

# do pineapples belong on pizza?

a.k.a. the most important question ever



# Flow of Presentation

- Big Data problem
- Source dataset and how you processed into the MongoDB
- Expected output

---

**Given tweets regarding pineapple on pizza, how many users believe that pineapple should or should not belong on pizza?**



**“I like pineapples  
on pizza”**

**“I don’t like  
pineapples on pizza”**  
**“I do not like  
pineapples on pizza”**



## Obtaining the data set

1. We used a twitterscraper found on github by taspinar. It can be found here:  
<https://github.com/taspinar/twitterscraper>
2. Next, go to twitter and search for the query that you wish to see. Copy paste a part of the url on the terminal and run the twitterscraper. The data will be saved in a json file.
3. Each like on a tweet is also be recorded to see how many people agree with the tweet.
4. Each retweet on a tweet is also be recorded to see how many people agree with the tweet
5. Import the JSON files to the database:  

```
mongoimport --db "databaseName"  
--collection "collectionName"  
--type json --file "fileName.json"
```



# Obtaining the data set

Each tweet recorded will have the following data:

1. Username and Full Name
2. Tweet-id
3. Tweet text
4. Tweet timestamp
5. No. of likes
6. No. of replies
7. No. of retweets



# Expected output

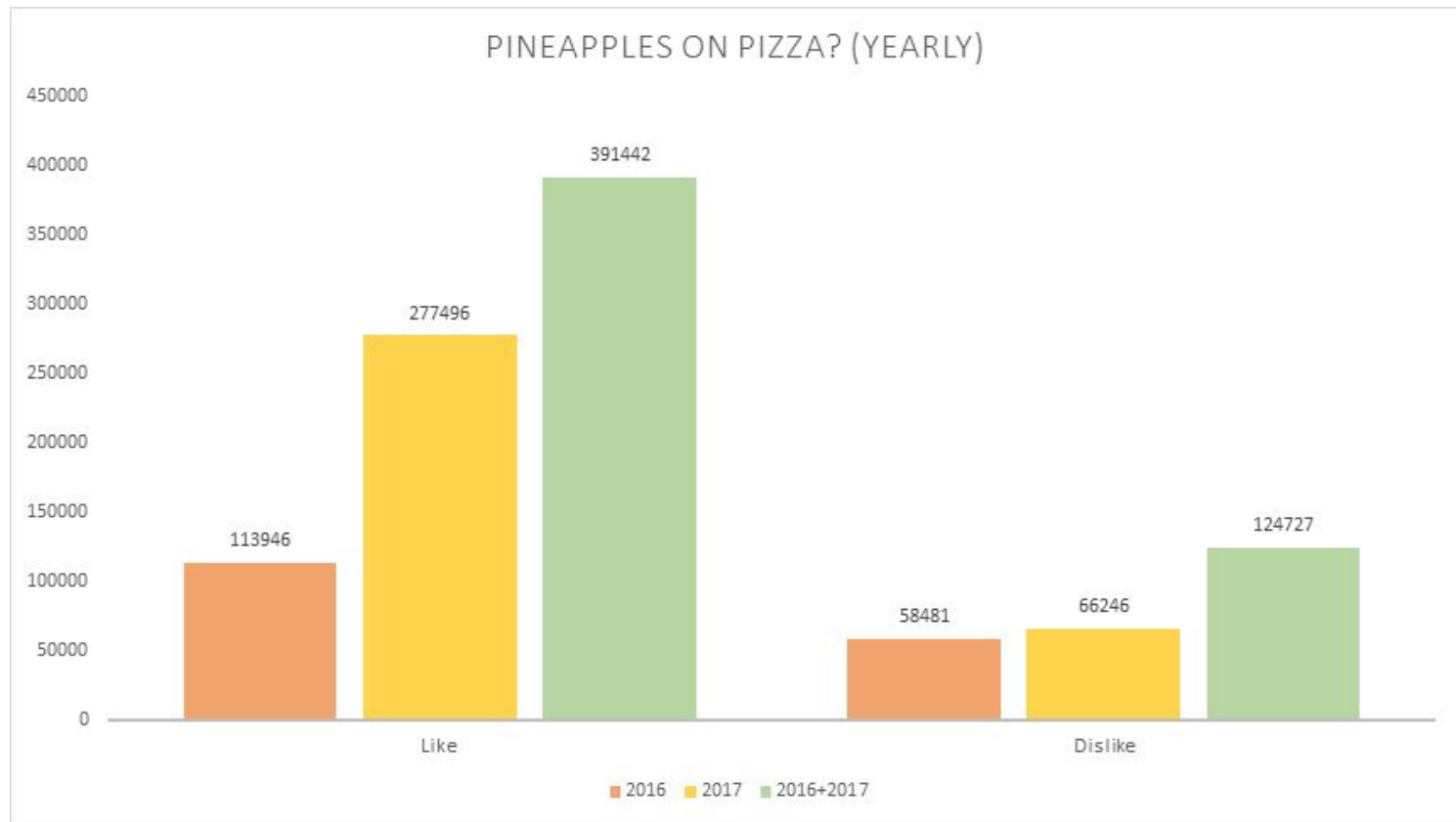
Reports of the following:

1. Number of likes (the tweet+the likes+the retweets) per year
2. Number of dislikes (the tweet+the likes+the retweets) per year
3. Number of likes (the tweet+the likes+the retweets) in the same month
4. Number of dislikes (the tweet+the likes+the retweets) in the same month
5. Number of likes (the tweet+the likes+the retweets) monthly from 2016-2017
6. Number of dislikes (the tweet+the likes+the retweets) monthly from 2016-2017

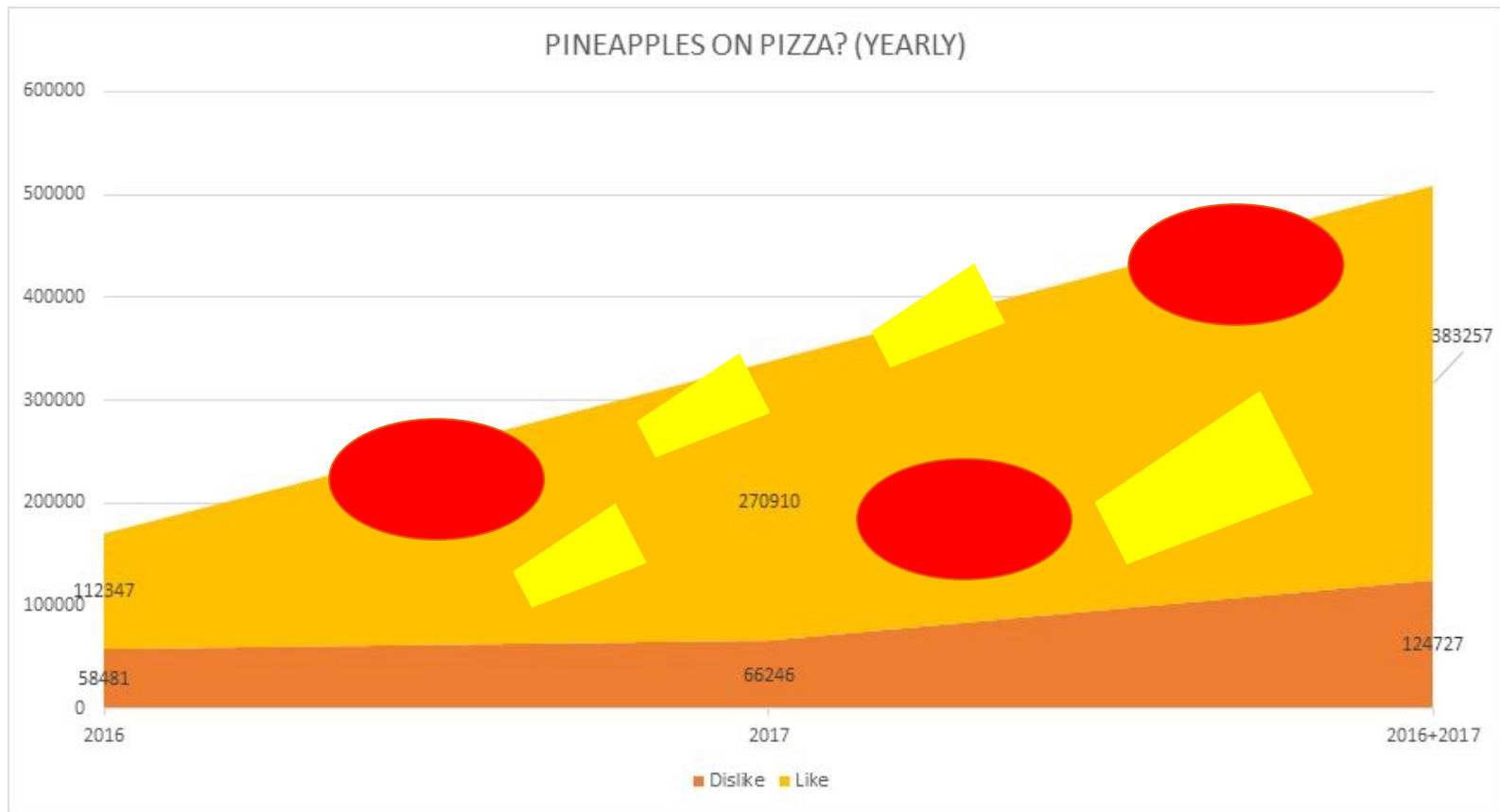
# Output

---

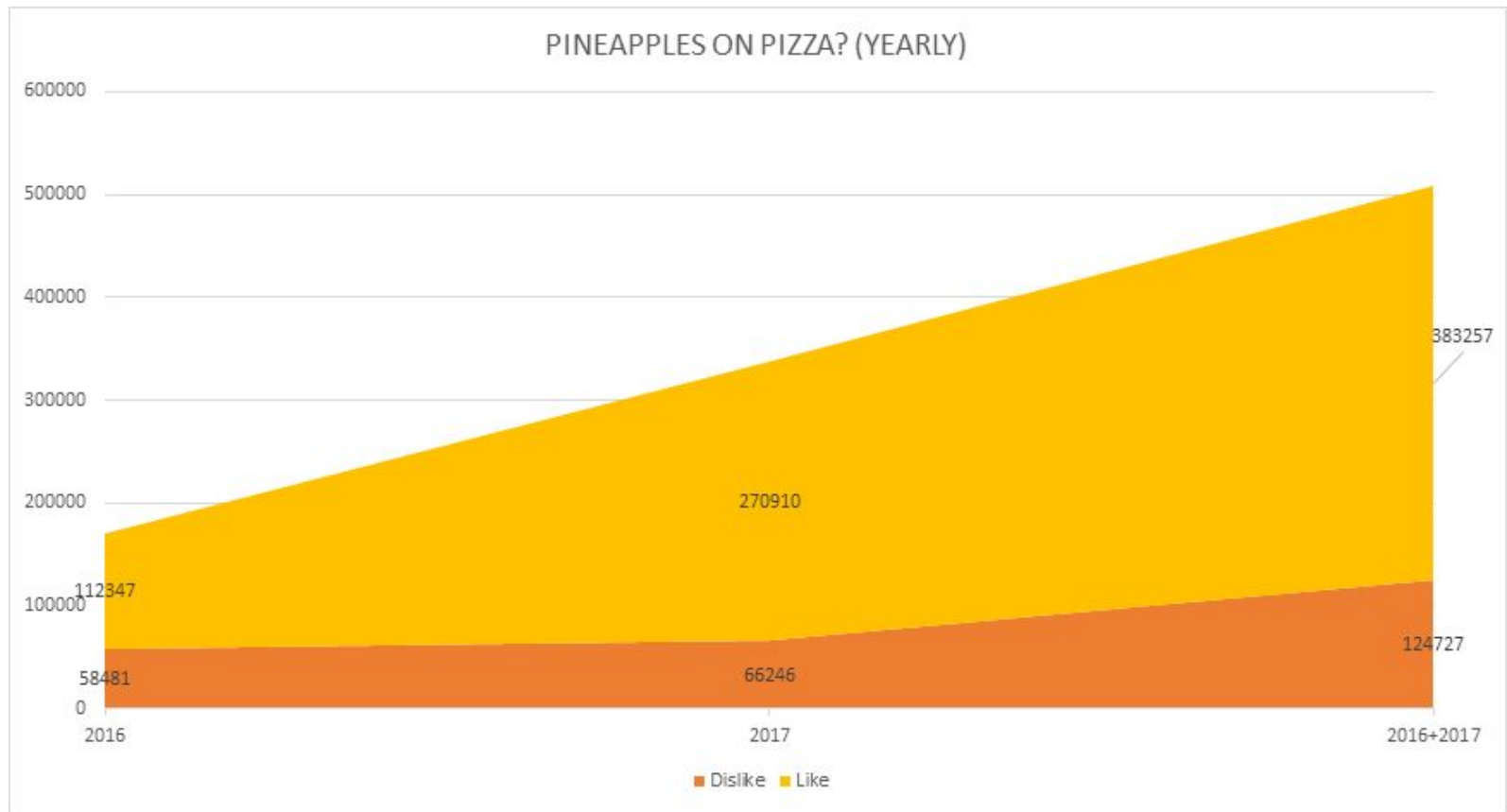




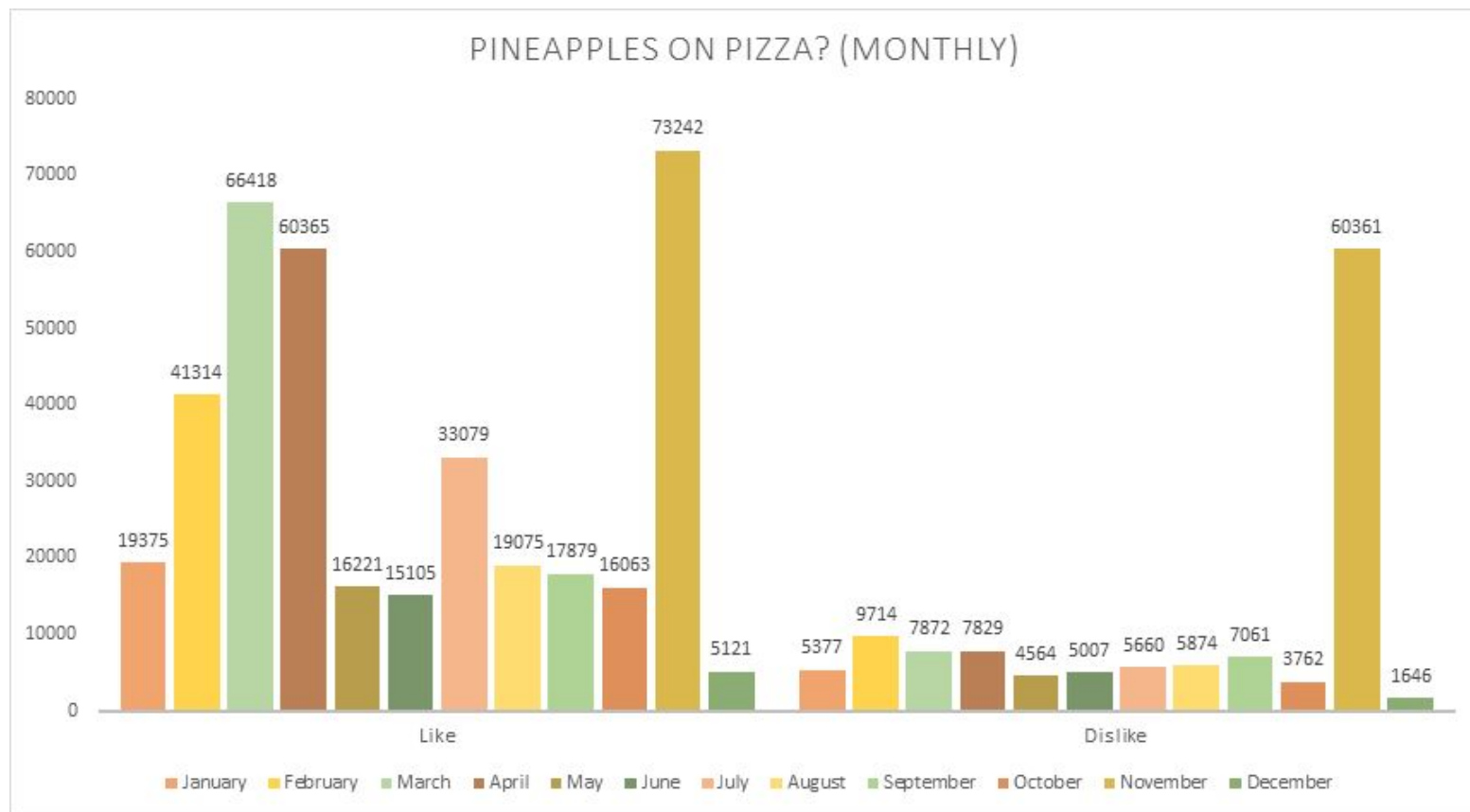
More people have tweeted about their pineapple on pizza opinion in 2017 compared to 2016. But in total, more people like pineapple on pizza.



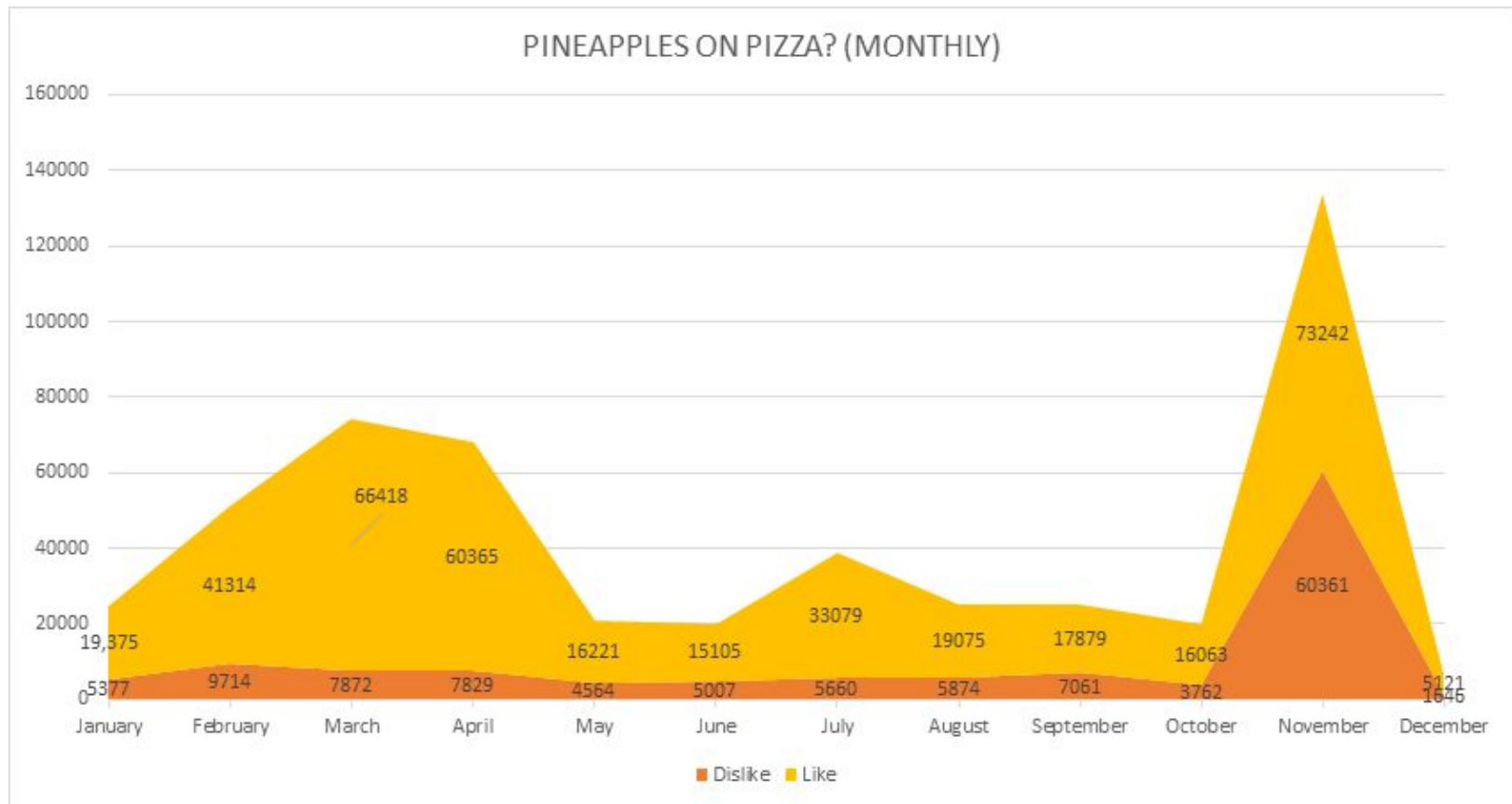
This looks like a slice of pizza, but it just contains the same data on the previous slide



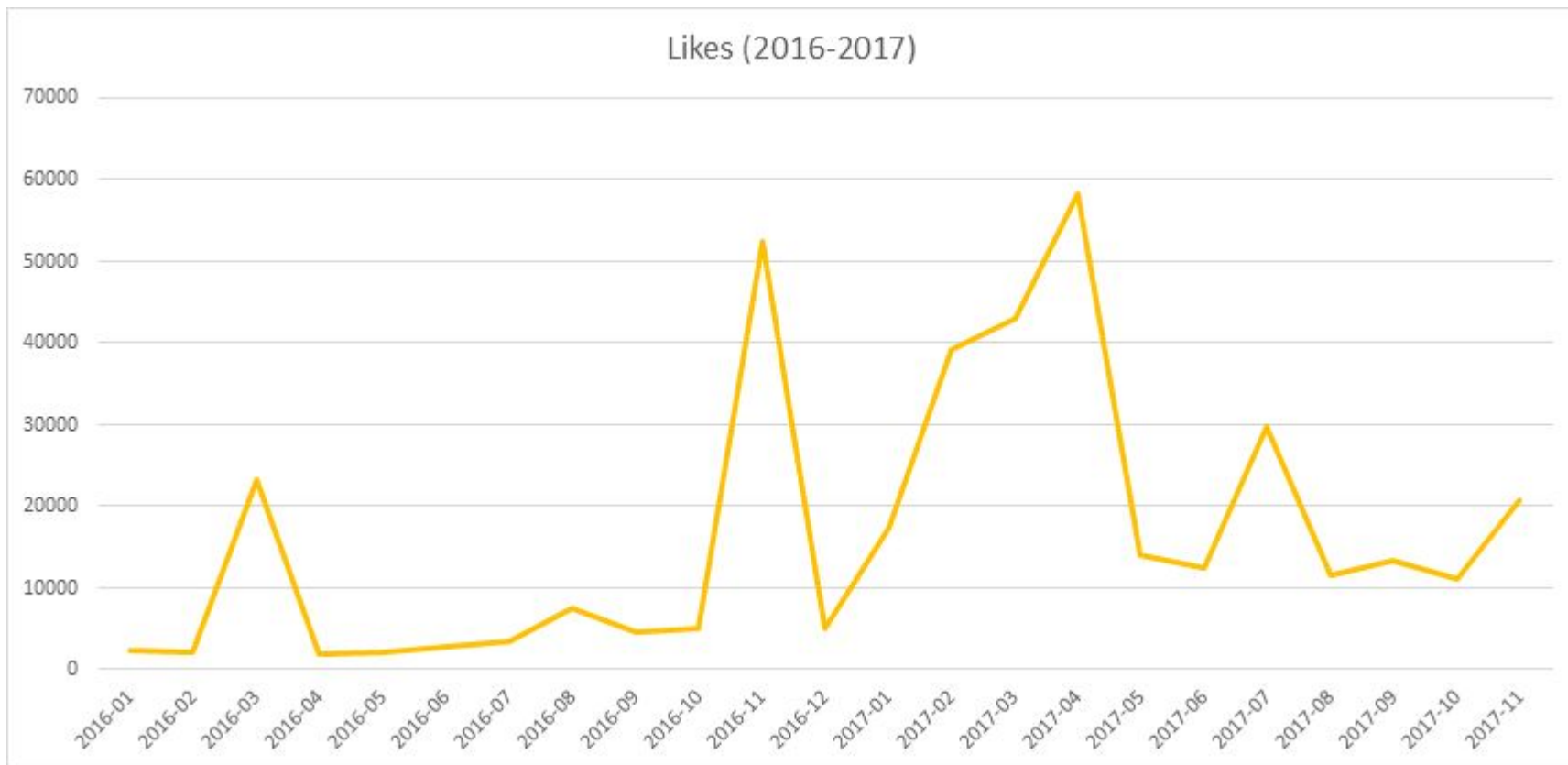
Overall, a lot of people like pineapples on their pizza.



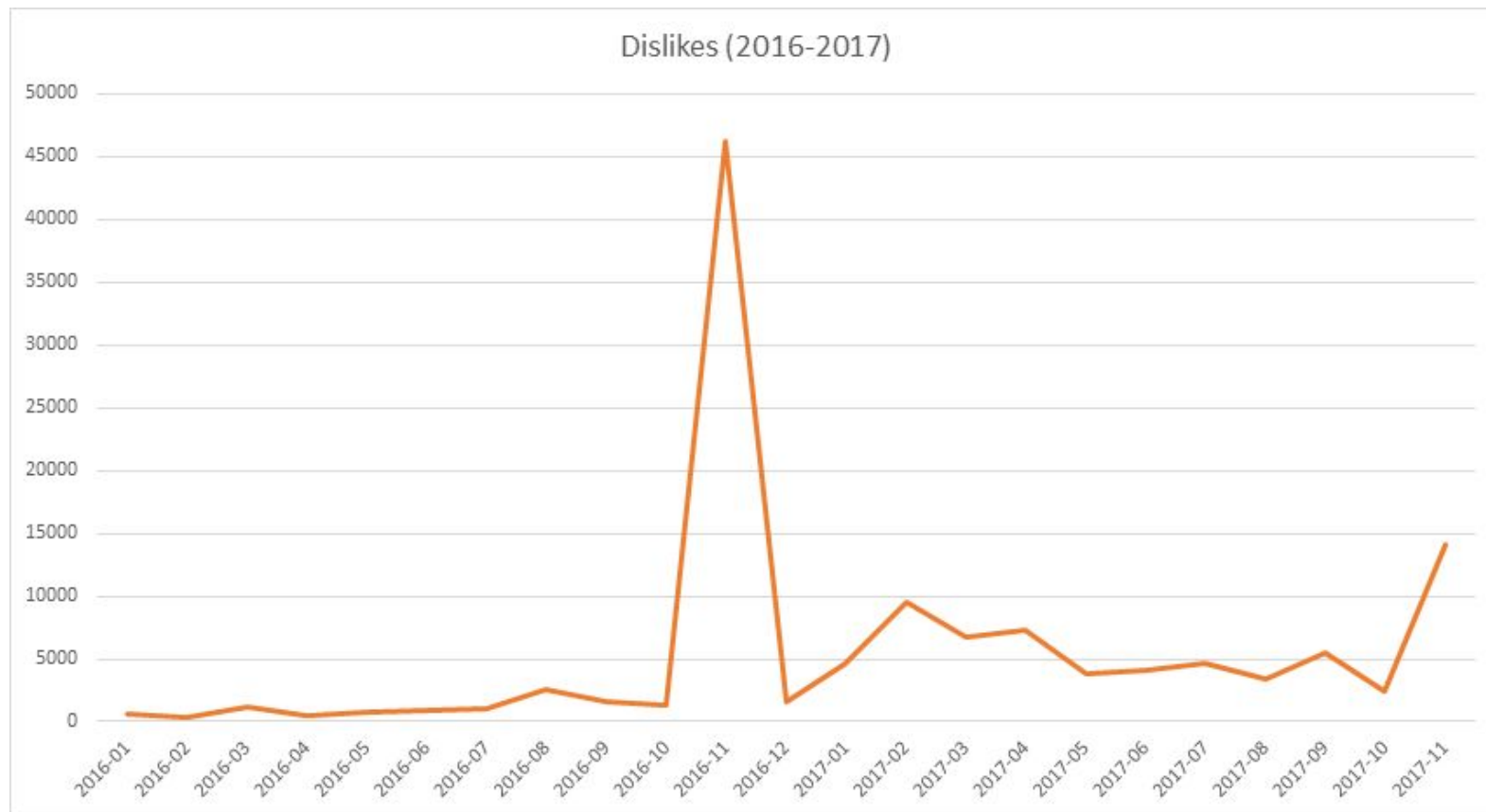
Over the past 2 years, people have voiced out the most during Novembers. The next month, December, is the month they don't voice out the most.



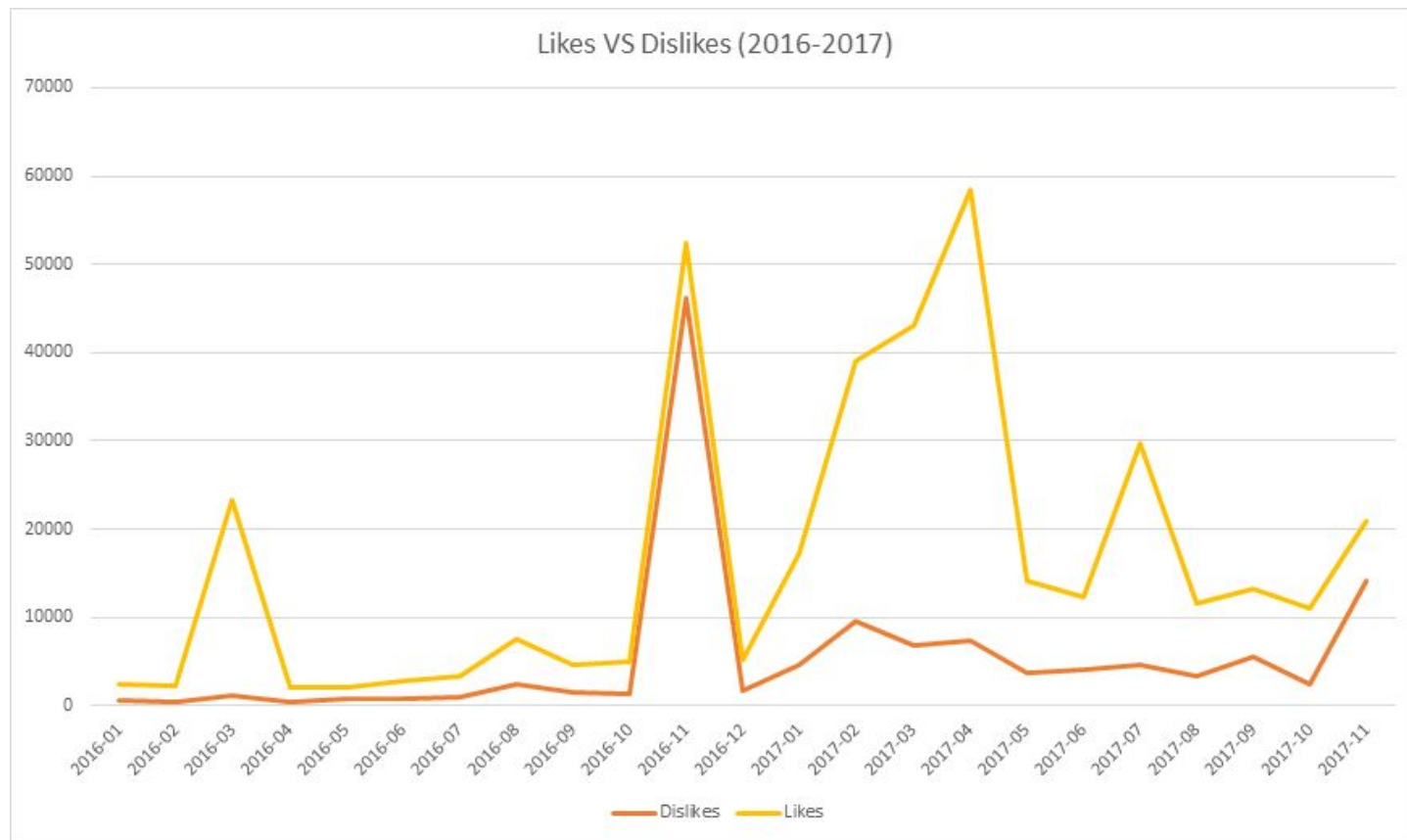
There are more significant spikes in the number of people who agree that pineapple belongs on pizza.



From 2016 to 2017, this is the record of likes.



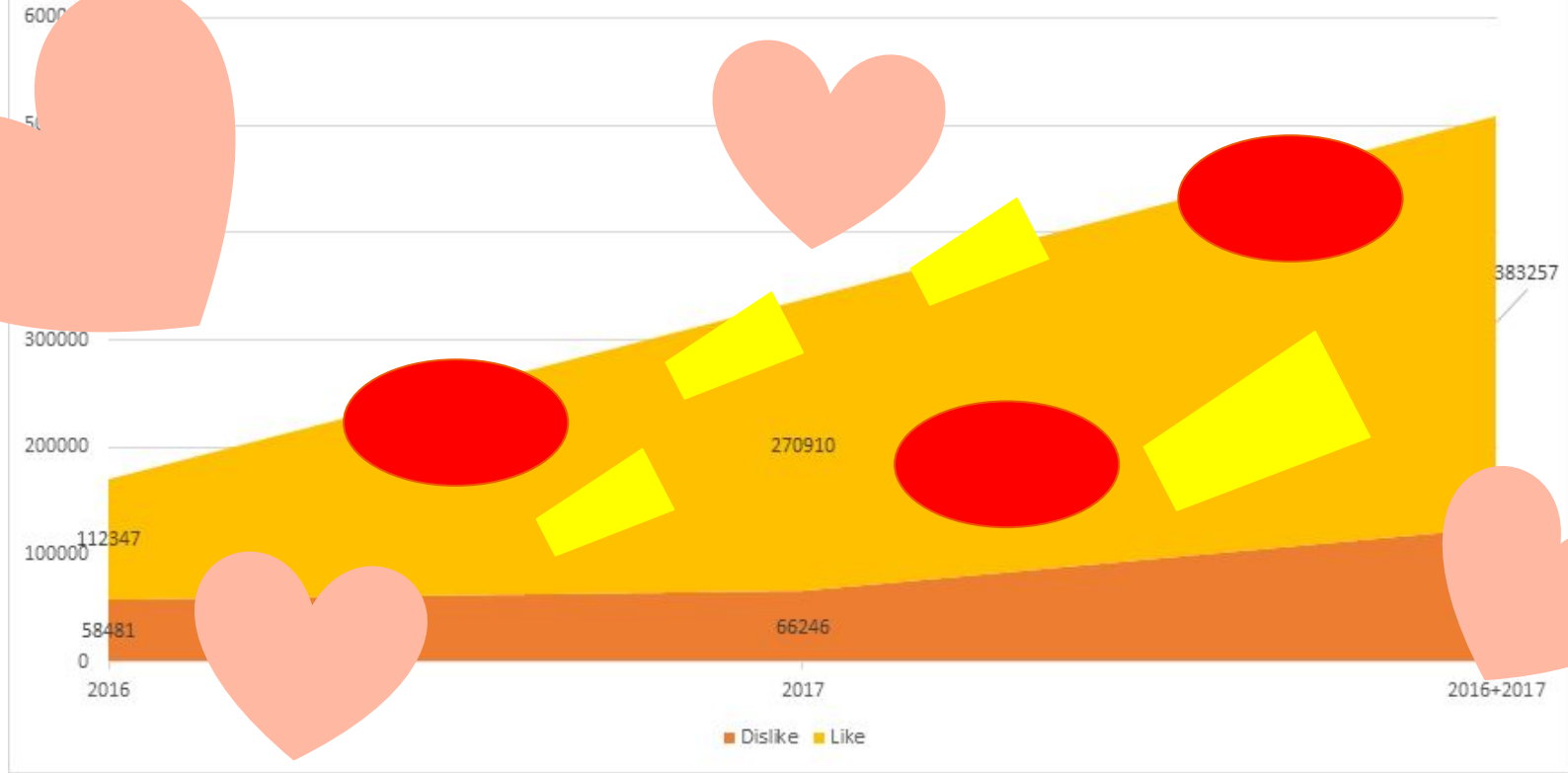
From 2016 to 2017, this is the record of dislikes.



From 2016 to 2017, this is the record of likes vs dislikes. A lot of people must have debated on this issue during Nov. 2016.



## PINEAPPLES ON PIZZA? (YEARLY)



*Therefore, most people agree that pineapples do belong on pizza.*