

THANK YOU DAN, BRENT, JOE, AND MY FELLOWS

PREDICTING ENDORSEMENT OF
USER-GENERATED CONTENT

GO FOR GOLD



"EARLY ENDORSEMENT" OF USER-GENERATED CONTENT:

- * PERCEIVED AS AUTHENTIC,
WHEREAS LATER RESPONSES
BECOME MORE LIKELY TO INVOKE
THE SUSPICION OF
BANDWAGONISM

- * A BRAND "SPENDING"
ATTENTION ON SOMETHING THAT
DOES NOT HAVE AN IMMEDIATE
OBVIOUS PAYBACK IN EXPOSURE

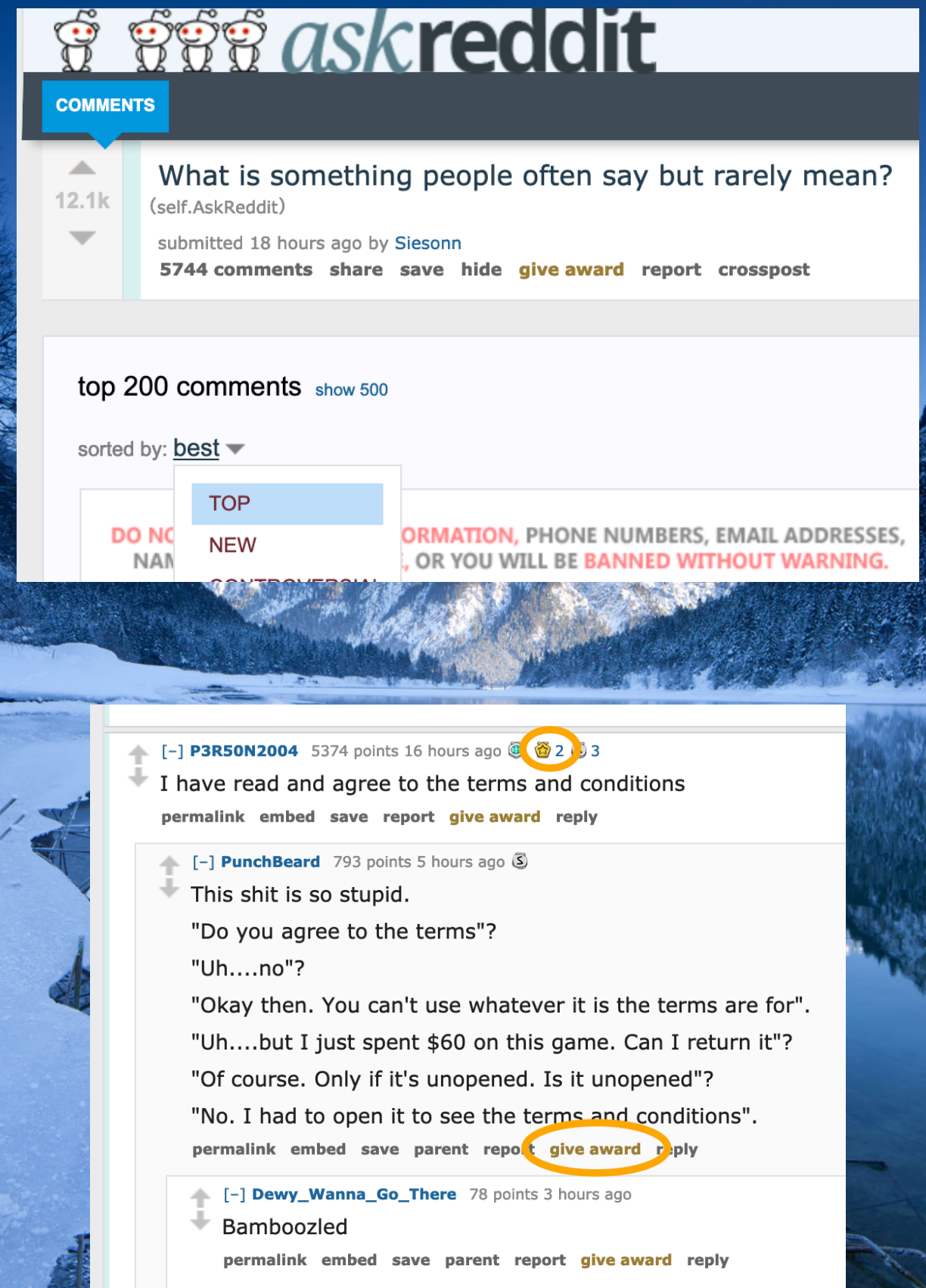
- * CORRELATES WITH,
SYNERGISTIC WITH LATENT VIRAL
POTENTIAL



#MeToo wasn't originally called a campaign.
Early joiners got praise, later ones got
suspicion.

TEXT CLASSIFIER THAT PREDICTS EARLY ENDORSEMENT

- * SUBFORUM "ASKREDDIT" POSES QUESTIONS, EXPECTS MEANINGFUL COMMENTS
- * PAY \$10 TO GIVE A COMMENT "GOLD" - STAMPS IT FOR ALL TO SEE
- * GOLD IS ANONYMOUS, SO A GOOD MEASURE OF ACTUAL POTENTIAL

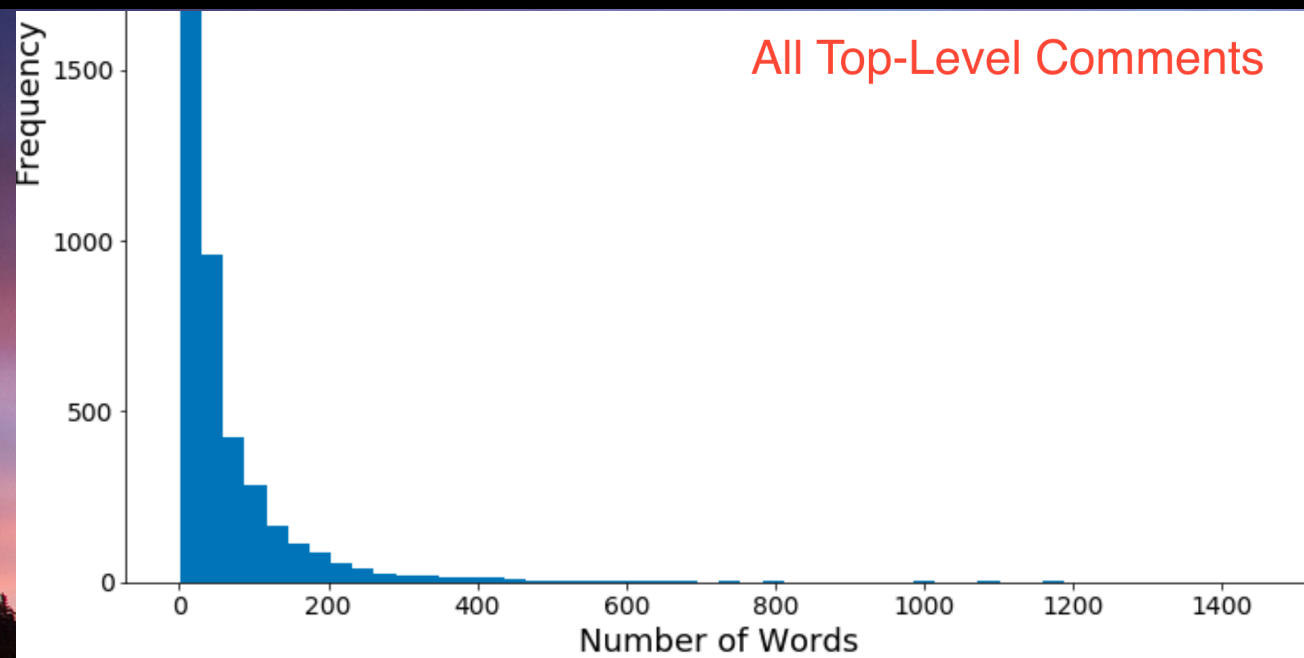


The screenshot shows the AskReddit subreddit interface. At the top, the "askreddit" header is visible with four Reddit avatars. Below it, a "COMMENTS" tab is selected. The main post is titled "What is something people often say but rarely mean?" by user "self.AskReddit", with 12.1k upvotes and 5744 comments. The post is sorted by "best" comments. A dropdown menu is open, showing options: "TOP", "NEW", and "CONTROVERSIAL". A warning banner is visible: "DO NOT POST PERSONAL INFORMATION, PHONE NUMBERS, EMAIL ADDRESSES, OR YOU WILL BE BANNED WITHOUT WARNING." Below the post, the "top 200 comments" section is shown. The first comment is by user "P3R50N2004" (5374 points, 16 hours ago) with the text "I have read and agree to the terms and conditions". This comment is highlighted with a yellow circle around the "give award" button. The second comment is by user "PunchBeard" (793 points, 5 hours ago) with the text "This shit is so stupid. 'Do you agree to the terms'? 'Uh....no'? 'Okay then. You can't use whatever it is the terms are for'. 'Uh....but I just spent \$60 on this game. Can I return it'? 'Of course. Only if it's unopened. Is it unopened'? 'No. I had to open it to see the terms and conditions'." This comment is also highlighted with a yellow circle around the "give award" button. The third comment is by user "Dewy_Wanna_Go_There" (78 points, 3 hours ago) with the text "Bamboozled".

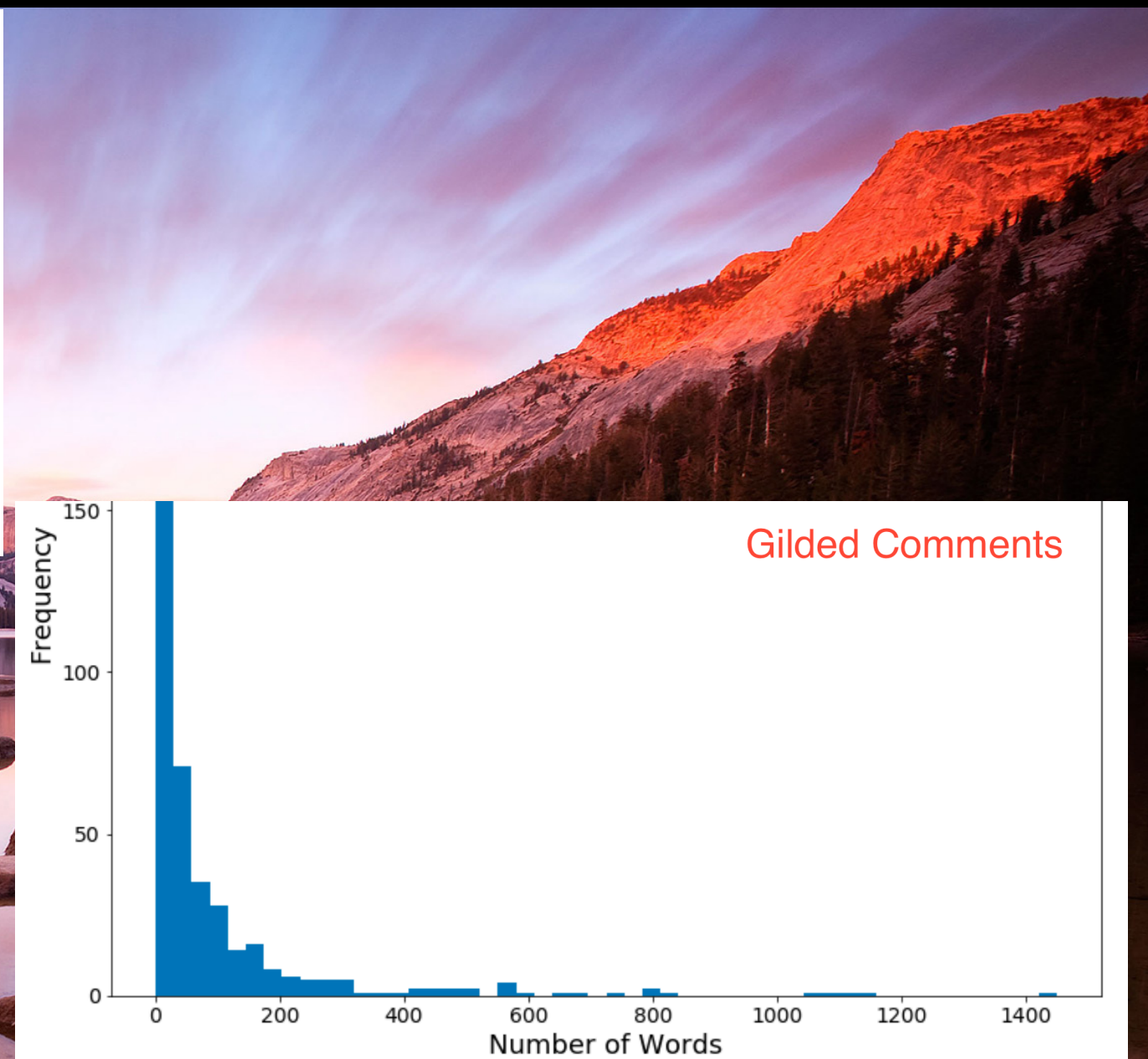
Twice "gilded" top-level comment

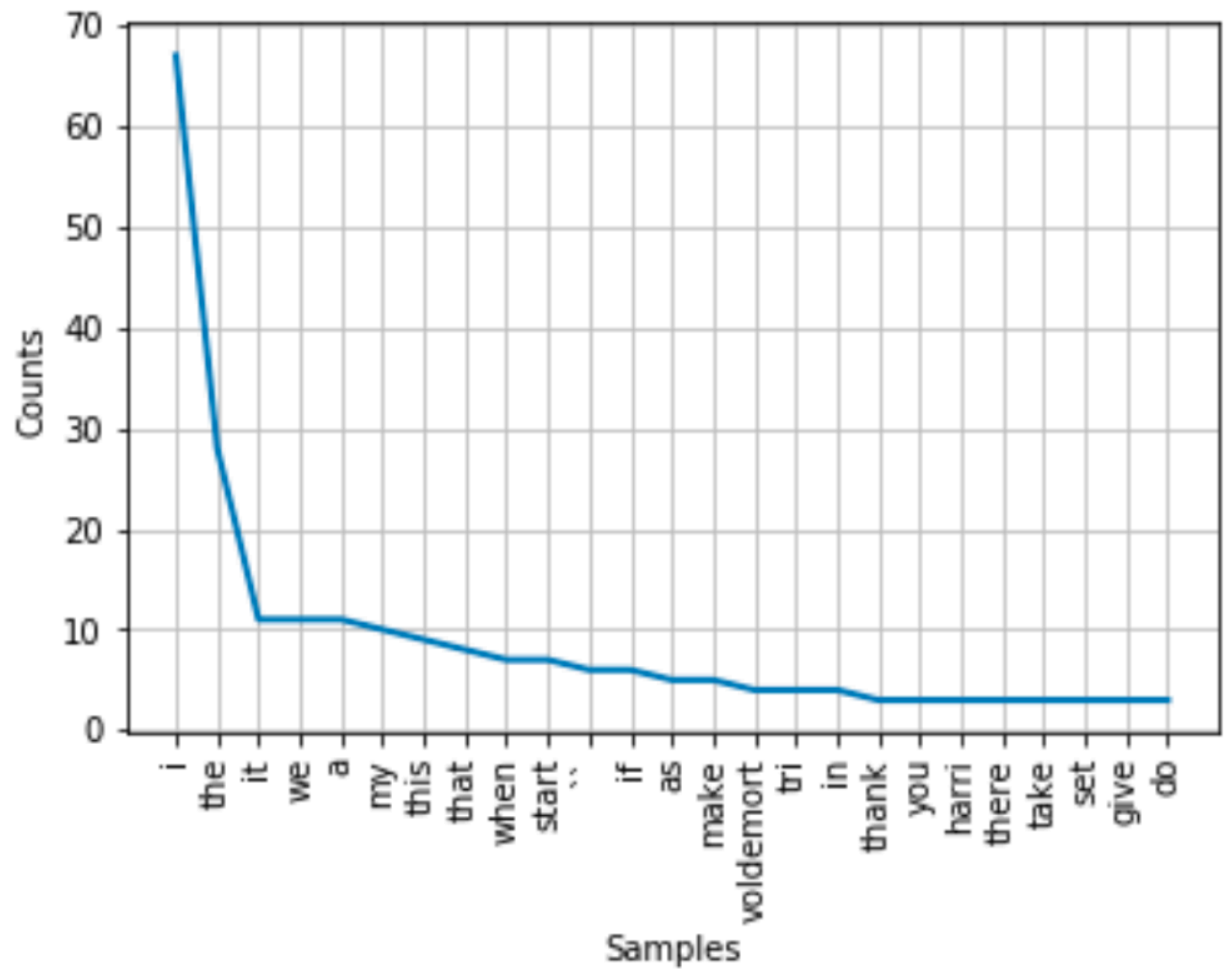
“Three carefully stringed words are worth more than a book of gibberish. It's not the word count but the impact of those words that counts.”

—RICHELLE E. GOODRICH

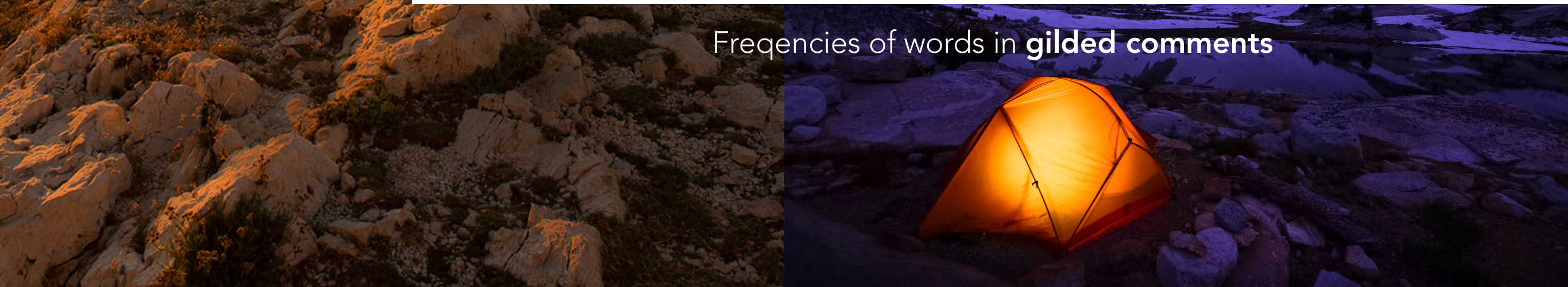


T-Test gave $pvalue = 7.862637227785603e-09$





Frequencies of words in **gilded comments**



		Predicted	
		Negative	Positive
Actual	Negative	True Negative	False Positive
	Positive	False Negative	True Positive

True Positive + False Positive = Total Predicted Positive

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}$$

$$= \frac{\text{True Positive}}{\text{Total Predicted Positive}}$$

Performance Metrics: Precision, Recall, and most practically, FP:TP ratio

```

                precision    recall  f1-score   support

     0       0.910387    0.983498    0.945531       909
     1       0.166667    0.032967    0.055046        91

   accuracy          0.897000       1000
  macro avg       0.538527    0.508233    0.500289       1000
 weighted avg       0.842708    0.897000    0.864497       1000

```

```

Classes predicted:  [0 1]
Accuracy:    0.897
ROC AUC Score:  0.6487324556631487
True Pos:  3 False Pos:  15
FP to TP Ratio::  5.0
True Neg:  894 False Neg:  88

```