Data tech task

Limitation on data –

1. The data does not show which products the users are upvoting. The post table only has poster user id
2. We have data for only a two-month period in 2014. There could be a bias when comparing tagline trends across different years.

Q1. Who are the real “spotters” (people that continually upvote products that end up to be very popular) on Product Hunt?

The table and graph display the top 10 users with the highest number of votes, who I consider as the real spotters. However, we cannot discern the connection between the popularity of a product and actual 'spotters' because we lack visibility into which products the spotters are upvoting.

A graph of blue bars

Description automatically generated with medium confidence

|  |  |  |  |
| --- | --- | --- | --- |
| user\_id | **name** | **username** | **votes\_count** |
| 115829 | Jonathan Levinson | jonathanlevins1 | 15682 |
| 345043 | Eduard Ionuţ Ichim | eeschimosu | 11092 |
| 48793 | Jamie Davidoff | davidoff36ch2o | 9731 |
| 2 | Ryan Hoover | rrhoover | 9140 |
| 76196 | Jack Shalom | jeshalom | 8886 |
| 151665 | Kate | katesegrin | 8642 |
| 144770 | Zac Nielson | zacherynielson | 8257 |
| 85429 | Zeeshan Mughal | \_zeeshan\_mughal | 7965 |
| 57468 | Kendall Holliday | hi\_kendall | 7915 |
| 946 | Niv Dror | nivo0o0 | 7577 |

Q2. How many products does the average user upvote? In which categories? How often?

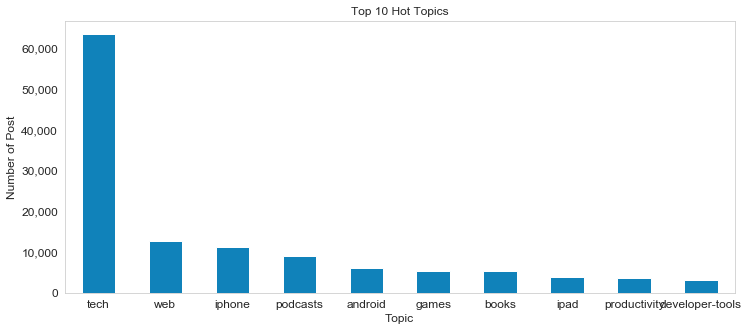
On average, users upvoted approximately 7 products. However, due to data limitations, we do not have visibility into which product categories users were upvoting or the frequency of their upvotes.

Additionally, it's important to exercise caution when solely considering the average number of votes. The median number of votes was just 1, and the majority of customers had either 0 or 1 vote, accounting for 78% of the total.

Q3. What are some commonalities between the products or the PH upvotes behavior?

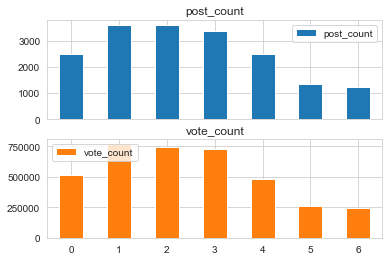
The most popular topic was ‘tech’, with 63,552 posts, nearly five times more than the second most popular topic, 'web.'

|  |  |
| --- | --- |
| topic name | **num\_post** |
| tech | 63552 |
| web | 12724 |
| iphone | 11276 |
| podcasts | 9043 |
| android | 6008 |
| games | 5389 |
| books | 5379 |
| ipad | 3802 |
| productivity | 3696 |
| developer-tools | 3234 |



Analyzing the number of products and votes based on the day and time of day reveals that most products were generated during the nighttime and on weekdays, especially Tuesday, Wednesday, and Thursday. These products also tend to receive a higher number of votes.

|  |  |  |
| --- | --- | --- |
| **day** | **post\_count** | **vote\_count** |
| Tuesday | 3596 | 772571 |
| Wednesday | 3591 | 746637 |
| Thursday | 3357 | 729216 |
| Monday | 2507 | 514791 |
| Friday | 2497 | 485656 |
| Saturday | 1335 | 261788 |
| Sunday | 1246 | 241387 |



|  |  |  |
| --- | --- | --- |
| **day** | **post\_count** | **vote\_count** |
| Night | 9250 | 2235455 |
| Morning | 8028 | 1437399 |
| Afternoon | 724 | 71172 |
| Evening | 127 | 8020 |

In addition, about 98.6% of products were created with image, and they tend to have more votes.

|  |  |  |
| --- | --- | --- |
| **thumbnail\_type** | **post\_count** | **vote\_count** |
| image | 17878 | 3701655 |
| video | 239 | 48421 |
| audio | 11 | 1911 |
| book\_preview | 1 | 59 |

Q4. Is there a core group of users that count for most (80%?) of all upvotes on all products, or on the top products?

The top 4% of users generated 80% of votes. I am not able to see if those votes were on the top products due to data limitation.

Q5. Do products in collections get more upvotes on average?

We do not have collection in post/product table.

However, I found collection in user table, so I tried my best to answer this question. Based on the user table, there is a correlation of approximately 0.226 between 'collection' and 'votes\_count,' suggesting a relatively weak relationship between these two dimensions. There is no clear evidence to suggest that users with a higher number of collections do upvote more.

Q6. What are 2016's tagline trends vs. 2015 and 2014 (e.g., are we seeing less "uber for" and more "AI for")

Taglines related to electronic products and operating systems like iPhone, Mac, and iOS consistently trend in popularity.

In 2016, several taglines linked to messaging, including chat, messenger, and iMessage, showing a significantly increase compared to previous years. Slack also emerged as a popular tagline in 2016.

Conversely, taglines related to gaming were less prevalent in 2016. In addition, "Twitter" was only a prominent tagline in 2015.

Q7. Do animated GIF thumbnails provide an advantage?

There is no type called ‘GIF’ in thumbnail dimension. With all available types, ‘image’ took the advantage.

However, I found a topic called ‘gifs’, so I used it. The average vote counts across all products was 207, and the average vote counts of products on gifs topic was 246, indicating the gif topic might provide an advantage.

Q8. What is the relationship between comments and upvotes?

The correlation between comments and upvotes is about 0.736, indicating a relatively strong positive relationship between these 2 dimensions. The more comments, the more upvotes.

