

# Do you know yourself well?

o o o o

An analysis for  
Facebook Live  
Sellers to  
understand their  
business and  
know what their  
viewers truly  
thinks.

# Background

o o o o

We are the data science team for Facebook Watch.

Do you know

400%

o o o o

Growth Rate of Facebook Live Viewership from 2020 to 2021

Source: <https://blog.hootsuite.com/facebook-live-video/>

# Problem Statement

o o o o

We wish to conduct a training to the e-commerce food businesses to help them to better **understand their business** and **understand viewer sentiments and discussion themes**.

As such, they will be able to provide more **appropriate products** and **improved services** to their viewers over their live streaming videos, which will in turn increase their revenue from their live streaming videos.

# Content

o o o o

- 1 Data Collection
- 2 Data Cleaning
- 3 Exploratory Data Analysis
- 4 Modelling: LDA
- 5 Conclusion, Insights & Recommendation

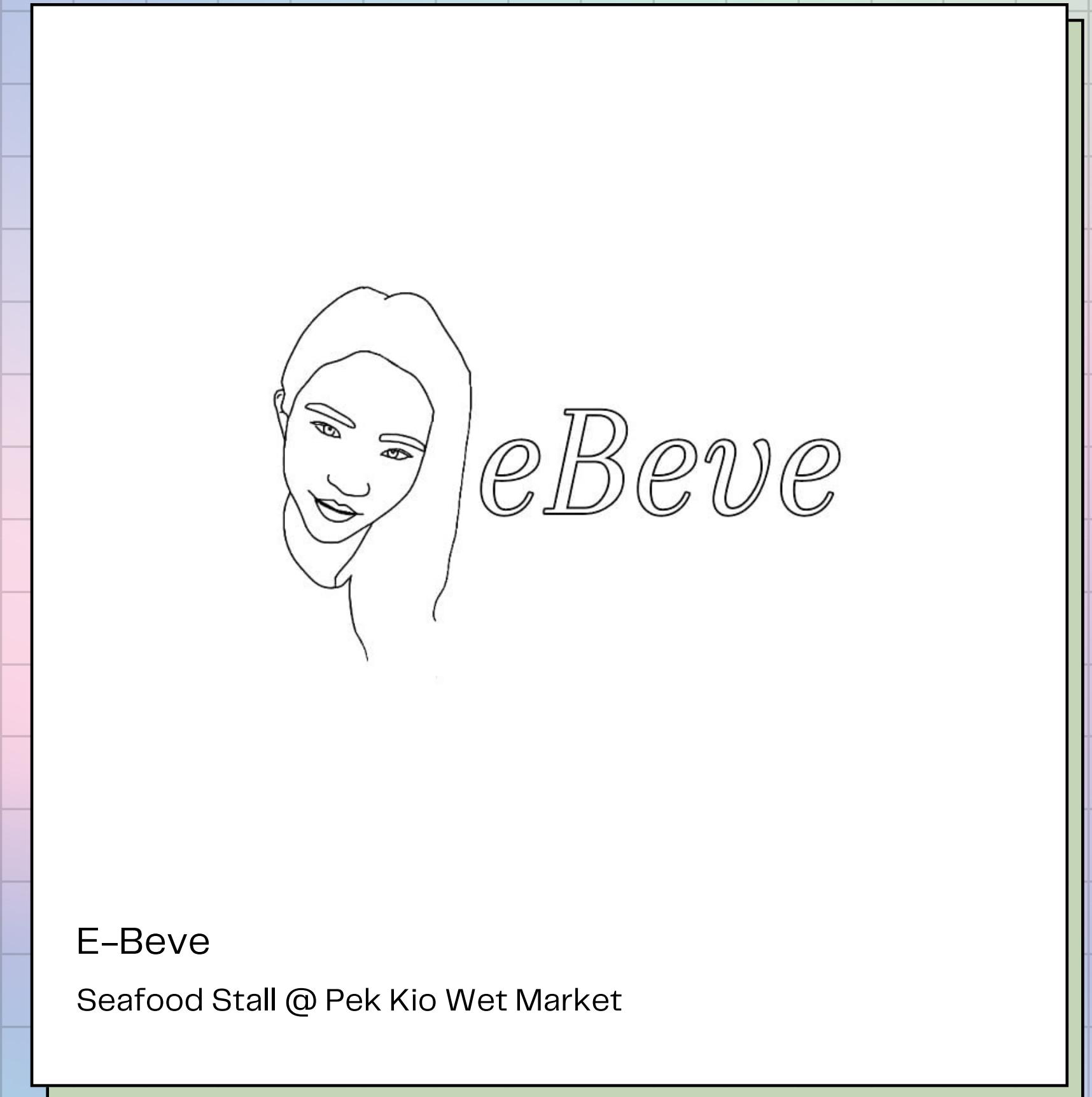


# Data Collection

**Let's look at our  
sellers!**



3 Sellers







Seller

# First Market Online Bidding



**FIRST MARKET<sup>®</sup>**

**ONLINE BIDDING**

**First Market Online Bidding**

Online Seafood Market

# Scrapping Comments

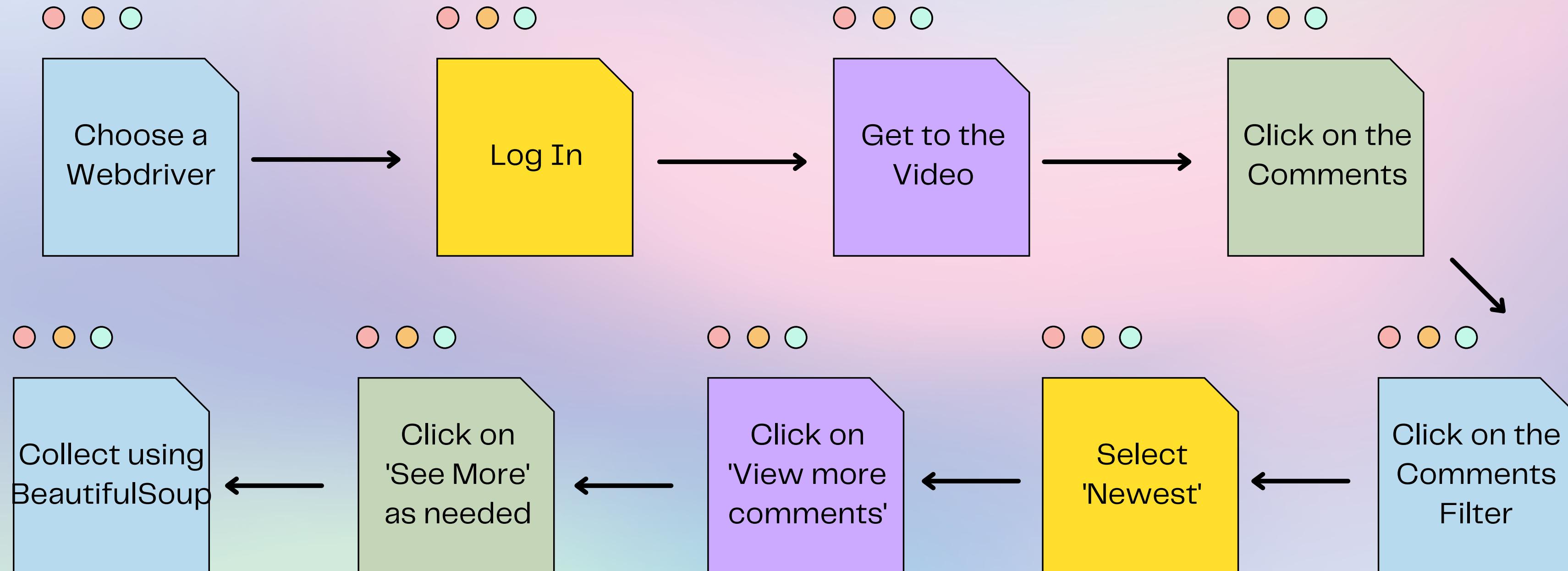
o o o o

1

Using Selenium

2

Using BeautifulSoup



# Scrapping

o o o o

1

Choose a Webdriver



Webdriver

Using the Firefox  
Webdriver



When using other  
webdrivers, Facebook  
would limit the number of  
actions our scrapping  
function is able to do.

facebook

Facebook helps you connect and share  
with the people in your life.

  
  
  
[Forgotten password?](#)  
  
[Create a Page for a celebrity, brand or business.](#)

# Scrapping

o o o o

2

Log In



Log In

**To View all comments,  
we need to log in!**

If not, Facebook will  
prompt you to log in to  
see more, when Selenium  
clicks on the comments

See more on Facebook

Email address or phone number

Password

Log In

[Forgotten password?](#)

or

[Create New Account](#)

# Scrapping

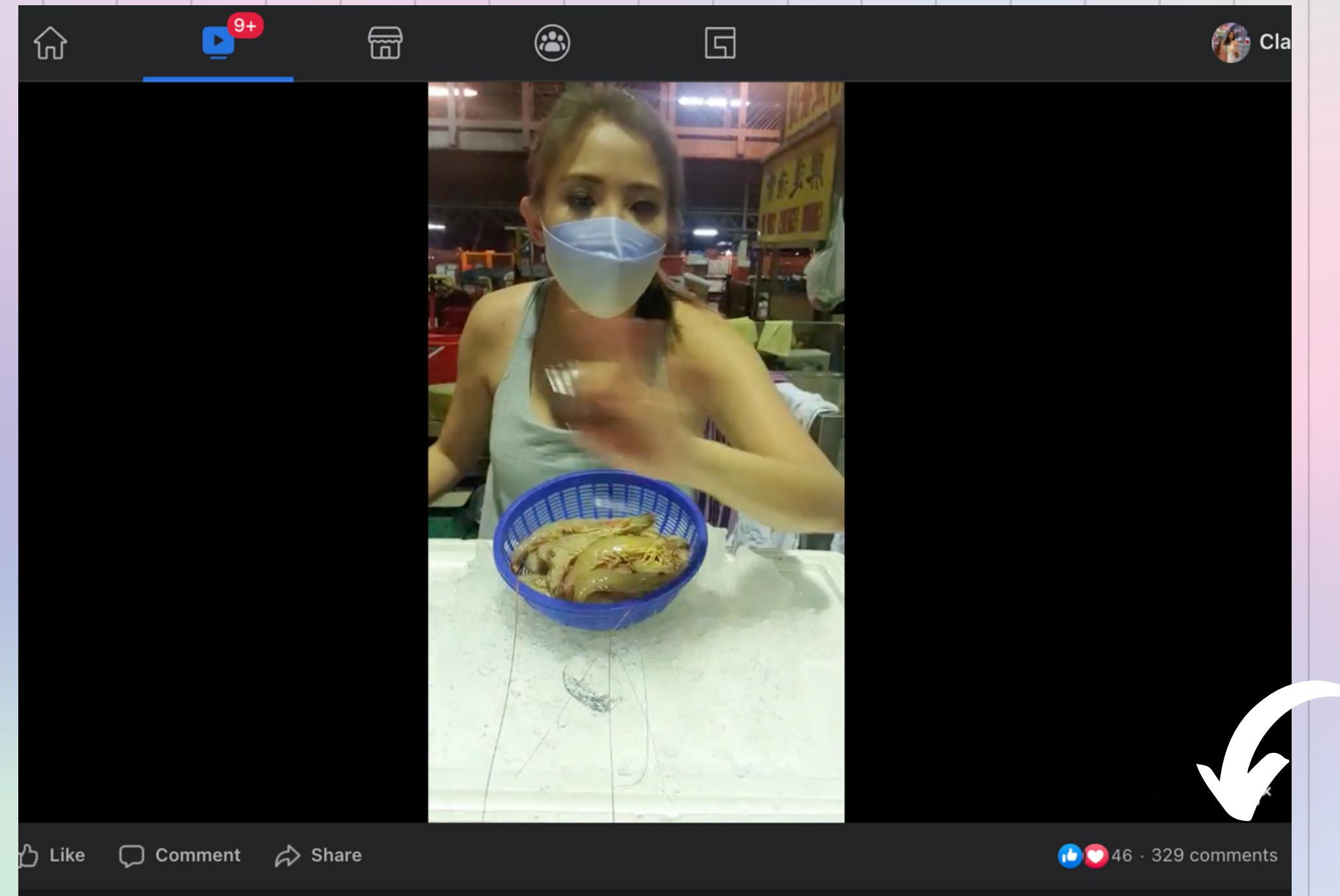
o o o o

3

Get To the Video

4

Click on the Comments



# Scrapping

o o o o

5

Click on the Comments Filter

Like Comment Share

46 · 329 comments

Most relevant ▾

Pinned comment

E-Beve · 18:39  
[Product] Prawn Maw (1pkt) - S\$2.50 | Keyword: PM

Like · 6 w

E-Beve · 0:55  
Hello everyone! Please remember to check out your cart!

Like · Reply · 6 w

E-Beve · 6:14  
[Product] Wild Angkah Prawns (1kg/mix sizes) - S\$20.00 | Keyword: WAK

Write a comment...

Smiley face icon

Comment icon

# Scrapping

o o o o

6

Select 'Newest'



Select  
'Newest'

**The filter 'Newest'  
has to be selected so  
that all comments  
can be loaded.**

The default setting is  
to show only the 'Most  
Relevant' comments.

Most relevant ▾

Real-time comments

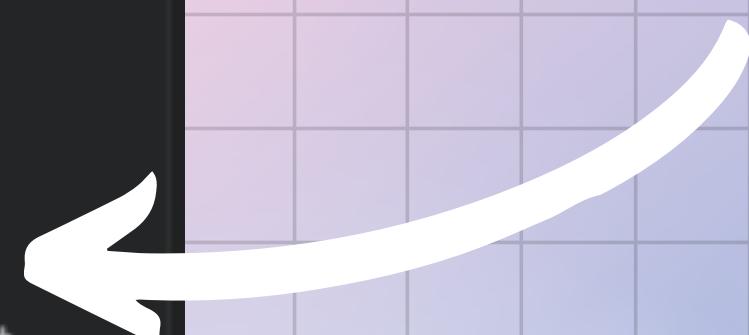
Comments shown at the moment when they appeared  
in the video.

Most relevant

Show friends' comments and the most engaging  
comments first.

Newest

Show all comments, with the newest comments first.



# Scrapping

o o o o

7

Click on 'View more comments'

Most relevant ▾

【Product】 Wild Black Pomfret (400g) - S\$6.50 | Keyword: BP

Like · 6 w

E-Beve · 0:55  
Hello everyone! Please remember to check out your cart!

Like · Reply · 6 w

E-Beve · 6:14  
【Product】 Wild Angkah Prawns (1kg/mix sizes) - S\$20.00 | Keyword: WAK

Like · Reply · 6 w

[View more comments](#)

2 of 225

 Write a comment...

# Scrapping

o o o o

8

Click on 'See More' as needed



FirstMarketOnlineBidding · 14:00

Bid Code : KCT13+1 // Kampung Chicken Thigh 1kg \$13

Bid Code : NC12+1 // Nutri Chicken \$12

Bid Code : BC65+1 // Black Chicken \$6.5... [See more](#)

Like · Reply · 21 w



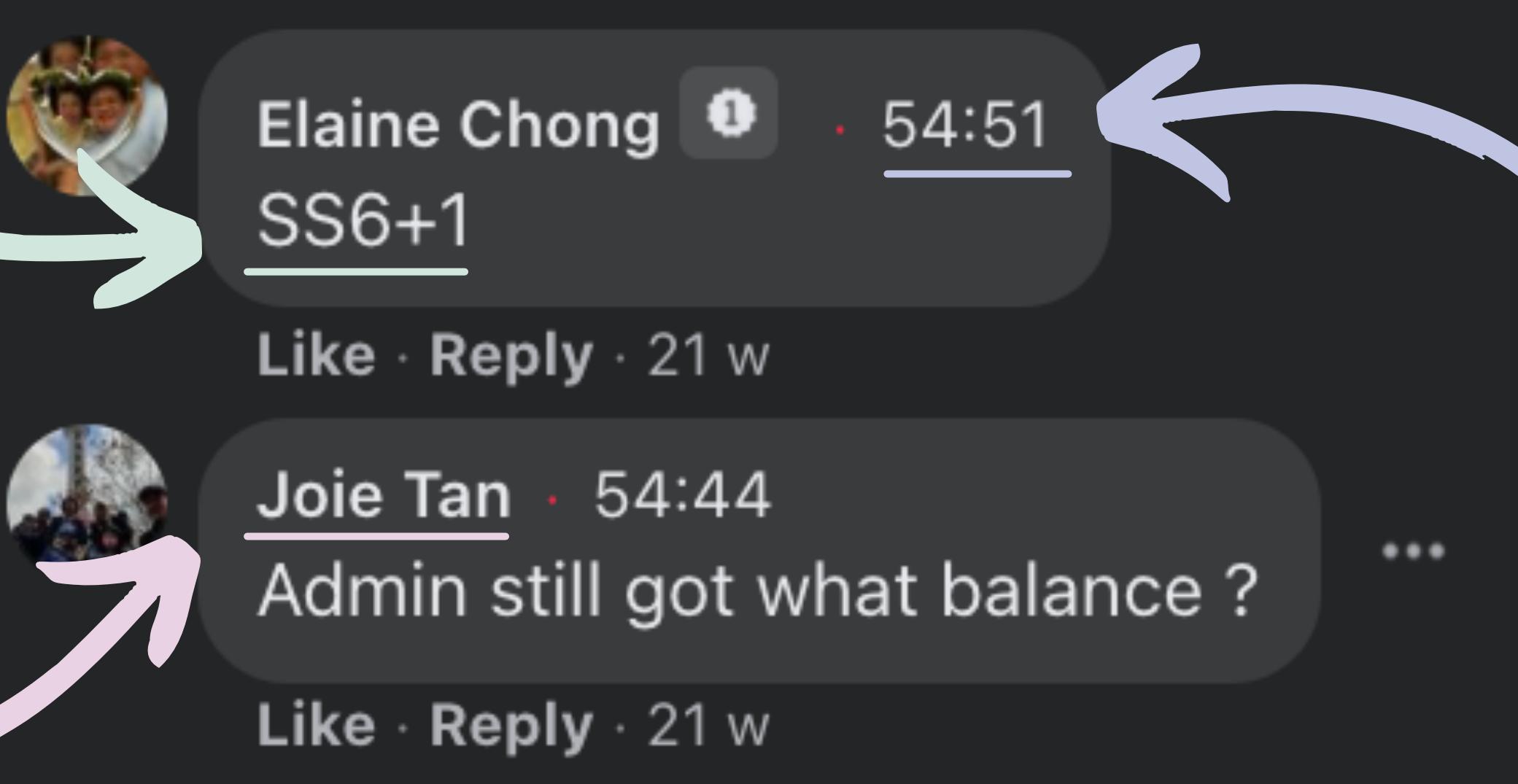
# Scrapping

o o o o

9

Collect using BeautifulSoup

Comment



Author of  
Comment

Comment  
Time

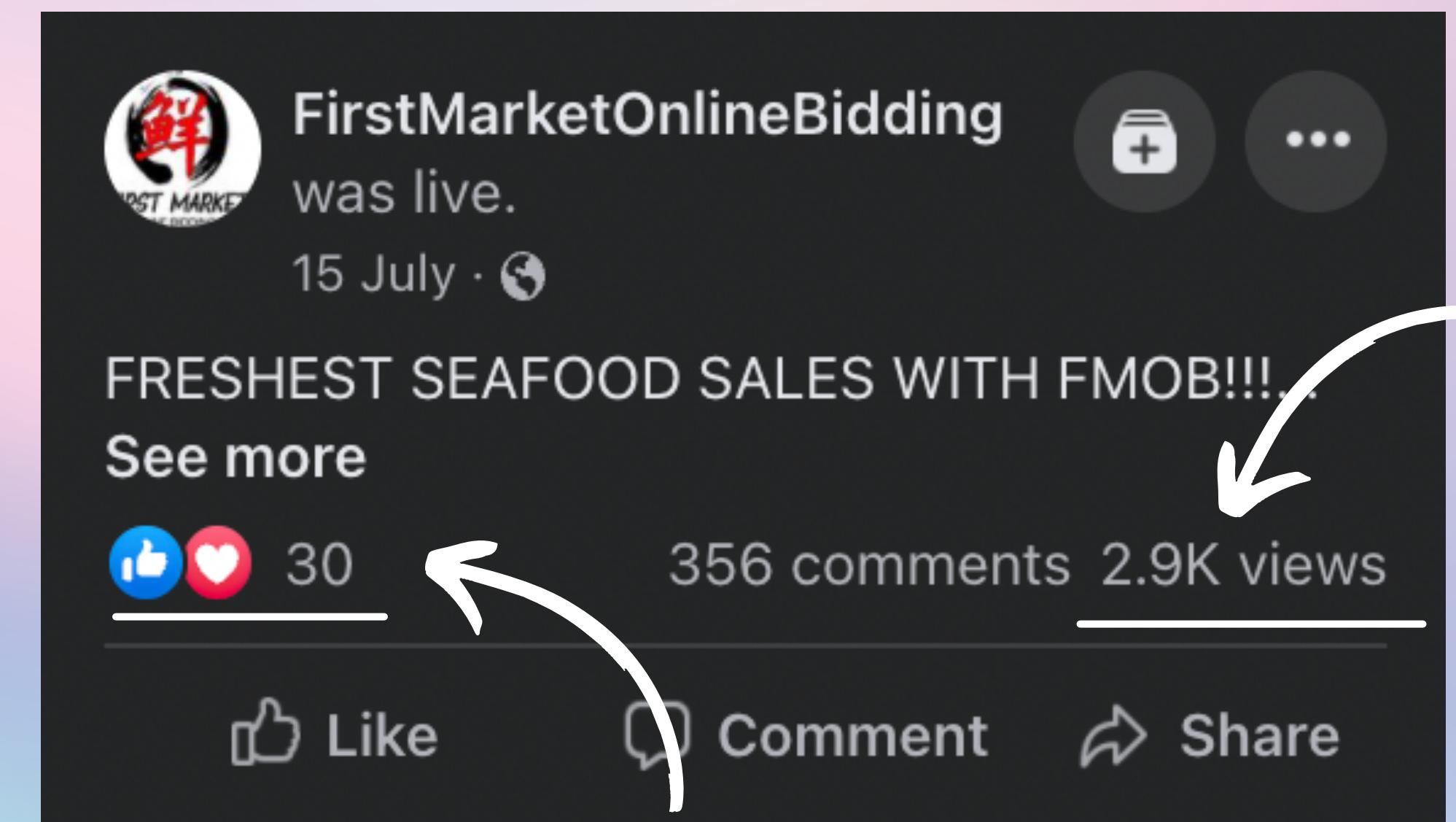
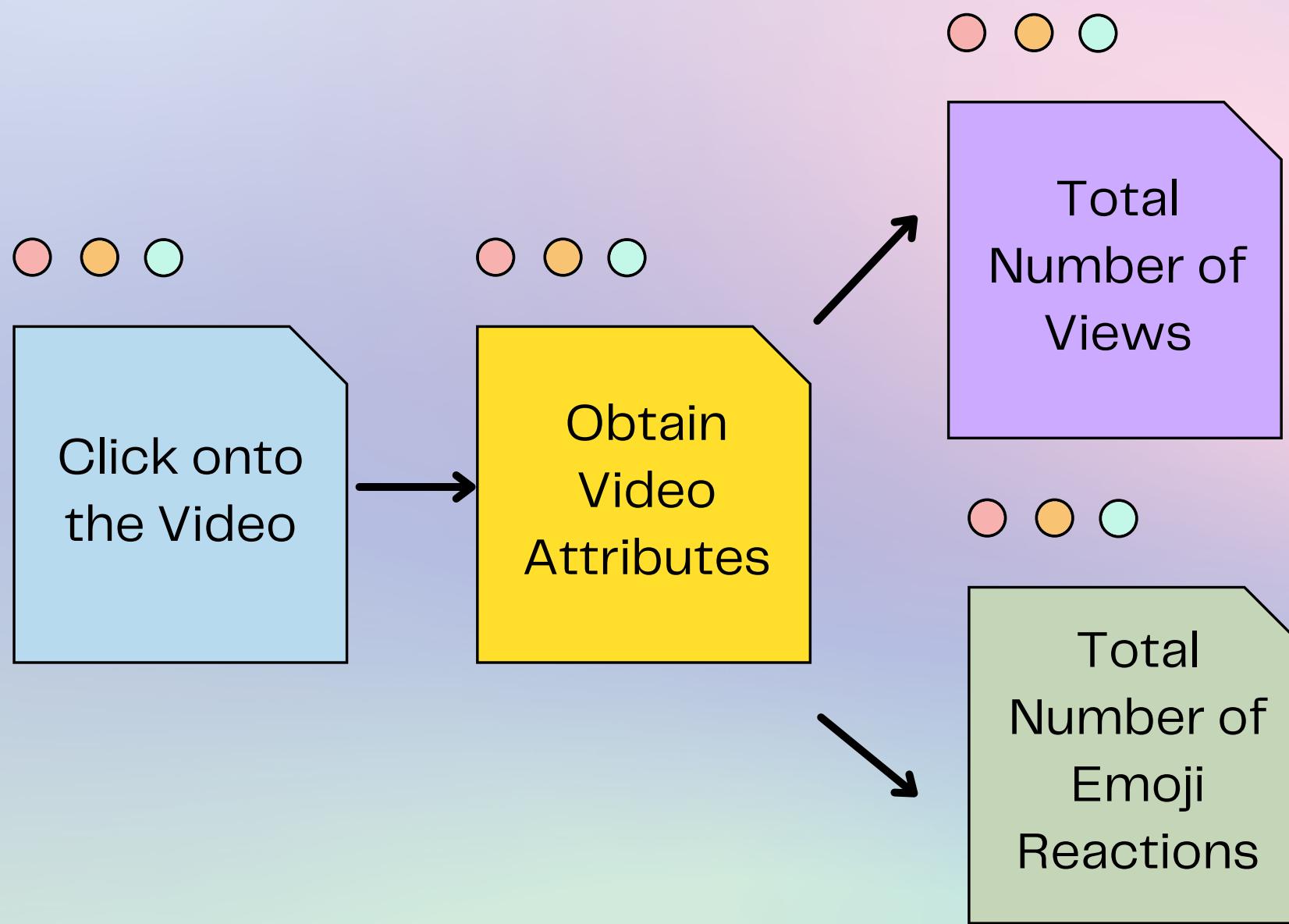
# Scrapping Video Attributes

1

Using Selenium

2

Using BeautifulSoup



# Scrapping Challenges

o o o o

## You're Temporarily Blocked

It looks like you were misusing this feature by going too fast.  
You've been temporarily blocked from using it.  
If you think that this doesn't go against our Community Standards,  
let us know.

OK



## Challenges

1

**Preventing a timeout exception before all the comments have been loaded**  
Increase the implicit wait time

2

**Ending the while loop for loading all the comments**  
Allow the function to loop through 5 more times before breaking the while loop

3

**Preventing Selenium from clicking into another user's account**  
Increasing the implicit wait time helps, but this still occurs when there are too many 'See More'

4

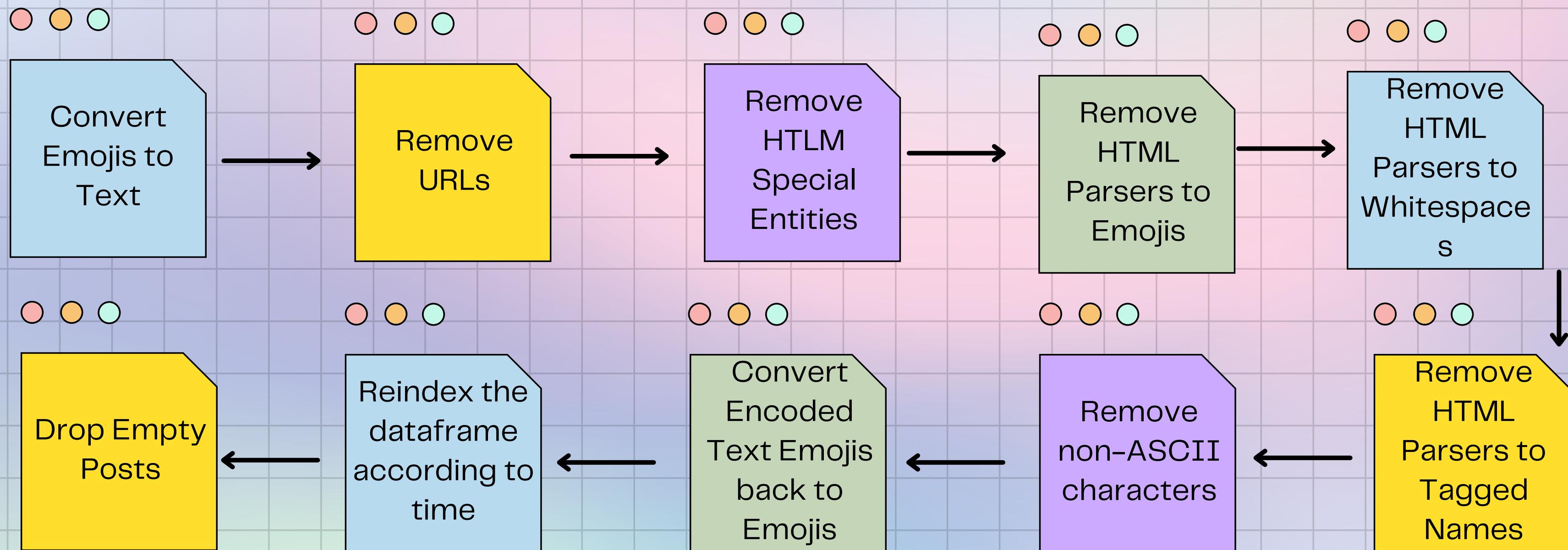
**Being blocked by Facebook**  
Create a new email and Facebook account.  
Or use VPN.

2

# Data Cleaning

# Data Cleaning

## 1 Preprocessing



# Data Cleaning

1

## Preprocessing



Examples

### Comments with Emojis

Original

Ok settle liao! 4kg of 3 type prawns prawns <span class="pq6dq46d tbxw36s4 knj5qynh kvgmc6g5 ditlmg2l oygrvhab nvdbi5me sf5mxxl7 gl3lb2sf hhz5lgdu"></span> good night.

After Demojizing

Ok settle liao! 4kg of 3 type prawns prawns <span class="pq6dq46d tbxw36s4 knj5qynh kvgmc6g5 ditlmg2l oygrvhab nvdbi5me sf5mxxl7 gl3lb2sf hhz5lgdu"></span> good night.

# Data Cleaning

1

## Preprocessing

Examples

Tagged Names

```
<a class="oajrlxb2 g5ia77u1 qu0x051f esr5mh6w e9989ue4 r7d6kgcz rq0escxv nhd2j8a9 nc684nl6 p7hjln8o kvgmc6g5 cxmmr5t8 oygrvhab hcukyx3x jb3vyjys rz4wbd8a qt6c0cv9 a8nywdso i1ao9s8h esuyzwwr f1sip0of lzcic4wl oo9gr5id gpro0wi8 lrazzd5p" href="https://www.facebook.com/tingwen?__tn__=R" role="link" tabindex="0"><span class="nc684nl6"><span>Ting  
Wen</span></span></a> yes i wake up le
```

# Data Cleaning

1

## Preprocessing

Examples

Non-ASCII Characters

Ok我拿

4Kg多少钱

# Data Cleaning

1

## Preprocessing



Examples

### Reindexing according to Time

|   | postComment               | postCommentAuthor | postCommentTime           |
|---|---------------------------|-------------------|---------------------------|
| 0 |                           | Arigato           | き リーサン 0:00               |
| 1 |                           | Goodnite          | き リーサン 0:00               |
| 2 |                           | Live+1            | Tan Poh Kim Irene 1:00:23 |
| 3 | Code For ang kar hia pls! | E-Beve            | 1:00:27                   |

### Standardize Timestamp format using TimedeltaIndex

|   | postComment               | postCommentAuthor | postCommentTime           | postCommentTime_final |
|---|---------------------------|-------------------|---------------------------|-----------------------|
| 0 |                           | Arigato           | き リーサン 0:00               | 0 days 00:00:00       |
| 1 |                           | Goodnite          | き リーサン 0:00               | 0 days 00:00:00       |
| 2 |                           | Live+1            | Tan Poh Kim Irene 1:00:23 | 0 days 01:00:23       |
| 3 | Code For ang kar hia pls! | E-Beve            | 1:00:27                   | 0 days 01:00:27       |

# Data Cleaning

1

## Preprocessing



Examples

### Remove the number of days from the Timestamp

|   | postComment               | postCommentAuthor | postCommentTime   | postCommentTime_final |          |
|---|---------------------------|-------------------|-------------------|-----------------------|----------|
| 0 |                           | Arigato           | き リーサン            | 0:00                  | 00:00:00 |
| 1 |                           | Goodnite          | き リーサン            | 0:00                  | 00:00:00 |
| 2 |                           | Live+1            | Tan Poh Kim Irene | 1:00:23               | 01:00:23 |
| 3 | Code For ang kar hia pls! |                   | E-Beve            | 1:00:27               | 01:00:27 |

### Sort dataframe using time

|     | postComment         | postCommentAuthor | postCommentTime | postCommentTime_final |          |
|-----|---------------------|-------------------|-----------------|-----------------------|----------|
| 0   |                     | Arigato           | き リーサン          | 0:00                  | 00:00:00 |
| 1   |                     | Goodnite          | き リーサン          | 0:00                  | 00:00:00 |
| 64  | Helloooo ❤️         | Ernest Tan        | 1:20            | 00:01:20              |          |
| 104 | Hi got crab today?? | Veon Veon         | 1:40            | 00:01:40              |          |

# Data Cleaning

1

## Preprocessing



Examples

### Remove the number of days from the Timestamp

|   | postComment               | postCommentAuthor | postCommentTime   | postCommentTime_final |          |
|---|---------------------------|-------------------|-------------------|-----------------------|----------|
| 0 |                           | Arigato           | き リーサン            | 0:00                  | 00:00:00 |
| 1 |                           | Goodnite          | き リーサン            | 0:00                  | 00:00:00 |
| 2 |                           | Live+1            | Tan Poh Kim Irene | 1:00:23               | 01:00:23 |
| 3 | Code For ang kar hia pls! |                   | E-Beve            | 1:00:27               | 01:00:27 |

### Sort dataframe using time

|     | postComment         | postCommentAuthor | postCommentTime | postCommentTime_final |          |
|-----|---------------------|-------------------|-----------------|-----------------------|----------|
| 0   |                     | Arigato           | き リーサン          | 0:00                  | 00:00:00 |
| 1   |                     | Goodnite          | き リーサン          | 0:00                  | 00:00:00 |
| 64  | Helloooo ❤️         | Ernest Tan        | 1:20            | 00:01:20              |          |
| 104 | Hi got crab today?? | Veon Veon         | 1:40            | 00:01:40              |          |

# Data Cleaning

2

## Feature Engineering



New Columns

- Comments Attributes

|   | postComment | postCommentAuthor | postCommentTime_final | isSeller | postCommentLength | Ins | salesQuantity | revenue | seller         |
|---|-------------|-------------------|-----------------------|----------|-------------------|-----|---------------|---------|----------------|
| 0 | Arigato     | き リーサン            | 00:00:00              | 0        |                   | 1   | 0             | 0       | 0.0 ebeveadmin |

- Video Attributes

|  | totalEmojiReaction | views | videoLength | numSellerComments | numComments | InsQuantity | salesQuantity | numProducts | totalRevenue | frequencySeller | averageCompound |
|--|--------------------|-------|-------------|-------------------|-------------|-------------|---------------|-------------|--------------|-----------------|-----------------|
|  | 46                 | 1800  | 5746        | 50                | 1255        | 24          | 114           | 38          | 1409.00      | 114.92          | 0.003145        |

3

# Exploratory Data Analysis

# Exploratory Data Analysis

o o o o

1

Video Attributes

2

Comments Attributes

# EDA for Video Attributes

1 Total Revenue



Range of Total Revenue for each Video



500

1000

1500

2000

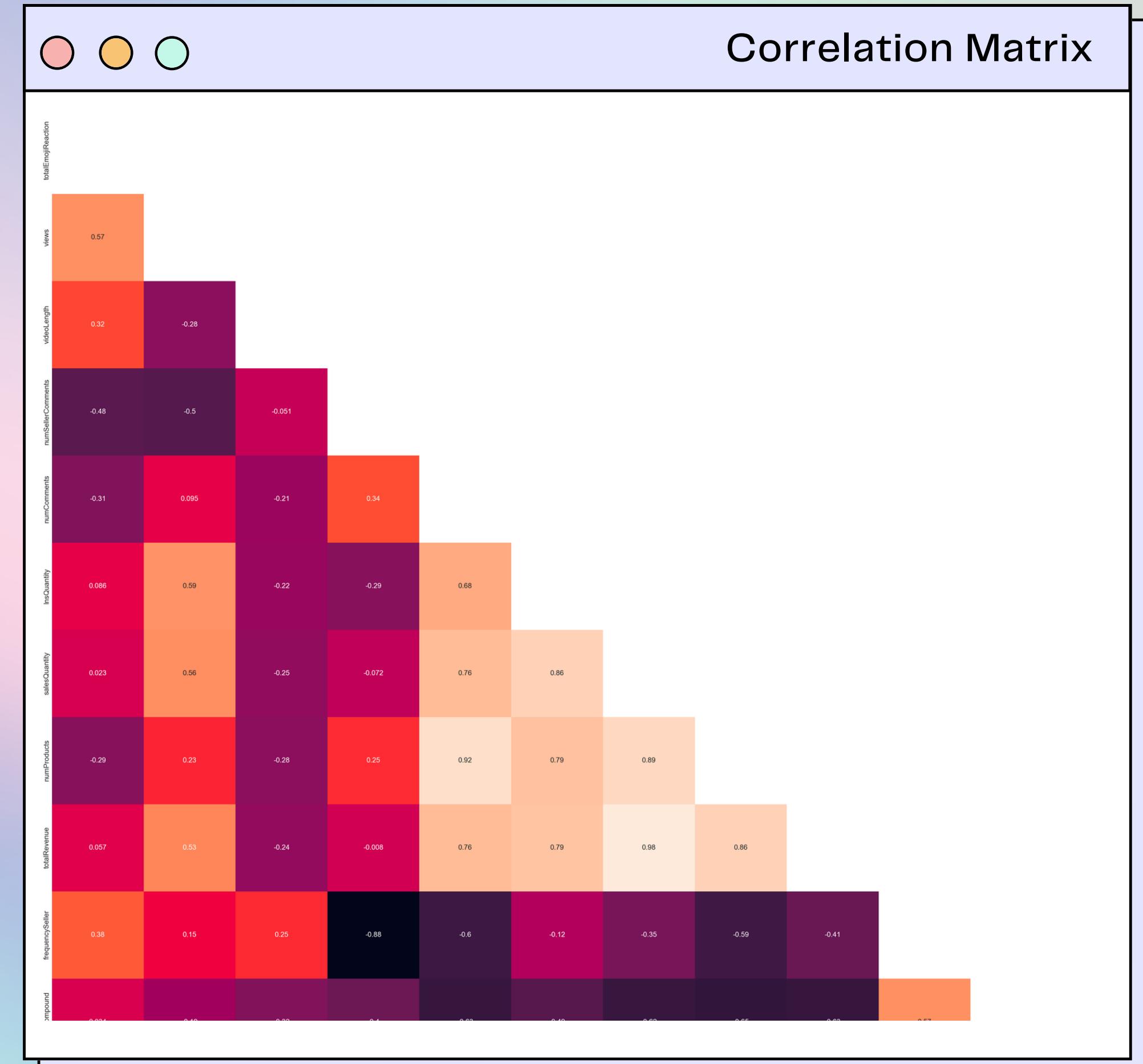
2500

3000

totalRevenue

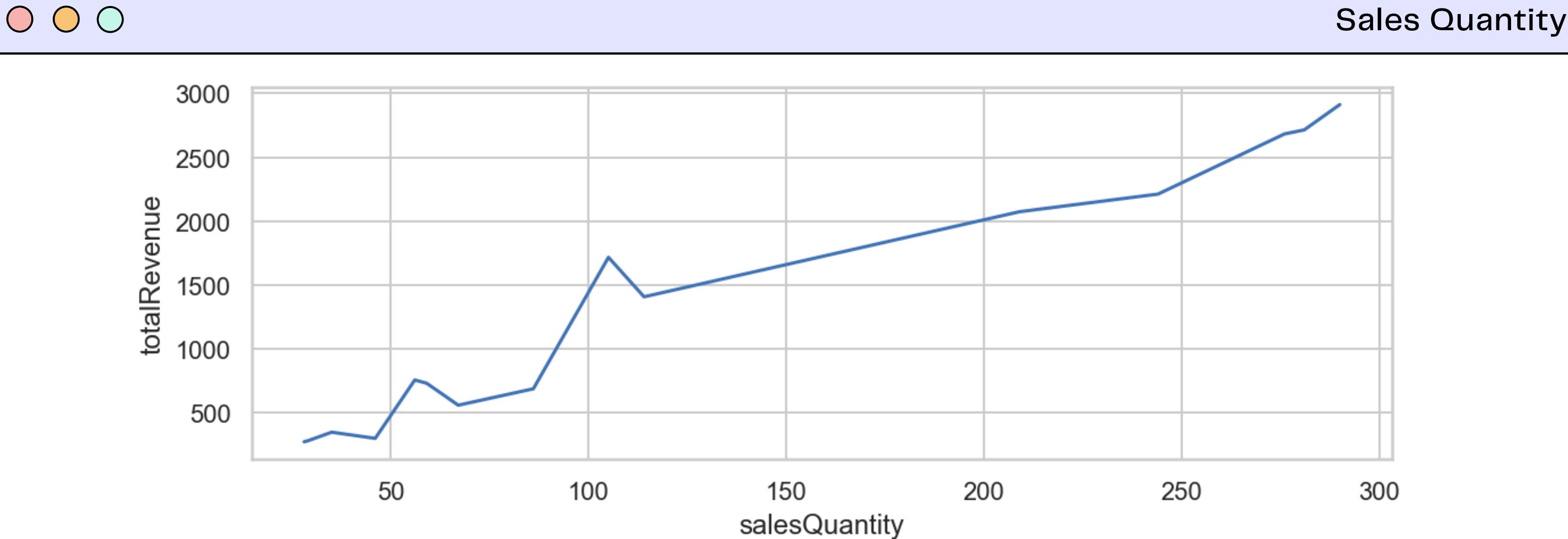
# EDA for Video Attributes

2 Correlation of All Variables against Total Revenue



# EDA for Video Attributes

3 Sales Quantity against Total Revenue



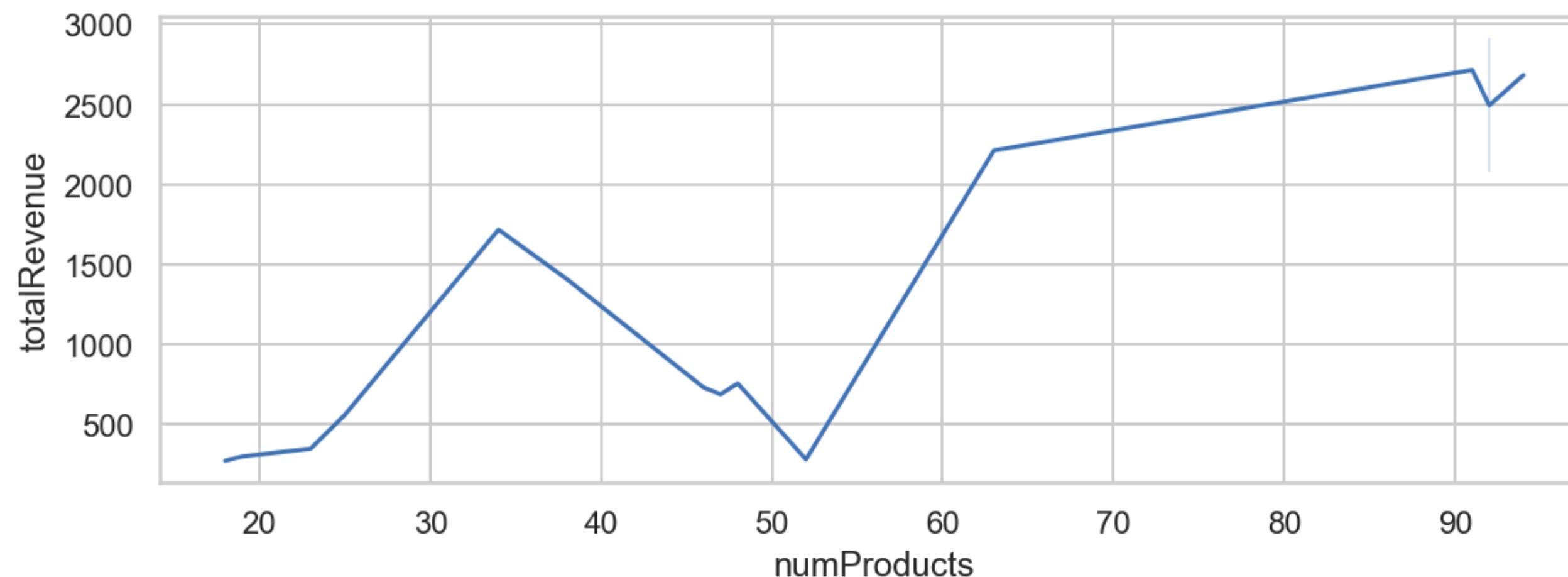
# EDA for Video Attributes

4

Number of Products  
against Total Revenue



Number of Products



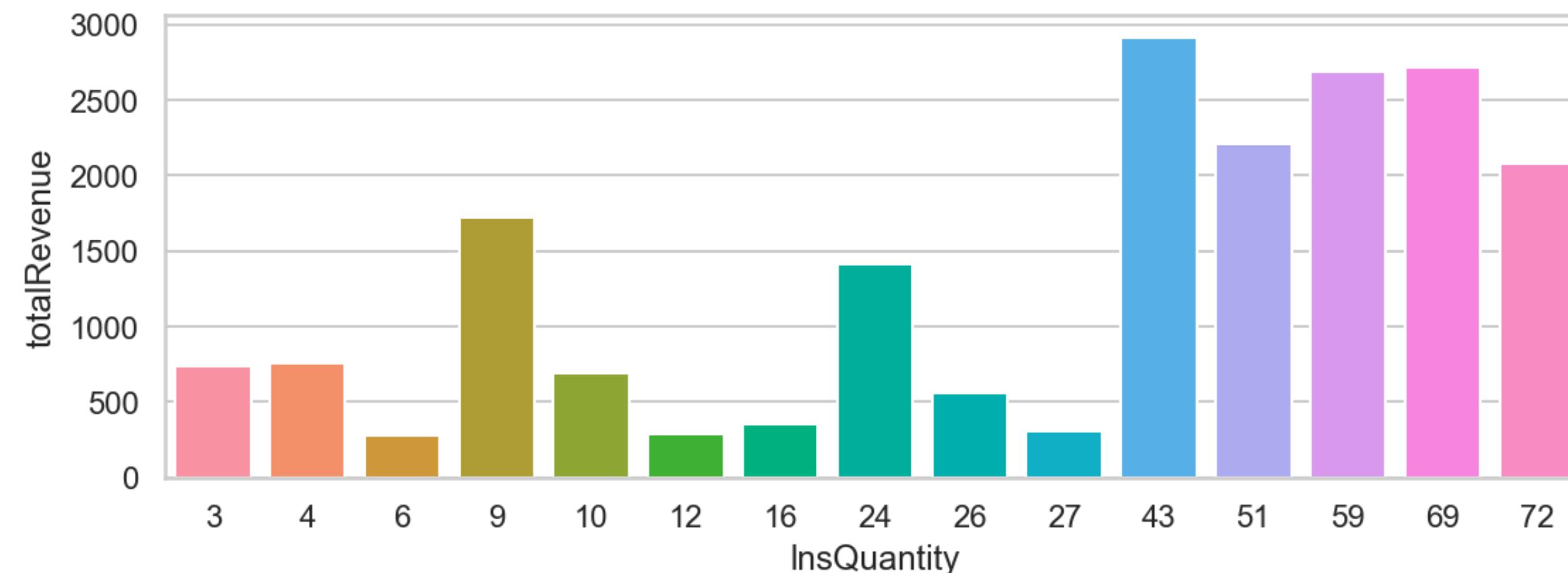
# EDA for Video Attributes

4

Quantity of LNS against  
Total Revenue

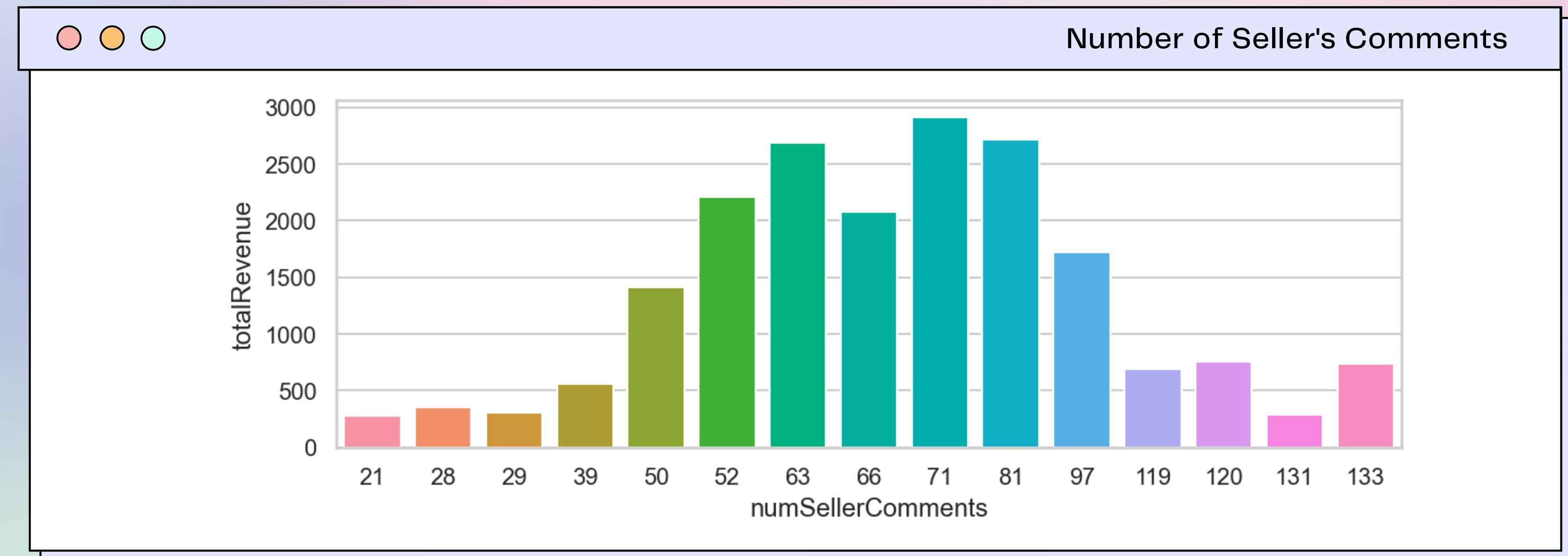


Quantity of LNS



# EDA for Video Attributes

Number of Seller's  
Comments against Total  
Revenue



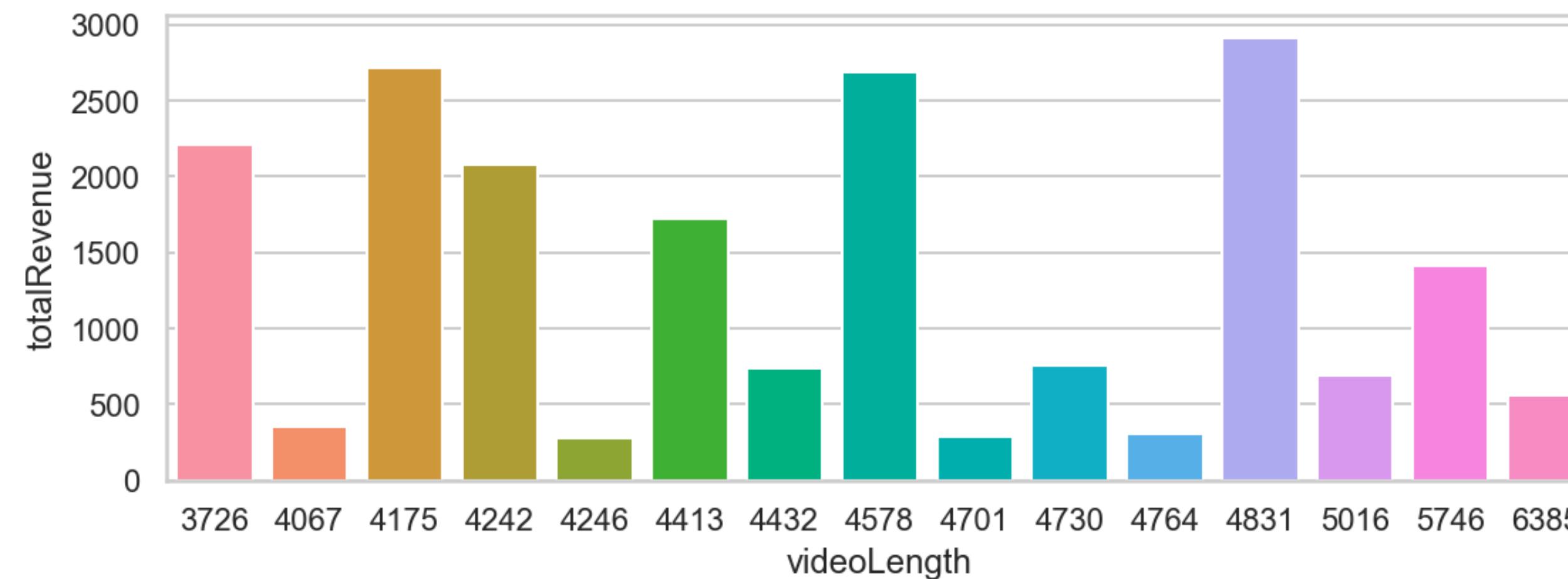
# EDA for Video Attributes

6

Length of Video against  
Total Revenue



Length of Video



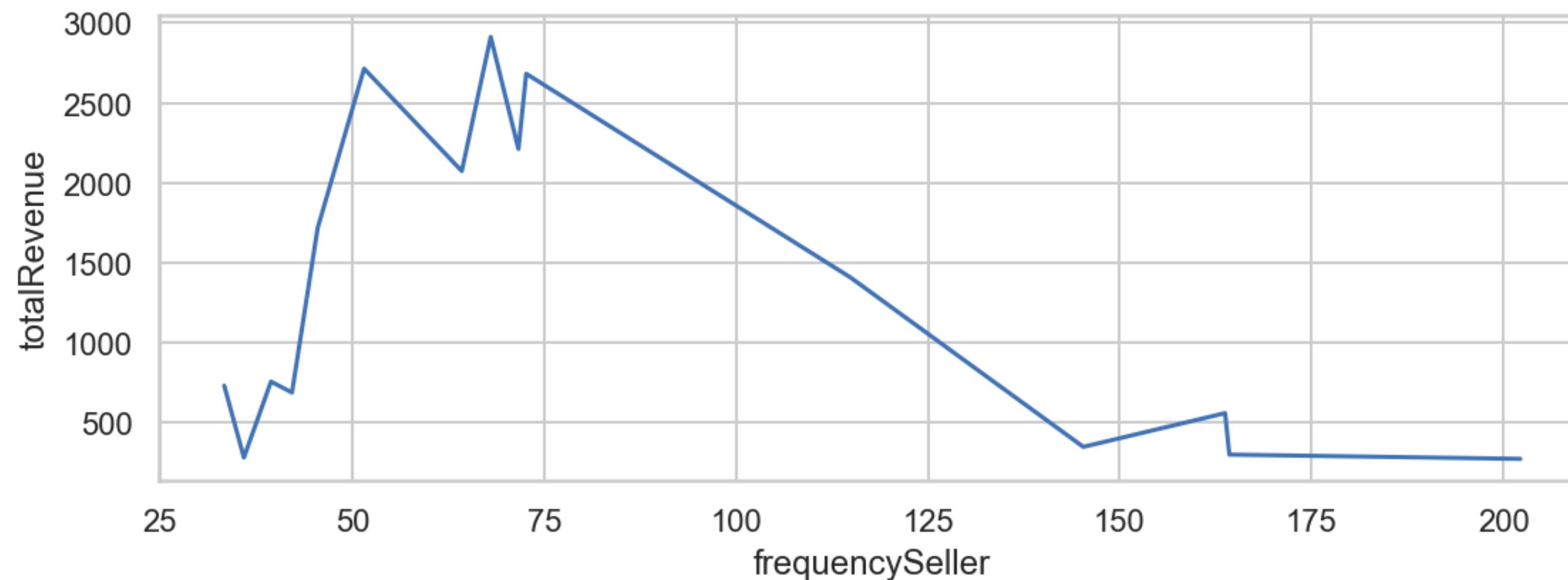
# EDA for Video Attributes

Frequency of Seller's  
Comments against Total  
Revenue

7



Frequency of Seller's Comments



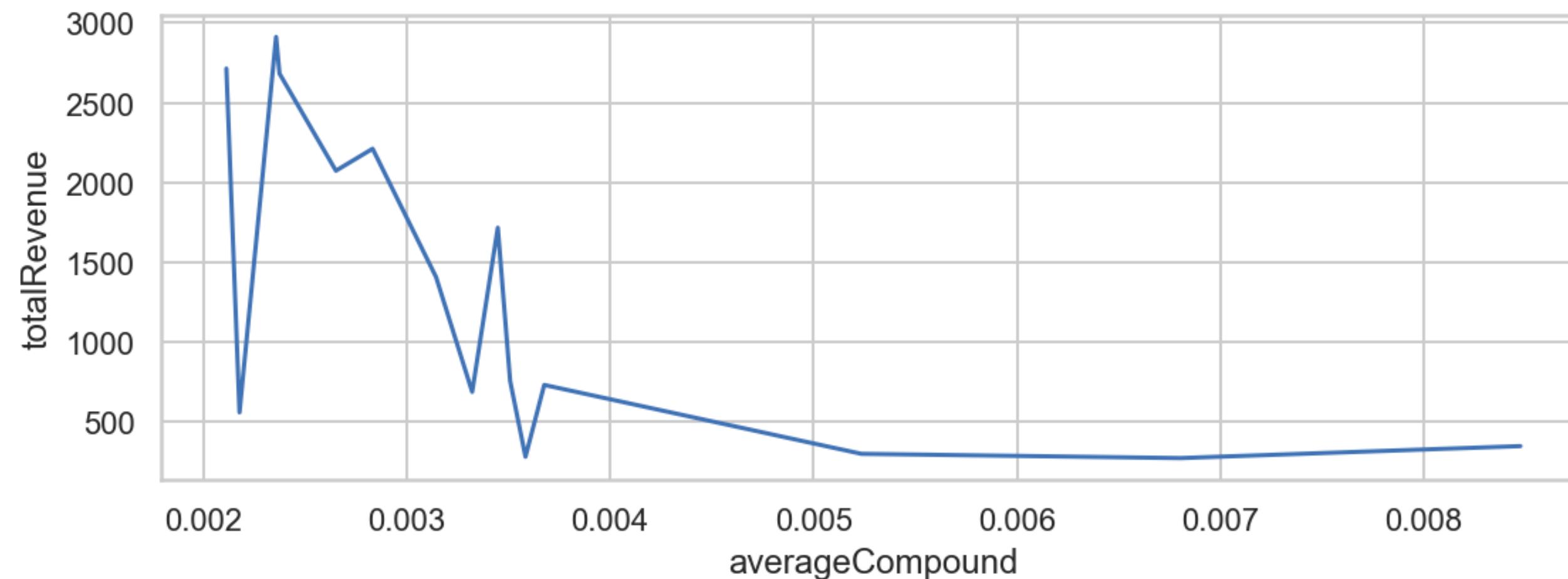
# EDA for Video Attributes

8

Average Compound Score  
for the entire Video against  
Total Revenue



Average Compound Score for the entire Video



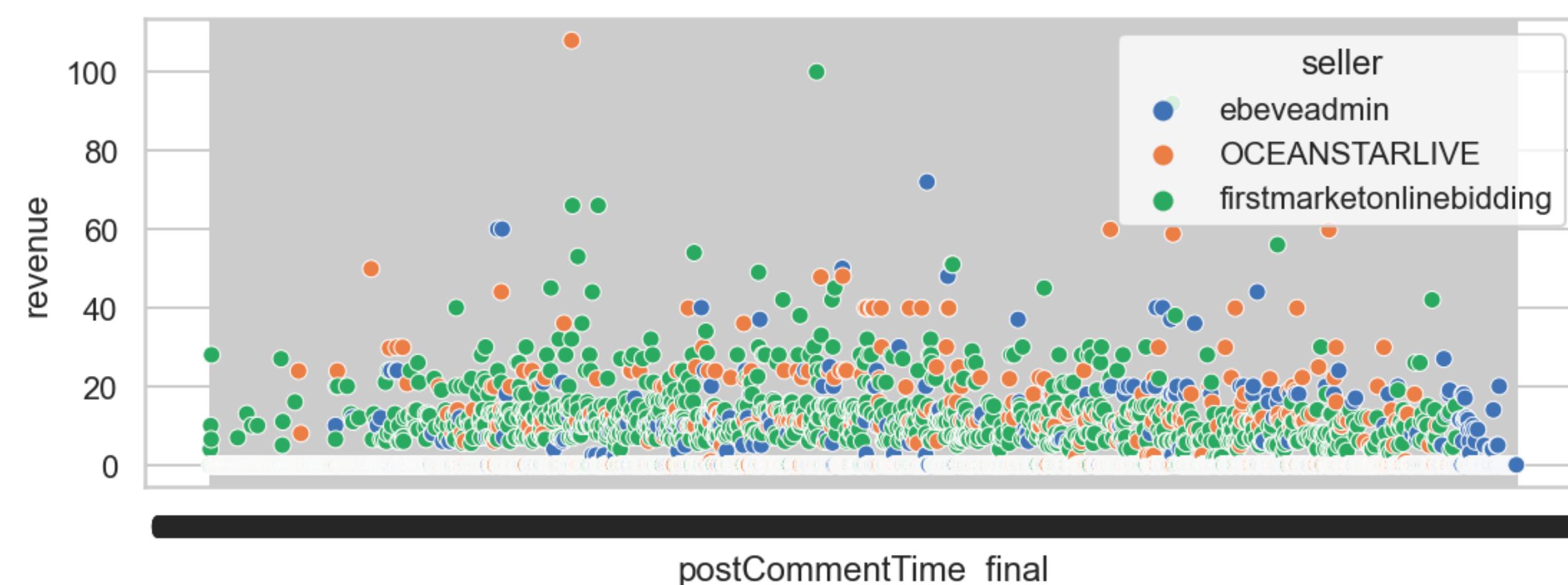
# EDA for Comments Attributes

1

Revenue against Time



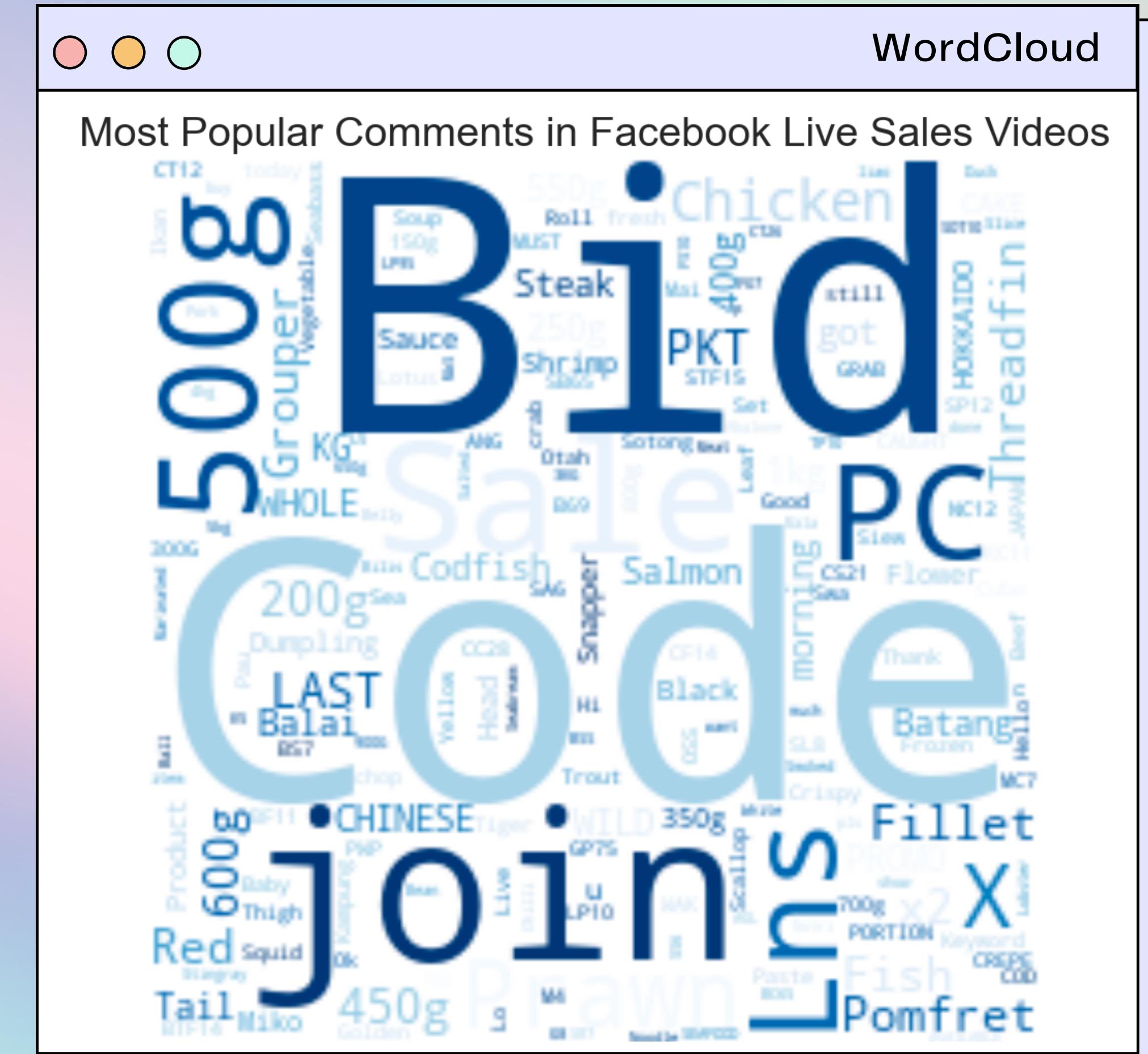
Revenue against Time



# EDA for Comments Attributes

2

# WordCloud



# EDA for Comments Attributes

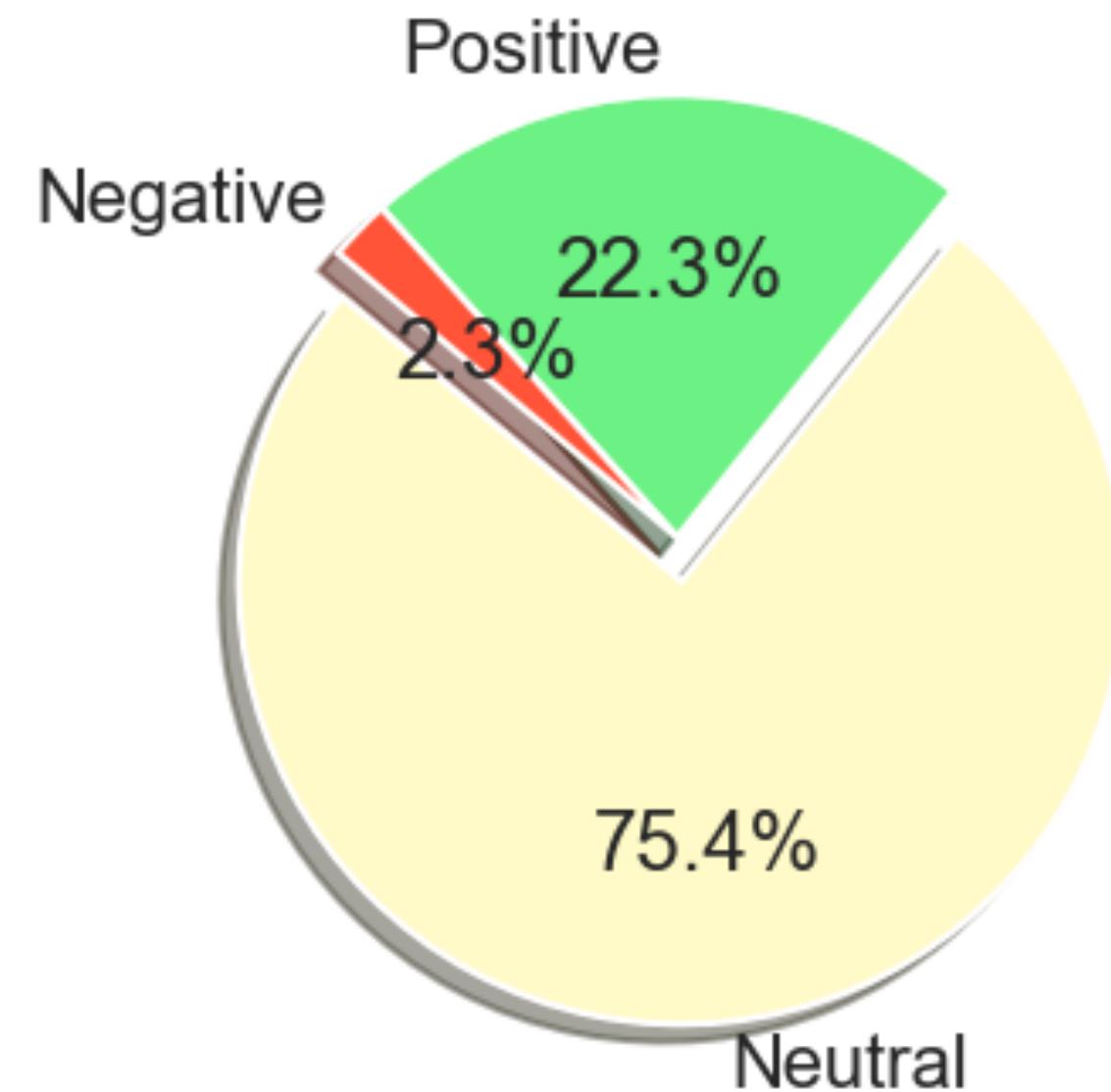
3

Sentiment Analysis



Sentiment Analysis for All  
Comments

Sentiment Category for Comments



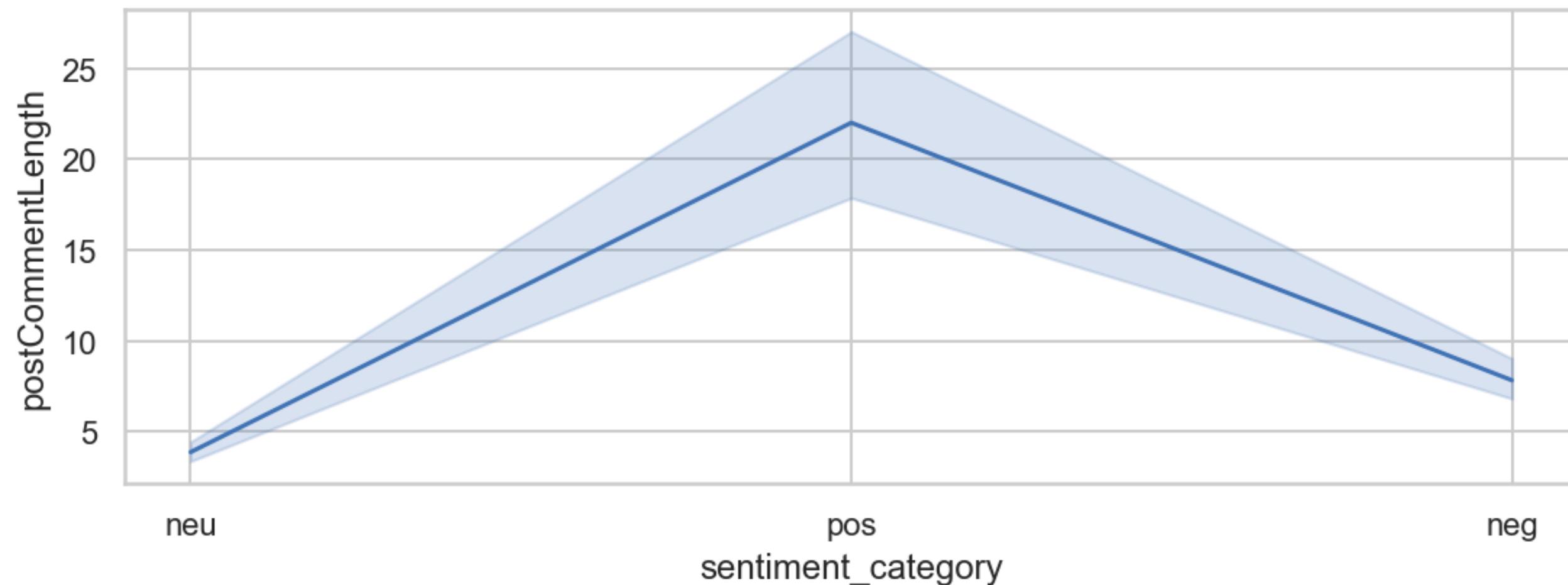
# EDA for Comments Attributes

4

Sentiment Category against  
Length of Comment



Sentiment Category against Length of Comment

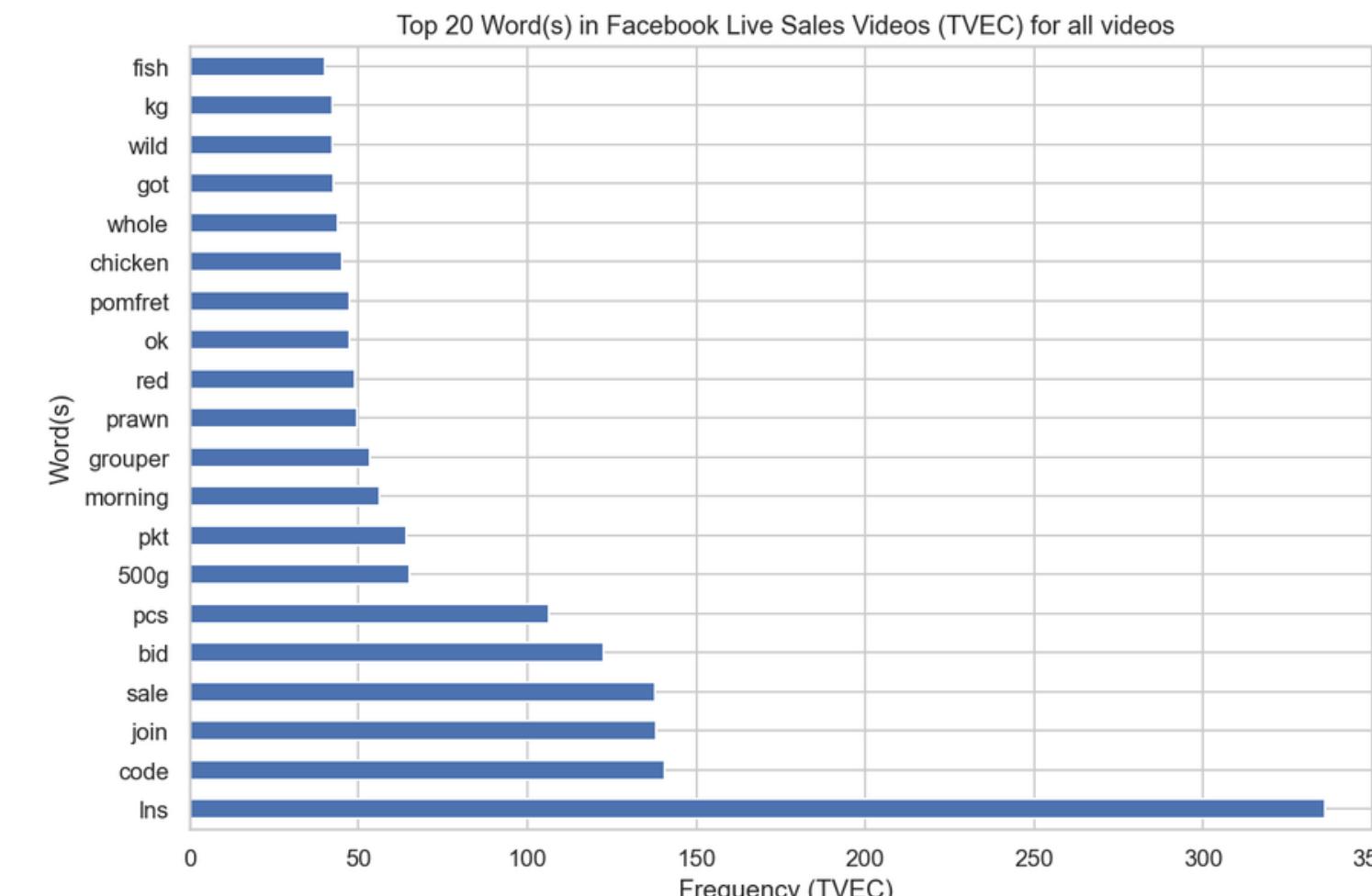
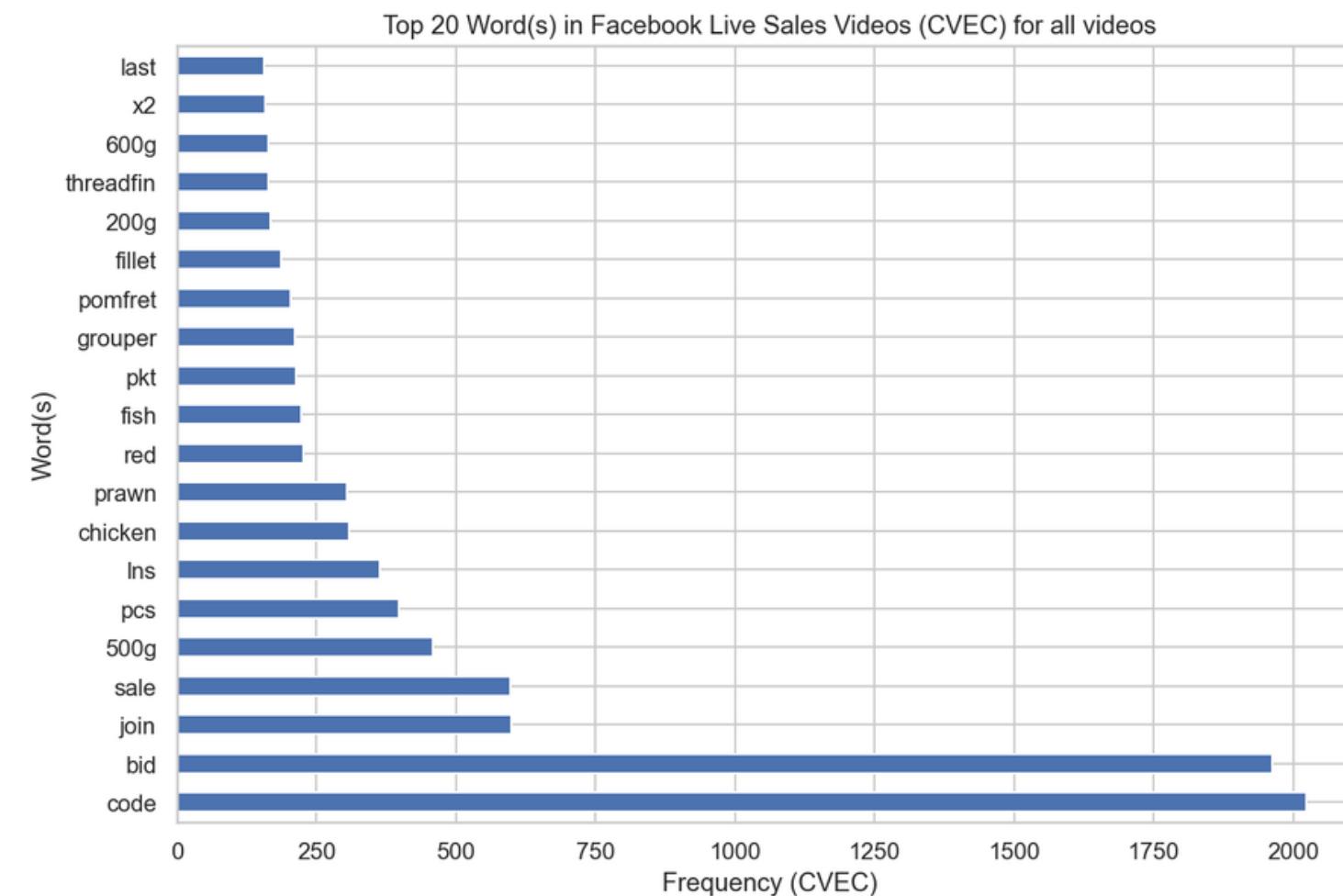


# EDA for Comments Attributes

## 5 Ngrams Visualization for All Comments



### Unigrams for All Comments

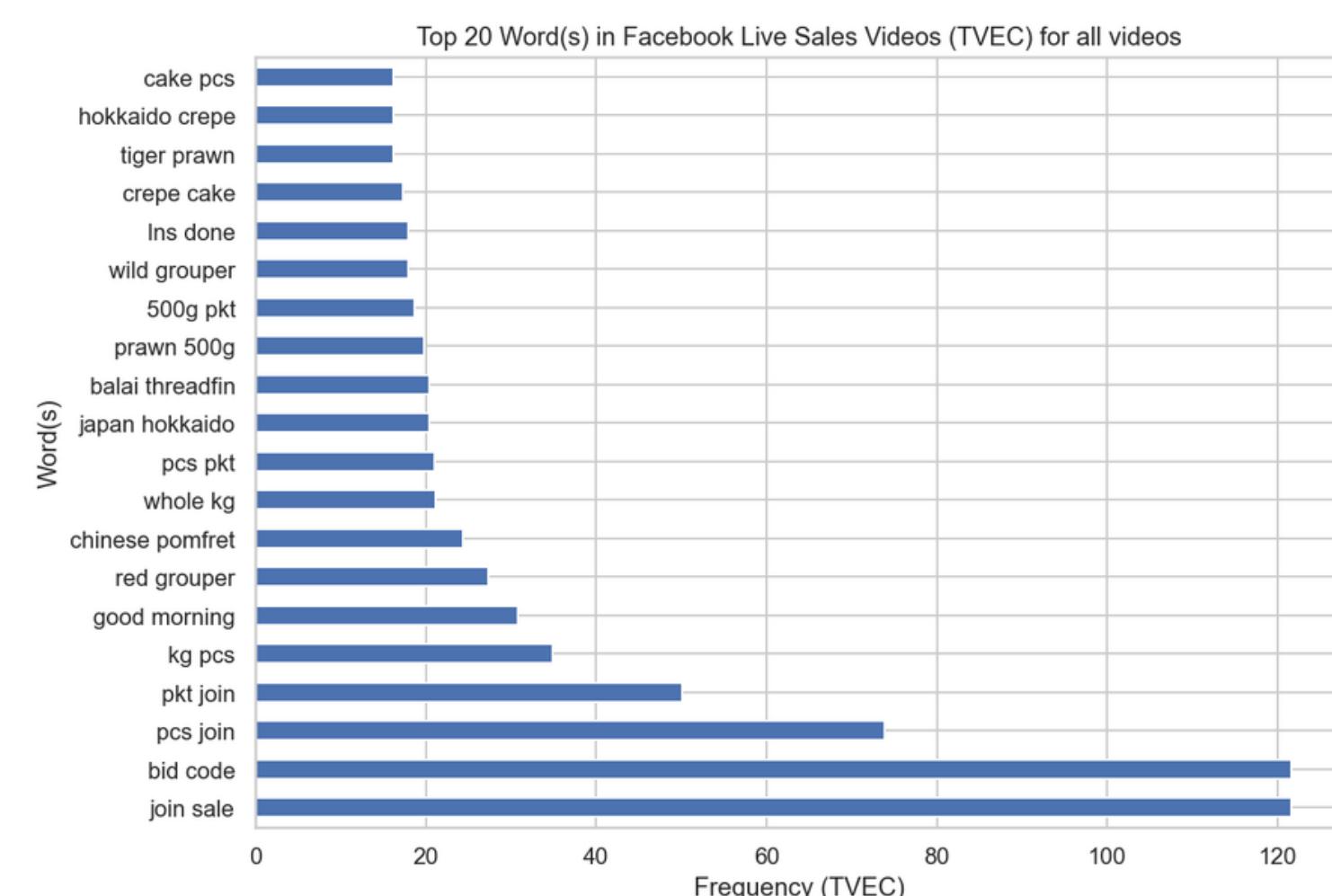
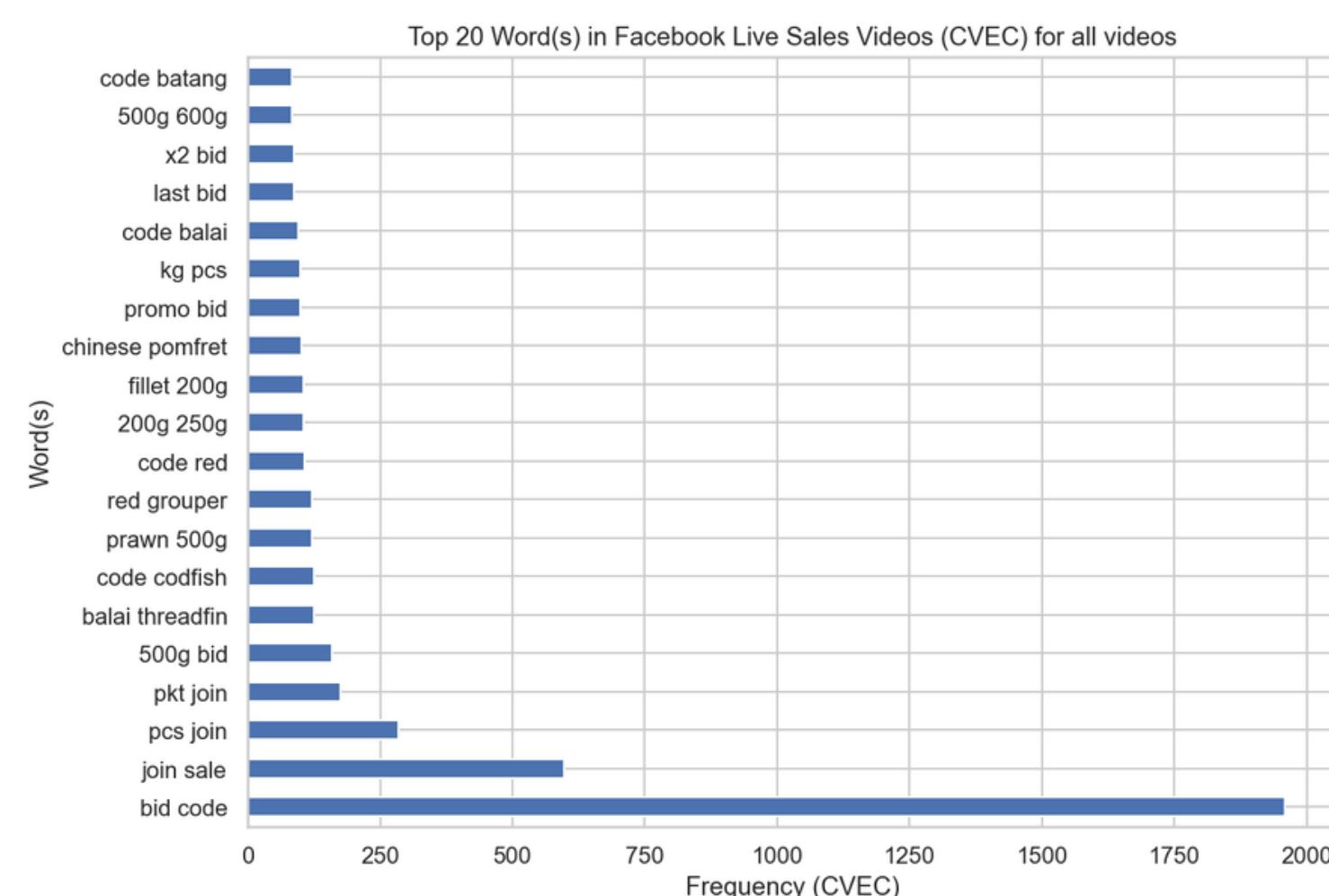


# EDA for Comments Attributes

## 5 Ngrams Visualization for All Comments



### Bigrams for All Comments

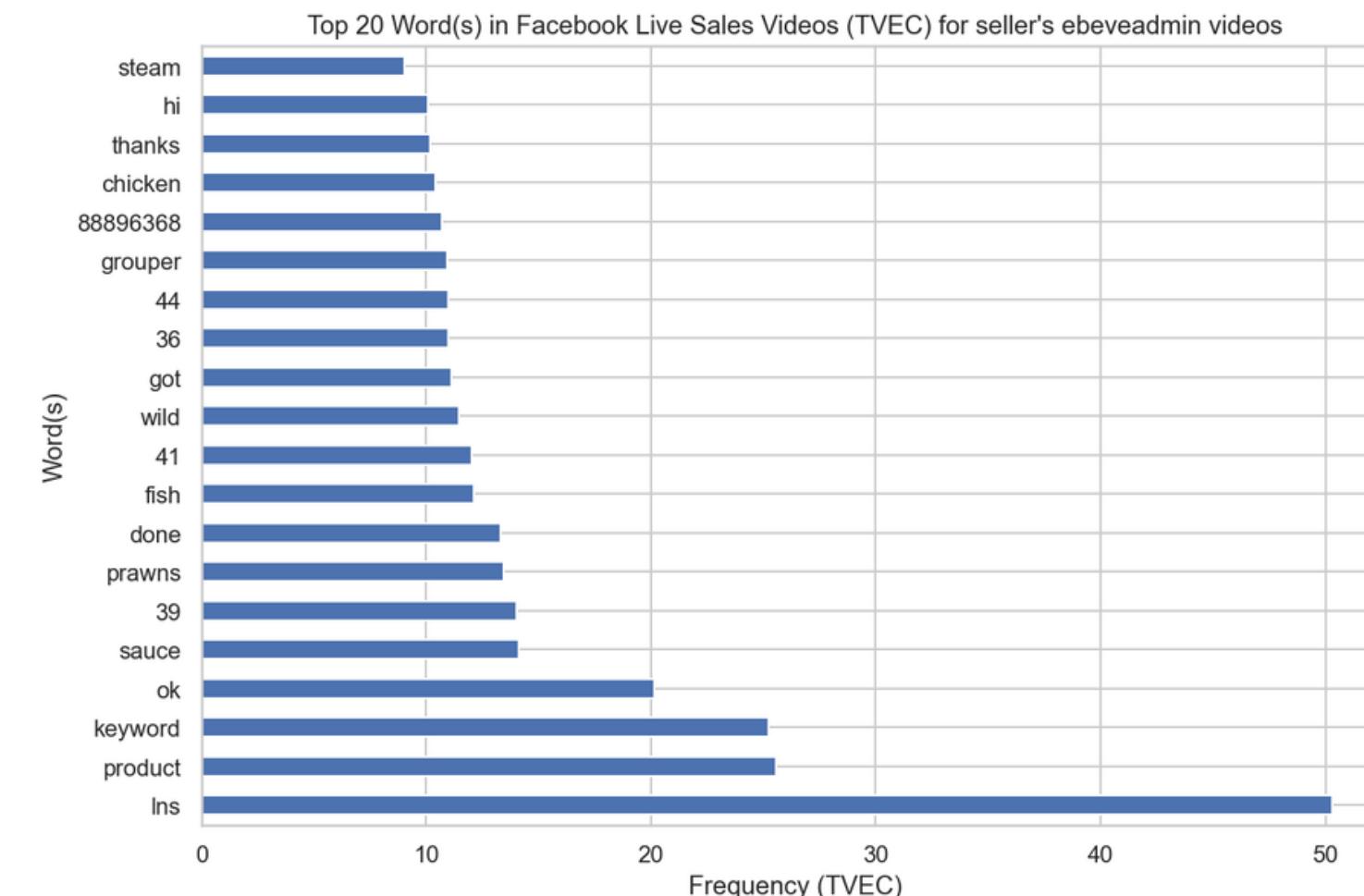
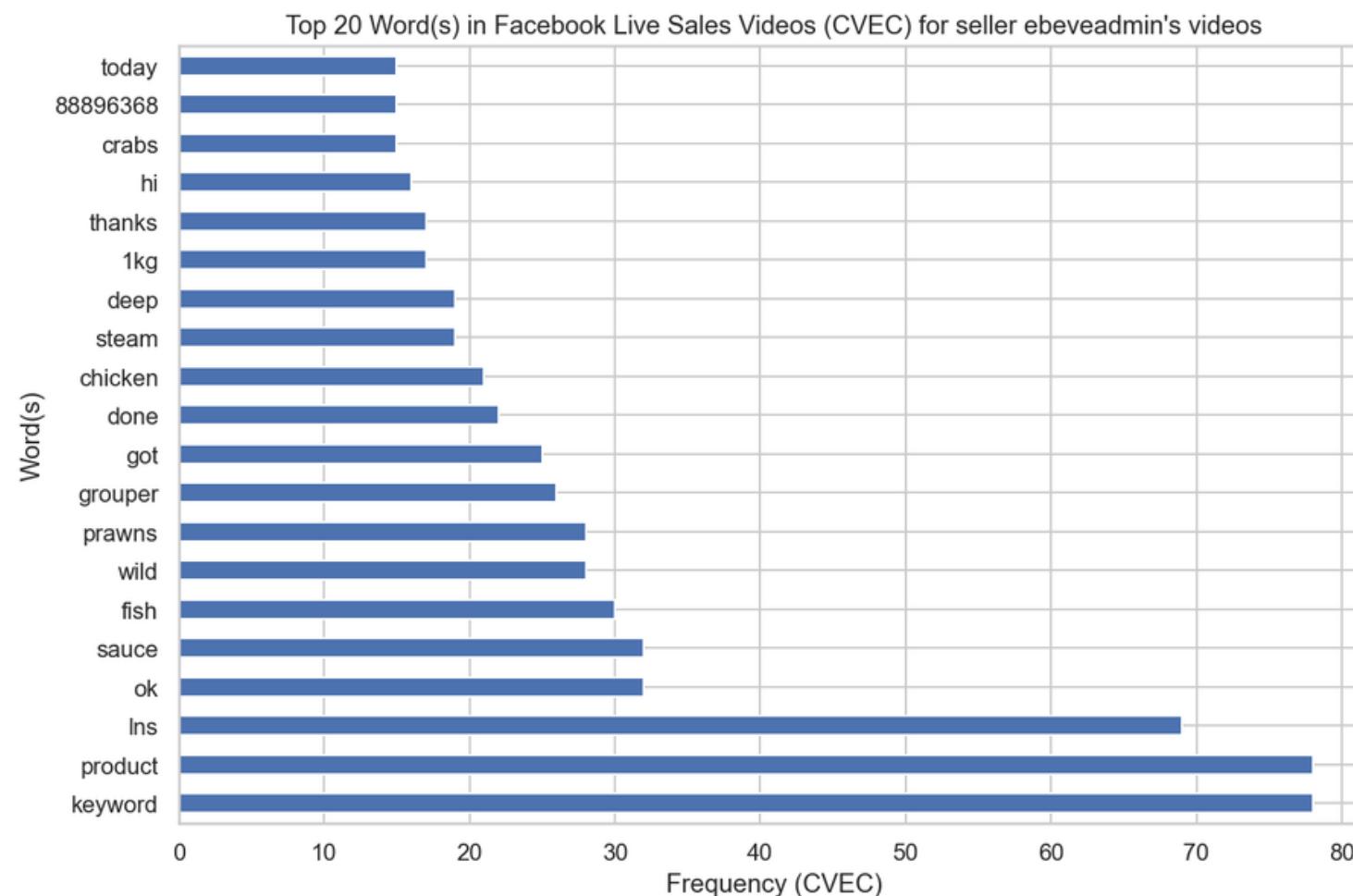


# EDA for Comments Attributes

## 5 Ngrams Visualization for All Comments for Seller E-beve



### Unigrams for All Comments for Seller E-beve

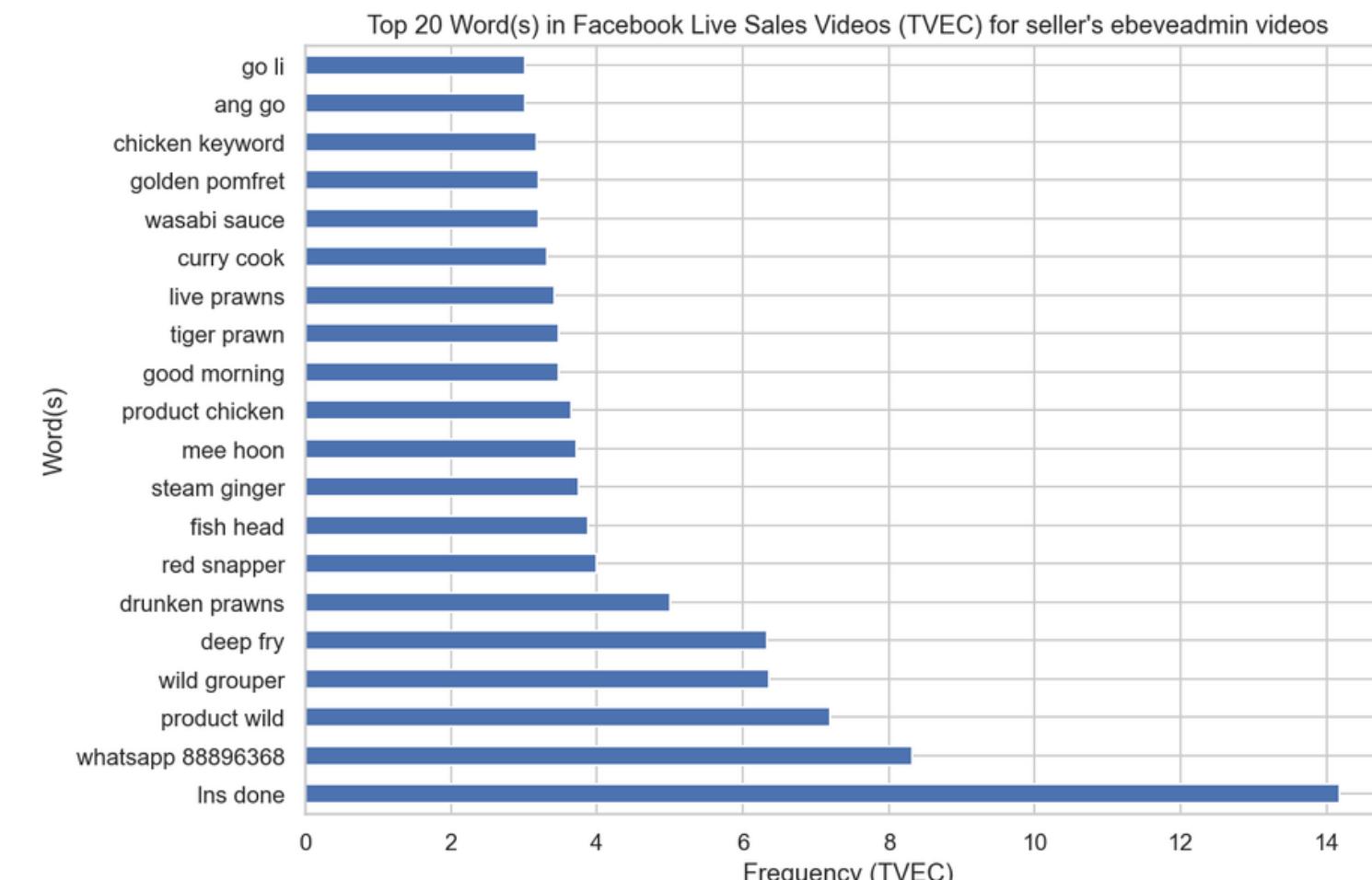
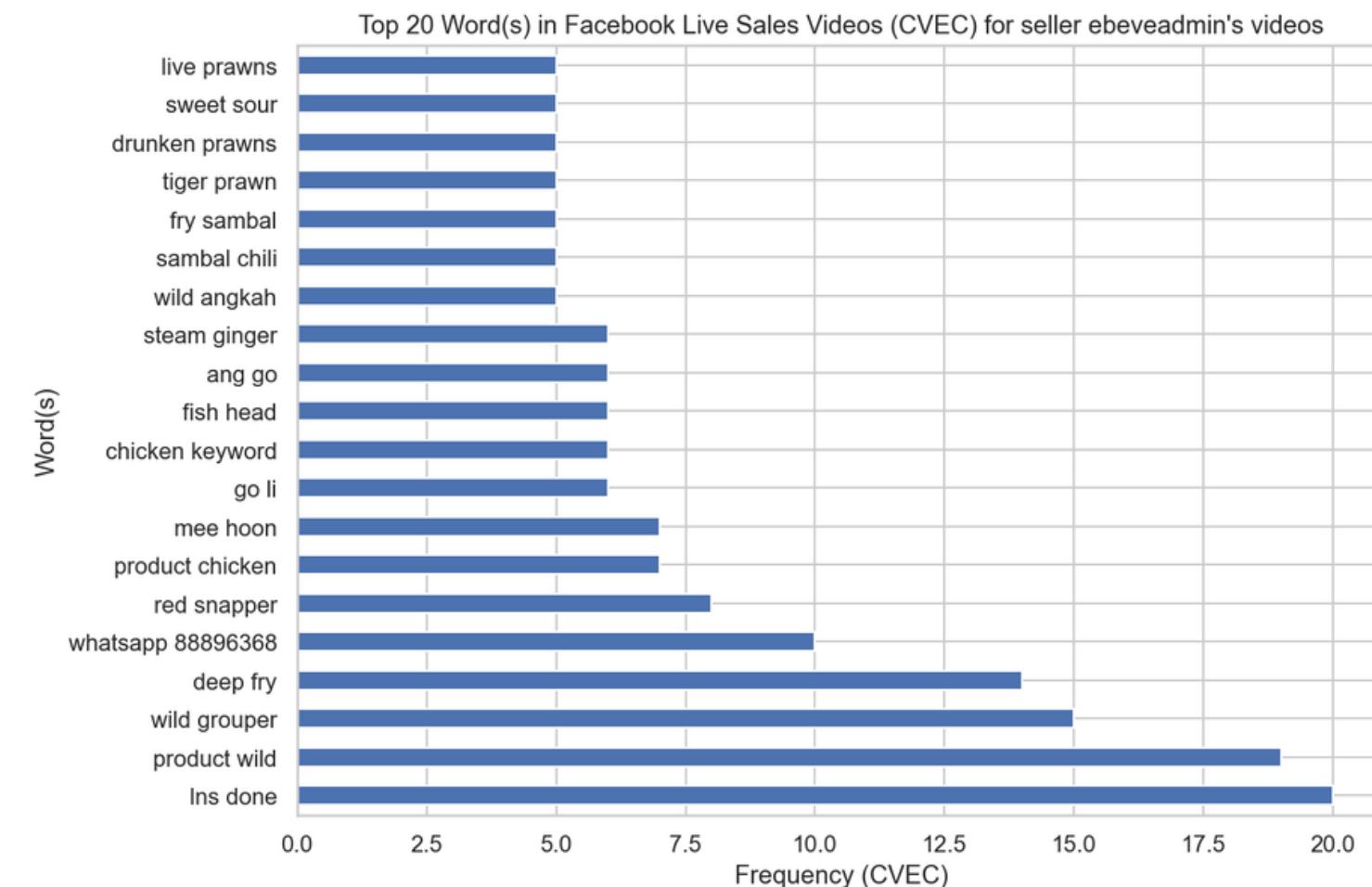


# EDA for Comments Attributes

## 5 Ngrams Visualization for All Comments for Seller E-beve



### Bigrams for All Comments for Seller E-beve

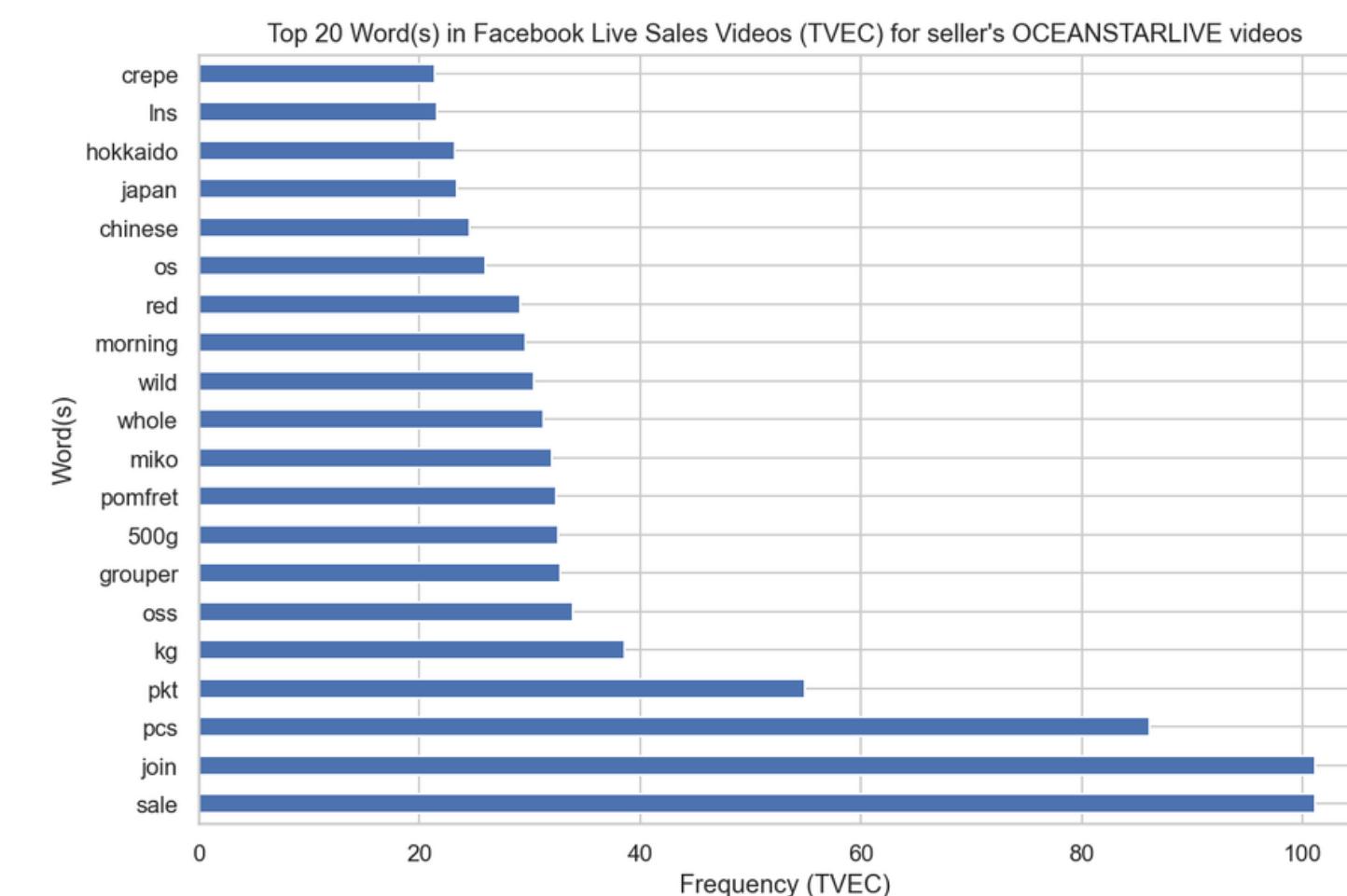
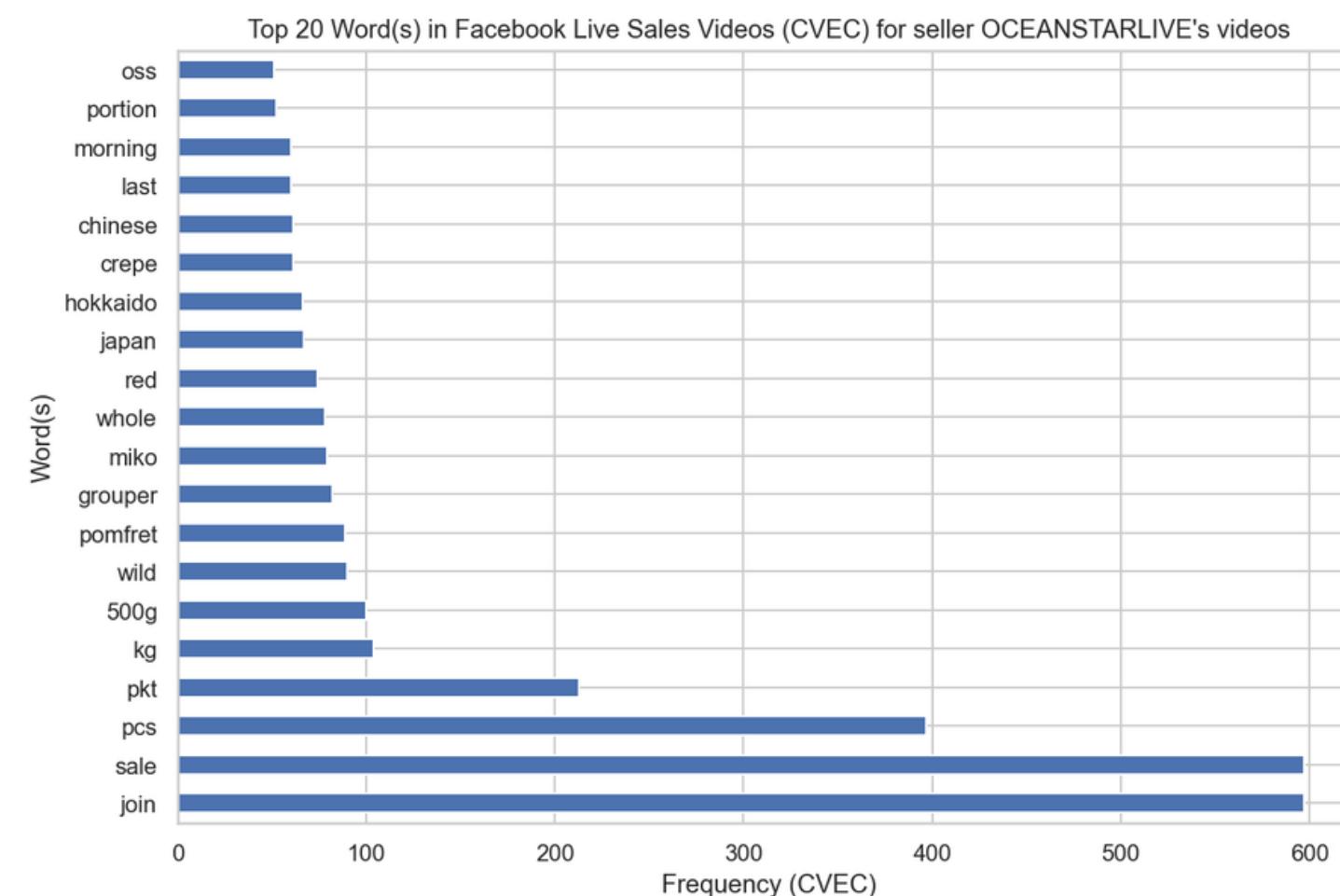


# EDA for Comments Attributes

## 5 Ngrams Visualization for All Comments for Seller OCEANSTARLIVE



### Unigrams for All Comments for Seller OCEANSTARLIVE

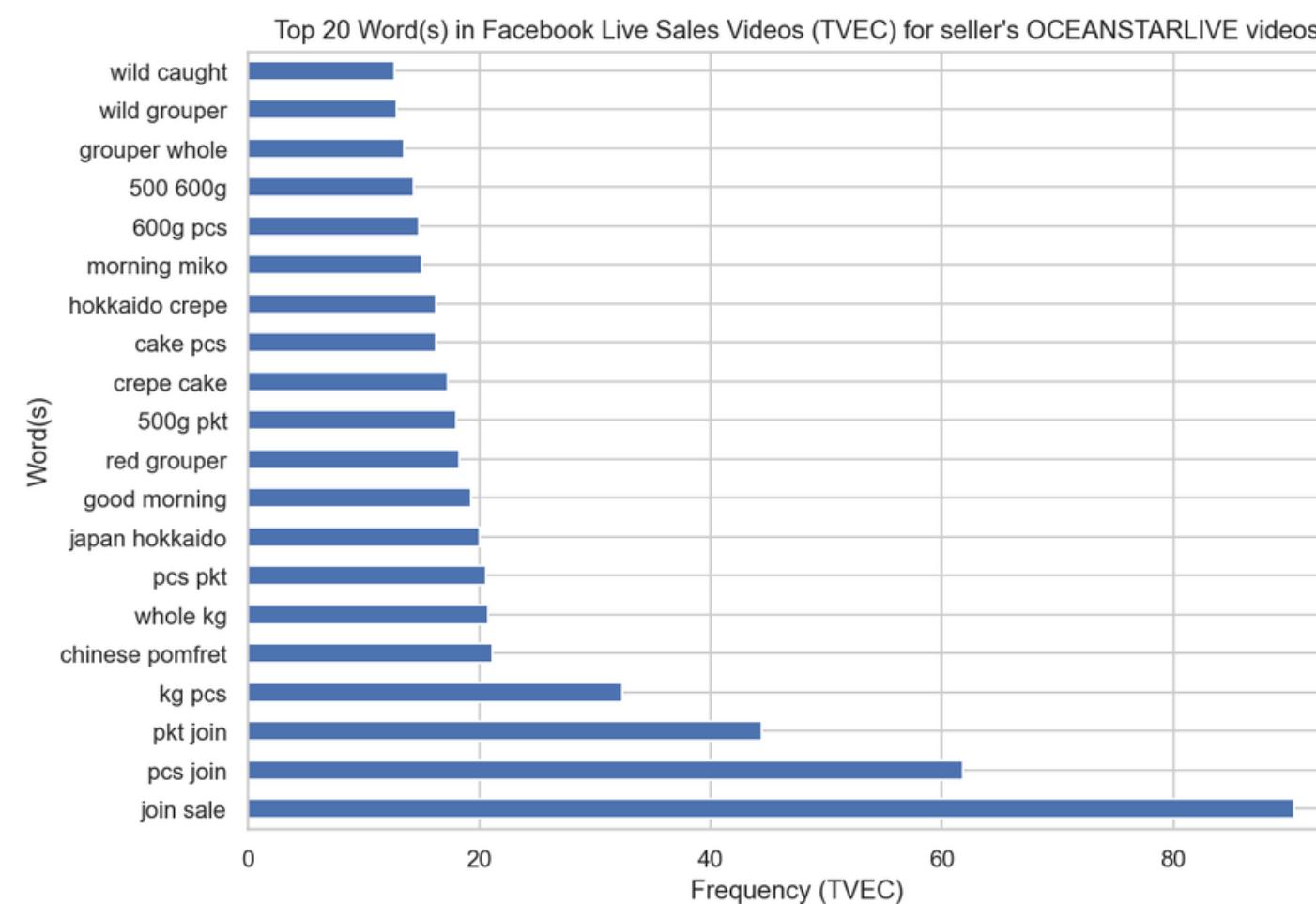
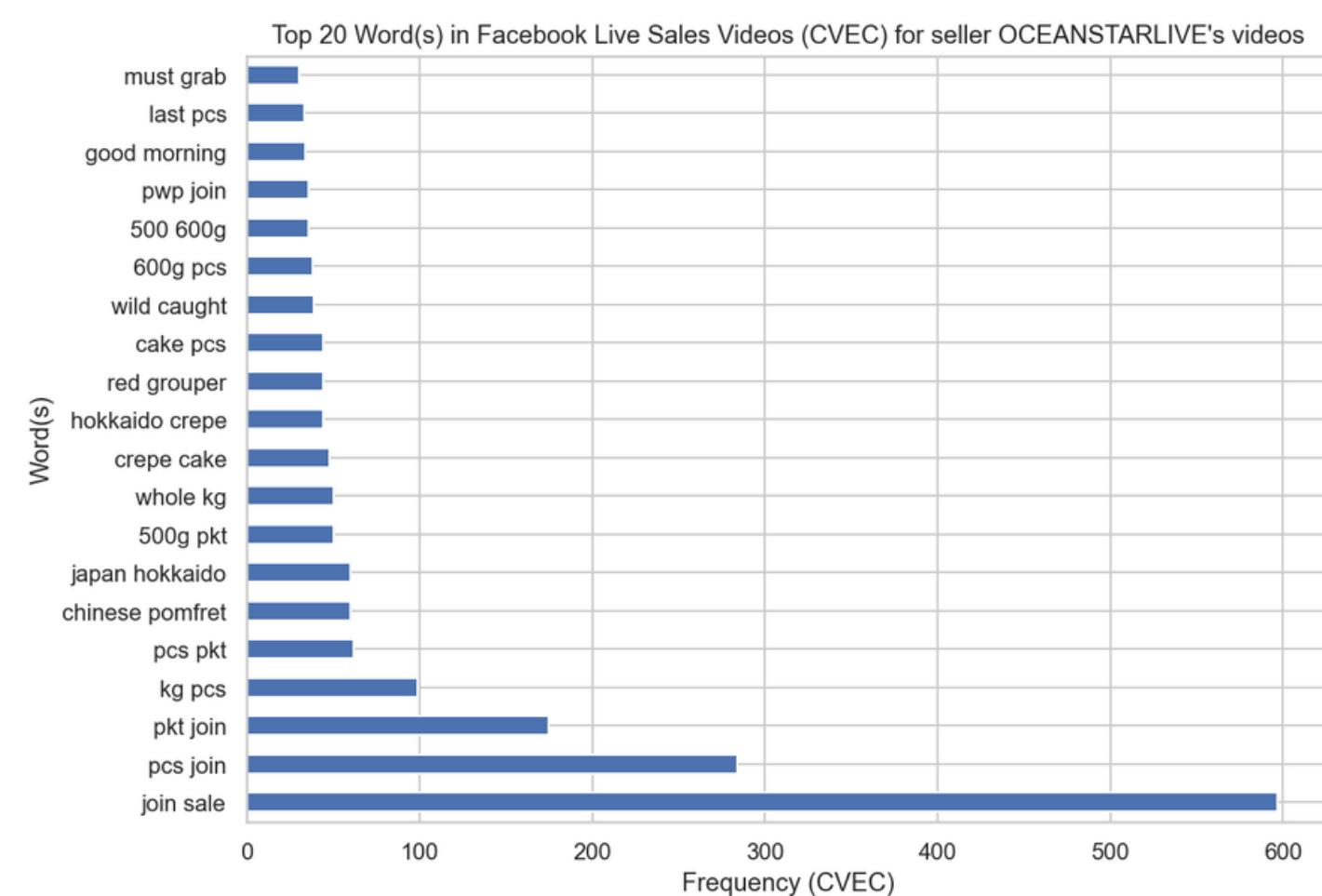


# EDA for Comments Attributes

## 5 Ngrams Visualization for All Comments for Seller OCEANSTARLIVE



### Bigrams for All Comments for Seller OCEANSTARLIVE

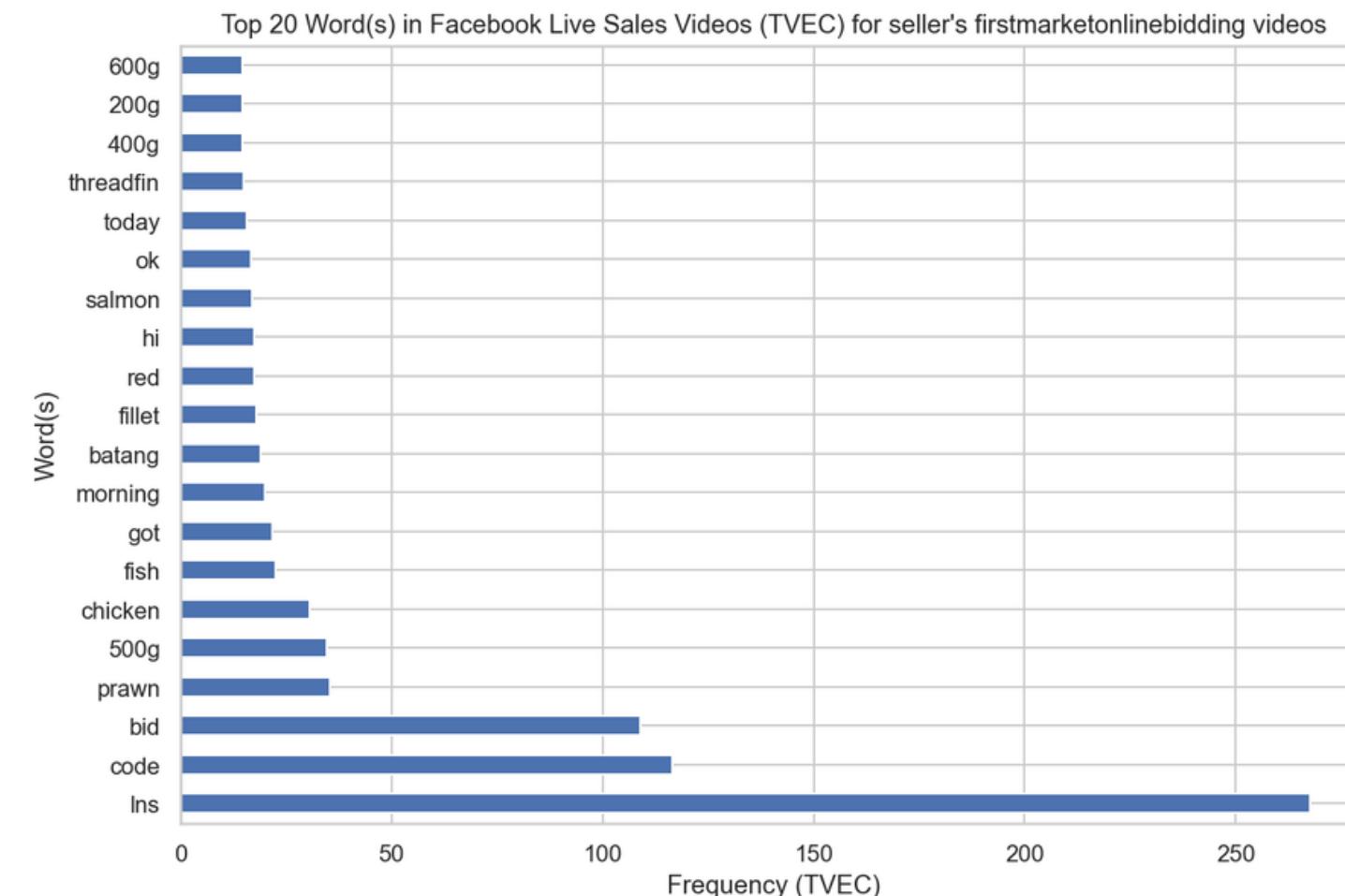
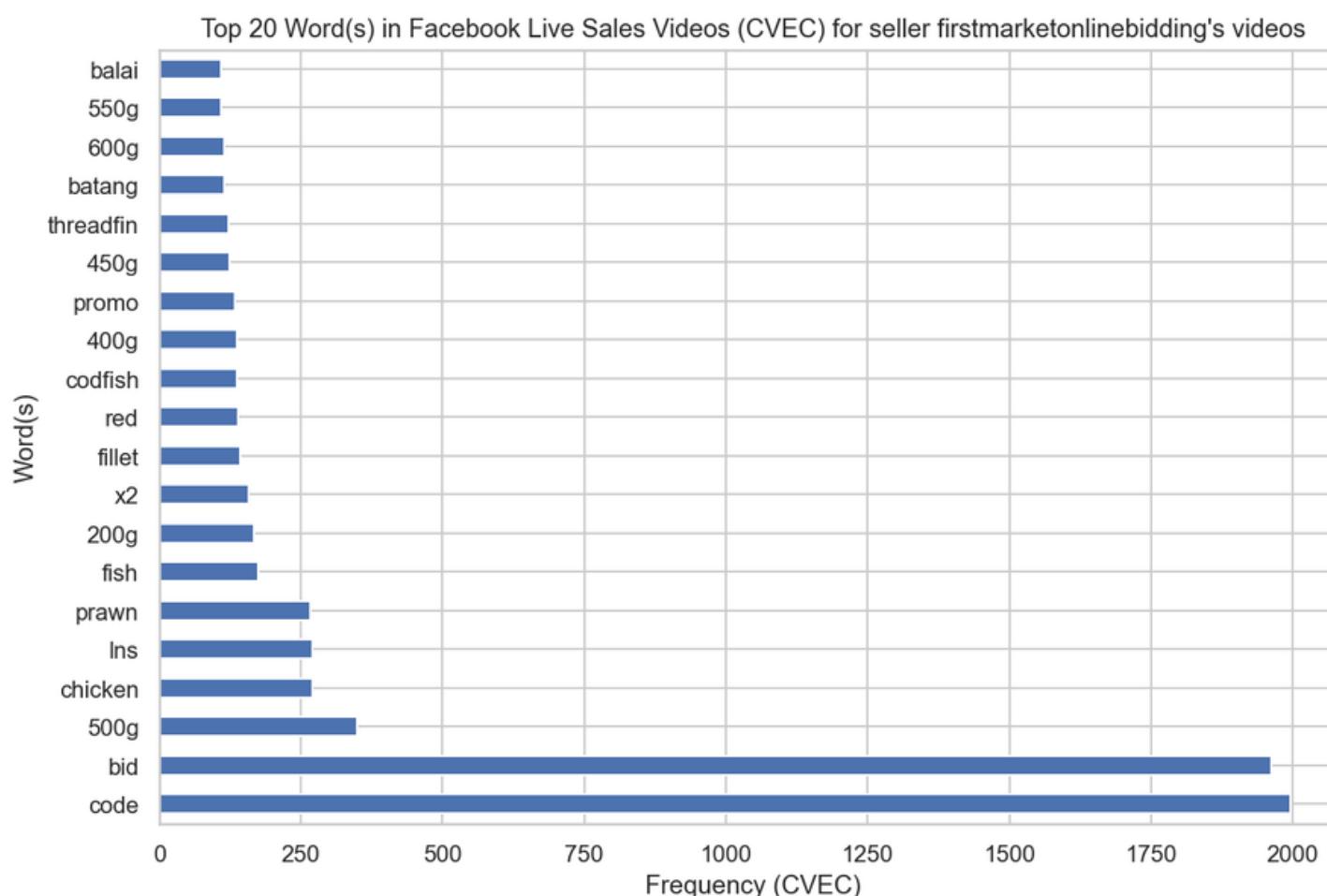


# EDA for Comments Attributes

## 5 Ngrams Visualization for All Comments for Seller firstmarketonlinebidding



### Unigrams for All Comments for Seller firstmarketonlinebidding

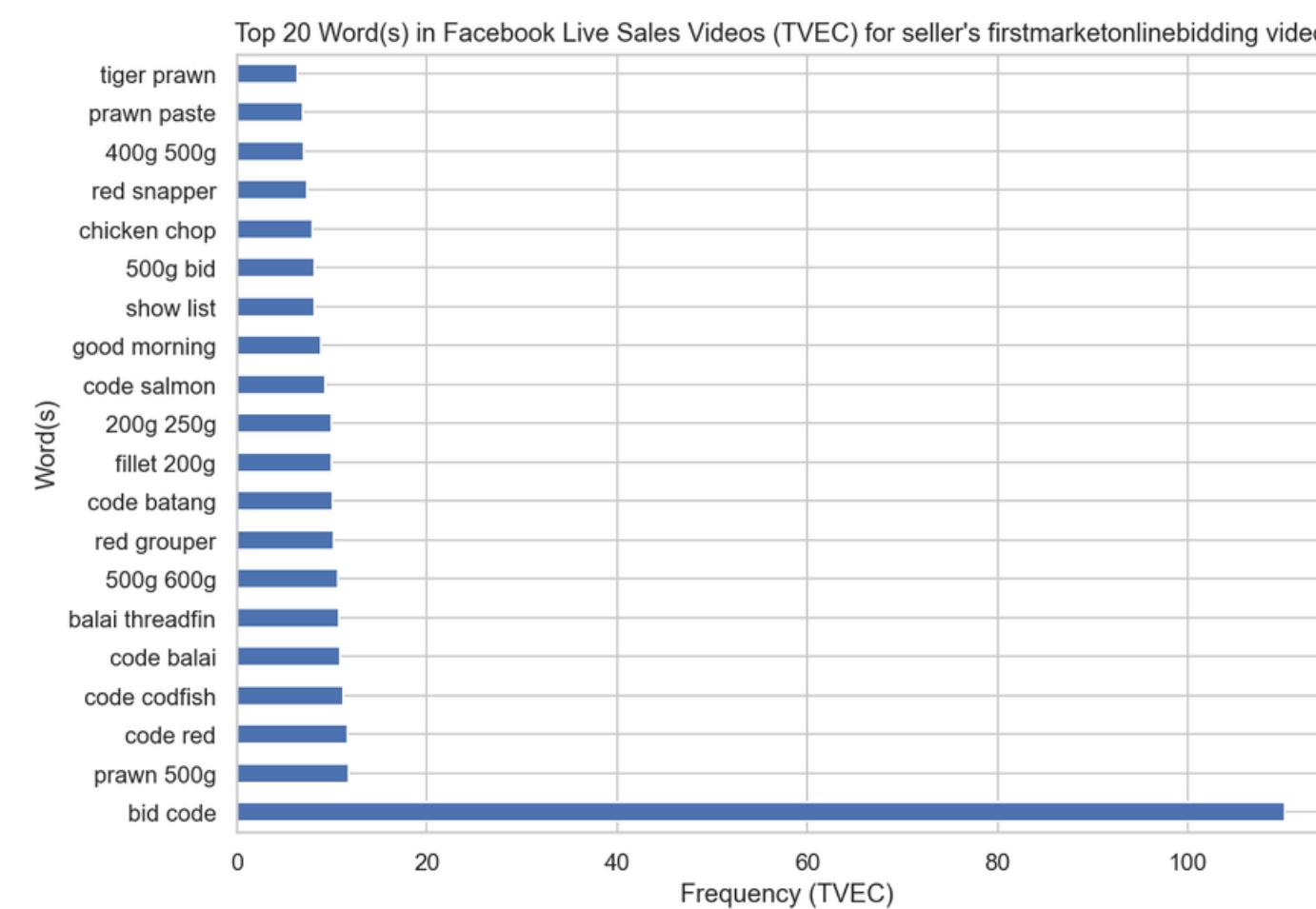
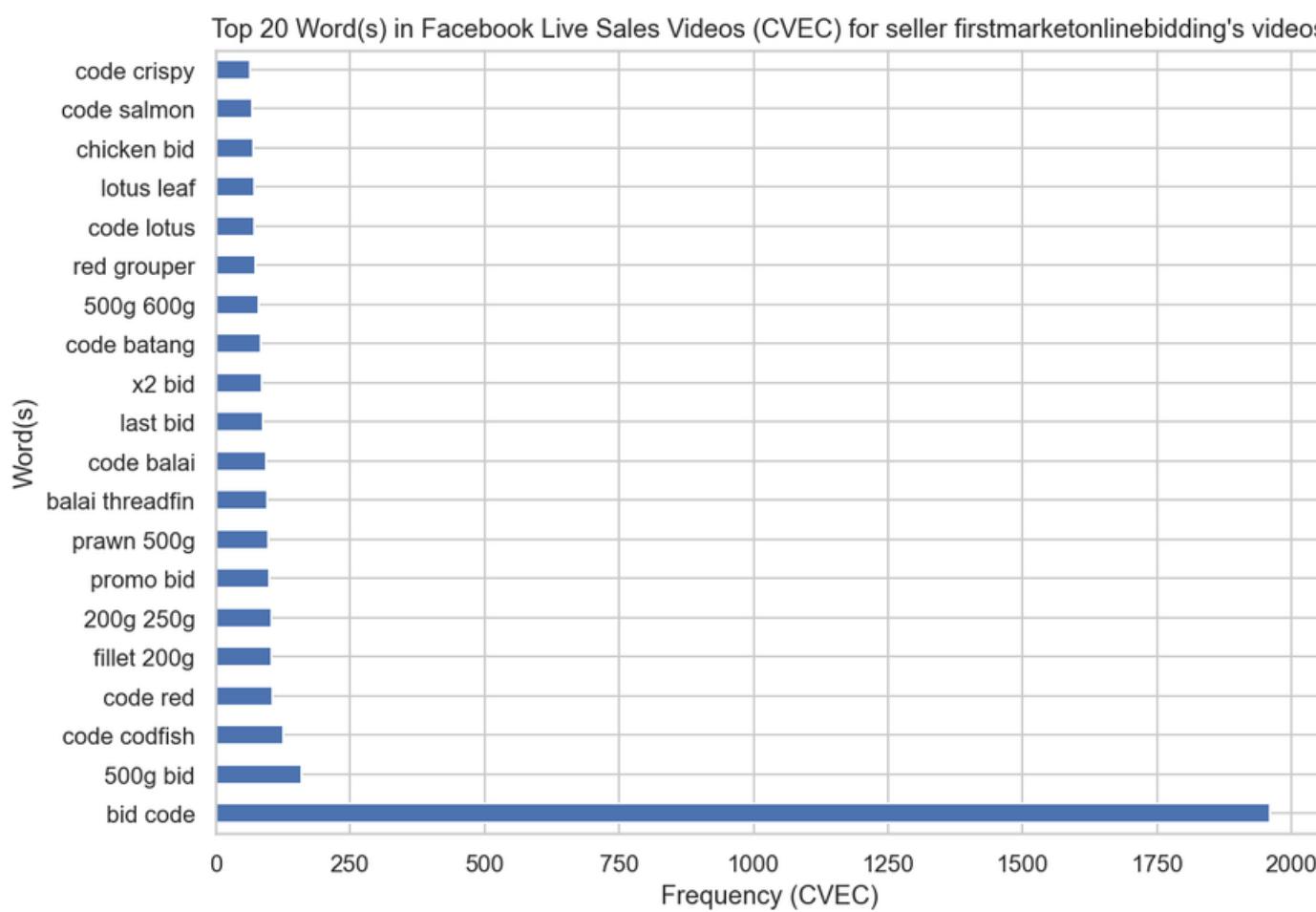


# EDA for Comments Attributes

## 5 Ngrams Visualization for All Comments for Seller firstmarketonlinebidding



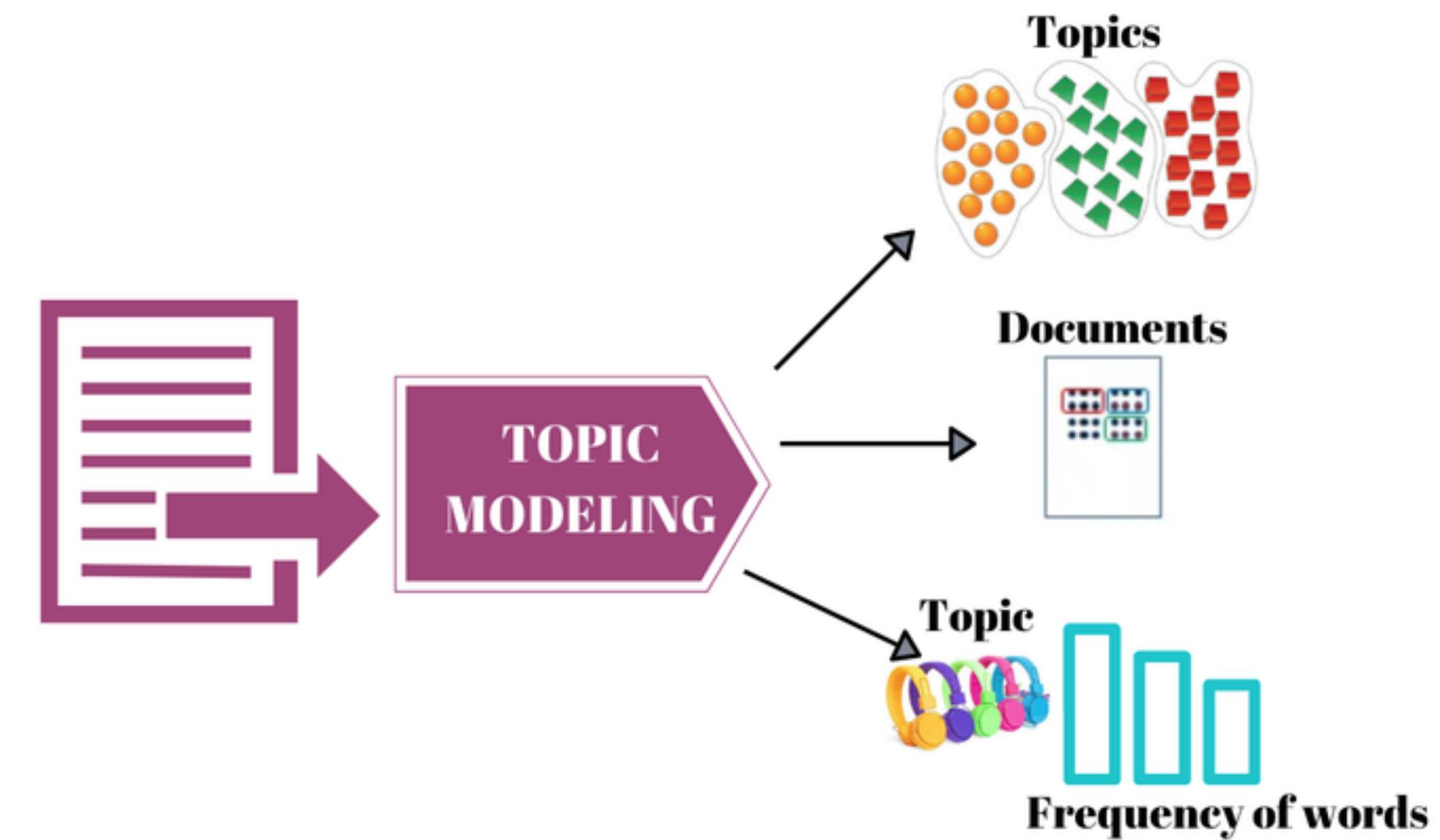
### Bigrams for All Comments for Seller firstmarketonlinebidding



4

# Modelling: LDA

# Latent Dirichlet Allocation



# **Sentiment Analysis of each Seller**

o o o o

1

2

3

Neutral Sentiments Themes

Positive Sentiments Themes

Negative Sentiments Themes

# Seller E-Beve: Neutral Sentiments



0 nan sauce steam crabs garlic ginger hoon mee soup 44  
• **main ingredients to cook crab mee hoon soup**

1 Ins done 88896368 sis hi 41 36 whatsapp 47 46  
• **seller's contact details.**  
◦ "hi sis, Ins done. whatsapp 88896368"

2 congrat seabass go ang tiger li see duri 50 babe  
• **common topic of the fishes the viewers are most interested to see**  
◦ **Seabass, Ang Go Li (aka white snapper) & Duri fish (aka Dory)**

3 gongxi emily wong grouper pomfret black wild much 38 hello  
• **congratulations to a viewer who has possible won a giveaway and popular fishes that viewers enquire about**  
◦ "gongxi emily wong", "hello black pomfret & wild grouper how much"

4 product keyword wild prawns fish 1kg got chicken live deep  
• **product keywords enquiries for specific products & weights.**  
◦ "live wild prawns", "deep fish", "1kg chicken".

## > Theme:

- **different types of dishes that the viewers would be most interested in cooking,**
- **product enquiry & general enquiry.**

# Seller E-Beve: Positive Sentiments



- 0 ok thanks yes want steak cut xie later pls congrats
  - **viewers' responses to the sellers and the giveaway winner**
- 1 sweet fresh slice also cook fish super tender mussels got
  - **brainstorming ideas on what ingredients they could possible purchase to cook together**
    - **"can cook the super tender sliced fish, and also got fresh & sweet mussels"**
- 2 pls belly want hahaha yr lala white number wan fresh
- 3 good wow ok thank tq nice lah de ph service
- 4 pls winner prawns lol code liao cut rat descale kham

> **Theme:**

**Trend that the viewers want fresh food and are pleased with the services provided by the seller like cutting the products and descaling the fishes.**

# Seller E-Beve: Negative Sentiments



- 0 pay oh liao bath ill winner hardtail today sleep flower
  - **viewers' responses to the sellers and the giveaway winner**
- 1 cut collect wrong wonder really steak items lea pls head
  - **wrong item could possibly be collected wrongly by the viewer or that the product had been cut in a way that displeases the customer.**
- 2 incoming pathetic buy ya fillet cut steak lazy crab call
  - **viewers' responses to the sellers and the giveaway winner**
- 3 worries die well sotong see bit collect cos limited need
- 4 message much rules kill dunno order bought monitor quantity keyword
  - **viewers are perhaps being overwhelmed by the keywords for ordering the products and have a difficult time following & monitoring which keyword to enter in the comments section for ordering**

> **Theme: Viewers being worried and receiving dissatisfactory service.**

# Seller OCEANSTARLIVE: Neutral Sentiments



- 0 today offer fry tml lj red miko liao hello grouper
- 1 bye pork marinated code cook work come chicken pau qm
- 2 nan miko oss morning lns team gd mrg ohaiyo everyone
  - **greetings**
- 3 os lns try friday black king salmon soon tried tail
  - **product has been out of stock ('os') and they will try to purchase the tail of the king salmon during black friday then.**
- 4 got fillet fish buy soup give salmon one leh still
  - **when they purchase fish fillet soup, they are given the salmon as well**

> **Theme: random remarks & greetings**

# Seller OCEANSTARLIVE: Positive Sentiments



- 0 iqf ok chicken fish ribs prime mid joint hahaha miko
- 1 pkt japan hokkaido 500g crepe cake join sale portion 200ml
  - **Sale for Hokkaido Crepe Cake**
- 2 sale join wild 500g caught good morning ang miko pcs
- 3 buy crab male mud beef aus 450g meltique 350 steak
  - **male mud crab & australian beef**
- 4 sale join pcs pkt kg pomfret grouper whole last chinese
  - **Sale for whole fishes like pomfret and grouper**

> **Theme: Interest in the sale of particular products that might be a crowd favourite or rare**

# Seller OCEANSTARLIVE: Negative Sentiments



- 0 time crepe link trout send batang find mala regret wait
  - **regret waiting for the products like the trout fish, mala batang fish and crepe**
- 1 shift problem la steak rainy red still cod tail son
- 2 space pkt buy sui mths qiao acc cut pregnant recently
  - **viewers got pregnant recently**
- 3 pkt sale offer leh join pwp sorry max crazy week
  - **viewers had a crazy week and are unable to join the sale for the purchase with purchase offer.**
- 4 pkt jackfruit sale pwp join crazy acc deal 500g max

> Theme: viewers' regret and their general mood

# Seller first market online bidding: Neutral Sentiments



- 0 code bid 500g 200g balai fillet threadfin salmon batang codfish
  - **different kinds of fishes**
- 1 got show wing list otah prawns hokkaido muar popiah get
  - **overseas imported food**
    - **hokkaido prawns, and muar otah, popiah & wings**
- 2 chicken prawn paste chop crispy marinated cake pork crab squid
  - **enquiring of specific marinated products**
    - **prawn paste chicken chop, squid & crab cake, and marinated pork**
- 3 fish still ball sauce much today frozen japanese tofu crispy
  - **enquiring of premade & processed food li**
    - **frozen food & sauces: fish ball, frozen japanese crispy tofy, sauce**
- 4 nan Ins hi today sashimi got red grouper lobsters see
  - **enquiring of seafood.**

> **Theme: product enquires**

# Seller first market online bidding: Positive Sentiments



- 0 good brother said saw job heard list oh admin haha
  - **job listing possibly posted by the seller.**
    - "admin, heard from my brother that he saw a job list haha".
- 1 bid code ok chicken thanks 500g prawn fish x2 chilli
- 2 code bid last chicken x2 dumpling prawn 450g leaf lotus
- 3 thank prawn share list grey always fried want tail pork
- 4 code bid 500g prawn fish chicken 600g red x2 450g

> **Theme: products that the customers are pleased with, especially prawn**

# Seller first market online bidding: Negative Sentiments



- 0 rambutan cut batang want today become soup time red slice
- 1 red fillet grouper got scallop sorry want pse around small
- 2 cut batang soup steak mean tmro watch bid yday get
- 3 snapper crab grouper confirmation today cover angry minced salmon dont
  - **certain products were not provided in the previous live sales video**
  - **"don't angry, today confirm cover snapper, crab, grouper"**
- 4 missed times prawn sorry notification tel provide yr ok think
  - **viewer missing the notification time, and being asked to provide contact details**

> **Theme: customers' disappointment, and some words with no connection.**

# Neutral Sentiments Summary



## Common Neutral Sentiment Themes:

- Product Enquiries



Neutral Sentiments

### Seller E-Beve:

- different types of dishes that the viewers would be most interested in cooking,
- types of fishes,
- product enquiry & general enquiry.

### Seller OCEANSTARLIVE:

- greeting the seller,
- random remarks such as:
  - remarking that they will try to purchase a product that has been out-of-stock on another day, and
  - the ingredients they are given when purchasing fish soup.

### Seller firstmarketonlinebidding:

- product enquiries

# Positive Sentiments Summary



## Common Positive Sentiment Themes:

- Satisfactory products and services provided by the seller



### Positive Sentiments

#### Seller E-Beve:

- congratulating giveaway winners,
- being pleased with the Seller's service, &
- being satisfied with the quality of the products provided by the Seller.

#### Seller OCEANSTARLIVE:

- viewers' interest in specific products, and
- viewers' interest in last piece sales for certain products.

#### Seller firstmarketonlinebidding:

- specific products, and
- random remarks on a job listing offered by the seller.

# Negative Sentiments Summary



## Common Negative Sentiment Themes:

- viewers' general mood
- viewers' regrets towards the sale



Negative Sentiments

### Seller E-Beve:

- viewers' worries,
- viewers' having received unsatisfactory services previously,
- being confused at the keywords.

### Seller OCEANSTARLIVE:

- viewers' regret in waiting for particular products during the Live Video,
- viewers' general mood

### Seller firstmarketonlinebidding:

- customers' disappointment, and some words with no connection.

5

# **Conclusion , Recommendations & Further Steps**

# Conclusion



- |  |   |   |  |
|--|---|---|--|
| 1  | 2   | 3   | 4  |
| <b>Confidence pays!</b> <ul style="list-style-type: none"><li>Genuine customer engagement does have a positive impact on the total revenue of the sales.</li></ul> | <b>Take things in moderation.</b> <ul style="list-style-type: none"><li>Too many unique products could confuse the viewers.</li><li>Commenting too often can be a turned-off.</li></ul> | <b>Consistently provide adequate product information, and deliver quality products &amp; services</b> | <b>Pay attention to the viewers' moods, and negative remarks</b> |

# Recommendations

• • • •

- Be bold in verbally engaging with the customers
- Provide recipes on how to cook the products offered by the seller
- Suggest products & recipes in accordance to the viewers' mood
- Continually provide satisfactory products & services as feedback by the customers
- Mixing the order of the sale of their products
- Use promotional words both verbally and in the comments section
- Provide logical product codes

# Further Steps

o o o o

- 1 Collect more data from more Facebook Live Sales Video
- 2 Collect data from videos of varying lengths, total number of comments and start time
- 3 Transcribe the videos
- 4 Use topic modelling models that are transformer-based as well

**Thank  
you!**

o o o o

Have a great  
day ahead.