

**Project Dissertation**  
**Uncovering trends in Airline Industry using**  
**Twitter Sentiment Analysis**

**Submitted By:**

**Akshay Darmwal**

**Roll No: 2K15/MBA/07**

**Under the Guidance of:**

**Dr Rajan Yadav**

**Associate Professor, DSM**



**DELHI SCHOOL OF MANAGEMENT**

**Delhi Technological University**

**Bawana Road Delhi 110042**

**January -May 2017**

## **CERTIFICATE**

This is to certify that the Project Report titled “Uncovering trends in Airline Industry Using Twitter Sentiment Analysis”, is a bonafide work carried out by Mr. Akshay Darmwal of MBA 2015-17 and submitted to Delhi School of Management, Delhi Technological University, Bawana Road, Delhi-42 in partial fulfillment of the requirement for the award of the Degree of Masters of Business Administration.

Signature of Guide

(Dr. Rajan Yadav)  
(Associate Professor and Head of Department(DSM))

Place: New Delhi

Date:

## **DECLARATION**

I, Akshay Darmwal, student of MBA 2015-17 of Delhi School of Management, Delhi Technological University, Bawana Road, Delhi-42 declare that project dissertation report on “Uncovering trends in Airline Industry Using Twitter Sentiment Analysis” submitted in partial fulfillment of Degree of Masters of Business Administration is the original work conducted by me.

The information and data given in the report is authentic to the best of my knowledge.

This report is not being submitted to any other University for award of any other Degree, Diploma and Fellowship.

Place: New Delhi

Akshay Darmwal

Date:

(2K15/MBA/07)

## **ACKNOWLEDGEMENT**

I would like to express my sincere gratitude to Dr. Rajan Yadav, professor, Delhi School of Management, Delhi Technological University, for encouraging me to work on this project and helping me understand it. I am extremely grateful to him for his kind support and guidance throughout the project.

The completion of this project would not have been possible without the valuable contribution of Ms. Khushbu Madan. I am sincerely thankful for her guidance, invaluable constructive criticism and friendly advice during the project work.

I would also like to thank my college, Delhi School of Management, Delhi Technological University for providing me with the facilities and making me capable enough to complete this project successfully for future development.

## **EXECUTIVE SUMMARY**

Micro-blogging websites have evolved to become a source of all kinds of information. People post all kinds of real time messages on micro-blogs including their experience of a service they use, opinions on a variety of topics and current issues, complains and positive sentiments about the product they use.

Twitter offers a unique dataset in the world of brand sentiment. Brands receive sentiment messages directly from their customers in real time on twitter. These brands have the opportunity to analyze these messages to determine the consumer sentiment. Airline industry being a high volume service industry receives hundreds of comments on their social media pages daily from their customers regarding their experiences, complaints and opinions on the services provided.

The aim of this study is to analyze the sentiments of a corpus of tweets posted with hashtags and twitter handles of major Airline players in India, including Jetairways and AirIndia. The study aims to classify the tweet sentiments as positive, negative and neutral. The purpose of this study is to identify key service areas of these companies which require further improvements and the areas which provide positive experience to the customers. The study further discovers trends from the data which may generate actionable insights.

# TABLE OF CONTENTS

1.	Introduction	1
2.	Literature Review	3
3.	Social Media Trends In India	4
	3.1 Measuring Success of Social Media Platform	5
	3.2 Measuring Brand Sentiment	6
	3.3 Using Twitter in Sentiment Analysis	9
4.	Industry Profile	12
	4.1 Airline Companies in India	12
	4.2 Growth of Airline Industry in India	14
	4.3 Major Players: Facts and Figures	16
5.	Methodology	21
	5.1 Significance of the project	21
	5.2 Scope of the study	21
	5.3 Research Methodology	21
	5.4 Sample Size	21
	5.5 Data Extraction	21
	5.6 Data Cleaning and Pre-Processing	23
	5.7 Tools Used	24
	5.8 Steps involved in sentiment analysis and category creation	27
6.	Analysis, Inferences and Recommendations	30

6.1 Tweet Analysis of JetAirways	30
6.1.1 Inferences and Recommendations for JetAirways	33
6.2 Tweet Analysis of AirIndia	34
6.2.1 Inferences and Recommendations for AirIndia	38
7. Limitations of the study	39
8. Conclusion	40
References	41
Annexures	43

## LIST OF FIGURES

Figure	Title	Page No
Figure 3.1	Primary objectives for being present on Social Media	5
Figure 3.2	Companies measuring Brand Sentiment using Social Listening Tools	6
Figure 3.3	Impact on Brand Sentiment after using Social Listening Tools	7
Figure 3.4	Frequency of Updates on Social Media	8
Figure 3.5	Average Response Time on Social Media	8
Figure 3.6	Tweets addressed to Jetairways and Spicejet Handles	10
Figure 4.1	Major Airlines operating in India	12
Figure 4.2	Major airports operating in India	13
Figure 4.3	Passenger traffic growth in FY16	14
Figure 4.3	Domestic Passenger traffic growth(millions)	15
Figure 4.4	Aircraft Movement growth	16
Figure 4.5	Twitter profile overview(April14 –May13)	16
Figure 4.6	Change in number of followers of JetAirways	17
Figure 4.7	Evolution of interaction for @jetairways(April14 –May13)	18
Figure 4.8	Change in number of followers of AirIndia	18



Figure 4.9	Evolution of interaction for @airindian(April14 –May13)	19
Figure 4.10	Change in number of followers of SpiceJet	20
Figure 4.11	Evolution of interaction for @flyspiceje(April14 –May13)	20
Figure 5.1	Tweets Extracted using Rapidminer	22
Figure 5.2	Display table of searched twitter results in Rapidminer	23
Figure 5.3	SPSS Text Analytics for survey Libraries	25
Figure 5.4	SPSS negative words library	26
Figure 5.5	SPSS positive word library	26
Figure 5.6	Flow chart for sentiment analysis and category creation	29
Figure 6.1	Sentiment analysis and category creation for JetAirways data in SPSS TAS	30
Figure 6.2	Distribution of positive and negative tweets across various categories for JetAirways	31
Figure 6.3	Word cloud for high frequency Positive and Negative words for JetAirways	32
Figure 6.4	Trend analysis of tweets over different time slots for JetAirways(Apr15- Apr30)	32
Figure 6.5	Response time segment for jetAirways tweets	33
Figure 6.6	Sentiment Analysis and Category creation for AirIndia data in SPSS TAS.	35
Figure 6.7	Distribution of positive and negative tweets across various categories for JetAirways	36

Figure 6.8	Word cloud for high frequency Positive and Negative words for AirIndia data	36
Figure 6.9	Trend analysis of tweets over different time slots for AirIndia(Apr15- Apr30)	37
Figure 6.10	Response time segment for AirIndia tweets	37

## LIST OF TABLES

Tables	Title	Page no
Table 5.1	Category wise negative and positive response for JetAirways	31
Table 5.2	Category wise negative and positive response for AirIndia	35

# CHAPTER - 1

---

## INTRODUCTION

The growth of social media over the last decade has revolutionized the way individuals and companies interact with each other and the way companies conduct business. With millions of people spending countless hours on social media to communicate, share, interact, and thereby create user-generated data at an unprecedented rate, social media has become a source of big data. (Reza Zafarani, Huan Liu et al, Cambridge Univeristy Press, 2014) Consumers use different types of online forums for social engagement including blogs, social media networking sites such as Facebook and Twitter. Through social media, consumer engagement with the industry occurs in real time. This type of interaction offers a unique opportunity for marketing intelligence.

Social media data mining is the process of extracting, analyzing and representing actionable patterns from social media data to fetch meaningful insights. A study indicates that out of 98 percent of customers who contact companies to raise issues about customer service, 30 percent do so via online social networking sites (Chandra, International Journal of Business Research, 2011). To the marketer, this data provides deep insights into consumer behavioral tendencies and presents an opportunity to learn about customer feelings and perceptions in real time, as they occur without any intrusion or provocation (Shintaro Okazaki et al, 2014). However, recent explosions of user generated content on social media sites are presenting unique challenges in interpreting, harnessing, and analyzing textual content as the data is fragmented and disorganized. Sentiment analysis is a technique that can overcome these challenges by systematically extracting and analyzing data from these sources.

Sentiment analysis is a data mining technique that evaluates textual content using machine learning techniques. As a research method in the field of marketing, sentiment analysis presents an effective and efficient evaluation of consumer opinions in real time. (SearchBusinessAnalytics, 2016). It allows data extraction and analysis from a very large corpus without obstructions and time delays. Through sentiment analysis, marketers collect rich data on attitudes and opinions in real time, without compromising generalizability, reliability and validity. Marketers also gain information on attitudes and opinions of

consumers as they occur, without having to invest in lengthy and costly market research activities.

In this study, sentiment analysis techniques were applied to gauge the sentiments of the customers regarding the services provided by Airline companies. The objective of this study is to analyze the sentiments of a corpus of tweets posted with hashtags and twitter handles of major Airline players in India, including Jet airways and AirIndia. The study aims to classify the tweet sentiments as positive, negative and neutral. Further, the purpose of this study is to identify key service areas which require further improvements and the areas which provide positive experience to the customers.

## CHAPTER-2

---

### LITERATURE REVIEW

Kaplan and Haenlein (2010) define social media as “a group of Internet-based applications that build on the ideological and technological foundations of Web 2.0, and that allow the creation and exchange of user-generated content”. Social media include online social communities (e.g. Twitter, Facebook, LinkedIn etc.), blogs, content communities (e.g. Flickr, YouTube, etc.), wikis (e.g. Wikipedia), social bookmarking, and news sites (e.g. Delicious, Digg, Reddit) (Boyd and Ellison, 2007; Constantinides and Fountain, 2008). The primary difference between social media and the “old Web” is the availability of free, accessible and easy-to-use online tools for content publishing, editing, sharing, bookmarking, and classification that enable interactive exchanges between individuals and groups (O’Leary, 2011).

Sentiment analysis is regarded as valuable technique in the field of “customer experience analysis” (Kho, 2010). Unlike traditional surveys of customer opinions (often time-consuming, cumbersome, formal, and costly), automated text-mining is more efficient alternative for opinion analysis, especially when dealing with the instantaneous nature of social media (Koppel and Schler, 2006; Deshpande and Sarkar, 2010; Arnold, 2011).

With a characteristic matter-of-fact honesty, genuineness and originality, consumer commentaries in social media are one of the most sought types of information online (Deshpande and Sarkar, 2010). Consumers’ own reviews of products, services, and brands are highly valued among online audiences (Burton and Khammash, 2010). Customers are also eager and open to voice their dissatisfaction with products and services via social media (Weiet al., 2008; Geho et al., 2010).

Research on opinion mining or sentiment analysis started with identifying opinion (or sentiment) bearing words, e.g., great, amazing, wonderful, bad, and poor. Many researchers have worked on mining such words and identifying their semantic orientations or polarity determination (i.e., positive, negative and neutral).

Although blogs, review sites and forums are some of the richest sources of content about customer experiences, recently research efforts in field have pointed at the opportunities that Twitter opens for opinion-mining studies.

## CHAPTER-3

---

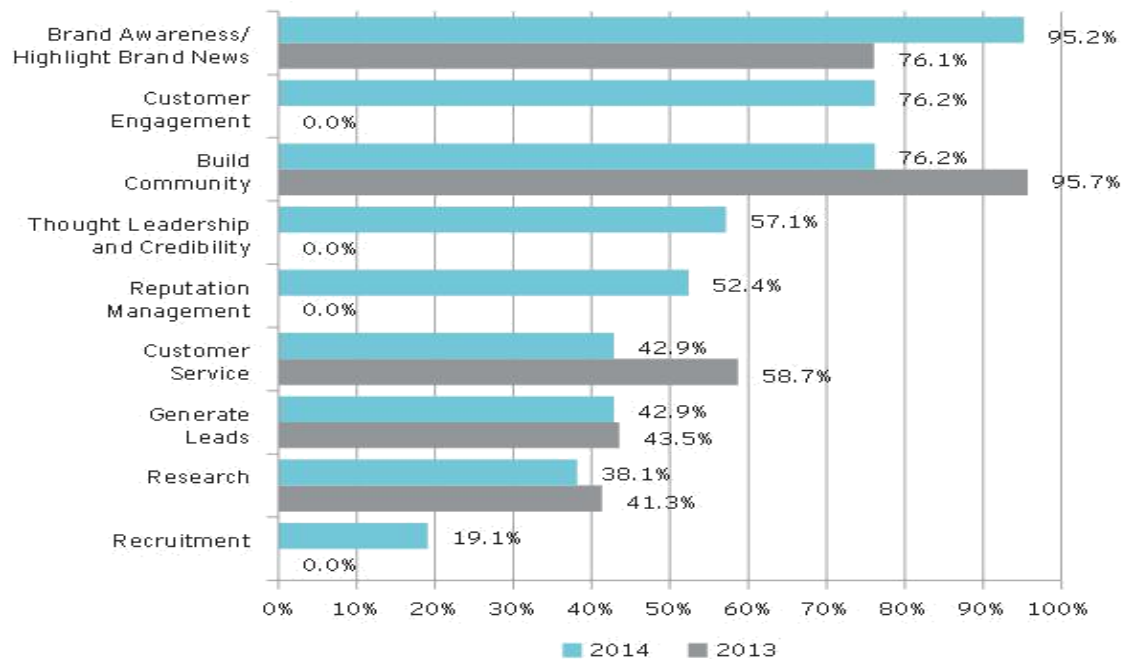
### SOCIAL MEDIA TRENDS IN INDIA

Social media usage in India increased manifold, as the number of internet users in India reached 375 million users. (Statista.com, 2016). India is a key market for social media giants as India has mobile social media penetration of about 9% (Statista.com, 2016) and the active social media users in India has grown to around 106 million. India is among the top three countries in terms of number of people using Facebook (100m+ users), whereas Twitter is seeing an increased user base of over 33 million (Ernst Young, 2015). The increased mobile web penetration is also seen as a key contributor to increased growth in social media usage. About 84% of Facebook's 100 million users in India access it using their mobile devices (Ernst Young, 2015).

Almost 81% of the brands surveyed have considered Facebook to be the most important platform among others, while about 48% of surveyed brands think that Twitter is the second-most important channel to have presence on, followed by YouTube (43% surveyed brands have considered it to be the third-most important channel). (Ernst Young, 2015)

75% of India's online population is digital consumers i.e. those who use digital media for purchasing goods and services (Ernst Young, 2015). Trust in a brand and company is no longer dependent only on company-controlled, traditional mass media channels, but rather on peers and communities through social media and other digital channels. Therefore, brands need to allocate a significant proportion of marketing budget to social media marketing and digital in order to successfully market to these customers.

The top 3 objectives to be present on social media for the brands are Building Brand Highlight Brand News/ Awareness, Customer Engagement and Building a Community.



**Fig 3.1: Primary objectives for being present on Social Media**  
 (Source: Social Media Marketing-India Trends Study, EY, 2015)

### 3.1 Measuring Success of Social Media Platform

For measuring the success at social media platform, marketers use various metrics like

#### Brand sentiment analysis

It refers to the emotion behind a social media post. Sentiment analysis is a way to measure the tone of the conversation and adds important context to social conversations. Sentiment is usually measured as a ratio of positive to negative comments.

#### Customer engagement

It measures how often and how much people interact with a brand and its content in social media. When a visitor or customer takes the time to like, re-tweet or comment on something a brand has posted, they're actively engaging with the brand.

#### Brand mentions

This is the total number of times a brand is mentioned on a webpage or social media network over a particular period of time.



## Social reach

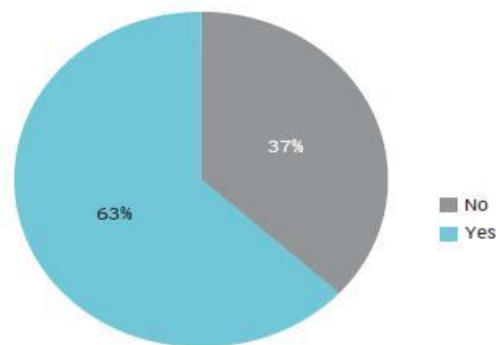
It measures the total number of people a brand is able to reach via various social media networks it is present in. It is a measure of range of influence and takes into account the number of, retweets, shares, repins, visitors (fans/ subscribers/ followers), click-through rates, referral rates etc.

### **3.2 Measuring Brand Sentiment**

Brand sentiment can be used to evaluate brand health, to evaluate campaigns and other initiatives or for competitive research.

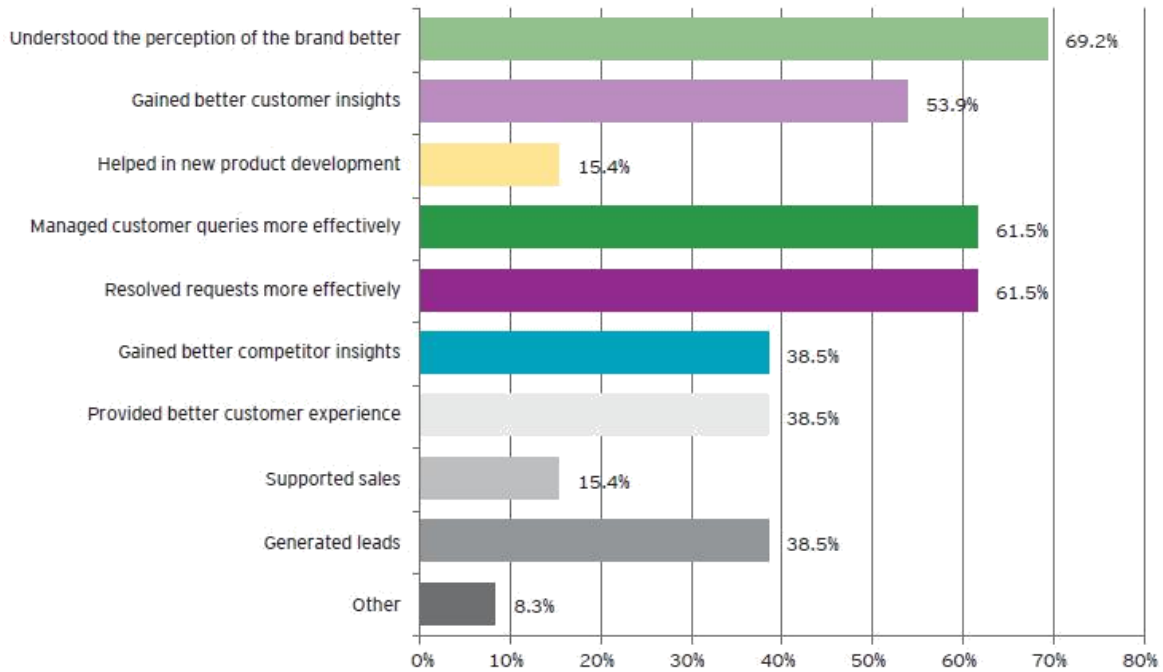
As a company's online footprint increases manifolds, it becomes almost impossible to track the sentiment manually. There are a host of tools both free and paid that can help in tracking and evaluating the same.

A majority of organizations monitor their brand sentiment through social listening tools. Tools used for social listening purpose include Meltwater , SAS text analytics, Radian6, Simplify 360, Iristrack, Social Mention, Hootsuite and Netbase.



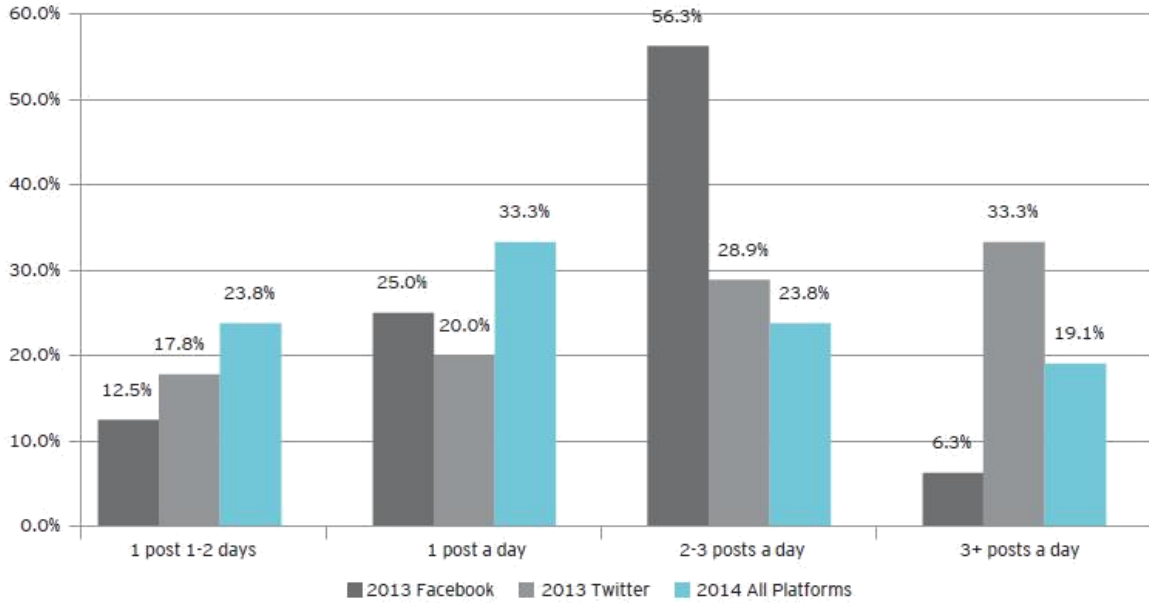
**Fig 3.2: Companies measuring Brand Sentiment using Social Listening Tools**  
(Source: Social Media Marketing-India Trends Study, EY, 2015)

Marketers believe that sentiment analysis has reaped benefits for their companies. Some of the major benefits that organizations have realized by combining social listening tools with social media marketing include effective management of customer queries, better understanding of the perception of the brand and effective resolution on requests.



**Fig 3.3: Impact on Brand Sentiment after using Social Listening Tools**  
 (Source: Social Media Marketing-India Trends Study, EY, 2015)

The time taken to respond via social media channels will be seen as critical in the way a company handles customers as brands start using social media platforms for more than marketing (e.g., Customer Service and Crisis Management). From a marketing point of view, well stated responses that result in meaningful conversations on social media will prove to be more valuable than a quicker response time.

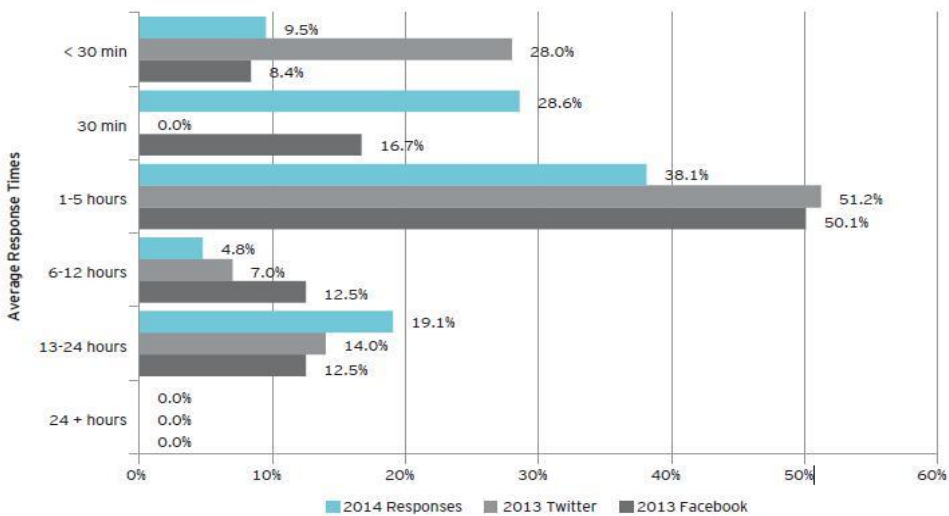


**Fig 3.4: Frequency of Updates on Social Media**  
 (Source: Social Media Marketing-India Trends Study, EY, 2015)

Average response time of 30 min or less was recorded by 38% of brands in 2014 (compared to average response times stated by 25% brands for Facebook and 28% of brands in 2013). (Ernst Young, 2015).

### Average Response Time

What is your average response time in which you attempt to reply to users on a social networking platform?



**Fig 3.5: Average Response Time on Social Media**  
 (Source: Social Media Marketing-India Trends Study, EY, 2015)

### **3.3 Using Twitter in Sentiment Analysis**

Micro-blogging site Twitter is among the world's top ten social media channels in terms of daily number of visitors and registered users (Garg et al., *Journal of Management Information Systems*, 2011). With over 160 million registered users, an average of 55 million tweets per day and 600 million search queries, Twitter is one of the most popular social media channels for receiving real-time updates and sharing information (Savage, 2011; Thelwall et al., *Journal of American Society for Information Science and Technology*, 2011). Despite its concise format that allows posting messages of only 140 characters at a time and a simplistic purpose i.e. to answer the question "What are you doing?", Twitter has proved to be one of the most favorite platforms for information sharing on social media (Geho et al., 2010; Gayo-Avello, 2011; O'Leary, *Communications of the ACM*, 2011). Its popularity has appealed to many companies which are seeking to not only reach masses of customers for sharing promotional material, but to also interact with them in real-time and address issues they face with their products and services. A branch of sentiment analysis field focuses on analyzing twitter feeds to analyze the effects of distinct events, to monitor public opinions on various topics and to make predictions about future customer behavior (Fotis Misopoulos et al, 2014).

The reason for using Twitter feeds in sentiment analysis is that when a user responds to an event by tweeting, she/he demonstrates an "information behavior". This contains an affective component (e.g. judgments or intentions), and therefore reveals important information about users' sentiments and on the tweeted topic. Savage (2011) observed that while individual tweets might not be of great value, "taken together they open a surprising window into the moods, thoughts, and activities of society at large" allowing researchers to get meaningful insights about the customer opinions.

While the number of studies using twitter for sentiment analysis is still modest, the interest in the channel as a source of customer information is growing. Researchers are still experimenting with various techniques for analyzing Twitter feeds and extracting meaningful information (Thelwall et al., *Journal of American Society for Information*

Science and Technology 2011). Based on the aforementioned, the text that follows presents a study conducted within the framework of sentiment analysis and provides an example of how twitter information can be used to assess customer experiences in the Airline industry.

Consumers of Airline services actively post their experiences in the form of opinions, complaints, suggestions etc. on a daily basis.

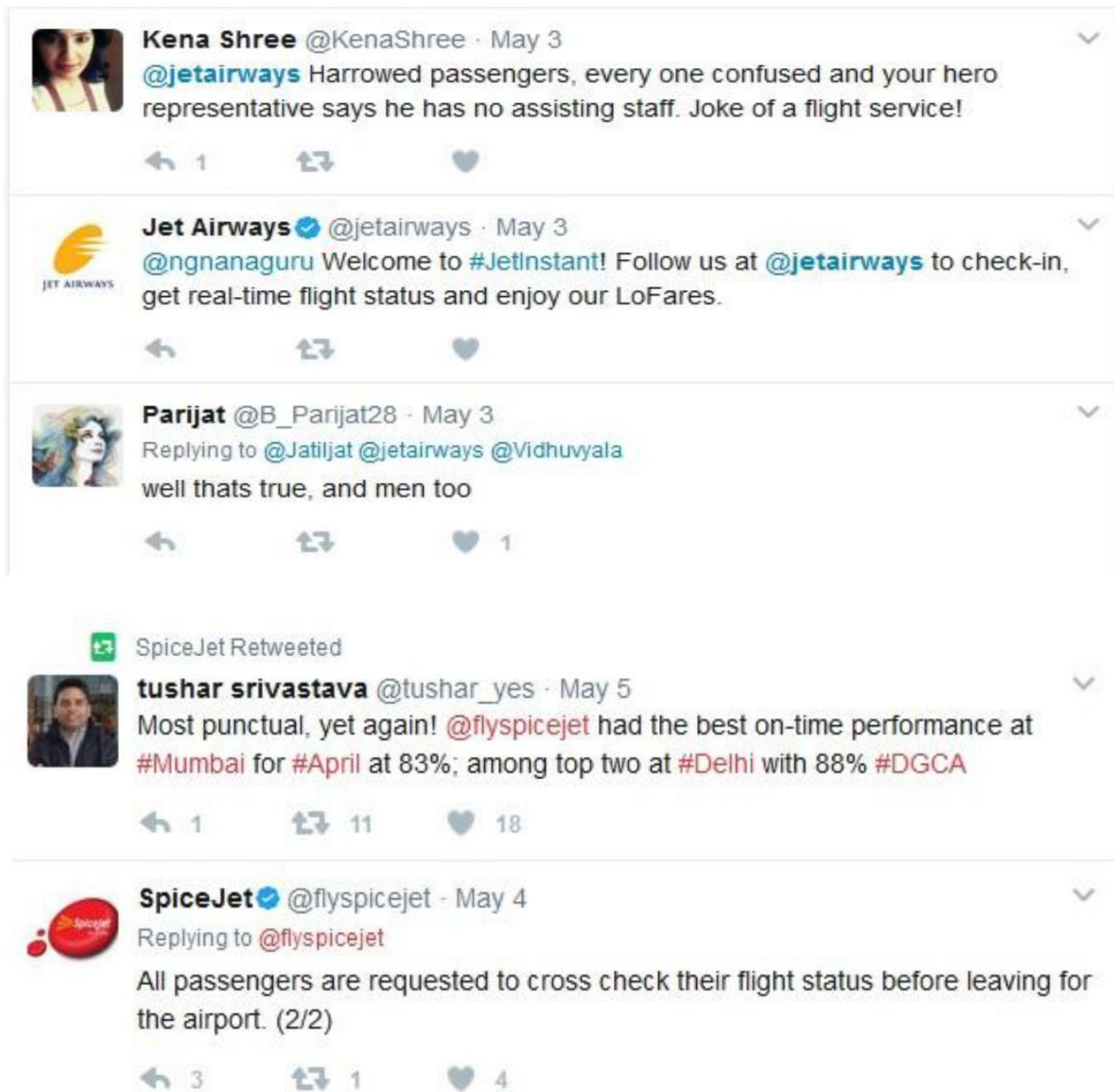


Fig 3.6: Tweets addressed to Jetairways and Spicejet Handles

(Source: Twitter.com, 2017)

These aggregators can mine the customer comments and analyze them to understand the general customer sentiment. They can further analyze areas which particularly require immediate attention for improvement and areas which delight the customer.

For example: Multiple posts on a particular day about the delay in Flight can make the company look into its operational issues and rectify them to deliver better services and ensure customer satisfaction.

Customer sentiment can be measured over a long duration to analyze the impact of a newly adopted strategy. These companies can mine for trends in their feed to understand the customer better and thereby deliver a better service.

# CHAPTER-4

## INDUSTRY PROFILE

### 4.1 Airline Companies in India

India is one of the fastest-growing aviation market (International Air Transport Association (IATA)). 65% of the aviation manufacturing is done in Bangalore and it is emerging as an aviation manufacturing hub in India.

Air India is India's national flag carrier after merging with Indian (airline) in 2001 and plays a major role in connecting India with the rest of the world. IndiGo, Jet Airways, Air India, Spicejet and GoAir are the major carriers in order of their market share. The above mentioned airlines connect more than 80 cities in India and also operate overseas routes after the liberalization of Indian aviation. Several other foreign airlines connect Indian cities with other major cities across the globe.

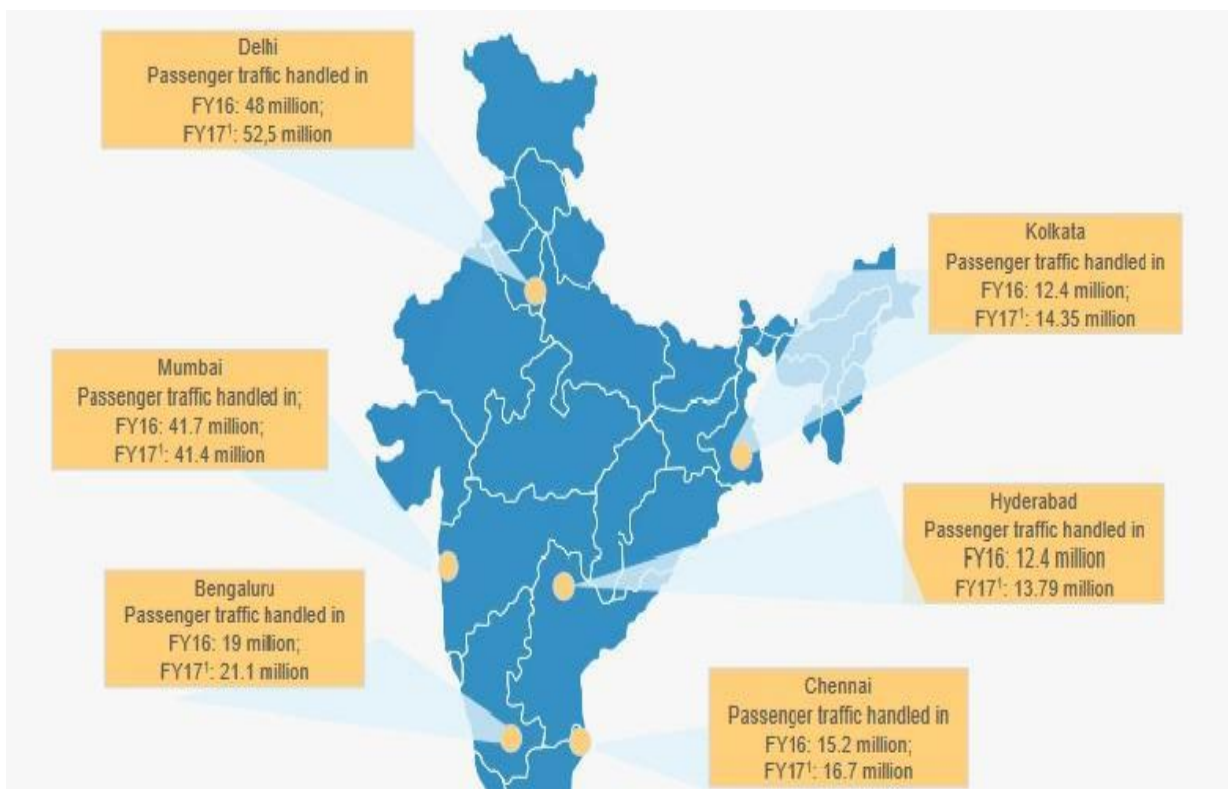


Fig 4.1: 6 Major Airlines operating in India

(Source: TechSci Research, ibef.org, 2016)

India has become the third largest aviation market in terms of domestic passenger traffic, beating Japan. India's domestic air passenger traffic stood at 100 million in 2016 and was behind only the US (719 million) and China (436 million).

The market is also estimated to have 800 aircraft by 2020. In 2015, Boeing projected India's demand for aircraft to touch 1,740, valued at \$240 billion, over the next 20 years in India. This would account for 4.3 per cent of global volumes. Airbus is expecting an annual growth rate of over 11 per cent for the domestic market in India over the next ten years, while the combined growth rate for domestic and international routes would also be more than 10 per cent.



**Fig 4.2: 6 Major Airports operating in India**

(Source: TechSci Research, ibef.org, 2016)



## 4.2 Growth of Airline Industry in India

### Passenger Traffic Growth

- Witnessing a growth of 17.62 per cent over the previous year, total passenger traffic stood at a 223.6 million in FY16, which was recorded at 190.1 million in FY15 in India.
- Growth in passenger traffic has been strong since the new millennium, especially with rising incomes and low-cost aviation; during FY06-16, passenger traffic grew at a CAGR of 11.8 per cent in the country.

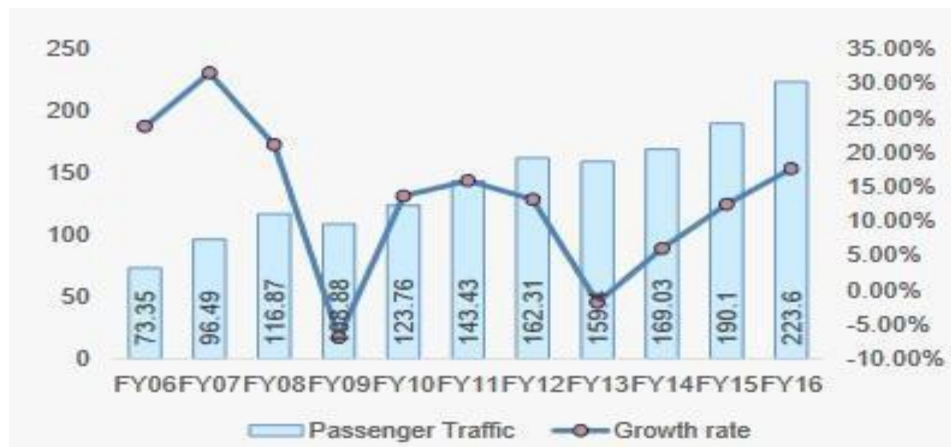


Fig 4.3: Passenger traffic Growth in FY16 (millions)

(Source: Association of Private Airports Operator Authority of India, TechSci Research, ibef.org, 2016)

### Growth in Domestic Passenger Traffic

- Domestic passenger traffic expanded at a CAGR of 11.46 percent over FY06–17
- According to Directorate General of Civil Aviation, domestic passenger traffic witnessed growth at a rate of 22 per cent, till February 2017, in comparison 21.24 per cent in FY16.
- International passenger traffic registered growth at a CAGR of 8.33 per cent over FY06-17.

- During February 2017, domestic airlines carried over 8.23 million passengers, showing a growth of more than 23 per cent compared to the same period last year.
- During FY17, domestic passenger traffic increased by 22 per cent in comparison with growth rate of 21.24 per cent in FY16.
- During FY16, international passenger traffic increased by 7.72 per cent



Fig 4.4: Domestic Passenger traffic Growth (millions)

(Source: Airports Operator Authority of India, Ministry of Civil Aviation, TechSci Research, ibef.org, 2016)

### Aircraft movement growth

- During FY07-17, domestic aircraft movement increased at a CAGR of 5.59 per cent, while international aircraft movement expanded at 5.39 per cent CAGR over the same period.
- In FY17, domestic aircraft movement increased at a YoY of 1.43 per cent, while international aircraft movement declined at a YoY of 0.68 per cent, in comparison to FY16.
- During FY17, the total number of domestic aircraft movement increased to 1.86 million, as compared to FY16

- In FY17, the total number of international aircraft movement increased to 0.365 million, in comparison to FY16.



Fig 4.5: Aircraft movement growth

(Source: Airports Operator Authority of India, Ministry of Civil Aviation, TechSci Research, ibef.org, 2016)

### 4.3 Major Players: Facts and Figures

#### Overview

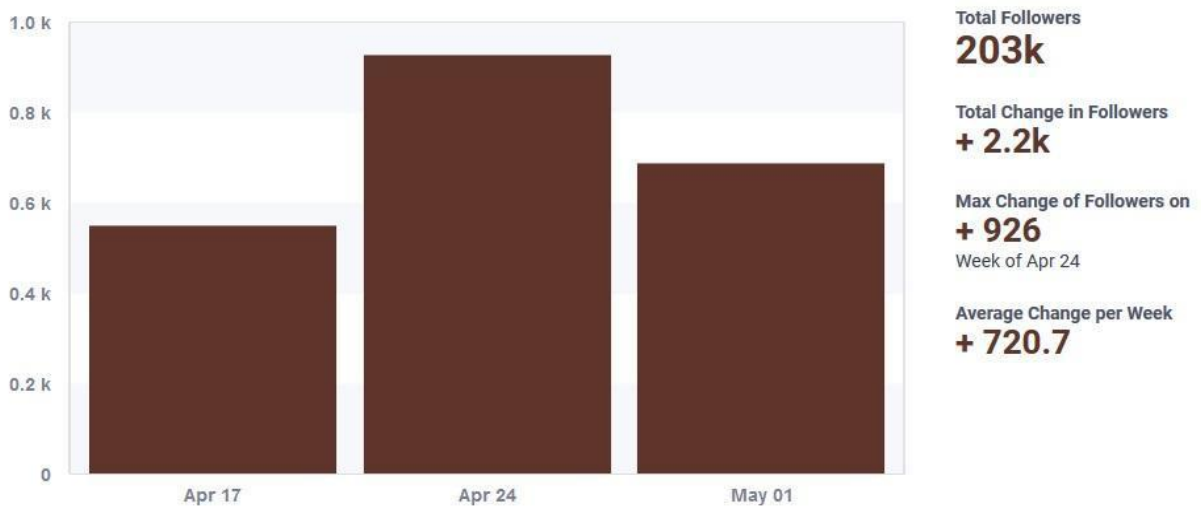
Profile Name	Total Followers	Total Change in Followers	Relative Change in Followers	Number of Interactions per 1000 Followers
 Jet Airways / jetsairways	204 419	+ 3 706	1.85 %	43.06
 Air India / airindia	120 573	+ 2 264	1.91 %	31.82
 SpiceJet / flyspicejet	90 004	+ 1 801	2.04 %	87.26

Fig 4.6: Twitter profile overview (Apr14- May13)

(Source: Socialbakers.com, 2017)

## Jetairways (Twitter handler: @jetairways)

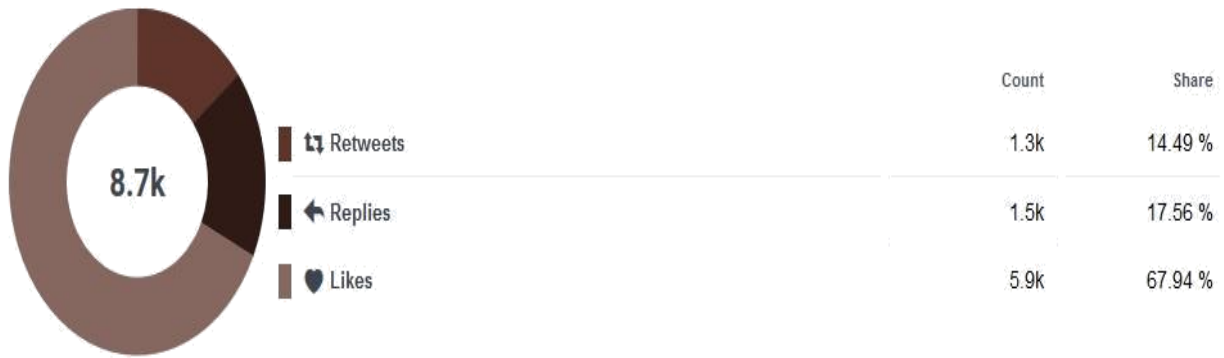
Jet Airways is an Indian airline based in Mumbai. It operates over 300 flights daily to 68 destinations worldwide from its main hub at Chhatrapati Shivaji International Airport. Incorporated in April 1992 as a limited liability company, the airline began operations as an air taxi operator in 1993. It began full-fledged operations in 1995 with international flights added in 2004. The airline went public in 2005 and in 2007, it acquired Air Sahara.



**Fig 4.7: Change in number of followers of JetAirways**

(Source: Socialbakers.com, 2017)

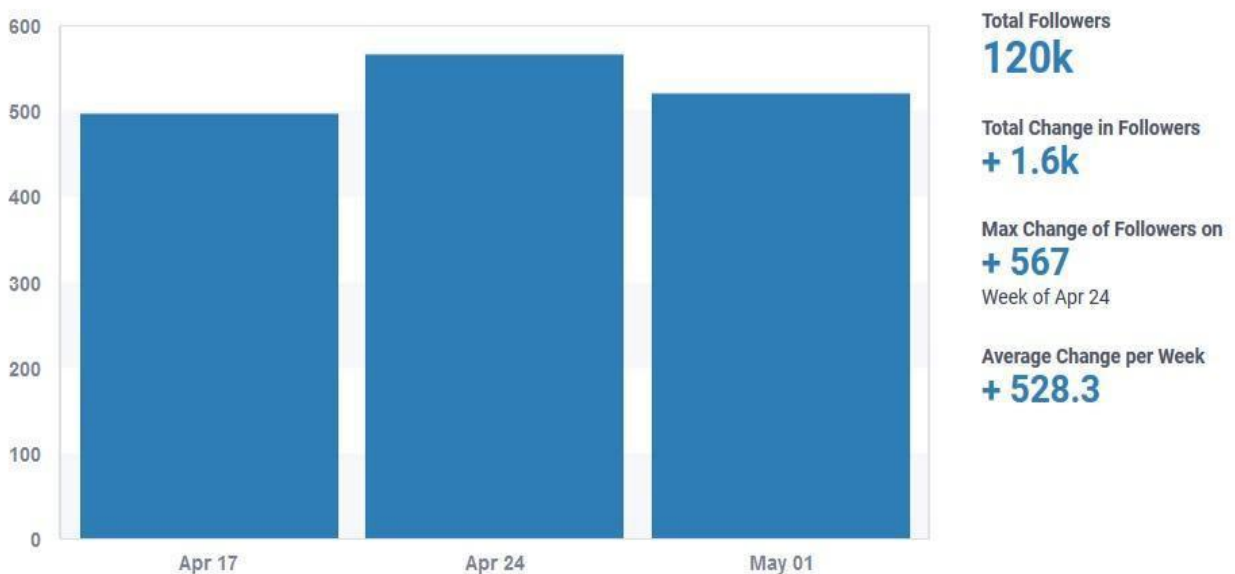
- Change in followers in week of Apr17 is +549
- Change in followers in week of Apr24 is +926
- Change in followers in week of May01 is +687



**Fig 4.8: Evolution of Interactions for @jetairways(Apr14-May13)**  
 (Source: Socialbakers.com, 2017)

**AirIndia (Twitter handler: @airindiain)**

It is owned by Air India Limited, a Government of India enterprise, and operates a fleet of Airbus and Boeing aircraft serving 90 domestic and international destinations. It is headquartered in New Delhi. Air India has its main hub at Indira Gandhi International Airport, New Delhi and a secondary hub at Chhatrapati Shivaji International Airport, Mumbai.



**Fig 4.9: Change in number of followers of AirIndia**  
 (Source: Socialbakers.com, 2017)

- Change in followers in week of Apr17 is +497
- Change in followers in week of Apr24 is +561
- Change in followers in week of May01 is +521



**Fig 4.10: Evolution of Interactions for @airindiain (Apr14-May13)**  
 (Source: Socialbakers.com, 2017)

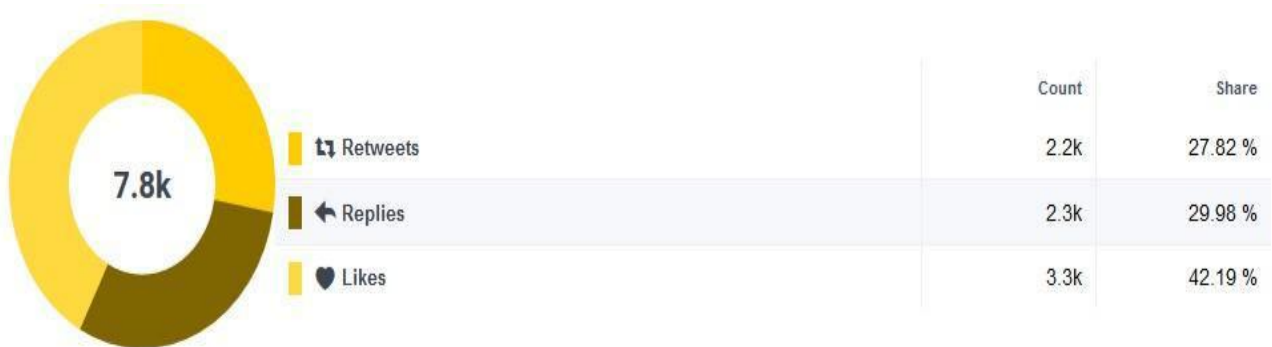
**Spicejet (Twitter handler: @flyspicejet)**

SpiceJet is a low-cost airline headquartered in Gurgaon, India. It is the third largest airline in the country by number of domestic passengers carried. The airline operates 312 daily flights to 55 destinations, including 45 Indian and 10 international destinations from its hubs at Delhi, Kolkata and Hyderabad. Established as air taxi provider ModiLuft in 1994, the company was acquired by Indian entrepreneur Ajay Singh in 2004 and re-christened as SpiceJet. The airline operated its first flight in May 2005.



**Fig 4.11: Change in number of followers of Spicejet**  
 (Source: Socialbakers.com, 2017)

- Change in followers in week of Apr17 is +310
- Change in followers in week of Apr24 is +546
- Change in followers in week of May01 is +494



**Fig 4.12: Evolution of Interactions for @flyspicejet (Apr14- May13)** (Source: Socialbakers.com, 2017)

## CHAPTER-5

---

### METHODOLOGY

#### 5.1 Significance of the study

The study is aimed at analyzing the customer sentiment of Airline companies (AirIndia and Jetairways) using their twitter handles and associated hashtags. The goal of the study was to investigate the possibilities that social media platform Twitter offers for identifying, monitoring, and analyzing customer experiences for the purpose of optimizing service performance. The study was conducted under the framework of sentiment analysis

#### 5.2 Scope of the study

This study is focused on the airline companies including JetAirways and AirIndia their customers and their management executives.

#### 5.3 Research Methodology

Descriptive research techniques were employed for the study. Tweets addressed to @jetairways and @airindiain were extracted and preprocessed to remove hashtags, URLs, @ symbols etc. These tweets were then analyzed using SPSS Text Analytics for Surveys and results were plotted in the same.

#### 5.4 Sample Size

A corpus of 1803 tweets of JetAirways and 780 tweets of AirIndia posted between the dates 14-April-2016 to 30-April-2016 was analyzed to determine their sentiment polarity.

#### 5.5 Data Extraction:

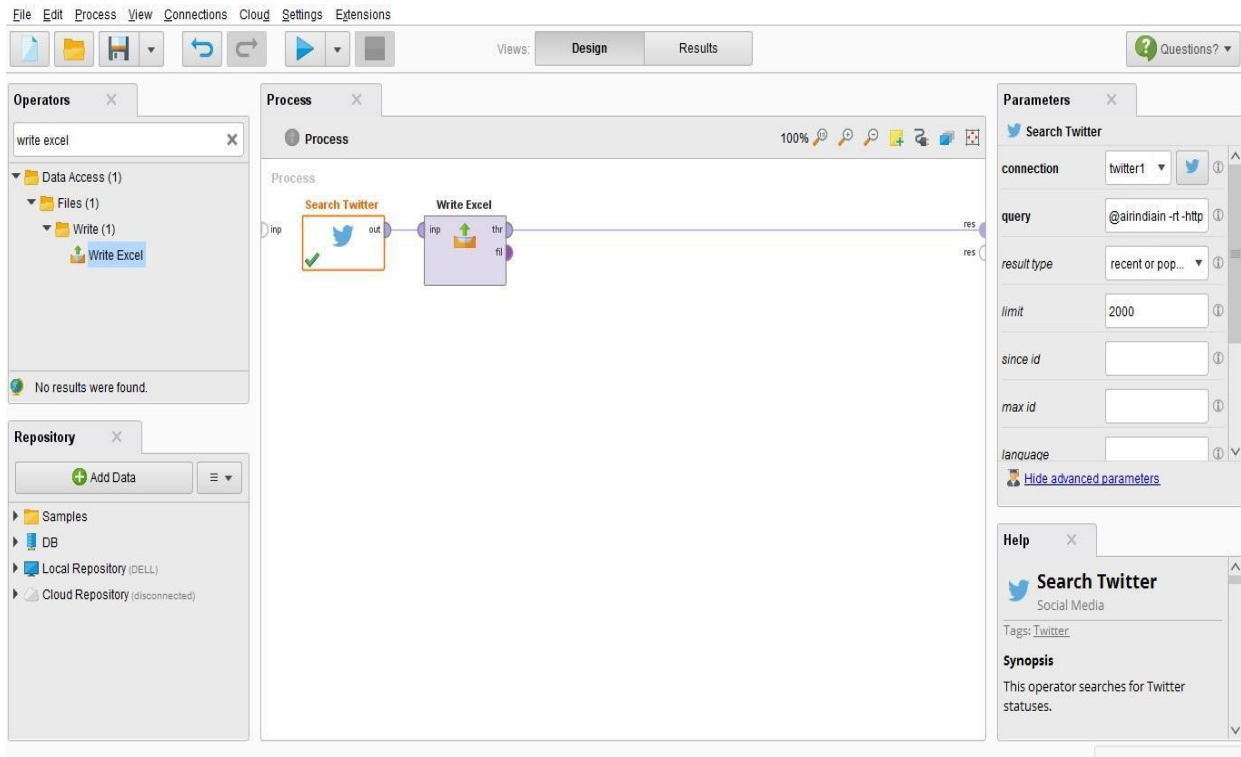
A number of free GUI tools like Facepager and Rapidminer etc. are available for data extraction which use the Twitter API in the background. However, a limitation of using such tools is that for a particular search term, tweets of past 7-8 days or a maximum of 2000 tweets can be extracted. This study involves analysis of a data collected over a period of 15



days, so a workaround had to be employed in order to extract more tweets. A semi-automatic approach was employed to extract the tweets (Tillkeyling.com, 2016).

For extraction, the following steps were taken:

1. Using Rapidminer's search twitter tool, tweets addressed to various airline companies were searched for using their handle (For Example: @jetairways) between the dates 15<sup>th</sup> April 2016 to 30<sup>th</sup> April 2016. The search results were scrolled to the bottom.



**Fig 5.1: Tweets are extracted using Rapidminer**

2. The extracted tweets are displayed in form of a table which can be exported to excel file using write excel command in the Rapidminer tool.

Id	Created-At	From-User	From-User-Id	To-User	To-User-Id	Language	Source	Text
8578008980...	Apr 28, 2017 ...	Pankaj Pach...	33464913	?	-1	en	<a href="http://...	UDAN faces t...
8564466626...	Apr 21, 2017 ...	Akash Soni	604430370	AkashV1	604430370	en	<a href="http://...	@airindian ...
8564440645...	Apr 21, 2017 ...	Akash Soni	604430370	?	-1	en	<a href="http://...	IF U CAN
8579128013...	Apr 28, 2017 ...	Snow leopard	3626015	airindian	387573617	en	<a href="http://...	@airindian b...
8579071186...	Apr 28, 2017 ...	Conspiracy T...	106324944	?	-1	en	<a href="http://...	Hey @airindi...
8579048233...	Apr 28, 2017 ...	IITL Summit...	8253161138...	?	-1	en	<a href="http://...	We acknowle...
8579035109...	Apr 28, 2017 ...	Racetherapy	396786833	timesofindia	134758540	en	<a href="http://...	@timesofindi...
8579024294...	Apr 28, 2017 ...	Cool1	75252600	ShereenBhan	96087281	en	<a href="http://...	@ShereenBh...
8579012278...	Apr 28, 2017 ...	Common man	3309762926	ShereenBhan	96087281	en	<a href="http://...	@ShereenBh...
8579000404...	Apr 28, 2017 ...	Ramya	8578900519...	airindian	387573617	en	<a href="http://...	@airindian i...
8578996182...	Apr 28, 2017 ...	SCM Consult...	286294527	ShereenBhan	96087281	en	<a href="http://...	@ShereenBh...
8578992554...	Apr 28, 2017 ...	Shereen Bhan	96087281	?	-1	en	<a href="http://...	@jayantsinh...
8578986659...	Apr 28, 2017 ...	Mohammed...	107937639	airindian	387573617	en	<a href="http://...	@airindian h...
8578966530...	Apr 28, 2017 ...	gavish george	87C94846	airindian	387573617	en	<a href="http://...	@airindian y...

**Fig 5.2: Display table of searched twitter results in Rapidminer.**

## 5.6 Data Cleaning and Pre-Processing

Before analyzing the records, they need to be pre-processed to remove different symbols and URLs which are irrelevant for sentiment analysis. Following steps were taken in data pre-processing.

1. Convert the tweets to lower case.
2. URLs –The URLs in the tweets don't fetch any information about the sentiment of the tweet, so eliminate all of these URLs and replace with generic word URL
3. @username - Replace "@username" with generic word AT\_username using Search and Replace tool in Excel.
4. #hashtag - hash tags can give us some useful information, so it is useful to replace them with the exact same word without the hash. (RavikantRajBlog, 2015) E.g. #jetairways was replaced with 'jetairways'

## **5.7 TOOLS USED**

### **1. IBM SPSS Text Analytics for Surveys 4.0.1**

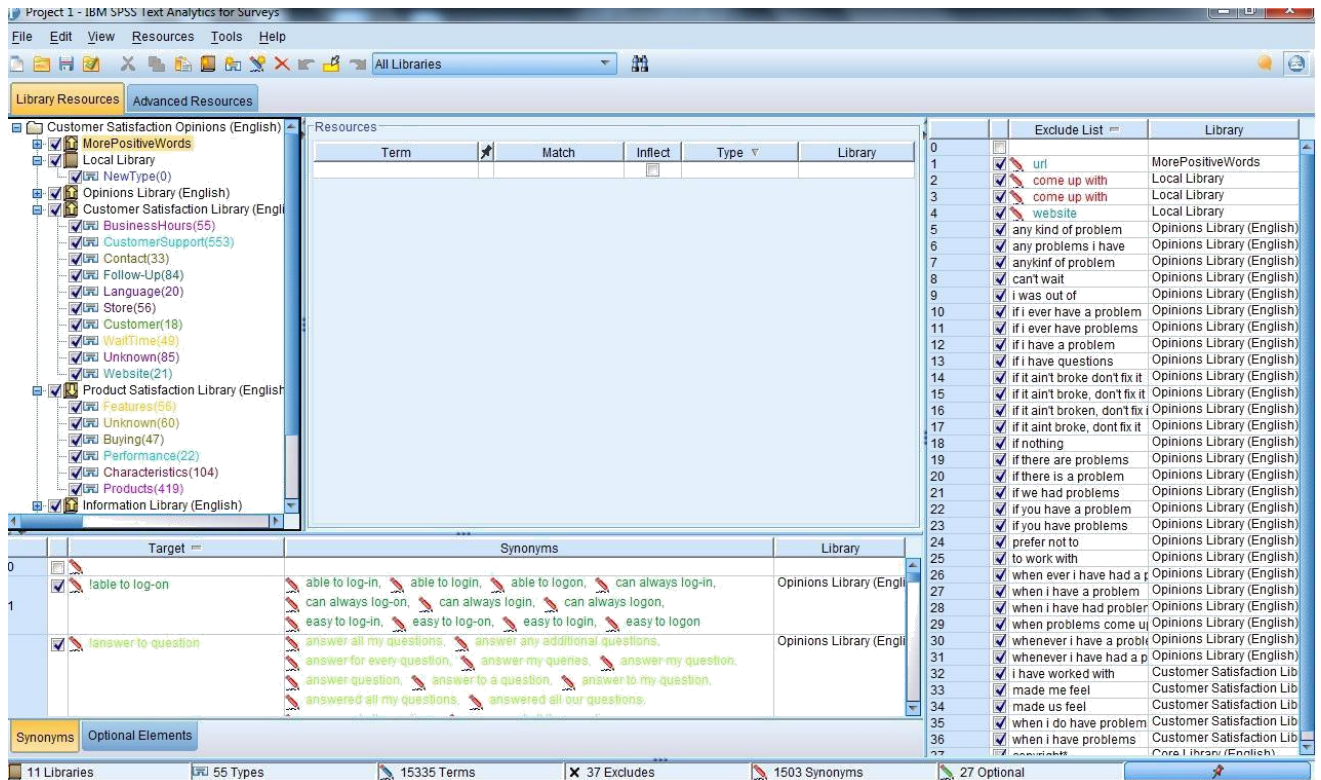
SPSS Text Analytics for Surveys combines advanced linguistic technologies to reliably extract and classify key concepts from open-ended survey responses with manual techniques. It has robust category building algorithms and simple drag-and-drop functionality, which enables one to create categories, or “codes,” into which responses will be categorized. The categories produced can also be reused to provide consistent results across the same or similar studies.

During the extraction of key concepts and ideas from your responses, it relies on linguistics-based text analysis. This approach offers the speed and cost effectiveness of statistics-based systems. And it also offers a far higher degree of accuracy. Linguistics-based text analysis is based on the field of study known as natural language processing, also known as computational linguistics.

For a thorough analysis, SPSS Text Analytics for Surveys provides a list of lexicon libraries including:

1. Customer Satisfaction Library
2. Product Satisfaction Library
3. Opinions Library
4. Slang Library
5. Information Library
6. Emoticon Library

These dictionaries provide an effective way of classifying text as expressing positive or negative sentiment (Apoorv Agarwal, 2011). In the figure 4.3, the left pane shows the different libraries present in the text analysis software and the types, the right pane shows exclude list (stop words), bottom pane shows synonyms and optional elements.



**Fig 5.3: SPSS Text Analytics for Surveys Libraries**

Each library has further certain types defined in it. For example: The Opinions library has Positive, Positive Attitude, Positive Budget, Positive Competence, Positive Feeling, Positive Functioning, Negative, Uncertain, Negative Attitude, Negative Budget, Negative Competence, Negative Feeling, Negative Functioning types defined under it.

Each type has a number of words defined under it. In figure 4.4, the center pane shows words belonging to 'Negative' Type in Opinions Library. Further, in figure 4.5, the center pane shows words belonging to 'Positive Attitude' Type in Opinions Library.

Term	Match	Inflect	Type	Library
a bit less than expected	Entire (no compounds)	<input type="checkbox"/>	Negative	Opinions Library (Engl)
a little doubt	Entire (no compounds)	<input type="checkbox"/>	Negative	Opinions Library (Engl)
a little tired of	Entire (no compounds)	<input type="checkbox"/>	Negative	Opinions Library (Engl)
abashed	Entire (no compounds)	<input type="checkbox"/>	Negative	Opinions Library (Engl)
abhor	Entire (no compounds)	<input type="checkbox"/>	Negative	Opinions Library (Engl)
abnormal	Entire (no compounds)	<input type="checkbox"/>	Negative	Opinions Library (Engl)
abominable	Entire (no compounds)	<input type="checkbox"/>	Negative	Opinions Library (Engl)
abrasive	Entire (no compounds)	<input type="checkbox"/>	Negative	Opinions Library (Engl)
abrupt	Entire (no compounds)	<input type="checkbox"/>	Negative	Opinions Library (Engl)
absolute nightmare	Entire (no compounds)	<input type="checkbox"/>	Negative	Opinions Library (Engl)
absurd	Entire (no compounds)	<input type="checkbox"/>	Negative	Opinions Library (Engl)
abused	Entire (no compounds)	<input type="checkbox"/>	Negative	Opinions Library (Engl)
abysmal	Entire (no compounds)	<input type="checkbox"/>	Negative	Opinions Library (Engl)
achilles heel	Entire (no compounds)	<input type="checkbox"/>	Negative	Opinions Library (Engl)
add more	Entire (no compounds)	<input type="checkbox"/>	Negative	Opinions Library (Engl)
adverse	Entire (no compounds)	<input type="checkbox"/>	Negative	Opinions Library (Engl)
adverse reaction	Entire (no compounds)	<input type="checkbox"/>	Negative	Opinions Library (Engl)
afraid	Entire (no compounds)	<input type="checkbox"/>	Negative	Opinions Library (Engl)
aggravating	Entire (no compounds)	<input type="checkbox"/>	Negative	Opinions Library (Engl)
aggravating	Entire (no compounds)	<input type="checkbox"/>	Negative	Opinions Library (Engl)
aimless	Entire (no compounds)	<input type="checkbox"/>	Negative	Opinions Library (Engl)
alarming	Entire (no compounds)	<input type="checkbox"/>	Negative	Opinions Library (Engl)
allergic	Entire (no compounds)	<input type="checkbox"/>	Negative	Opinions Library (Engl)
along time	Entire (no compounds)	<input type="checkbox"/>	Negative	Opinions Library (Engl)
alongtime	Entire (no compounds)	<input type="checkbox"/>	Negative	Opinions Library (Engl)
always late	Entire (no compounds)	<input type="checkbox"/>	Negative	Opinions Library (Engl)
always the same	Entire (no compounds)	<input type="checkbox"/>	Negative	Opinions Library (Engl)
ambig	Entire (no compounds)	<input type="checkbox"/>	Negative	Opinions Library (Engl)

Fig 5.4: SPSS Negative Words Library

Term	Match	Inflect	Type	Library
100 % commitment	Entire (no compounds)	<input type="checkbox"/>	PositiveAttitude	Opinions Library (Engl)
100% commitment	Entire (no compounds)	<input type="checkbox"/>	PositiveAttitude	Opinions Library (Engl)
accomadating	Entire (no compounds)	<input type="checkbox"/>	PositiveAttitude	Opinions Library (Engl)
accomidating	Entire (no compounds)	<input type="checkbox"/>	PositiveAttitude	Opinions Library (Engl)
accomidating	Entire (no compounds)	<input type="checkbox"/>	PositiveAttitude	Opinions Library (Engl)
accommodating	Entire (no compounds)	<input type="checkbox"/>	PositiveAttitude	Opinions Library (Engl)
accommodating	Entire (no compounds)	<input type="checkbox"/>	PositiveAttitude	Opinions Library (Engl)
adorable	Entire (no compounds)	<input type="checkbox"/>	PositiveAttitude	Opinions Library (Engl)
aim to please	Entire (no compounds)	<input type="checkbox"/>	PositiveAttitude	Opinions Library (Engl)
always ready to help	Entire (no compounds)	<input type="checkbox"/>	PositiveAttitude	Opinions Library (Engl)
always smile	Entire (no compounds)	<input type="checkbox"/>	PositiveAttitude	Opinions Library (Engl)
always taking care	Entire (no compounds)	<input type="checkbox"/>	PositiveAttitude	Opinions Library (Engl)
always there to help	Entire (no compounds)	<input type="checkbox"/>	PositiveAttitude	Opinions Library (Engl)
ambitious	Entire (no compounds)	<input type="checkbox"/>	PositiveAttitude	Opinions Library (Engl)
amiable	Entire (no compounds)	<input type="checkbox"/>	PositiveAttitude	Opinions Library (Engl)
amicable	Entire (no compounds)	<input type="checkbox"/>	PositiveAttitude	Opinions Library (Engl)
approachable	Entire (no compounds)	<input type="checkbox"/>	PositiveAttitude	Opinions Library (Engl)
approachable	Entire (no compounds)	<input type="checkbox"/>	PositiveAttitude	Opinions Library (Engl)
attended	Entire (no compounds)	<input type="checkbox"/>	PositiveAttitude	Opinions Library (Engl)
attentative	Entire (no compounds)	<input type="checkbox"/>	PositiveAttitude	Opinions Library (Engl)
attention to detail	Entire (no compounds)	<input type="checkbox"/>	PositiveAttitude	Opinions Library (Engl)
attentive	Entire (no compounds)	<input type="checkbox"/>	PositiveAttitude	Opinions Library (Engl)
bend backwards to help	Entire (no compounds)	<input type="checkbox"/>	PositiveAttitude	Opinions Library (Engl)
big smile	Entire (no compounds)	<input type="checkbox"/>	PositiveAttitude	Opinions Library (Engl)
care about	Entire (no compounds)	<input type="checkbox"/>	PositiveAttitude	Opinions Library (Engl)
care about my issues	Entire (no compounds)	<input type="checkbox"/>	PositiveAttitude	Opinions Library (Engl)
care about my problems	Entire (no compounds)	<input type="checkbox"/>	PositiveAttitude	Opinions Library (Engl)
care deeply about	Entire (no compounds)	<input type="checkbox"/>	PositiveAttitude	Opinions Library (Engl)

Fig 5.5: SPSS Positive Words Library

3. **Microsoft Excel** -Microsoft Excel is a spreadsheet developed by Microsoft for Windows, Mac OS X, Android and iOS. It features calculation, pivot tables, graphing tools. The twitter data is cleaned and analyzed using Microsoft Excel once sentiment analysis and categorization is done.
4. **Tableau** - Tableau, developed by Tableau Software, is software used for business intelligence and analytics. It produces family of interactive data visualization based on data.

## 5.8 Steps involved in sentiment analysis and category creation

1. Extraction of Concepts and Patterns: The extraction engine identifies candidate key terms. These key terms are grouped under a main concept. Concepts are then grouped into types, which are a collection of similar words such as positive opinion words, words related to customer support etc.

For example: Words and patterns such as overcharge, cheap fare, high pricing, inexpensive can all be grouped under the concept 'fare'.

Depending on the dictionary the word matches, it is categorized into different types. For Example: Overcharge, High Pricing come under 'NegativeBudget' type in Opinions Library.

Cheap Fare, Inexpensive come under 'PositiveBudget' type in Opinions Library.

2. Refine and fine tune extraction results: The automatic extraction results depend on the existing linguistic resources that come bundled with the software. The results need to be fine-tuned for more accurate results. Fine tuning can be done in the following ways:
  - a. Adding new libraries, types and words specific to the domain of words being analyzed. For example: Words related to airline industry can be added to the linguistic resources for generating more accurate 'concepts'.
  - b. Adding new synonyms to the existing words from the corpus being analyzed.
  - c. Excluding irrelevant words and patterns from further extractions.
  - d. Changing the 'type' of current words in the context of the text being analyzed. For example: The word 'cool' can be generally associated as a positive feature

when associated with humans, however if the comments being analyzed are from a feedback to a restaurant, 'cool' would be a negative word to associate with food.

3. **Build Categories:** The software uses linguistic techniques to build categories based on the concepts and types extracted.

For example: <Fare> + <Negative> may be a category with all the terms related to type <fare> and <negative> together.

4. **Refine Categories:** The categories built automatically need further refinement for more accurate results. Fine tuning of categories can be done in the following ways:

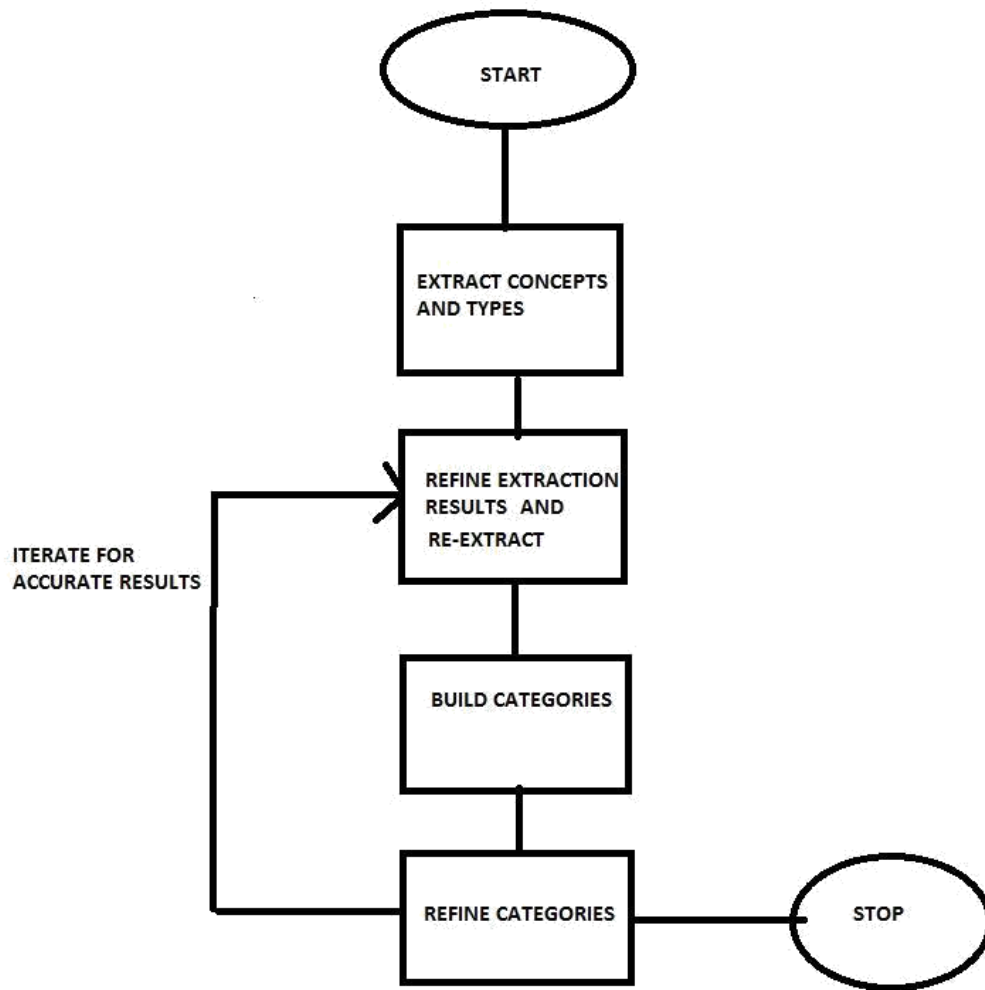
- a. **Define category rules:** Rules can be defined to make categories. One can combine the different 'types' and 'concepts' extracted in previous steps to make categories. Boolean operators AND, OR and NOT can be used to combine types.

For example:

Category 'Negative Service Quality' can be defined rules such as:

[waiting time] & <service> & (<Negative> | <NegativeFeeling Emoticon>)

- b. Manually forcing comments to a particular category.
- c. Combining similar categories.



**Fig 5.6: Flowchart for Sentiment Analysis and Category Creation**

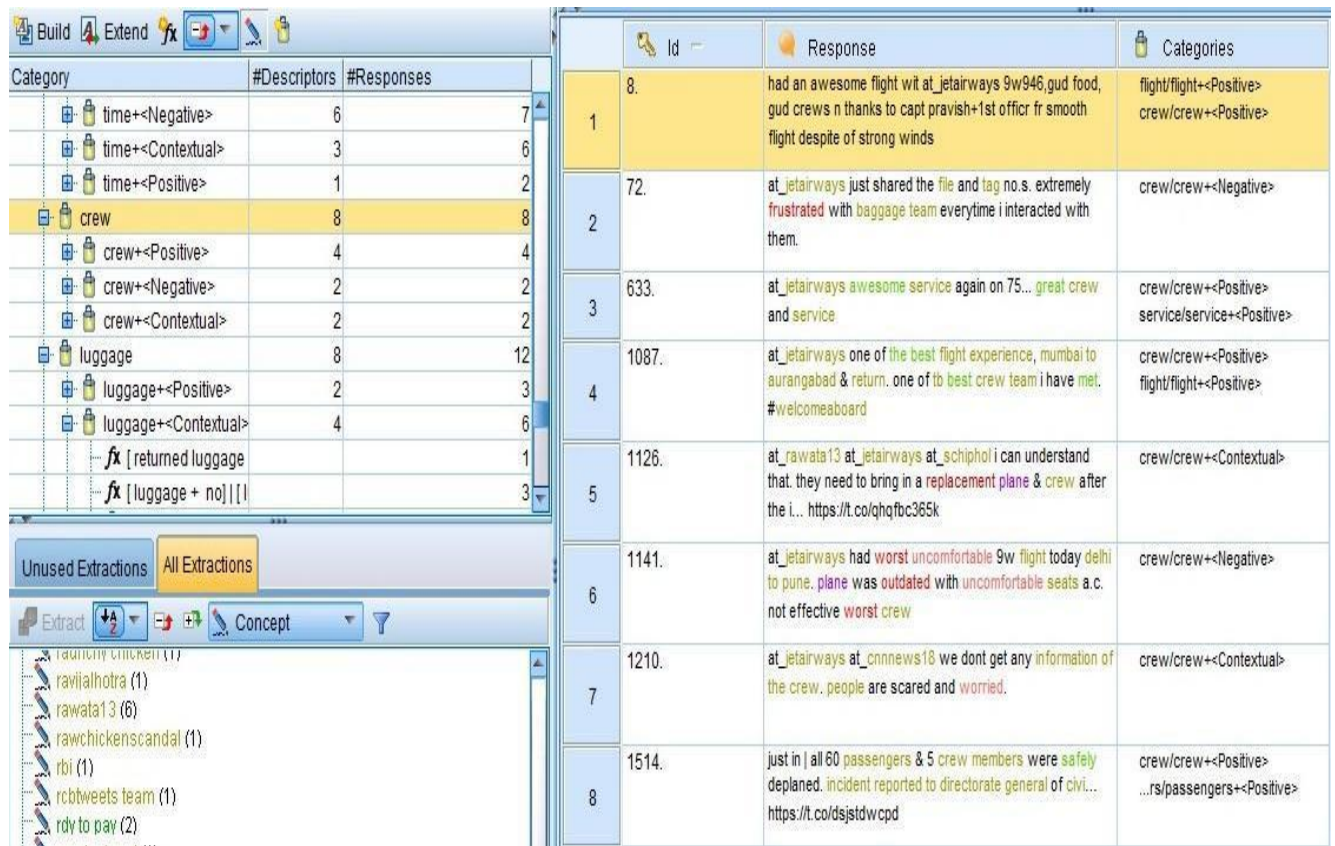


# CHAPTER-6

## ANALYSIS, INFERENCES AND RECOMMENDATIONS

### 6.1 Tweet Analysis of JetAirways

A corpus of 1803 tweets addressed to @jetairways was analyzed to identify customer sentiment polarity. The records were further classified into categories such as Customer Support, Service Quality, Fare Contextual to identify areas which attracted maximum positive and negative comments.



**Fig 6.1: Sentiment Analysis and Category creation for JetAirways data in SPSS TAS.**

Contextual attracted the maximum number of comments, followed by service quality, customer support, fare related issues.

Out of the total of 1803 tweets, 797 tweets corresponded to negative sentiment, 398 to positive sentiment and rest 608 were neutral/contextual.

	Positive	Negative	Total
Fare	61	175	236
Service Quality	181	330	511
Customer Support	156	292	448
Contextual			608

**Table 6.1: Category wise negative and positive responses for JetAirways India**



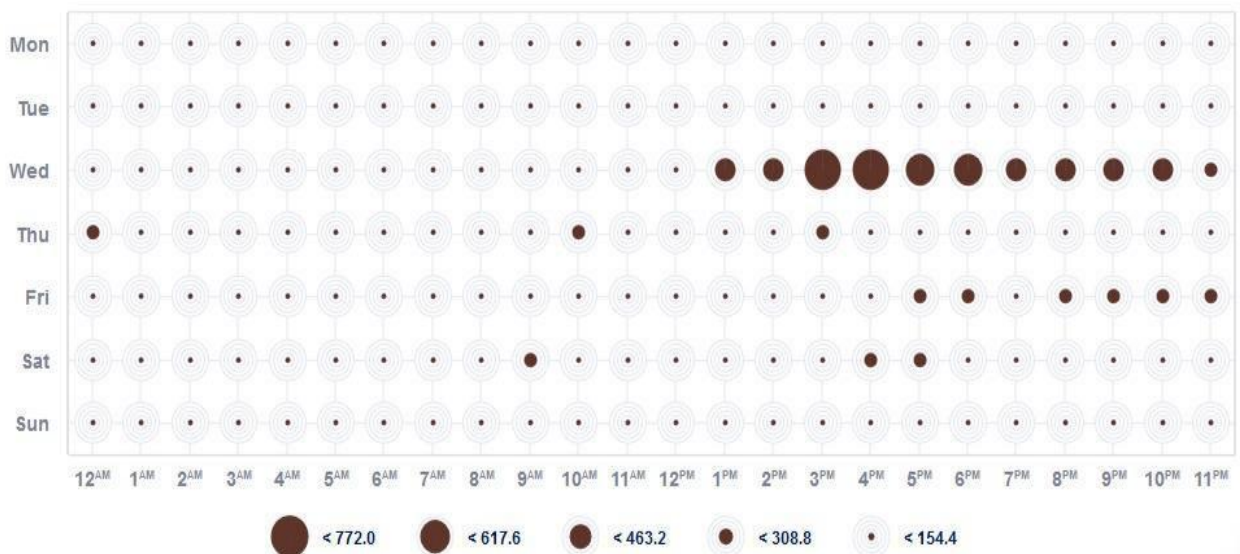
**Fig 6.2: Distribution of positive and negative tweets across various categories for JetAirways**



Word. Color shows details about Type. Size shows sum of Frequency.

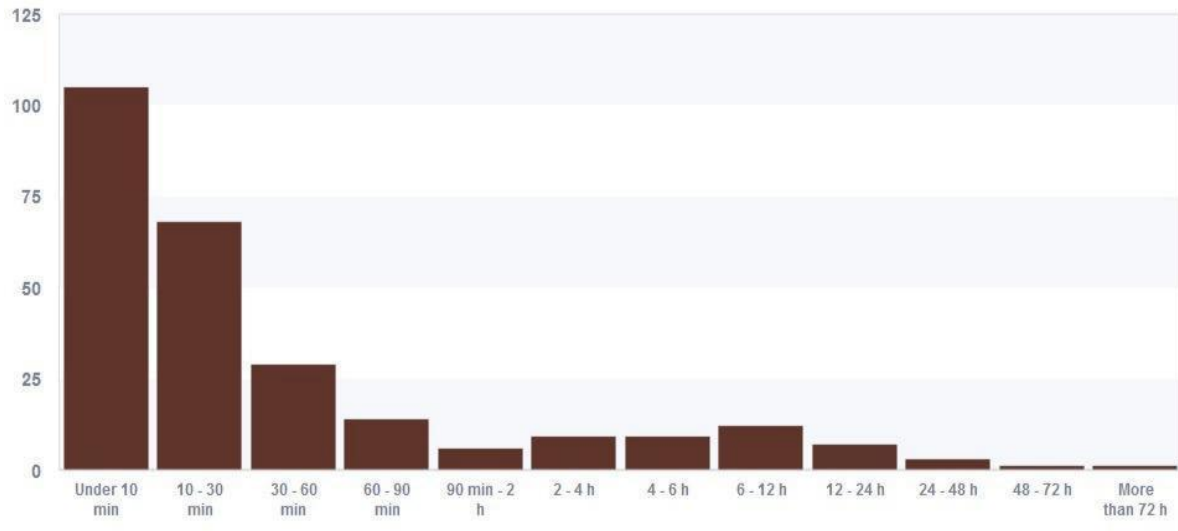
**Figure 6.3: Word Cloud of High Frequency Positive and Negative Words for Jet Airways Cabs**

As seen from the above figure, high frequency negative words such as cancel, problem, bad, complaint, declined, pathetic, wrong, unable, frustrating, expensive, overcharge, not fair, loot etc. were extracted. Amongst the positive words, thanks, fast, excellent, available, problem resolved, answered properly, resolved were common which mainly reflect effective customer support system of Jet Airways.



**Fig 6.4: Trend analysis of tweets over different time slots for Jet Airways (Apr 15-Apr 30)**

The above graph shows the aggregated values for interaction spread over the week. From the above graph, it is observed that maximum number of tweets were captured in 1 pm-4 pm slot, followed by 5 pm-9pm.



**Fig 6.5: Response time segment for JetAirways tweets**

From the above graph, it is observed that most of the queries are responded under 60 min time.

### **6.1.1 Inferences and Recommendations for JetAirways:**

1. Based on the above data, service quality has attracted maximum share of negative comments and the company needs to focus in this area. The customers also faced issues related to high prices, service quality, delayed flight, bad customer support and crew being rude and unprofessional.
2. Customer support attracted more positive than negative comments, positive being related to problem being resolved, helpful; answering queries properly, knowledgeable executives.
3. “Bad service quality” was the main captured theme from the analysis which can be because people tend to tweet more about negative experience rather than positive ones.

4. Issues with luggage handling were also common. The company needs to focus on overall better services including passengers and their luggage.
5. The number of tweets was highest in the 1 pm – 4 pm slot followed by 5 pm – 9pm slot which shows that maximum tweets were posted during office hours, followed by the tweets during the time of travel, followed by post-midnight tweets. Wednesday observed the maximum number of tweets with the number gradually dipping throughout the week. Customer support can be done managed better using this data to handle heavy volume of tweets addressed to customer care for providing quick and effective response.
6. The difference in number of positive and negative posts is clearly visible which is not good for the health of the brand.

## **6.2 Tweet Analysis of AirIndia**

A corpus of 780 tweets addressed to @airindiain was analyzed to identify customer sentiment polarity. The records were further classified into categories such as Customer Support, Fare Related, Service Quality, Contextual to identify areas which attracted maximum positive and negative comments.

Out of the total of 780 tweets, 284 tweets corresponded to negative sentiment, 200 to positive sentiment and rest 296 were neutral/contextual.

Id	Response	Categories
33.	at_airindian take a bow, perfect service for a member of star alliance <a href="https://t.co/qe0qxrvtzu">https://t.co/qe0qxrvtzu</a>	service/service+<Positive>
35.	pathetic service and absolutely miserable experience with at_airindian today. was denied boarding on bom-del flight coz it was overbooked	experience+<Negative> service/service+<Negative>
73.	at_airindian very disappointing service, left one of the bags in mumbai i was standing in check-in queue for almost an hour!	service/service+<Negative>
105.	at_airindian it's 30 minutes that i'm placed on hold. no response from a rep. horrible service!	service/service+<Negative>
156.	thanks for living up to my expectations at_airindian .poor service as always. have to travel more out of compulsion than choice. #bagdogra	service/service+<Negative> travelers/travel+<Contextual>
251.	at_airindian your customer case on this case has been absolutely helpless and frustrating too. your overall service literally sucks.	service/service+<Negative>
277.	at_jayantsinha at_airindian sir pls do look at this. how horrifying ur services r. nobody cn even dare 2 travel wth #ai... <a href="https://t.co/wxulwytwj">https://t.co/wxulwytwj</a>	service/service+<Negative>
308.	at_airindian traveling in ai 608 from blr to mum pathetic service no air-conditioning in cabin ground staff confused unable to manage queue.	service/service+<Negative>
393.	at_airindian i'm still waiting for your revert or call. . amazing service...	...ain/airindian+<Negative> service/service+<Positive>
431.	really poor service at_airindian - ai 641 . 50 minutes plus for baggage drop. almost zero ground support. no guidance.c'mon you can do better	service/service+<Negative>

Fig 6.6: Sentiment Analysis and Category creation for AirIndia data in SPSS TAS.

	Positive	Negative	Total
Fare	31	52	83
Service Quality	84	124	208
Customer Support	85	108	193
Contextual			296

Table 6.2: Category Wise Positive and Negative Responses for AirIndia



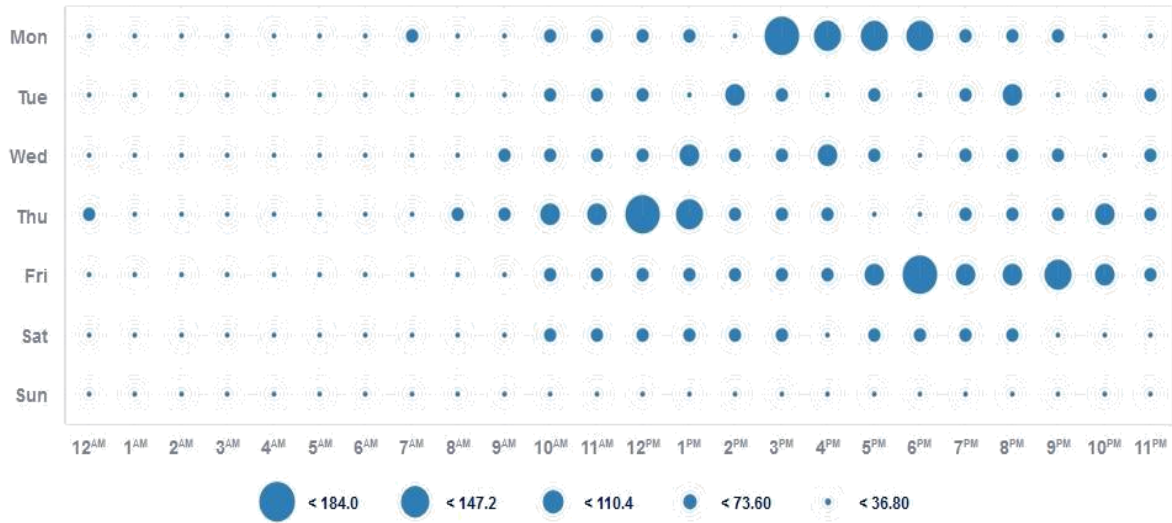
**Fig 6.7: Distribution of positive and negative tweets across various categories for AirIndia**

Customer support attracted maximum number of comments followed by service quality, fare related issues, driver experience, app and contextual.



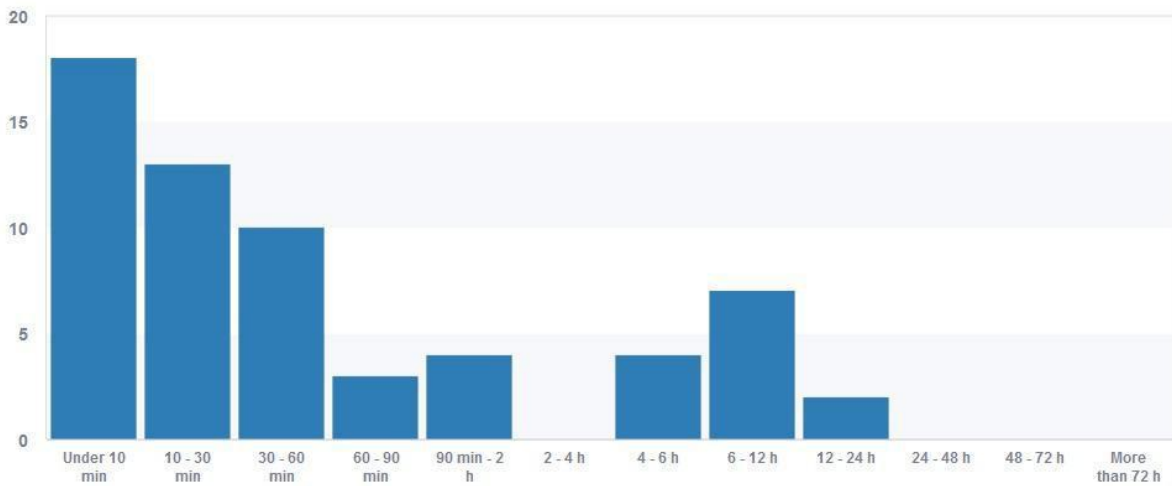
**Figure 6.8: Word Cloud of High Frequency Positive and Negative Words for AirIndia**

High frequency positive words like better, updated, 😊 (happy smiley), great, hope, quick, best etc. were extracted from the data. Negative words include unprofessional, pathetic, problem, worst, ridiculous, missed etc.



**Fig 6.9: Trend analysis of tweets over different time slots for AirIndia (Apr15-Apr30)**

From the above graph, it is observed that Monday received the maximum number of tweets followed by Thursday and Friday. Maximum number of tweets were captured in time slot 12pm – 4pm followed by 6pm-9pm.



**Fig 6.10: Response time segment for AirIndia tweets(Apr15-Apr30)**

From the above graph, it is observed that response time is around 90 minutes and in some cases even more than 6 hours.



### **6.1.1 Inferences and Recommendations for AirIndia:**

1. Service quality issues, fare related issues and customer care received similar share of negative comments with service quality issues leading slightly than others. Bad customer service, pathetic service and delayed flight were the main themes extracted which reflect the same. The company needs to focus on these areas.
2. Service Quality was again the main theme in a lot of tweets which is common in all the airlines under investigation.
3. Words such as unprofessional, pathetic, problem, worst, ridiculous, missed etc. dominated the list of negative words which again reflect on service quality issues and crew behavior.
4. No category received more positive than negative comments.
5. The number of tweets was highest in the 12 pm – 4 pm slot followed by 6 pm – 9 pm slot which shows that maximum tweets were posted during office hours, followed by the tweets during the time of travel, followed by post-midnight tweets. Monday observed the maximum number of tweets followed by Thursday and Friday. Customer support can be done managed using this data to handle high volume of tweets for providing quick and effective response.
6. There is small difference between the number of positive and negative posts which is good sign for the health of the brand.
7. Response time for the queries is not good which also leads to negative customer support and poor service quality.

---

## CHAPTER-7

---

### LIMITATIONS OF THE STUDY

1. Sarcasm can't be detected by automated sentiment analysis tools as it involves use of positive words to express a negative meaning or vice versa.
2. The precision of sentiment analysis tool is not 100% and the tool can report false positives and false negatives.
3. The recall of sentiment analysis tool is not 100% and the tool can possibly report less number of positives and negative comments than the actual number.
4. All the tweets associated with the taxi aggregators during the period for which the analysis is done could not be captured as users can use any expression in hashtags to express their opinions. For example: The tweets addressed to ola cabs can be addressed with @jetairways, @jetairlines, #jetsucks, #ola\_airways,etc. There is no standard convention of addressing companies in tweets and hence it is impossible to consider all the cases.
5. Only one month's data was used to conduct the study which could be a limiting factor in determining the trends.

## **CHAPTER-8**

---

### **CONCLUSION**

Tweets addressed to twitter handles of Jetairways and AirIndia for the month of April were extracted and analyzed. The tweets were analyzed for their sentiment polarity in terms of positive, negative and neutral polarities. Further, the study categorized the tweets into various service KPIs including Service Quality, Fare related issues, Customer Support etc and found areas which required improvements and areas of customer delight. The study also extracted various entities to understand the general trends in the tweets.

Response time Analysis, time slot trend analysis and weekday-volume of tweets trend analysis was done to uncover insights in the tweet patterns. The analysis graphs were plotted to determine positive and negative effect of tweets to the brand reputation. Recommendations based on the data and its analysis were made to various departments of the aggregators.

The study concludes that the social media data can indeed be a rich source of information which, if harnessed by the marketers, can lend organisations an upper edge over its competitors. Analysis of this huge chunk of unstructured data can lead to actionable insights and help marketers in better understanding of customer's behaviour, perceptions and feelings. It can help organisations in improving their products and services by listening to their customers in real time. Organisations should build suitable talent and invest in social media mining to reap the benefits of this data.

## REFERENCES

Shintaro Okazaki, Ana M. Diaz Martin, Hector David et al (2014). Using Twitter to engage with customers: a data mining approach. Retrieved from: [www.emeraldinsight.com/1066-2243.htm](http://www.emeraldinsight.com/1066-2243.htm).

Fotis Misopoulos, Miljana Mitic, Alexandros Kapoulas et al (2014). Uncovering customer service experiences with Twitter: the case of airline industry. Retrieved from: [www.emeraldinsight.com/0025-1747.htm](http://www.emeraldinsight.com/0025-1747.htm)

Apoorv Agarwal, Boyi Xie, Iia Vovsha et al (2011), SENTIMENT ANALYSIS OF TWITTER DATA. Retrieved from: <http://www.cs.columbia.edu/~julia/papers/Agarwaletal11.pdf>.

SocialBakers.com. (2016). Social Media Marketing, Statistics and Monitoring Tools. Retrieved on 30<sup>th</sup> April 2017, from <http://www.socialbakers.com/>.

Ernst Young (2015), Social Media Marketing India Trends Study 2<sup>nd</sup> Edition. Retrieved on 30<sup>th</sup> April 2017, from [http://www.ey.com/Publication/vwLUAssets/EY-social-media-marketing-india-trends-study-2014/\\$FILE/EY-social-media-marketing-india-trends-study-2014.pdf](http://www.ey.com/Publication/vwLUAssets/EY-social-media-marketing-india-trends-study-2014/$FILE/EY-social-media-marketing-india-trends-study-2014.pdf)

Ravikant Raj Blog (2015), How to build a twitter sentiment analyzer? Retrieved on 30<sup>th</sup> Jan 2016, from: <http://ravikiranj.net/posts/2012/code/how-build-twitter-sentiment-analyzer/>

Statista.com(2016). The statistics Portal | Statistics and facts on internet usage in India. Retrieved on 20<sup>th</sup> April 2017, from: <http://www.statista.com/topics/2157/internet-usage-in-india/>

What is opinion mining (sentiment mining)? - Definition from WhatIs.com. (2016). SearchBusinessAnalytics(2016). Retrieved 2 May 2017, from <http://searchbusinessanalytics.techtarget.com/definition/opinion-mining-sentiment-mining>

Sentiment analysis. (2016). Wikipedia. Retrieved 2 May 2017, from [https://en.wikipedia.org/wiki/Sentiment\\_analysis](https://en.wikipedia.org/wiki/Sentiment_analysis)

Opinion Mining. (2016). swshar. Retrieved from  
<https://swshar.wordpress.com/2016/04/14/56/>

# ANNEXURE A

Screenshot of twitter Extraction using Rapidminer

The screenshot displays the Rapidminer software interface. At the top, there is a menu bar with options: File, Edit, Process, View, Connections, Cloud, Settings, and Extensions. Below the menu is a toolbar with various icons for file operations and process execution. The main workspace is divided into several panels:

- Operators:** A list of available operators, including 'write excel' and 'Write Excel' under the 'Files (1)' category.
- Repository:** A panel for data storage, showing 'Samples', 'DB', 'Local Repository (DELL)', and 'Cloud Repository (disconnected)'. An 'Add Data' button is present.
- Process:** The central workspace showing a workflow diagram. It starts with an input 'inp' leading to a 'Search Twitter' operator (orange box with a Twitter icon), which outputs to a 'Write Excel' operator (purple box with a document icon). The output of 'Write Excel' is labeled 'res'.
- Parameters:** A panel for configuring the 'Search Twitter' operator. It includes fields for:
  - connection: twitter1
  - query: @airindia-in-rt-htp
  - result type: recent or pop...
  - limit: 2000
  - since id: (empty)
  - max id: (empty)
  - language: (empty)
- Help:** A panel providing information about the 'Search Twitter' operator, including its tags and a synopsis: 'This operator searches for Twitter statuses.'

# ANNEXURE B

Screenshot of Sentiment Analysis by SPSS Text Analytics for Surveys

The screenshot displays the IBM SPSS Text Analytics for Surveys interface. The main window is titled "Project 1 - IBM SPSS Text Analytics for Surveys". The interface is divided into several panes:

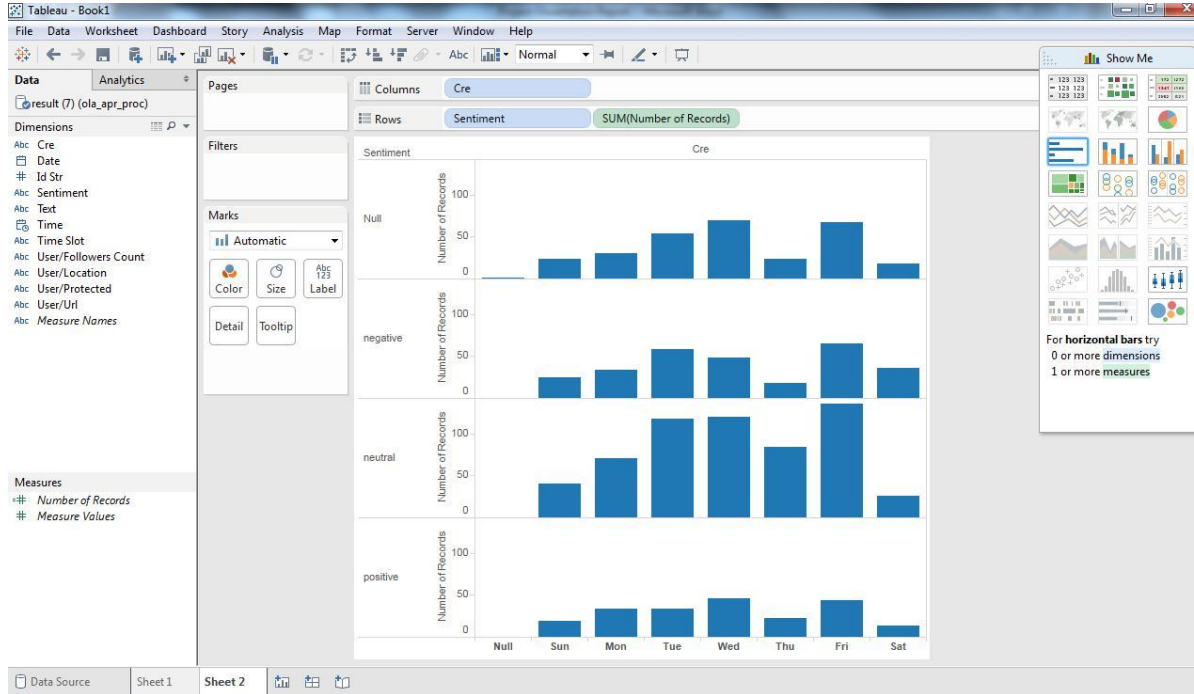
- Category List (Left):** A tree view showing categories and their associated counts. The "service" category is highlighted, with 5 descriptions and 17 responses. Other categories include "airindian", "flight", "air india", "airlines", "passengers", and "jayantsinha".
- Unused Extractions (Middle-Left):** A list of extractions that have not been used in the analysis, such as "business class (3)", "bussterz (1)", "duszy (2)", "cabin ground staff (1)", etc.
- Main Table (Right):** A table showing the results of the sentiment analysis. It has columns for "Id", "Response", and "Categories". The table lists 12 rows of data, each representing a survey response and its assigned sentiment categories.

Id	Response	Categories
1	at airindian take a bow, perfect service for a member of star alliance <a href="https://t.co/qe0qrvuzu">https://t.co/qe0qrvuzu</a>	service/service+<Positive>
2	pathetic service and absolutely miserable experience with at airindian today, was denied boarding on bom-del flight coz it was overbooked	experience+<Negative> service/service+<Negative>
3	at airindian very disappointing service, left one of the bags in mumbai!! i was standing in check-in queue for almost an hour!!	service/service+<Negative>
4	at airindian it's 30 minutes that i'm placed on hold, no response from a rep, horrible service!!	service/service+<Negative>
5	thanks for living up to my expectations at airindian, poor service as always, have to travel more out of compulsion than choice. #bagdogra	service/service+<Negative> travelers/travel+<Contextual>
6	at airindian your customer case on this case has been absolutely helpless and frustrating too, your overall service literally sucks.	service/service+<Negative>
7	at jayantsinha at airindian sir pls do look at this, how horrifying ur services r, nobody on even dare 2 travel with #ai... <a href="https://t.co/wxulwytwjg">https://t.co/wxulwytwjg</a>	service/service+<Negative>
8	at airindian traveling in ai 608 from bnr to mumbai pathetic service no air-conditioning in cabin ground staff confused unable to manage queues.	service/service+<Negative>
9	at airindian i'm still waiting for your revert or call... amazing service...	...ian/airindian+<Negative> service/service+<Positive>
10	really poor service at airindian - ai 641. 50 minutes plus for baggage drop, almost zero ground support, no guidance, whom you can do better	service/service+<Negative>
11	at jayantsinha at airindian i fly 100k miles a year on different airlines, i can say for sure air india offers the best meal service among all.	service/service+<Positive>
12	at airindian you guys only know how to harass people, pathetic service and fraud airport	airlines/airlines+<Negative> service/service+<Negative>

The bottom status bar shows "30 (56) Categories" and the system tray displays the time as 7:39 PM on 4/30/2017.

# ANNEXURE C

## Screenshot of Tableau Worksheet





# Akshay darmwal plag report

## ORIGINALITY REPORT

12%

SIMILARITY INDEX

10%

INTERNET SOURCES

2%

PUBLICATIONS

## PRIMARY SOURCES

1

[www.socialbeat.in](http://www.socialbeat.in)

Internet Source

2

[www.ibef.org](http://www.ibef.org)

Internet Source

3

Submitted to Vrije Universiteit Brussel

Student Paper

4

Submitted to Laureate Higher Education

Student Paper

5

[dmml.asu.edu](http://dmml.asu.edu)

Internet Source

6

[www.financialexpress.com](http://www.financialexpress.com)

Internet Source

7

[www.spssmr.com](http://www.spssmr.com)

Internet Source

8

[ravikiranj.net](http://ravikiranj.net)

Internet Source

Publication