# Social Media Analytics on Canadian Airlines

DS8006 – Social Media Analytics – Winter 2017

Group #1: Aliasghar Arabi, Saeede Sadat Asadi Kakhki, Bernardo Najlis

April 5th, 2017

AIR CANADA

WESTJET

sunwing AIRLINES

porter

newleaf
Low fare. Here to there.

jetlines
Flying Your Way

Ryerson University

# Outline

**01**

Problem Statement and Dataset Selection

**03**

Data Collection & Preprocessing

**04**

Analysis

**05**

Conclusions

**06**

References, Previous Work and Literature Review

# Problem Statement and Dataset Selection

- Social Media Analytics on Canadian Commercial Airlines Industry to:
  - Find Insights about social media usage
  - How companies use social media to communicate with passengers
  - Types of conversations
- Identify patters or behaviours to make recommendations and improve customer engagement

| Selected Airline | Why? |
|---|---|
| Air Canada | Flag Carrier and largest airline in Canada |
| Westjet | Second largest Canadian airline with national / international reach |
| Porter | Regional airline based in Toronto, Ontario |
| Sunwing Airlines | Low-cost airline with destinations in US, Mexico and the Caribbean |
| Jetlines | Ultra low-cost airline based in Vancouver, BC |
| NewLeaf | Virtual Airline selling tickets for Flair Airlines |

# Data Collection and Preprocessing

- Twitter data collected using *'twitter'* R package search API
  - Collect as many tweets as possible (up to 3,200 maximum allowed by search API)
  - All tweets and retweets from / to the official airline accounts
  - All tweets that mentioned the airline's name (capital/lowercase) or hashtag
- Data Collected on March 9, 11 and 18, 2017
- Collected, saved and shared using CSV files
- Data scrubbing to
  - remove duplicate observations
  - Parse date/time fields (using *'lubridate'*)
  - Added metadata columns to filter replies, retweets and official account tweets

# Analysis – Exploratory Data Analysis / Metadata

- For all airlines the most used term is their own brand

| Airline | Destination |
|---|---|
| Air Canada | Toronto, Denver |
| WestJet | Toronto, St Marteen |
| JetLines | Vancouver |
| Porter | Toronto |
| NewLeaf | N/A |
| SunWing | Tenerife, Beach, Resort, Vacation |
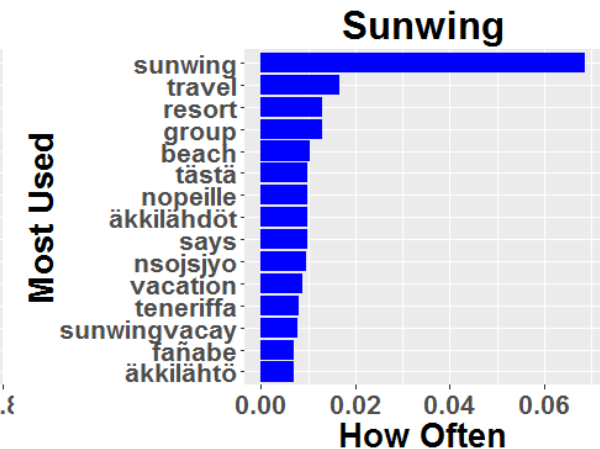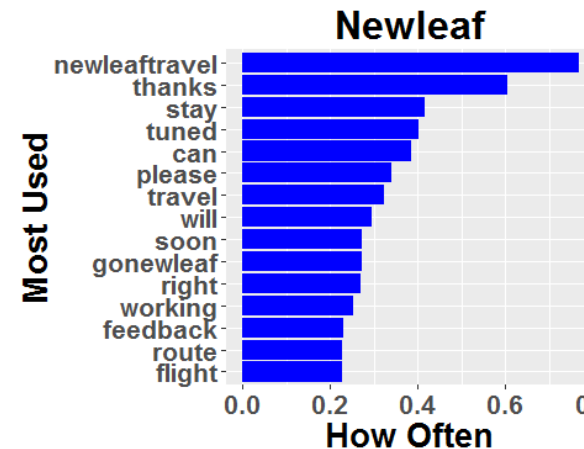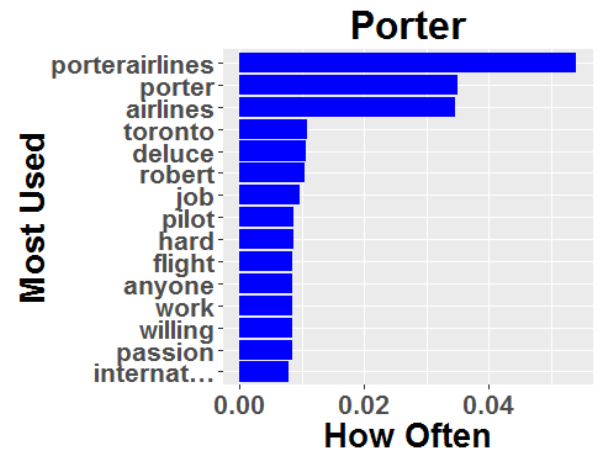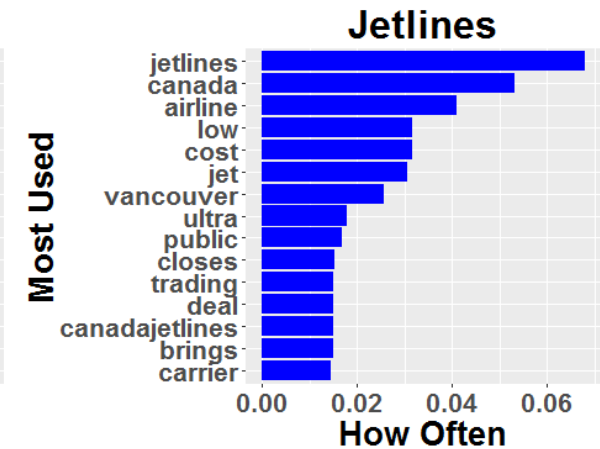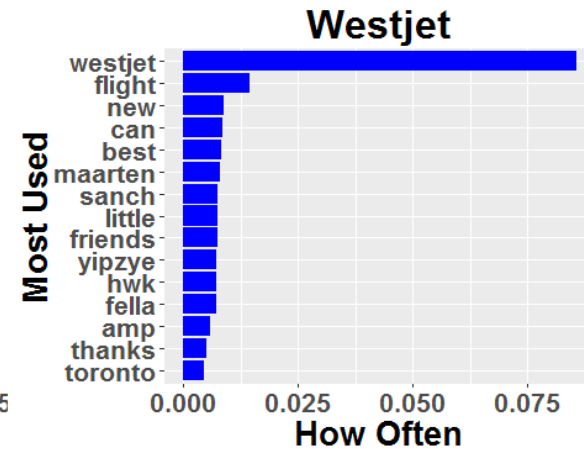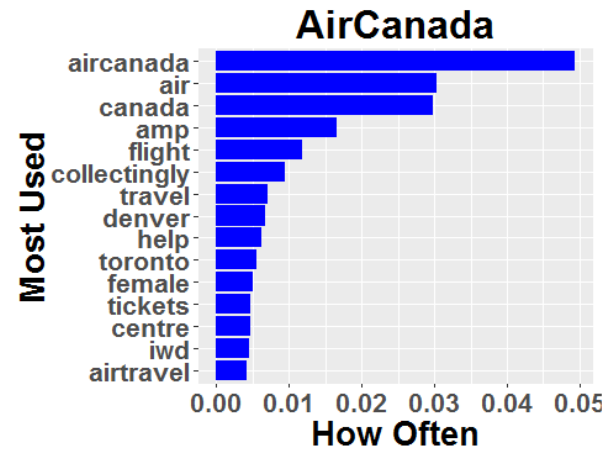
# Analysis – Exploratory Data Analysis / Metadata

- 90% of NewLeafTravel tweets come from its official account vs only 2% for Sunwing and 3% for JetLines

- NewLeaf has the higher reply rate at 60% vs JetLines 5% and SunWing 9%

- **NewLeafTravel has the highest "engagement" in their official account vs lowest in SunWing and Jetlines**
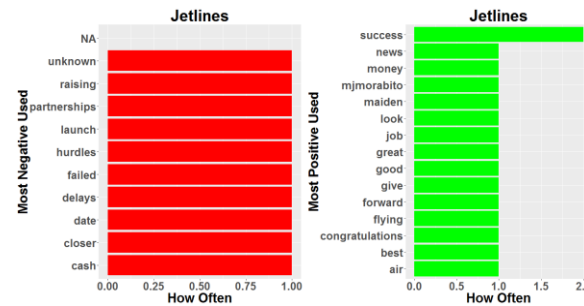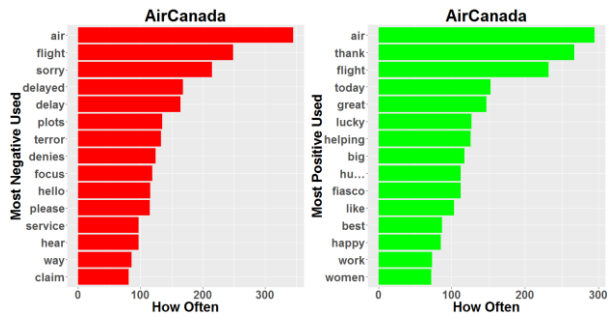


Canadian Airlines - Twitter Metadata Analysis
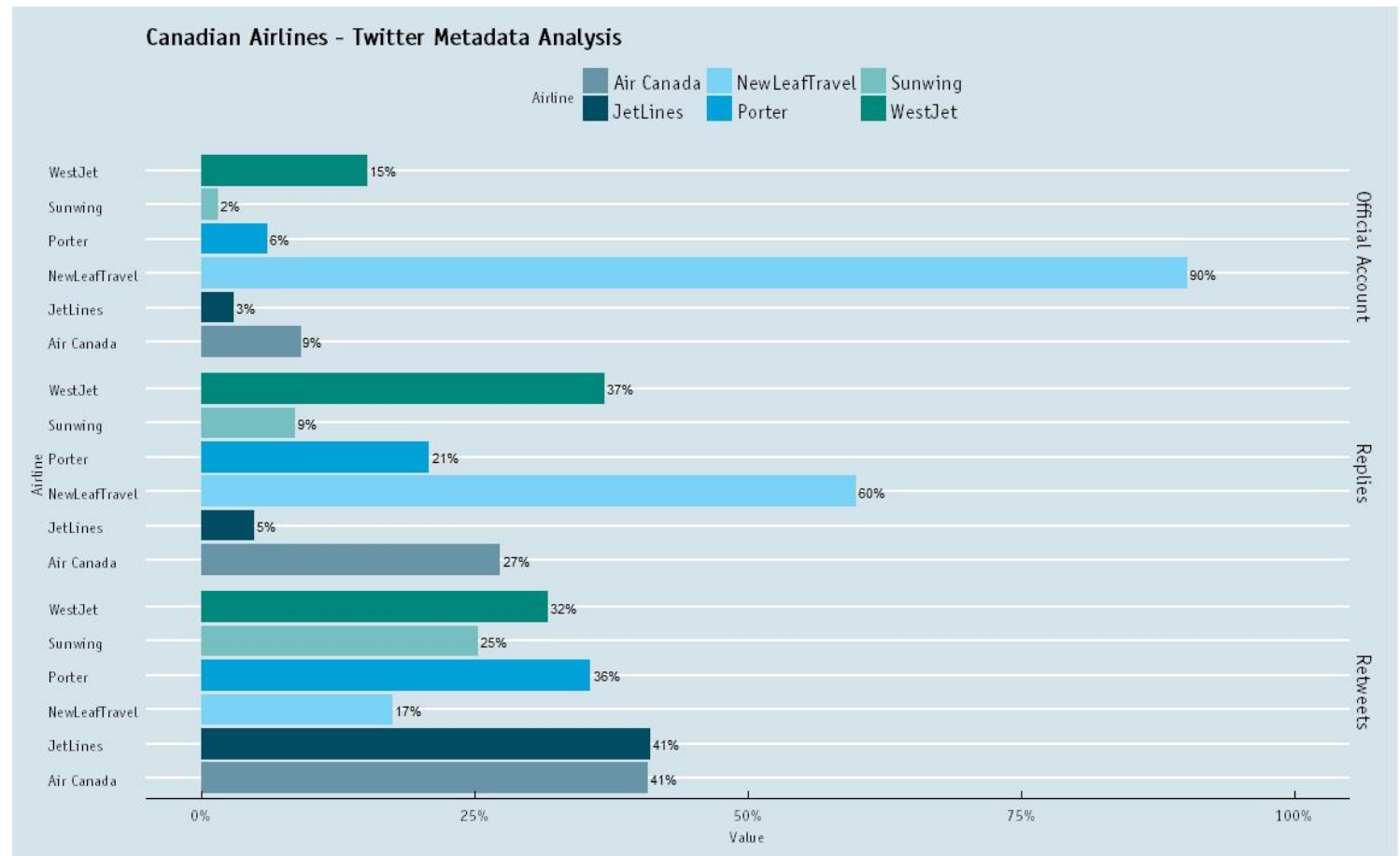
# Analysis – Topic Modelling
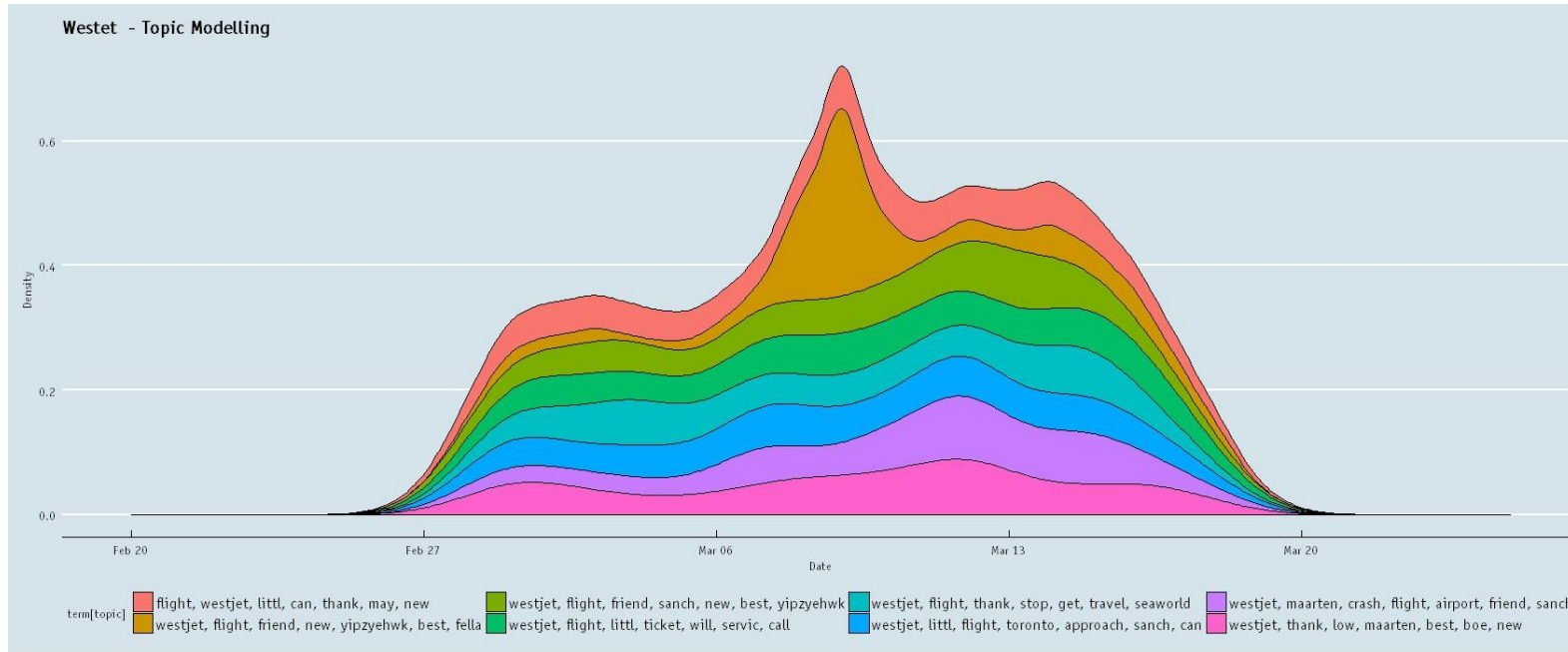


- Topic Modelling isolated the top tweet peak on March 8 from all other conversation topics

- **Messages with emotional content generate more response and echo**

# Analysis – Sentiment Analytics



| group | pos.count | neg.count | all.count | score |
|---|---|---|---|---|
| AirCanada | 1258 | 1122 | 2380 | 53 |
| WestJet | 866 | 465 | 1331 | 65 |
| Newleaf | 851 | 207 | 1058 | 80 |
| SunWing | 35 | 69 | 104 | 34 |
| Porter | 160 | 11 | 171 | 94 |
| Jetlines | 3 | 1 | 4 | 75 |
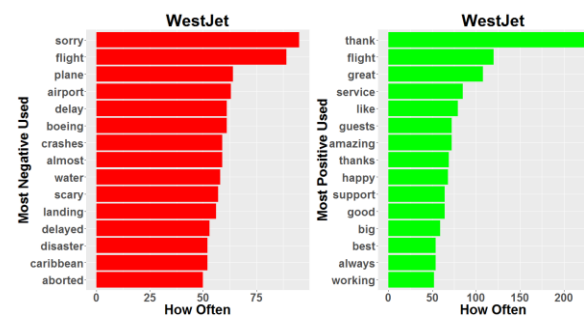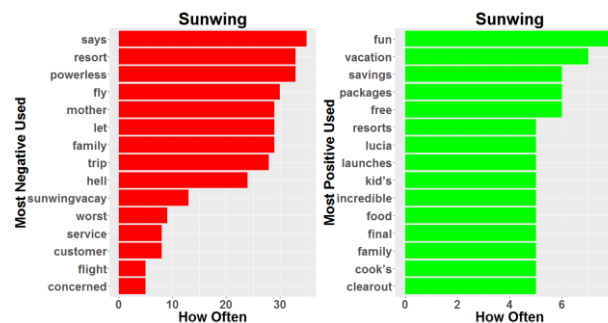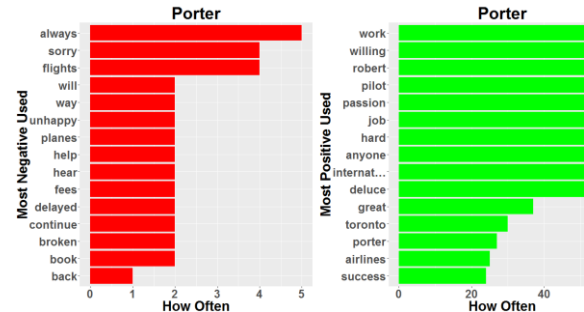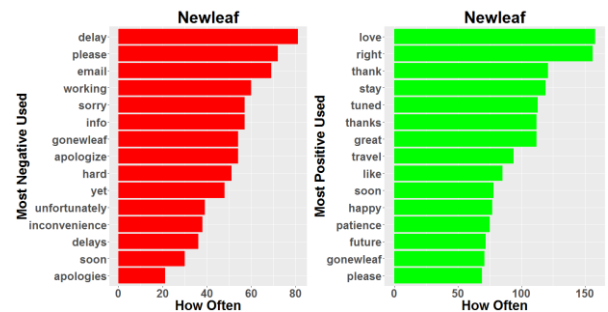
- References to **delays** are the most negative used terms across all airlines
- AirCanada had a terror threat RT: **"#US Homeland Secretary: Terrorists threatened #aircanada flights"**

# Analysis – Sentiment Analysis
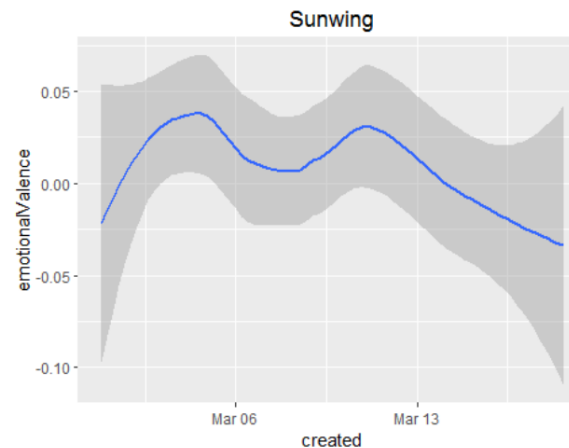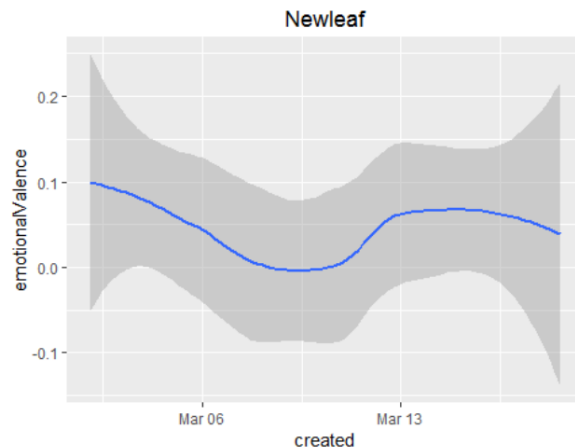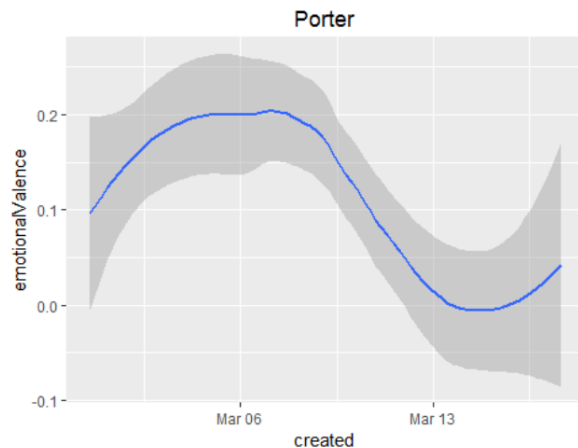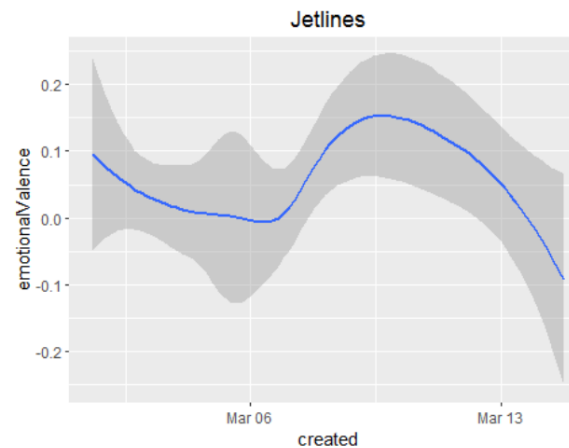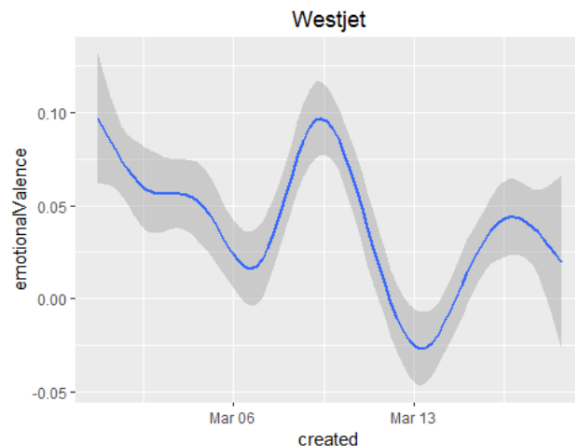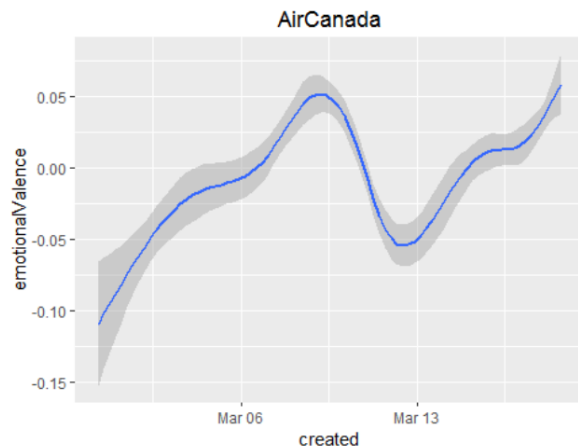


- Air Canada and Westjet have the most variability (and the highest number of total tweets overall)

- **NewLeaf, who controls 90% of the tweets from their official account, is the most stable and stays always above positive**

# Analysis – Network Analytics
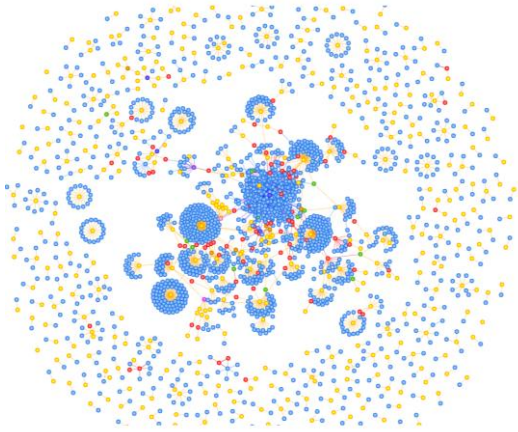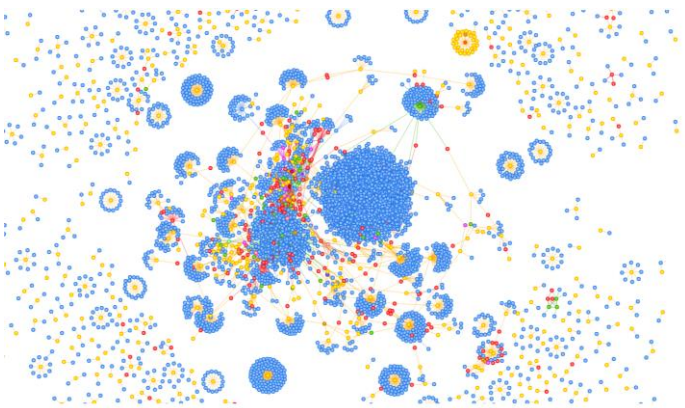
| Airline | Density | Diameter | Reciprocity | Centralization | Modularity |
|---------|---------|----------|-------------|----------------|------------|
| AirCanada | 0.000141809 | 8 | 0.007968127 | 0.1509167 | 0.8468935 |
| Porter | 0.00474389 | 3 | 0 | 0.09467456 | 0.8195001 |
| Jetlines | 0.01603146 | 2 | 0 | 0.1175746 | 0.7178711 |
| Sunwing | 0.004352031 | 3 | 0 | 0.06551517 | 0.9156307 |
| WestJet | 0.000400569 | 4 | 0.009920635 | 0.08149005 | 0.909061 |
| NewLeaf | 0.0119863 | 3 | 0 | 0.2624421 | 0.5748299 |

- Reciprocity can show the responsivity of airlines to customers.

- Modularity values in average are high for these airlines. It shows these networks consist of different conversations.

# Conclusions and Lessons Learned

- Airlines need social media: all of their services are used by individuals

- Knowing positive/negative terms helps understand what the company does right or wrong: most negative terms are related to flight delays

- Be like NewLeaf, don't be like SunWing

- How to combine different types of social media analysis to make conclusions and recommendations

- **A:** igraph, vizNetwork, Sentiment Analytics

- **S:** Influencers analysis, Association analysis

- **B:** Usage of ggthemes() and LDA() for topic modelling

# References, Previous Work and Literature

- 1.   Social Media Analytics for American Airlines [online]. Available: https://www.talkwalker.com/industry-research/airlines/american-airlines-social-media-analytics

- 2.   Social Media for Airline Industry [online]. Available FTP: niit-tech.com Directory: sites/default Files: files/Social%20Media%20in%20Airline%20Industry(ISG)_0.pdf

- 3.   Leveraging Customer Analytics: The Airline Industry [online]. Available: http://knowledge.wharton.upenn.edu/article/leveraging-customer-analytics-the-airline-industry/

- 4.   Twitter US Airline Sentiment [online]. Available: https://www.kaggle.com/crowdflower/twitter-airline-sentiment

- 5.   Twitter Topic Modeling Using R [online]. Available: https://gist.github.com/bryangoodrich/7b5ef683ce8db592669e

- 6.   Sien Chen, Yinghua Huang, Wengqiang Huang, "Big Data Analytics on Aviation Social Media: The Case of China Southern Airlines on Sina Weibo", 2016 IEEE Second International Conference on Big Data Computing Service and Application. doi:10.1109/BigDataService.2016.51

- 7.   Pouti Isidore, Moll More Borja, "Airline companies in social media: a case study of HOP", Halmstad, 2013

- 8.   Dolan, Rebecca and Shahbaznezhad, Hamidreza, Social Media Strategy in Services: The Effect of Airline Brand Posts on Engagement (February 7, 2017). Available at SSRN: https://ssrn.com/abstract=2913352

- 9.   Impact Of Social Networking In The Airline Industry Marketing Essay [online]. Available: https://www.ukessays.com/essays/marketing/impact-of-social-networking-in-the-airline-industry-marketing-essay.php

- 10.  Xiaotong Duan,TianshuJi,WanyiQian, "Twitter US Airline Recommendation Prediction". Available: http://cs229.stanford.edu/proj2016spr/report/042.pdf

- 11.  List of Canadian Airlines [online]. Available: https://en.wikipedia.org/wiki/List_of_airlines_of_Canada

- 12.  Text Mining with R – Twitter Data Analysis [online]. Available: http://www.rdatamining.com/docs/text-mining-with-r

- 13.  Finding and Removing Duplicate Records [online]. Available: http://www.cookbookr.com/Manipulating_data/Finding_and_removing_duplicate_records/