Mini Project 1
Analysis of Twitter Feed
Brand Comparison: Comparing with ELLE,
what is the value of VOGUE
Yue Liu

Product Mission

For women,
Who want to know the latest news about fashion,
The Brand Comparing is a recommendation system,
That need to know the differences between the two magazines,
Unlike a blog written by an editor,
Our product makes a through comparison via social media.

We want to give the users an insight about their brand with the main competitors. In this case, we served the largest fashion magazine in the world, Vogue(real data, but a different story). Vogue is a monthly magazine with 126 year history, it is the most popular fashion media globally. What can we do to have a digital magazine about worldwide fashion like Vogue?

Target User(s)

Our target users are young females as well as whoever likes news about fashion, they can be women in their 50s or males who want to send gifts to their families. Knowing the users is very important, we will start from the user stories, and go back to the user stories.

User Stories

As a twitter user, I want to know the like and follow of the different brands, so that I can find the topic I like, with correct hashtag. As a twitterer likes the topic of fashion(content consumer), I want to get the products I like in the coming quarter/ year.
In order to receive benefit as an administrator, I want to compare the brand between the Vogue and Elle from sentiment analysis. It will be beneficial to conduct the topic selections to the magazine editors(content creators).

See my github webpage about this project. For me, the MVP should include two modules: (1) using twitter API, get the tweets; (2) sentiment analysis by using Google API.

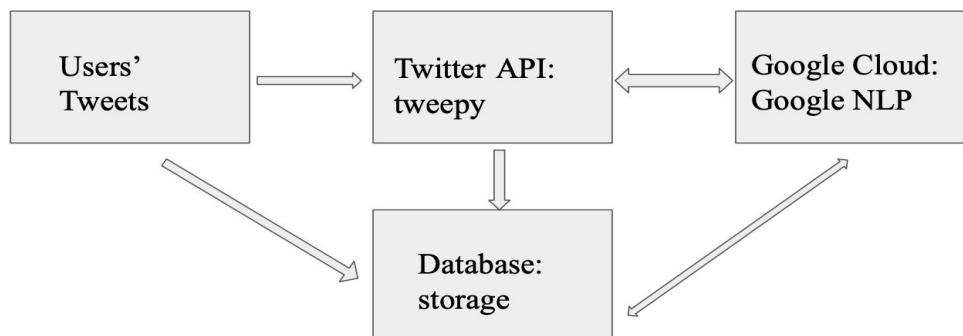**User Interface Design for main user story if required**

To make sure the product is competed, we should provide the users at least a main menu in the program in the future.

System Design and how your design addresses your user stories.

The system design is based on the framework of the two APIs. When you invoke a method from the API class, it will return to For more information, see the former drafts. I will focus on the system architecture in this part.

Getting the tweets	Analyzing the tweets
Streaming	Visualize the tweet data(tweepy)
Get the published tweets	Sentiment Analysis(google nlp)

(1) Streaming;
(2) Get the published tweets;
(3) Analyze the tweet data;
(4) Visualize the tweet data;
(5) Sentiment Analysis.

READM of how to build your system.

Related concepts

(1) Software Development Kit (SDK): A software development kit (SDK) is a collection of software development tools in one installable package. They ease creation of applications by having compiler, debugger and perhaps a software framework. They are normally specific to a hardware platform and operating system combination.

(2) Application Programming Interface (API): An **application programming interface (API)** is an interface or communication protocol between a client and a server intended to simplify the building of client-side software. It has been described as a "contract" between the client and the server, such that if the client makes a request in a specific format, it will always get a response in a specific format or initiate a defined action.

(3) Library (LIB): In computer science, a **library** is a collection of non-volatile resources used by computer programs, often for software development. These may include configuration data, documentation, help data, message templates, pre-written code and subroutines, classes, values or type specifications.

(4) Framework: In computer programming, an **application framework** consists of a software framework used by software developers to implement the standard structure of application software.

For more information, please see my github webpage.

Testing

Error conditions

(1) What happens if twitter is not responding?
401 Unauthorized (RFC 7235)Similar to *403 Forbidden*, but specifically for use when authentication is required and has failed or has not yet been provided.

```
[YUEs-Air:MiniProject1 yue$ python3 tweepy_steamer.py
401
```

In this case the server does not respond, we can solve it by regenerating the keys and secrets.

(2) What happens if I pass the wrong handle?

```
25 ●              stream.filter(track=hash_tag_list)
26   💡           print("We are here!$")
27   ⌂            #stream.filter(track=['vogue','elle'])
```

In this case, we need to debug. If you are in the terminal, the print function will help you a lot. If you are in the IDE, you can build your own test function. If the program contains so many lines, please add breakpoints. Finally, we know that there is something wrong with the parameter "track".

(3) What happen if the images I pass to FFMPEG are not images and FFMPEG fails?
I did not use the images in the project so far, just like my classmates. I still put the question into the document as a reminder. FFmpeg is a free and open-source project consisting of a vast software suite of libraries and programs for handling video, audio, and other multimedia files and streams.

(4) What happens if Google Vision API respond back?
It is about authenticating to the Cloud Vision API, which I did not use in my project.

What do you want to expose in your own API and how to map it to modules you use?

(1) Google NLP API. What do you want to expose? How are you provide this data to the user of your module?
I want to use the structured texts, including tweets, id, length, date, source, likes, retweets. In order to get the clean data, the pandas and numpy library is applied.
After that, I will use the Google NLP API for sentiment analysis. The most important part is to get the correct credential.

(2) What if there are no images in the twitter feed? What will you do?
This is a case about missing value. In my case, the images are not as important as the texts, because we are evaluating the attitudes towards a brand, not measuring the creative photos from the users.

(3) What if there are 5 images for one feed, 100 for another? What is the video you want to provide to the user of your API? What information you want to pass to the user of your API.
This is about the inconsistency of the data. In this project, I only want to use the text information. For now, it seems that not all the image group solve this problem.
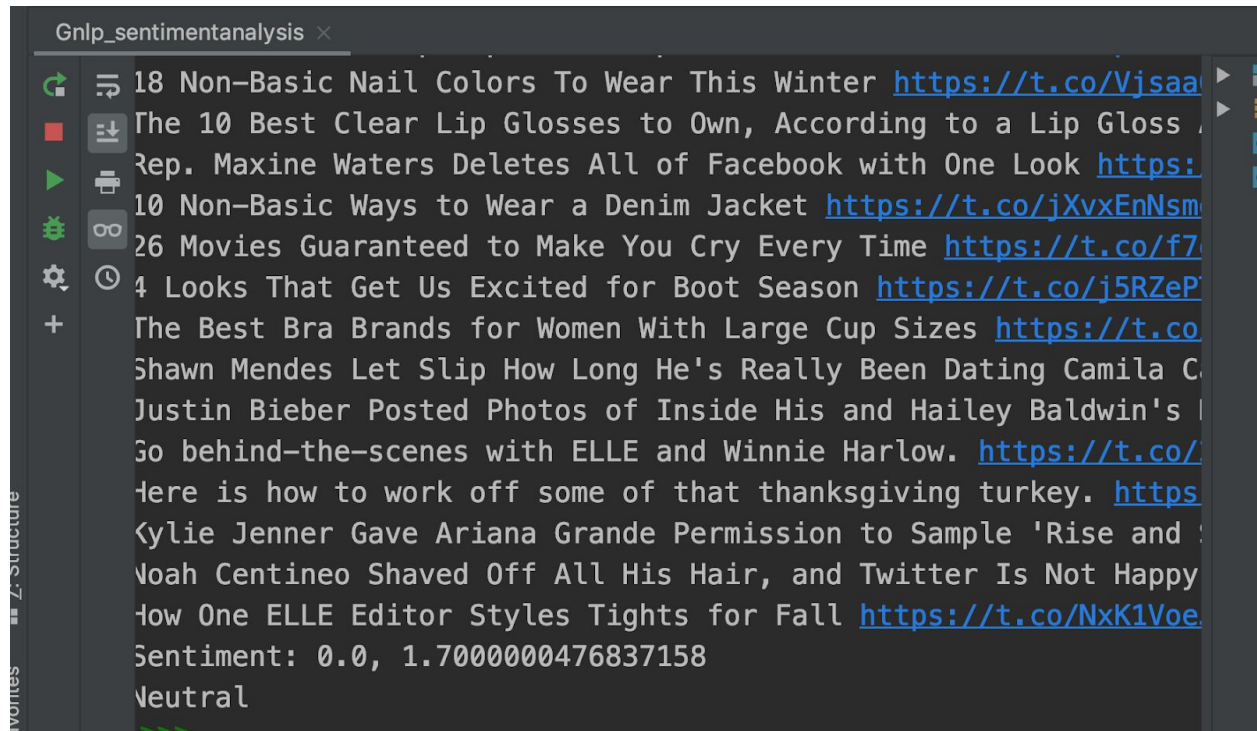(source: https://github.com/Bunny-Nora/EC601-Miniproject)

Test document providing how each test case addresses the user stories

	A	B	C	D	E	
		tweetE				
1		Tweets	id	len	date	source
2	0	20 of Lady Gaga's Best, Most Important Songs https://t.co/Atdr0Newc3	1194473004471848961	68	2019-11-13 04:32:11	SocialF
3	1	Everything We Know About the #BlackWidow Prequel Starring Scarlett Johansson https://t.co/ym1FurhMJg	1194467930412605441	100	2019-11-13 04:12:01	SocialF
4	2	All The Trends You'll Be Wearing Next Spring https://t.co/WFa74kapn1	1194463179931312128	68	2019-11-13 03:53:08	SocialF
5	3	9 Pairs of Luxury Sneakers That Are Well Worth the Money https://t.co/HeStEO4tKL	1194458383195279360	80	2019-11-13 03:34:05	SocialF
6	4	Rihanna on Being 'Happy' Dating Hassan Jameel and Why She Turned Down the Super Bowl Halftime Show https://t.co/uxkpJf5eh8	1194452963051397120	122	2019-11-13 03:12:32	SocialF
7	5	HBO Cancels Game Of Thrones Prequel Starring Naomi Watts https://t.co/FLKuClgPO3	1194447285696028672	80	2019-11-13 02:49:59	SocialF
8	6	7 Must-Know Makeup Tips for People with Sensitive Skin https://t.co/juEAkaWvnF	1194441846233149442	78	2019-11-13 02:28:22	SocialF
9	7	18 Non-Basic Nail Colors To Wear This Winter https://t.co/Vjsaa0PYgP	1194437165834493952	68	2019-11-13 02:09:46	SocialF
10	8	The 10 Best Clear Lip Glosses to Own, According to a Lip Gloss Addict https://t.co/BeRglUdfMV	1194432448580538369	93	2019-11-13 01:51:01	SocialF
11	9	Rep. Maxine Waters Deletes All of Facebook with One Look https://t.co/zeeQ2myhGG	1194428607348838401	80	2019-11-13 01:35:46	SocialF
12	10	10 Non-Basic Ways to Wear a Denim Jacket https://t.co/jXvxEnNsme	1194423901759127552	64	2019-11-13 01:17:04	SocialF
13	11	26 Movies Guaranteed to Make You Cry Every Time https://t.co/f7oRu4tNyL	1194418653812469760	71	2019-11-13 00:56:12	SocialF
14	12	4 Looks That Get Us Excited for Boot Season https://t.co/j5RZePT1vV	1194412847129939969	67	2019-11-13 00:33:08	SocialF
15	13	The Best Bra Brands for Women With Large Cup Sizes https://t.co/3VCmfYujlY	1194408255017115648	74	2019-11-13 00:14:53	SocialF
16	14	Shawn Mendes Let Slip How Long He's Really Been Dating Camila Cabello https://t.co/X9oidV4BXM	1194403708978835456	93	2019-11-12 23:56:49	SocialF
17	15	Justin Bieber Posted Photos of Inside His and Hailey Baldwin's House, Then Proclaimed 'I Wanna Sell It' https://t.co/bY1KqrZDsZ	1194398308640395266	127	2019-11-12 23:35:22	SocialF
18	16	Go behind-the-scenes with ELLE and Winnie Harlow. https://t.co/2HkSRh1MjS	1194394441169002497	73	2019-11-12 23:20:00	Twitter
19	17	Here is how to work off some of that thanksgiving turkey. https://t.co/PW12FFAN9y	1194393686278172672	81	2019-11-12 23:17:00	Twitter
20	18	Kylie Jenner Gave Ariana Grande Permission to Sample 'Rise and Shine' on One Condition https://t.co/XCpS7OD2rl	1194392446794452995	110	2019-11-12 23:12:04	SocialF
21	19	Noah Centineo Shaved Off All His Hair, and Twitter Is Not Happy https://t.co/0dYt1JVdkL	1194388346992287749	87	2019-11-12 22:55:47	SocialF
22	20	How One ELLE Editor Styles Tights for Fall https://t.co/NxK1VoeJfh	1194383466344984577	66	2019-11-12 22:36:23	SocialF

Result of testing (screen shots, etc.)

```
Gnlp_sentimentanalysis

Bella Hadid has set her sights on a new divisive accessory to bring back
The golden age of the silver screen siren may be long past, but Pat McGra
The bride approached her wedding look from the top down, basing everythir
Shorts were on all the major Spring 2020 runways, from Saint Laurent to V
The chic new advent calendars to countdown to the holidays with. https://
Katie Holmes's minimal kicks are packed with an ethical, eco-friendly mes
Self-care is a commitment, but this brand is making it easier to add a mu
"I feel as a mother, I just want to put my arms around her. I want to tel
A millennial twist on Princess Diana's iconic sporty style. https://t.co/
There were too many takeaways to count at the 2019 #GlamourWOTY Awards, b
Looks like baby Archie may be spending Christmas stateside! https://t.co/
Would you expect anything less from Captain Jean-Luc Picard? https://t.co
It's a bird. It's a plane. It's @Harry_Styles as the tender, well-dressec
Did we miss anything? https://t.co/KLdzDZw2F6
Australian photographer Simon Burstall shares a look into the '90s Sydney
"I've been using fragrance to help me define each character I've played s
Sentiment: 0.3000001192092896, 9.100000381469727
Clearly Positive
```

```
Gnlp_sentimentanalysis  ×

18 Non-Basic Nail Colors To Wear This Winter https://t.co/Vjsaa
The 10 Best Clear Lip Glosses to Own, According to a Lip Gloss
Rep. Maxine Waters Deletes All of Facebook with One Look https:
10 Non-Basic Ways to Wear a Denim Jacket https://t.co/jXvxEnNsm
26 Movies Guaranteed to Make You Cry Every Time https://t.co/f7
4 Looks That Get Us Excited for Boot Season https://t.co/j5RZeP
The Best Bra Brands for Women With Large Cup Sizes https://t.co
Shawn Mendes Let Slip How Long He's Really Been Dating Camila C
Justin Bieber Posted Photos of Inside His and Hailey Baldwin's
Go behind-the-scenes with ELLE and Winnie Harlow. https://t.co/
Here is how to work off some of that thanksgiving turkey. https
Kylie Jenner Gave Ariana Grande Permission to Sample 'Rise and
Noah Centineo Shaved Off All His Hair, and Twitter Is Not Happy
How One ELLE Editor Styles Tights for Fall https://t.co/NxK1Voe
Sentiment: 0.0, 1.7000000476837158
Neutral
```

As can be seen in the screenshots, the brand of Vogue is more positive than Elle. In the future, we can examine the tweets in a large-scale, by using the score of 1, 0, -1 for the streaming data.

Lessons learned

What you liked doing?

The important thing for me is to learn from the knowledge sharing from my teammates. That is because everyone has different opinions, it is efficient to learn from them after using the framework, even in the process of learning.

I want to hear more about their assignments by using about five to ten sentences. In this case, the user story is the most important part, because the engineers are not good at explaining themselves for the most of time. In our current status, the ability to add comments, debug and testing are not well-established. It must be a "horrible" thing if we do not talk about the coursework. Sometimes, it will be good if we can get familiar with each other.

What you could have done better?

If we learned more about setting the virtual machine, we can learn the basic procedure of using many more API in one platform.

The most important thing for me is the product mission. Unlike other groups, I keep doing the comparison by learning the source code from the tutorial. I believe it is the first thing we need to learn before working. Although I add some comments, it would be better if we can combine them into a series of file: some can be seen by the customers, while others are not.

I also want to compare the outcomes of visualization by using R and Python separately, which is mentioned in the "what's better part".

What you will avoid in the future?

Just keep finishing the requirement of the coursework, doing it only for ten times will be completely different from doing it only for one time.

Designing a different product doesn't mean to pick up a topic without enough data points. For me, what can differentiate my product is the uniform procedure of my project, it can compare the brands of fashion magazines, it can also compare the brand of two wholesalers.

Reference

[1] User Stories | Examples and Template
https://www.atlassian.com/agile/project-management/user-stories
[2] Minimum Viable Product (MVP)
https://www.agilealliance.org/glossary/mvp/#q=~
[3] Vogue
https://www.vogue.com/
[4] Elle
https://www.elle.com
[5] Interface
https://www.pcmag.com/encyclopedia/term/45137/interface
[6] Tweepy
http://docs.tweepy.org/en/latest/getting_started.html#api
[7] Google NLP: sentiment analysis
https://cloud.google.com/natural-language/docs/sentiment-tutorial