237.70277777777778
final 236.1863231213231
difficult 235.63730158730155

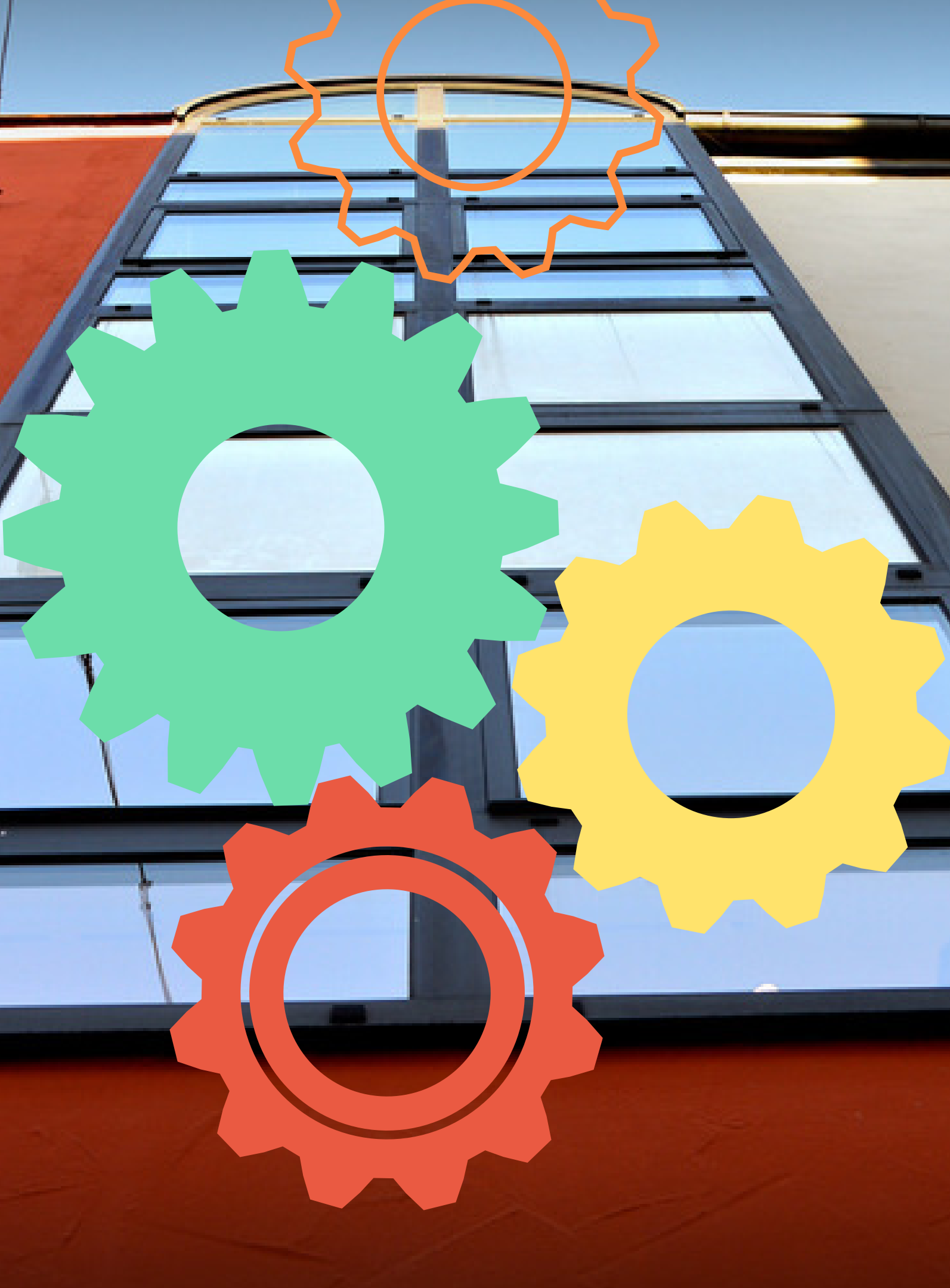
TAGS MAKES UP A HIGH- VALUE- PROF

HIGH VALUE PROFESSORS

tough grader	3.223427	tough grader	3.141844
skip class? you won't pass.	2.998856	get ready to read	2.888889
get ready to read	2.983139	lecture heavy	2.700935
respected	2.783544	hilarious	2.352941
hilarious	2.698630	skip class? you won't pass.	2.342593
lots of homework	2.673300	gives good feedback	2.212121
caring	2.618598	lots of homework	2.200000
lecture heavy	2.581267	participation matters	2.136364
amazing lectures	2.563607	respected	2.010753
gives good feedback	2.498199	group projects	2.000000
clear grading criteria	2.495105	extra credit	1.913043
extra credit	2.443299	clear grading criteria	1.885057
participation matters	2.425414	caring	1.880952
inspirational	2.257851	amazing lectures	1.877193
test heavy	2.174731	so many papers	1.833333
beware of pop quizzes	2.073298	test heavy	1.794872
group projects	2.000000	beware of pop quizzes	1.772727
graded by few things	1.851485	graded by few things	1.719298
accessible outside class	1.756691	inspirational	1.548387
so many papers	1.735632	accessible outside class	1.333333



FUTURE POTENTIAL



REAL WORLD APPLICATIONS

HR

INSTITUTIONAL

INDIVIDUAL

TECH STACK

NumPy

spaCy



Spark



seaborn

gensim

