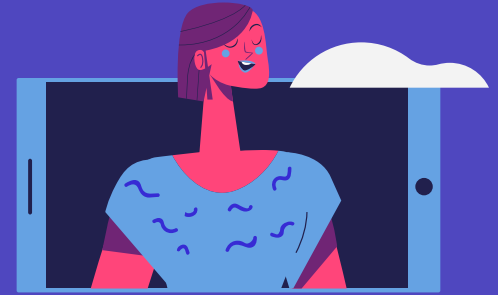


Airline Twitter Sentiment Analysis

University of Connecticut - MS BAPM

Chi-Hua Wu



Agenda

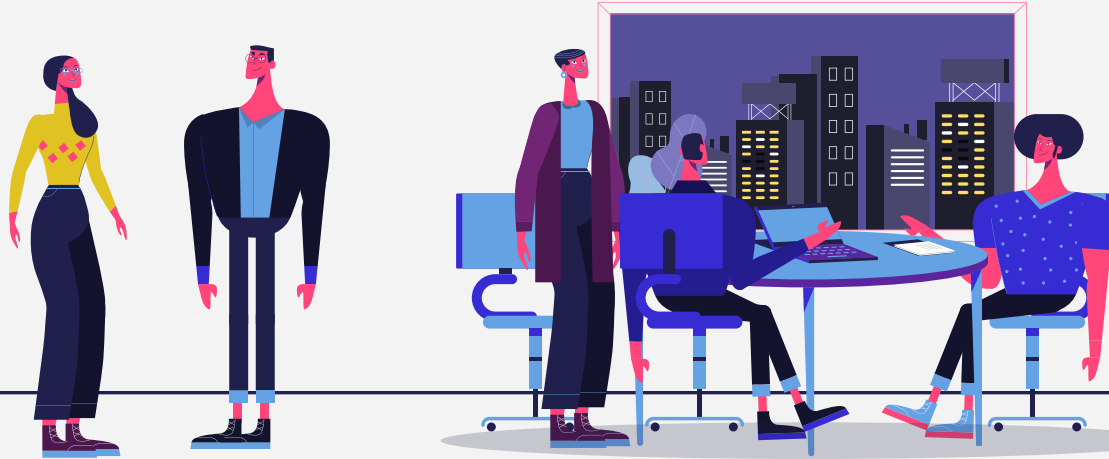
01 Objectives

02 Dataset

03 Sentiment &
Text Clustering

04 What's More?

01 Objectives





?

What insights can be draw
from tweets or text?

Objectives

→ Text Mining

- ◆ Techniques includes document classification, web mining, **concept extraction** and ...

→ Sentiment Analysis

- ◆ Understanding people's opinions
- ◆ Find out the reasons
- ◆ Improve certain services based on the results



02 Dataset



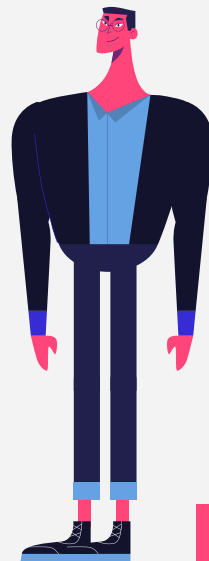
ABOUT DATASET

→ Dataset contains

- ◆ 14,640 rows
- ◆ From Crowdfunder's Data for Everyone library

→ Attributes

- ◆ Dates: From 02/16 to 02/24
- ◆ Airlines
- ◆ Tweets



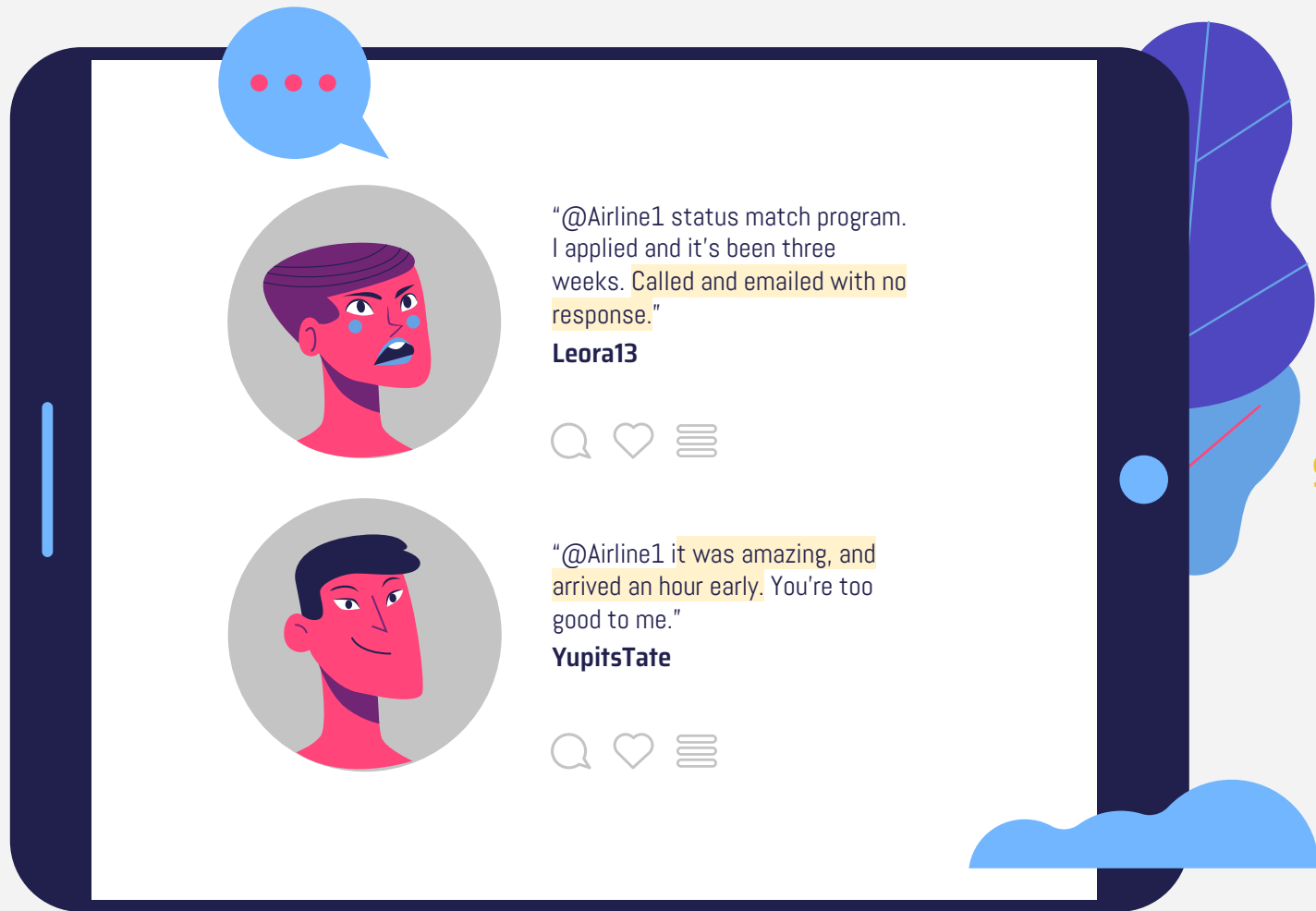
A screenshot of data

| | tweet_id | tweet_created | text |
|-------|--------------|---------------|---|
| 0 | 5.703060e+17 | 2/24/15 11:35 | @Airline 1 What @dhepburn said. |
| 1 | 5.703010e+17 | 2/24/15 11:15 | @Airline 1 plus you've added commercials to th... |
| 2 | 5.703010e+17 | 2/24/15 11:15 | @Airline 1 I didn't today... Must mean I need ... |
| 3 | 5.703010e+17 | 2/24/15 11:15 | @Airline 1 it's really aggressive to blast obn... |
| 4 | 5.703010e+17 | 2/24/15 11:14 | @Airline 1 and it's a really big bad thing abo... |
| ... | ... | ... | ... |
| 14635 | 5.695880e+17 | 2/22/15 12:01 | @Airline 5 thank you we got on a different fli... |
| 14636 | 5.695870e+17 | 2/22/15 11:59 | @Airline 5 leaving over 20 minutes Late Flight... |
| 14637 | 5.695870e+17 | 2/22/15 11:59 | @Airline 5 Please bring Airline 5 to |
| 14638 | 5.695870e+17 | 2/22/15 11:59 | @Airline 5 you have my money, you change my fl... |
| 14639 | 5.695870e+17 | 2/22/15 11:58 | @Airline 5 we have 8 ppl so we need 2 know how... |



03

Sentiment & Text Clustering



"@Airline1 status match program. I applied and it's been three weeks. Called and emailed with no response."

Leora13



"@Airline1 it was amazing, and arrived an hour early. You're too good to me."

YupitsTate



**What is
Sentiment
Analysis?**



How to do sentiment analysis?

1

Collect Data

@Airline1 The BEST Company EVER !!!

2

Data Preprocessing

best company ever

3

Tokenize

[best, company, ever]

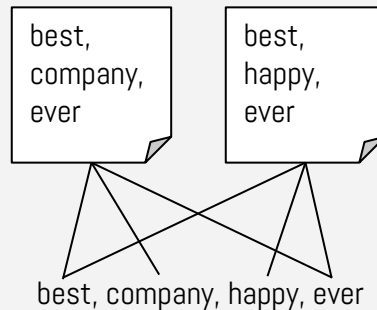
4

Sentiment Analysis

[best, company, ever]
0.7 0.1 0.1

5

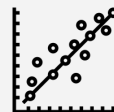
Text Clustering



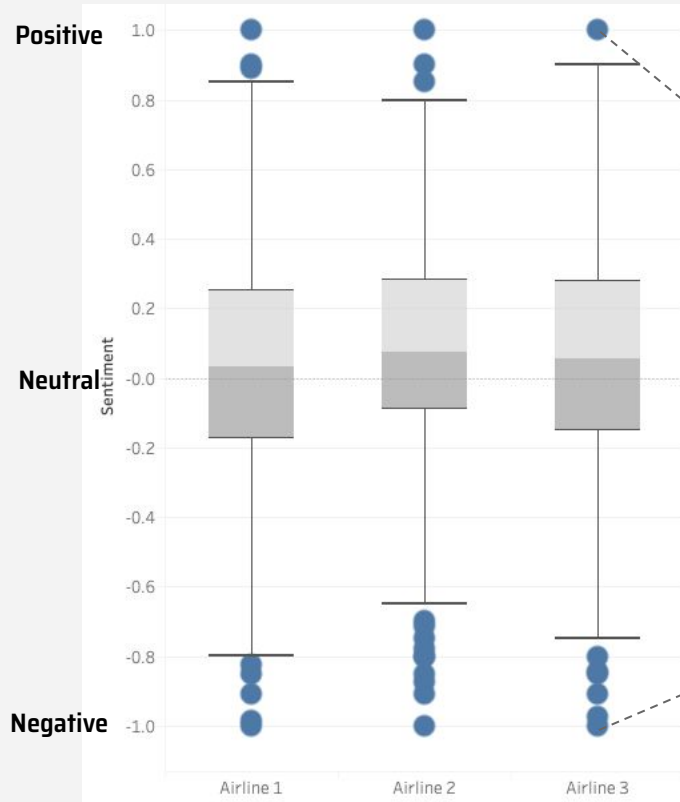
6

Modeling

Can use model to predict the future tweets



Results: Sentiments based on airline



Positive Tweets Example

You guys are **awesome** don't change a thing luv.

Negative Tweets Example

Please refund my ticket what **terrible** experience

Topic modeling on tweets

Positive Group

| Topic 3 words | Topic 3 weights |
|---------------|-----------------|
| Airline 1 | 901.1 |
| thank | 282.5 |
| fli | 260.1 |
| servic | 253.1 |
| Airline 2 | 226.2 |
| custom | 216.5 |
| love | 192.8 |
| Airline 3 | 156.9 |
| guy | 152.8 |
| flight | 148.4 |

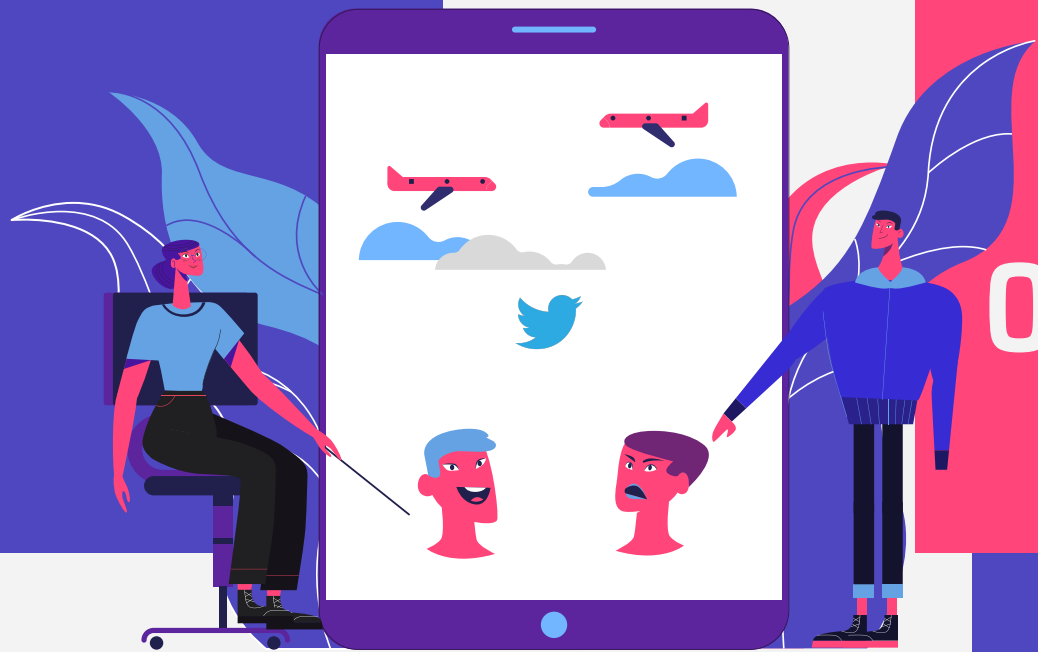
@Airline1 thx for your quick
response and action bf will make
good use of the voucher satisfied
happy customer

@Airline 1 flight cancelled flighted and
rebooked but agent made mistake.

Negative Group

| Topic 1 words | Topic 1 weights |
|---------------|-----------------|
| flight | 1193.2 |
| cancel | 311.2 |
| Airline 1 | 260.2 |
| delay | 232.3 |
| Airline 2 | 176.2 |
| hour | 165.4 |
| late | 149.5 |
| flightl | 147.2 |
| due | 143.4 |
| plane | 127.5 |





04 What's more?



Streaming Tweets

Instead of using the dataset online, we can extract the streaming tweets with certain hashtags or mentions



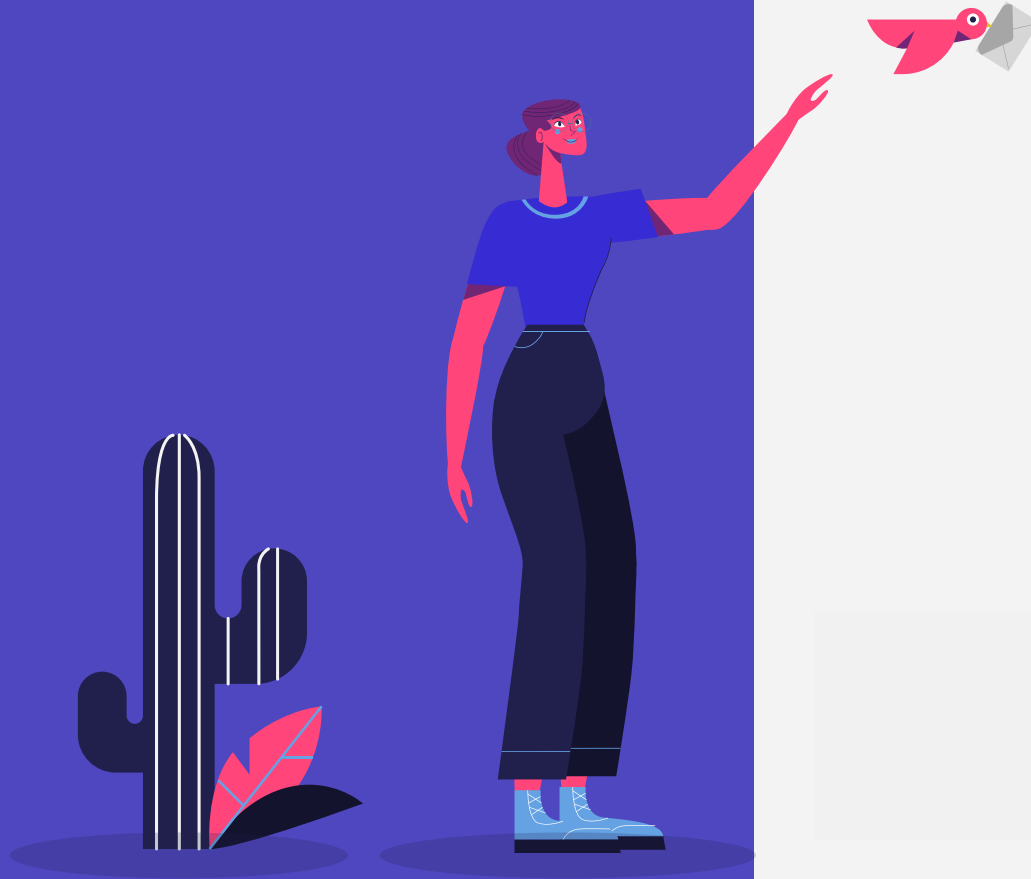
Improve models

Models could be improved by applying better data pre-processing. More models can be used to get the best accuracy



Internal Audit

Future potential Internal Audit used, such as culture review, social media review and the ability to identify areas for improvement and strengths



THANKS!