November 2022

Comparison Study of Consumer’s Perception toward Urban Air Mobility in the United States and Rest of the World Using Social Media Information

S M Toki Tahmid
University of South Florida

Comparison Study of Consumer's Perception toward Urban Air Mobility in the United States and Rest of the World Using Social Media Information

by

S M Toki Tahmid

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science in Civil Engineering
Department of Civil and Environmental Engineering
College of Engineering
University of South Florida

Major Professor: Yu Zhang, Ph.D.
Fred Mannering, Ph.D.
Qing Lu, Ph.D.
Pei-Sung Lin, Ph.D.

Date of Approval:
October 31, 2022

Keywords: UAM, EVTOL, Vertiport, Natural Language Process, Sentiment Analysis

Dedication

I would like to thank Dr. Yu Zhang for all her support and guidance throughout my master's program. I would also like to thank my committee members, and my friends and family who have always supported me.

Table of Contents

List of Figures

Abstract

Urban Air Mobility (UAM) is a wide concept that will enable access to on-demand air mobility, cargo delivery, medical applications, and emergency services that is run through a connected and integrated transportation system. This study aims to investigate social media data to compare the major topics discussed in the United States and rest of the world and perform a sentiment analysis to see if the comments are positive, negative, or neutral. Twitter was the primary source of information that was used in this thesis, as it is a place where people interact, create, and share information and ideas online. In this thesis I used an analytical approach to analyze the social media information within the context of social network theory and used the sentiment expressed in the content as a proxy to measure the performance. Text mining techniques and machine learning algorithms were employed to examine the social media information, to collect tweets using certain keywords, and do a comparison study between the United States and Rest of the World.

Chapter 1: Introduction

Urban Air Mobility (UAM) is defined as an alternative form of transportation to carry passengers and goods in urban areas by air. It is considered as an intercity and intracity air passenger service, with a focus on services enabled by novel aircraft types with the capability for vertical takeoff and landing (VTOL). The operational concept that utilizes VTOL technology is often referred to as flying taxis or flying autonomous vehicles. UAM intends to be a safe and efficient mode of aviation, using highly automated aircrafts that operate at lower altitudes within urban and suburban areas. This revolutionary mode will use new electric motor aircraft that are safer, greener, quieter, and more economical than traditional helicopters and airplanes. In terms of service, UAM will be on-demand similar to the other transportation network companies (e.g., Uber or Lyft).

UAM can be used for several purposes such as rapid last-mile delivery of packages from local distributers to customers and efficiently moving people using air metro or air taxi, which are expected to be in the market by 2030. The market value for UAM is estimated to be 4 billion USD by year 2030 and is predicted to increase from 74 billion USD in year 2030 to 641 billion USD in year 2035 globally (Herman et al., 2019). Mayor and Anderson (2019) estimated that air taxis could have 400 million enplanements, which will represent 4% of the domestic trips, by 2050. Lineberger et al. estimated in 2021 that the U.S. market for air taxi will be valued at 115 billion USD and employ more than 280,000 workers by the year 2035. While market forecasts provide valuable insights, their estimates will always be associated with uncertainty due to variations in study scope and assumptions related but not limited to geography, market segmentation, timeline, and involvement or not involvement of military applications of electric vertical takeoff and landing (eVTOL) aircraft. This study aims to investigate social media data to compare the major topics discussed in the United States and the rest of the world in order to perform a sentiment analysis to see if the comments are positive, negative, or neutral. Twitter was used a primary source of information, as it is a place where people interact, create, and share information and ideas online. This study provides an analytical approach to examine social media information within the context of social network theory and use the sentiment expressed in the content as a proxy to measure the performance. In this study, text mining techniques and machine learning algorithms were employed to examine different social media information, and tweets were collected from Twitter using certain keywords related to our questions.

This thesis is organized into five sections. The first section provides an overview of UAM and its market demand analysis. The next section is the literature review and is divided into three subsections: 1) The history and public perception towards UAM, 2) The market studies of UAM, and 3) Inferring Public Perceptions by Analyzing Social Media Information. The third section describes the methodology that was used to perform topic modeling and sentiment analysis. The fourth section discusses the results that were collected and how they were used to do the analysis and comparison. In the final section, I conclude with a discussion of how UAM can impact the evolution of transportation system and suggestions for future research.

Chapter 2: Literature Review

2.1 History and Public Perception Towards UAM

From 1947 to 1971, helicopters were used by Los Angeles Airways to transfer people in the LA area, which included major attractions such as Los Angeles International Airport (LAX) and Disneyland. The operation was ceased as the airways experienced two mechanical failures in 1968 (Harrison, 2017). New York Airways also used helicopters to transport passengers between Manhattan and the three major airports in New York, LaGaudia Airport (LGA), John F. Kennedy International Airport (JFK), and Newark Liberty International Airport (EWR), from 1953 to 1979. The service stopped later due to several accidents caused by mechanical failures as well (Witken, 1979, Thipphavong et al., 2018, Mayor and Anderson, 2019). Compared to other transportation modes, helicopters have higher cost and generate more noise, deficiencies that are expected to be addressed in UAM. UAM has received much attention in recent years because it is expected to be a fast, time-saving, safe, convenient, and affordable mode of transportation (Baur, Schickram, Homulenko, Martinez, and Dyskin, 2018). Given that UAM avoids congestion and takes direct routes to destinations, the travel time is expected to be lower compared to ground transportation. While UAM is anticipated to be less expensive than helicopter service, it still could be more expensive than ground transportation modes. Generally, public opinion considers UAM as a mode for the rich, similar to how they consider the toll roads implemented in highway systems. This new mode of transportation is demanded by the passenger, as Uber Elevate calls it the birthplace of UAM (Holden and Goel, 2016). Even though UAM is in high demand, it still has some drawbacks to the community. Vascik and Hansman

(2017) explained the community issues related to aircraft noise, takeoff, and landing areas. Airbus did a survey in 2020 regarding the public perception of UAM. The survey was circulated through four geographic locations including Los Angeles, Mexico City, New Zealand, and Switzerland. Los Angeles was chosen as it drives the city's urge to explore new modes of transportation for its congestion and pollution. The study surveyed 1,540 members from those four geographic regions to understand their opinions toward UAM from different aspects, including noise, safety, equity, visual pollution, and privacy. From that survey, 56% of respondents expressed safety concerns, 49% of them expressed noise concern, 44% supported the idea of being initiative, and 41% considered the aircraft would be safe. From this survey, respondents aged between 25 and 34 years had the highest positive reaction (55%) about this new mode of transportation, and they believe it's safe, while the older respondents believe it is not safe (Yedavalli and Mooberry, 2020).

2.2 Market Studies of UAM

NASA conducted one study to identify market size and possible barriers for UAM when using it as an airport shuttle, air ambulance, and air taxi. It was found from the research that in the U.S. airport shuttle and air taxi are more feasible, with a significant market value of $500 billion (Booz Allen Hamilton, 2018). This prediction indicates that last-mile parcel delivery and air metro will be profitable by 2030, provided that regulations are in place to support the market. However, air taxi would take longer to become profitable given the higher costs and lower passenger capacity of aircraft. In the NASA survey, an unconstrained scenario was used in that passengers would be able to fly using UAM at any time to any destination without being impacted by weather, traffic, or infrastructure. Air ambulance does not have a viable market if it uses electric vertical takeoff and landing (eVTOL), but it may be applicable if hybrid VTOL aircraft are employed. The barrier that was found from this study was the high cost of service, which will be an economic challenge, but this can be solved if there are enough vehicles and component efficiency. Some factors that tend to reduce the market potential include weather, certification, public concern, infrastructure, legal issues, and regulation. Another survey by NASA shows that 25% of 2,500 respondents reported that they feel secure with this new technology, and about 25% responded that they will not use this service (NASA, 2018). Most of the concern falls into five main categories: safety, job security, privacy, environmental threats, and visual and noise disruption (NASA, 2018).

In 2019, Deloitte conducted a market study on the intracity and regional market and estimated that UAM will reach a market of $1B by 2025 and $13.8B by 2040 for intracity, and it will reach $2.6B by 2025 and $3.9B by 2040 for the regional market (Lineberger et al., 2019). In 2020, Frost and Sullivan predicted that the market for passenger service globally will reach $3B by 2023 (Xu, 2020). According to Klynveld Peat Marwick Goerdeler (KPMG), the market for intracity and regional service globally will reach 12 million emplacements per year by 2040 and 400 million emplacements per year by 2050 (Mayor and Anderson, 2019). The market possibility for UAM is approximated within a wide range because it is nonexistent currently, but as this new technology is young and quickly growing, the market is providing hope. Porsche Consulting estimated in 2018 that the market would reach tens of billions of dollars by 2035.

2.3 Inferring Public Perceptions by Analyzing Social Media Information

Over the last few years, the research based on social media data has grown significantly. This section discusses social media theory and techniques that have been used to infer scientific conclusions about a variety of subject matters.

Social network theory deals with the interrelationship between components of a network composed of people, organizations, or groups (Williams and Durrance, 2008). The interaction within the network promotes collaborations among users, which could generate valuable information and insight into business development, marketing, research, municipal services, and so on. Twitter allows users to proactively express and exchange opinions, share knowledge and experiences, and develop their social networks. The use of social media such as Facebook, Twitter, YouTube, Instagram, and Weibo has grown dramatically in the past decade. These social media channels, which are Internet-based Web 2.06 applications, have provided a platform for people to freely express their observations, opinions, feelings, and frustrations. The number of users has increased exponentially over the years. For example, Twitter has more than 60 million monthly active users in the United States as of Q1 2019. People are constantly embracing these social media channels, sharing information and updates about their lives, community, political opinions, and frustrations about topics of immediate concern. To put things into perspective, Twitter generates more than 500 million tweets each day and Facebook has more than 4.75 billion posts per day (Dhaoui, Webster, and Tan, 2017). This massive crowdsourcing information is a desirable data source for researchers, business managers, auditors, and government agencies. Researchers are exploring the usage of social media, particularly in the accounting and auditing domains. In recent years, accounting and auditing are progressively embracing the new data environment, potentially transforming the nature, usage, and decision processes related to audit evidence (Brown-Liburd and Vasarhelyi, 2015).

New forms of audit evidence are progressively emerging to complement and replace old approaches (Moffitt and Vasarhelyi, 2013). Brown-Liburd and Vasarhelyi (2015) illustrate some examples of newly emerged evidence such as security recordings of arrivals and departures of trucks from parking lots, GPS tracks of truck trajectories, sentiment analysis of social media postings, RFID chips, and video streams. Textual analysis of this type of information can reveal rich details, including a person's political or religious affiliations, identity, location in a crowd, personal views, and opinions regarding various topics. Different parties are eager to retrieve and interpret this user-generated content for their use. As such, research on social media has grown exponentially in recent years. Many researchers conclude that Twitter data contains valuable information and has predictive power (Bollen, Mao, and Zeng, 2011; Risius, Akolk, and Beck, 2015; Sul, Dennis, and Yuan, 2017; Bartov, Faurel, and Mohanram, 2018). The analysis of sentiment in textual content often relies on simple sentiment annotation tasks during which annotators must determine whether a sentence is positive, negative, or neutral (Rosenthal et al., 2015; Mohammad et al., 2015). Despite this seeming simplicity, there are several challenges in applying sentiment analysis to unstructured and diverse content, such as hashtags, abbreviations, emojis, emoticons, misspellings, and slang. Data cleaning needs to be processed before performing further sentiment analysis.

Both machine learning and econometric modeling can be used for analizing social media data and inferring conclusions. Econometric models are considered as statistical models that are applied in econometrics, whereas machine learning is the scientific field that studies the analysis and formation of algorithm to learn from data. The main difference between machine learning and econometric model is that machine learning model deals with methods which is established to extract the details from data, and econometric model applies statistical methods to predict, and deals with insight and inference. Mannering et al. (2020) offered a profound discussion of the problems that are involved with different methodologies, which provides researchers a clear understanding of the trader-offs of the two methodologies and can help lead researchers' decisions while facing different research problems.

Chapter 3: Methodology

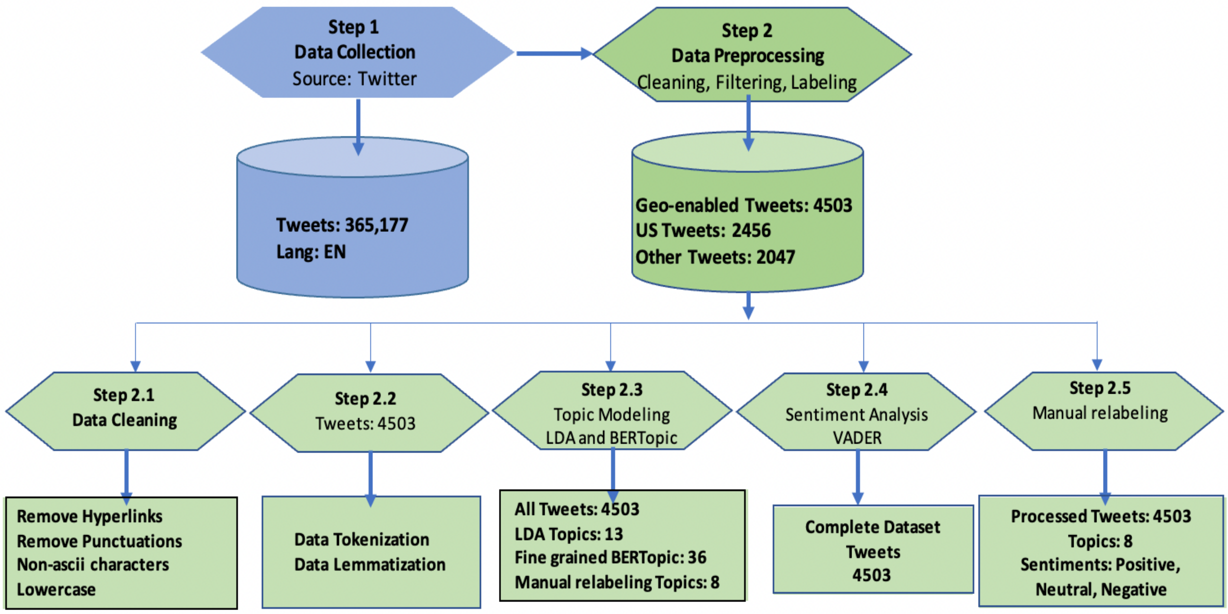


Figure 3.1 Illustration of Research Approach

3.1 Data Extraction

There are two types of application programming interface (API) provided by Twitter: REST API and Streaming API (Figure 3.2). The REST API is appropriate when users try to search for tweets that are authorized by a particular user or access their timeline. Streaming API is suitable when the user wants to apply certain keywords to download massive amounts of tweets (Bonzanini, 2016). REST API allows the users to go back in time and get tweets from only approximately one week (Bonzanini, 2016). The Streaming API collects more tweets and is usually the preferred method for collecting large numbers of tweets but is more time-consuming (Bonzanini, 2016). For this study, a Python script and several Python libraries were used to assemble the Twitter API and fetch the tweets related to UAM. A total dataset of 365,117 tweets were collected using the keywords (i.e., urban air mobility, aerial mobility, air mobility, air shuttle, eVTOL, VTOL, Vertiport, air taxi, flying vehicle, flying car) from January 01, 2017, to December 31, 2021. Since the target was to compare the study between the U.S. and rest of the world, I applied an additional location filter, which resulted in 4,503 tweets (i.e., 2,456 from the U.S. and 2,047 from rest of the world).

3.2 Data Preprocessing

In the data preprocessing stage, the first step was to clean data to remove all punctuations except for a few (e.g., "#", "_", and "-"). I did not remove the symbols that join or combine words (e.g., "_" and "–"). Next I removed all hyperlinks, non-ASCII characters, single characters, words mentioned with "@" and "&" symbols, and lowercase text. After data cleaning, I applied data tokenization to split text into proper word tokens, and applied lemmatization to map all inflected forms of words into their base word (e.g., playing → play).

3.3 Topic Modeling

For this study, I used Latent Dirichlet Allocation (LDA) and BERTopic as two unsupervised learning techniques for topic modeling to extract the topics in the collected tweets for all countries. LDA is one of the most popular methods in this field of research, which is a three-level Bayesian probabilistic hierarchical model that maps the text into separate topics by mapping different words in the text to a specific topic. BERTopic is another popular method that uses BERT (Bidirectional Encoder Representations from Transformers) to create clusters that determine topics easily and keep important words in the description.

I used LDA and BERTopic to extract topics from the tweets dataset. In LDA algorithm, each document is composed with a variety of corpus-wide topics and each topic can be viewed as a distribution over keywords. The main goal was to find hidden topics in each tweet by taking the small value of topics that cover only broad topics as too large a value may generate repeating topics or ambiguous results. To get deep insights, I kept a minimum number of topics to 2 and a maximum to 15. Results depict that using 13 topics has the highest coherence score. Later I trained a final LDA model with 13 topics to generate topic proportions in the documents. LDA implies that the number of topics is the same in all documents and that only the proportions of those topics vary. I then applied BERTopic, and during the training phase of both models, I tuned the hyperparameters to improve the model's performance and extract the most dominant topics. BERTopic performed better and gave an output of that generated 36 topics. To improve accuracy, I later went through the labels manually that were generated by BERTopic and the topics were mapped to 8 topics.

3.4 Sentiment Analysis

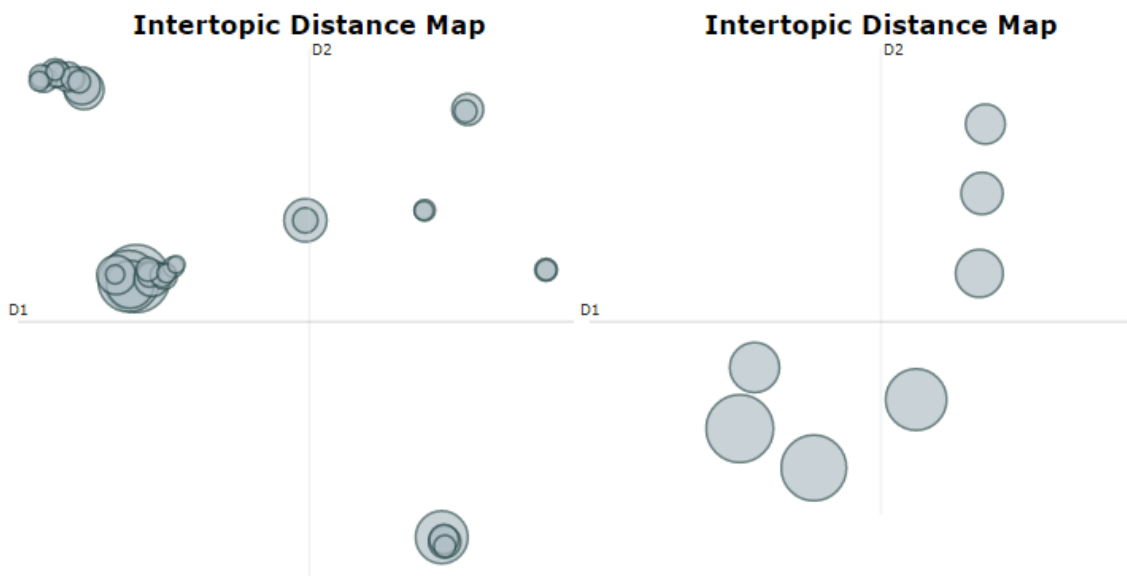
Sentiment analysis is a natural language process that is used to identify and quantify subjective information. There are several tools for sentiment analysis such as Linguistic Inquiry Word Count (LIWC), Monkeylearn, Lexalytics, Brandwatch, Social Searcher (SS), General Inquirer (GI), SentiWordNet (SWN), MeaningCloud, SenticNet (SCN), Social Mention (SM), Hootsuite Insights (HI), Word-Sense Disambiguation (WSD), and VADER, among others. For this study, we performed sentiment analysis using Valence Aware Dictionary and Sentiment Reasoner (VADER). Hutto and Gilbert (2016) studied VADER in detail and claimed that it is the most outperformed tool when it comes to dealing with social media texts. Hutto and Gilbert (2016) also found that VADER performed better because during its development it used a human rater from Amazon Mechanical Turk. Different raters might have different clarifications, some words being negative to one person while being neutral to another person, but VADER considers these factors and averages the ratings for each word and comes up with a rating (Hutto and

Gilbert, 2016). For example, “love”, “like”, “happy”, and “fun” are all positive sentiments, but VADER is smart enough to understand the basic situation of a sentence, such as “did not have fun” as a negative sentiment. It also recognizes the emphasis of punctuation (i.e., “!!!”), emoticons, acronyms (i.e., BRB, LOL), slang, and capitalization (i.e., ENJOY). The VADER sentiment analysis gives back a sentiment score ranging from -1 to 1. The scores are categorized into negative, positive, neutral, or compound. They are calculated by adding the sentiment scores of each word in the lexicon and giving a score between -1 (being extreme negative) and +1 (being extreme positive). Based on the output of the result, I checked the data manually to determine the accuracy and it resulted in a 95% accuracy rate from the score of VADER.

Chapter 4: Result and Discussion

4.1 Topic Modeling

In the topic modeling phase, I applied two unsupervised techniques (LDA and BERTopic) for topics extraction. As the small dataset was small, I applied unsupervised topic models on a complete dataset of 4,503 tweets. The analysis shows that BERTopic performs better than LDA as it generates more precise and discriminative topics. At first, LDA was trained with a custom embedding model that results in only two topics. All the documents were almost clustered into one topic (having keywords: air, mobility, shuttle, urban) and only 15 topics were clustered into another topic (having keywords: ultimate, Cincinnati, oh, shuttle). By manual evaluation, I realized that topic clusters were not good enough to distinct all documents. I then applied BERTopic with a default embedding model, and it extracted a total of 36 fine-grained topics in which many had identical characteristics, as shown in Figure 4.1 (a). From intertopic distance map visualization, it can be observed that many of the topics had high similarity level. I then classified the highly similar topics into single topics, as shown in Figure 4.1 (b), and this reduced the number of topics to 8 that were distinct from each other. BERTopic labeled -1 to the outlier documents that did not belong to any topic. It categorized 1,656 tweets as outliers that did not belong to any category. The 1,656 tweets that were classified as outliers, were labeled to the most significant topics in the corresponding documents, which resulted in 8 topics.



(a) (b)

Figure 4.1 Intertopic Distance Map

I tried to identify topic labels based on the keywords that were extracted by the model, and the topics included: 1) Public Acceptance, 2) Recreation, 3) Safety, 4) Infrastructure, 5) Promotion and service, 6) Mobility, 7) Technical Features, and 8) Regulations. From these topics and the result from the sentiment analysis, the main goal was to identify the perceptions of people in different areas toward UAM and its growth in the future.

4.2 U.S. Result

Figure 4.2 shows the topic distribution in the U.S. tweets data. The topic distribution is as follows: 1) Public acceptance (37.67%) was the topic most discussed in the U.S. having 925 tweets, 2) Recreation (15.46%) had 380 tweets; 3) Safety (13.21%) had 325 tweets; 4) Infrastructure (11.34%) had 278 tweets; 5) Promotion and Service (8.52%) had 209 tweets; 6) Mobility (5.34%) had 132 tweets; 7) Technical Featues (5.31%) had 130 tweets; and 8) Regulations (3.15%) had 77 tweets. Public acceptance is the most discussed topics in the US.

Public acceptance was the most discussed topic in the United States. The results show that people are excited and in favor of this new mode of transportation. This topic includes tweets like This topic includes tweets like (e.g., *“Urban air mobility will provide very rapid transit to people living in major cities, where there's a high trip density”*, “*UAM is a revolution, developing new capabilities across the industry”, “UAM will be safer, more effective, and more immediately deployable if it structurally resembles”*, *“The future is in the air. With Lilium announcing their future air taxi prices at $2.25 per mile, making air taxis a cheap option”*, *“Congestions in U.S. cost more than $87 billion in 2018, an average of $1,348 per driver. UAM is enabling the future of urban air mobility to help reduce this congestion”*). Recreation refers to tweets about leisure, fun, sight seeing, etc. Many tweets suggested that flying cars can be used for recreational use. Example tweets include (e.g., *“Can’t wait to see the beautiful views riding around in flying cars”, “The city lights at night are going to look amazing in an aerial mobility vehicle”, “Going on a sight seeing adenture would be fun when UAM is in the market”*). There were tweets related to safety such as (e.g., *“Safety and security are two challenges which will arise with this new technology”, “Thick low cloud, rain, turbulence and lightning can be problem for UAM if it not taken care properly”, “Flying in broad daylight, with clear sky, and very light wind can have a below-average probability of fatal accident”*). Tweets regarding infrastructure were such as, (e.g., “*There will no longer be a need to build infrastructure that cost billions of dollars and impact the environment”*)*, also concerns like* (e.g.,*“You need to get permission from the individual in order to hand on their rooftop, which can be an issue in terms of infrastructure”*). Tweets related to Promotions and Services were mostly regarding the companies that will bring UAM to the market such as Lilium, Airbus, Joby Aviation, etc. Some of the tweets were (e.g., “*Lilium announcing their future air taxi prices at $2.25 per mile which* *will make it cheaper than taxi", "There are 200 companies involved in the making of eVOTL for affordable, rapid, and sustainable intercity travel", "Airbus is working with different to give UAM a emerge of real transportation").*

4.3 Other Countries Result

For all other countries the topics that were discussed in the tweets are shown in Figure 4.2 and have the following distribution: the most widely discussed topic was 1) Public acceptance (39.45%) having 808 tweets, 2) Safety (16.65%) having 177 tweets, 3) Infrastructure (13.16%) having 341 tweets, 4) Recreation (8.66%) having 270 tweets, 5) Mobility (7%) having 97 tweets, 6) Promotions and services (4.75%) having 143 tweets, 7)Technical features (6.23%) having 127 tweets, and 8) Regulations (4.10%) having 84 tweets. When compared to the U.S. tweets, the issues with safety, infrastructure, technical features, and mobility were discussed more in other countries, and manual evaluation of these tweets showed that people mostly discussed about issues related to technical features and safety in both positive and negative ways. Examples of tweets related to safety and technical features include (e.g., *"Most of the aviation officials does not have defined certification categories required for UAM aircraft, and does not have for the new features in UAM and can be problem", "There can be few technical problems such asand combinations of features that are not typically found in other aircraft (distributed electric propulsion/tilt-wing propulsion, VTOL, autonomy hardware and software, and others)".*Analysis shows that the number of tweets for the topic "Public acceptance" is higher in both datasets, showing that people have positive behavior towards mobility.

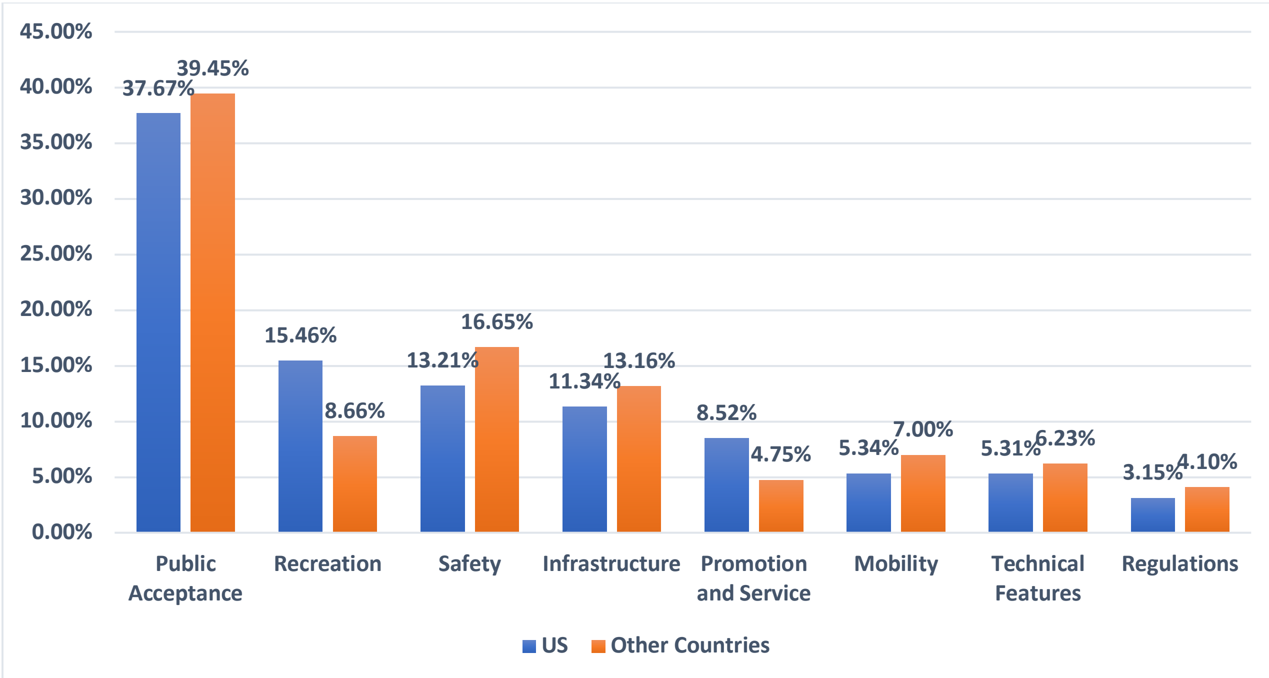


Figure 4.2 Distribution of Topics in U.S. and Other Countries

4.4 Sentiment Analysis

After labeling the complete dataset, we determined that most of the tweets were positive, followed by neutral, and a lower number of tweets had negative sentiment. People think UAM is a great mode of transportation, and the results indicate that most people have a positive belief toward mobility. Figure 4.3 shows the tweet length distribution in correspondence to countries and sentiments and Figure 4.4 shows the distribution of positive, negative, and neutral tweets in the United States and other countries. As can be seen from Figure 4.4, 47.35% of the dataset of tweets in U.S. belong to the positive class, 38.65% of tweets belong to the neutral class, and only 13.75% of tweets belong to the negative class. For other countries, 53.45% of the dataset was positive, 36.40% neutral, and 10.15% were negative. That depicts that most of the people are in favor of UAM. From manual evaluation it shows that most people are excited about how UAM will bring positive impacts, revolution, and expansion in the future.

The common technical issue with UAM is the battery technology, as UAM needs to perform vertical takeoff and landings which require significant amount of power to carry the load and to cover the distance, therefore the batteries need to provide significant power to operate the vehicle. Since this service will require vertical takeoffs and landings, this will require a location for these vehicles to take off and land which needs to be accessible and is the main problem in terms of infrastructure. The customers do not feel safe with autonomous technology, and some believe that the waste that would be built up from batteries and the energy used will impact the environment. The noise disturbance is another factor that concerns the people, and it is believed that this may be a pushback for the market of UAM if it is not taken care of properly. Most of the positive feedback for this new mode of transportation is because of the speed the service will provide and reduction in road traffic it will. UAM also tends to reduce the carbon emissions and air pollution that are associated with travel. Overall, people are in favor of UAM and think it as a great idea to have an alternative form of transportation to carry passengers and goods in urban areas by air.

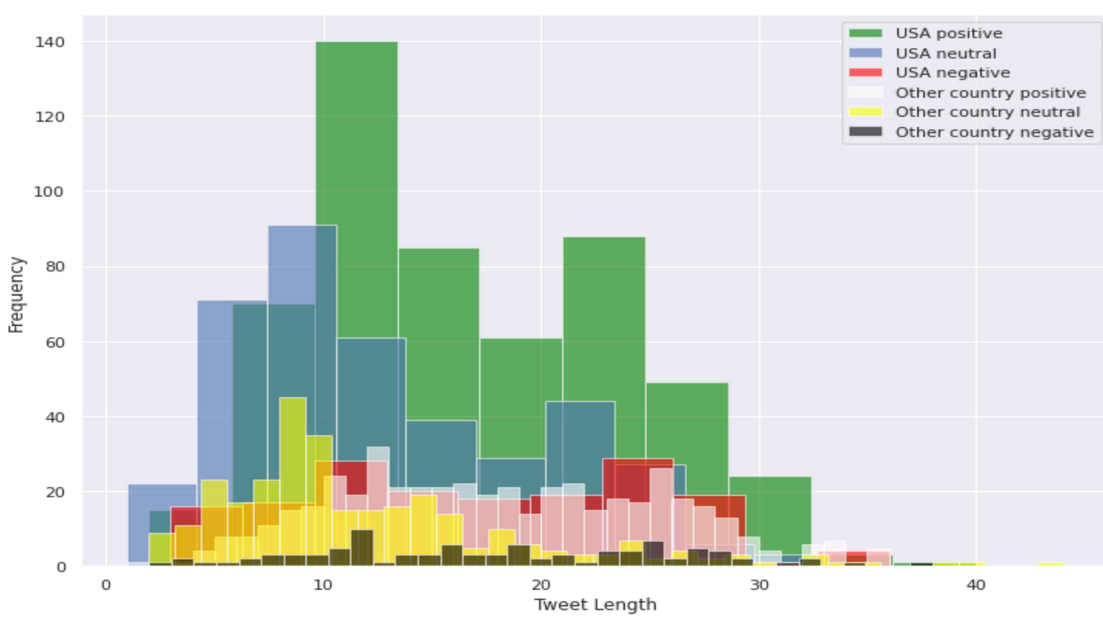


Figure 4.3 Tweet Length and Frequency for the Sentiments

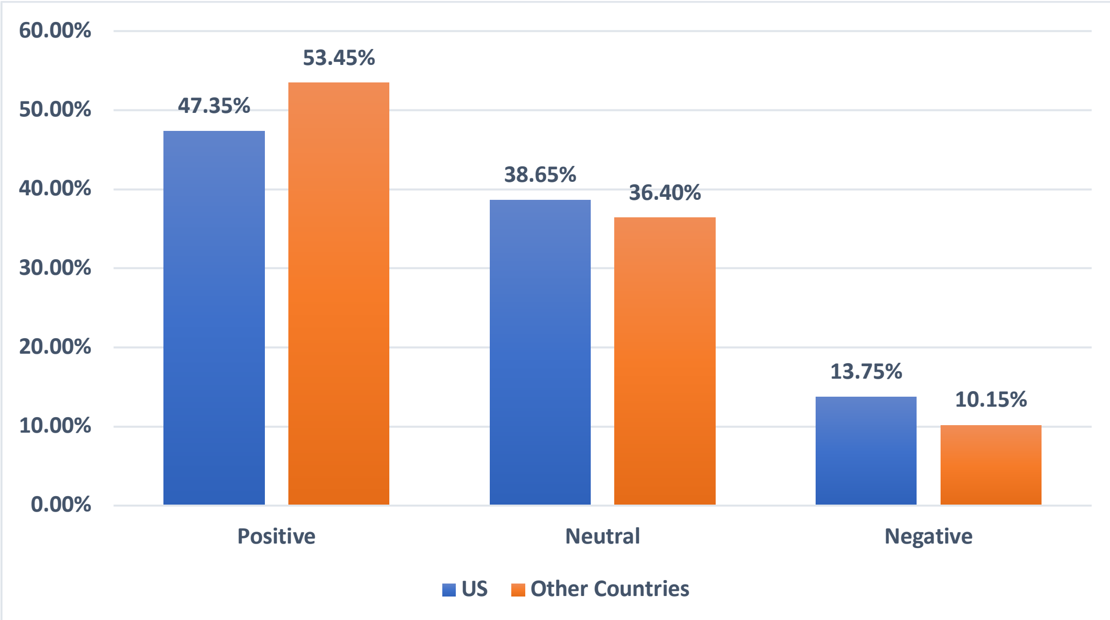


Figure 4.4 Distribution of Positive, Negative, and Neutral Tweets in the U.S. and Other Countries

4.5 Comparison between U.S. and Other Countries

In addition, I also conducted a Chi-Square analysis to determine the associations between sentiment polarities and topics. Chi-square test is a hypothesis test that is commonly used to find whether there is a relationship between two categorical variables. Figure 4.5 shows the topic and sentiment associations in each region based on pearson residuals shown in circles. In the complete dataset, the inter topic association is high with x^2 = 177.494 at p-value <0.0001, whereas the inter sentiment association is lower with x^2 = 11.298 at p-value = 0.004.

Analysis shows that for the U.S. data topics and sentiments are highly associated or dependent with x^2 = 56.184 at p-value < 0.0001. For the dataset of rest of the world topics and sentiments are also associated at a significance level with x^2 = 62.528 at p-value < .0001. The positive values, indicates positive association between topic and sentiment, and the larger the value is, the stronger the association was. For instance, regulations were strongly associated with negative sentiment and were weakly associated with the positive sentiment. Likewise, the negative values indicate the negative associations, and the size shows the strength of association.

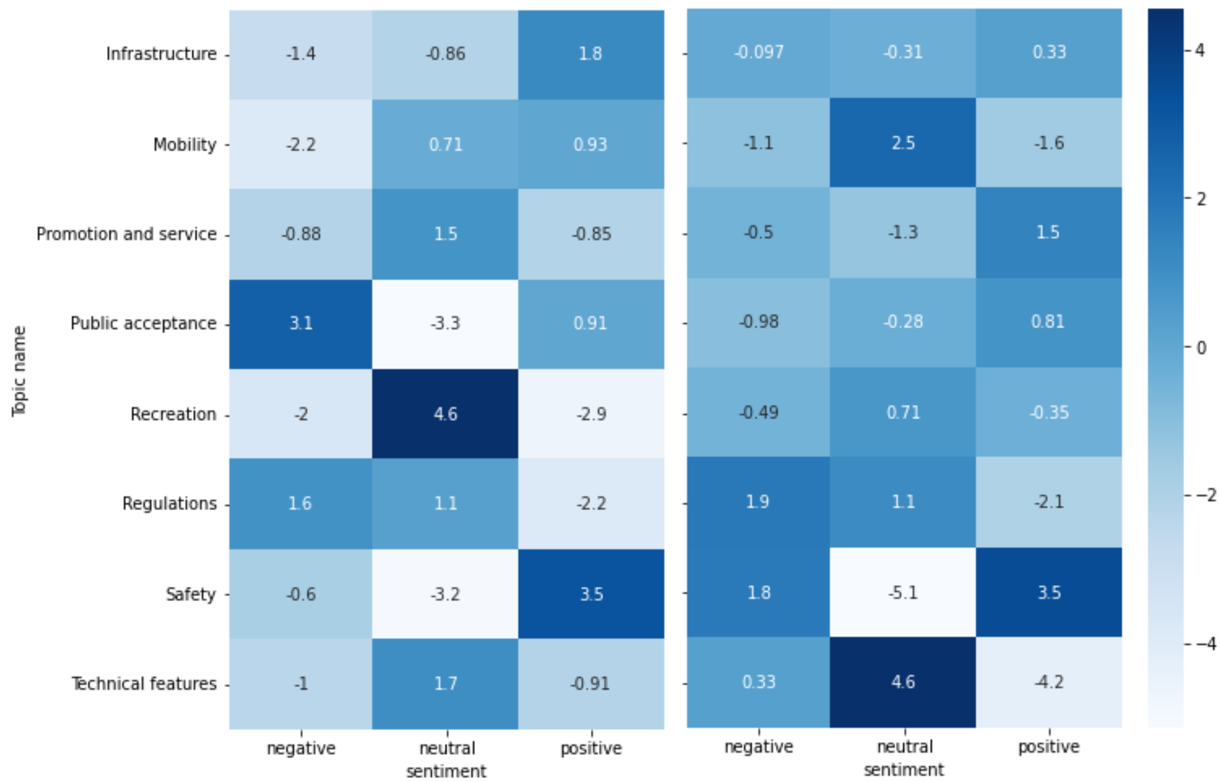


Figure 4.5 Association between topics and sentiments (left: U.S., right: Other Countries)

Chapter 5: Conclusion and Future Work

The purpose of this thesis study was to investigate the UAM trends and challenges in the U.S. and rest of the world using twitter data. NLP methods were applied to analyze the collected data and extract the main areas of people's interests in UAM and explore the polarity of their opinions related to particular topics. With BERTopic and manual labeling, eight major topics (i.e., Infrastructure, Mobility, Recreation, Safety, Promotion and service, public acceptance, Regulations, and Technical features) were identified. Then, sentiment polarity (i.e., negative, neutral and positive) of the tweets was explored to determine their distributions across the regions. It is found that data from rest of the world were dominated by more positive opinions about UAM compared to the U.S. For both the U.S. and the rest of the world, the results indicate that most users had a positive attitude toward UAM in terms of innovation. Most of the positive feedback for this new mode of transportation relates to the speed of this service and possible reduction in road traffic. UAM also tends to reduce the carbon emissions and air pollution that are associated with travel compared to other air mode. However, in both U.S. and the rest of the world the topics related infrastructure, regulations and technical features had negative opinions. These shows that the method I used showed similar trends and challeneded from the literature review, and gives better understanding from users' view that is expresses in social media.

This thesis had few limitations that can be examined in the future. The data that I collected was only from Twitter since other social media platforms do not provide geographic information of the data. Even for Twitter, only a small portion of tweets contained locations. Also, I only included tweets in English in this study, which further reduced the data sample size.

Therefore, for future research, we may consider 1) additional sources such as You-Tube videos, Facebook or articles from media agencies; 2) further data exploration process to determine the possible origin of the information, and 3) automated translation tools enabling the inclusion of non-English information in the study.

References

Al Haddad, Christelle & Chaniotakis, Manos & Straubinger, Anna & Plötner, Kay & Antoniou, Constantinos. (2020). Factors affecting the adoption and use of urban air mobility. Transportation Research Part A Policy and Practice. 132. Pages 696-712. https://doi.org/10.1016/j.tra.2019.12.020.

Bartov, Eli, Lucile Faurel, and Partha S. Mohanram. 2018. "Can Twitter Help Predict Firm- Level Earnings and Stock Returns?" Accounting Review 93 (3): 25–57. https://doi.org/10.2308/accr-51865.

Baur, S., A.Schickram, N. Homulenko, Martinez, and A. Dyskin 2018. "Urban air mobility: The rise of a new mode of transportation." Roland Berger Focus. https://doi.org/10.2514/1.D0275.

Beresnevicius, R. 2019. "Defining urban travel in 1950s: the story of New York Airways." Aerotime, July 8, 2019. https://doi.org/10.1017/aer.2021.110.

Brown-Liburd, Helen, and Miklos Vasarhelyi. 2015. "Big Data and Audit Evidence." American Accounting Association 12 (December): 1–16. https://doi.org/10.2308/jeta-10468.

Bollen, Johan, Huina Mao, and Xiaojun Zeng. 2011. "Twitter Mood Predicts the Stock Market." Journal of Computational Science 2 (1): 1–8.https://doi.org/10.1016/j.jocs.2010.12.007.

Booz Allen Hamilton, Ed.2018, "Urban Air Mobility (UAM) Market Study: Final Report," National Aeronautics and Space Administration (NASA). https://ntrs.nasa.gov/citations/20190001472.

Bonzanini, Marco. 2016. Mastering Social Media Mining with Python. Packt Publishing. Brown-Liburd, Helen, and Miklos Vasarhelyi. 2015. "Big Data and Audit Evidence." American Accounting Association 12 (December): 1–16. https://doi.org/10.2308/jeta-10468.

Cohen, A., S. Shaheen, and E. Farrar. 2021. "Urban Air Mobility: History, Ecosystem, Market Potential, and Challenges." IEEE Transactions on Intelligent Transportation Systems (9: 1–14. https://doi.org/10.1109/tits.2021.3082767.

Dhaoui, Chedia, Cynthia M. Webster, and Lay Peng Tan. 2017. "Social Media Sentiment Analysis: Lexicon versus Machine Learning." Journal of Consumer Marketing 34 (6): 480– 88. https://doi.org/10.1108/JCM-03-2017-2141.

Duan, Huijue. (2020). New York City Street Cleanliness: Applying Text Mining Techniques to Social Media Information. SSRN Electronic Journal. 10.2139/ssrn.3706506. https://dx.doi.org/10.2139/ssrn.3706506.

Fred Mannering, Chandra R. Bhat, Venky Shankar, Mohamed Abdel-Aty. (2020). Big data, traditional data and the tradeoffs between prediction and causality in highway-safety analysis, AnalyticMethods in Accident Research. https://doi.org/10.1016/j.amar.2020.100113.

Harrison, S. (2017). From the archives: Los Angeles Airways helicopter overturns. The Los Angeles Times. https://www.latimes.com/visuals/photography/la-me-fw-archives-airways-helicopter-overturn-20170221-story.html.

Holden, J., and N. Goel. (2016). Fast-forwarding to a future of on-demand urban air transportation. San Francisco, CA. https://doi.org/10.1016/j.trc.2021.103377.

Hutto, C. J., and Eric Gilbert. 2016. "VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text." International AAAI Conference on Web and Social Media. https://ojs.aaai.org/index.php/ICWSM/article/view/14550.

Johnston, T., Riedel, R., & Sahdev, S. (2020). To take off, flying vehicles first need places to land. McKinsey Center for Future Mobility. https://doi.org/10.1016/j.trc.2021.103377.

Lassen, Niels & Madsen, René & Vatrapu, Ravi. (2014). Predicting iPhone Sales from iPhone Tweets. Proceedings - IEEE International Enterprise Distributed Object Computing Workshop, EDOCW. 2014. https://doi.org:10.1109/EDOC.2014.20.

Lineberger, R., Hussain, A., and Rutgers, V. (2019). Change is in the air: The elevated future of mobility: What's next on the horizon. Deloitte Insights. https://www2.deloitte.com/content/dam/Deloitte/us/Documents/energy-resources/di-the-elevated-future-of-mobility.pdf.

Liu, Y., M. Kreimeier, E. Stumpf, Y. Zhou, and H. Liu. (2017). Overview of recent endeavors on personal aerial vehicles: A focus on the U.S. and Europe led research activities. *Progress in Aerospace Sciences*, 91, 53-66. https://doi.org/10.1016/j.paerosci.2017.03.001.

Mayor, T.; Anderson, J. Getting Mobility Off the Ground; KPMG: Atlanta, GA, USA, 2019; pp. 1–12. https://doi.org/10.3390/su13137421.

McKinsey & Company. Urban Air Mobility (UAM) Market Study; National Aeronautics and Space Administration: Washington, DC, USA, 2018; pp. 1–56. https://www.nasa.gov/sites/default/files/atoms/files/uam-market-study-executive-summary-v2.pdf.

Moffitt, Kevin, and Miklos Vasarhelyi. 2013. "AIS in an Age of Big Data." Journal of Information Systems 27 (2): 1–19. https://doi.org/10.2308/isys-10372.

Reiche, C., R. Goyal, A. Cohen, J. Serrao, S. Kimmel, C. Fernando, and S. Shaheen. 2018. Urban air mobility market study. https://doi.org/10.1016/j.trc.2021.103377.

Risius, Marten; Akolk, Fabian; and Beck, Roman, "Differential Emotions and the Stock Market - The Case of Company-specific Trading" (2015). ECIS 2015 Completed Research Papers. Paper 147. https://core.ac.uk/download/pdf/50528565.pdf.

Rosenthal, Sara & Nakov, Preslav & Kiritchenko, Svetlana & Mohammad, Saif & Ritter, Alan & Stoyanov, Veselin. (2015). SemEval-2015 Task 10: Sentiment Analysis in Twitter. 451-463. https://doi:10.18653/v1/S15-2078.

Schumaker, Robert P., Yulei Zhang, Chun Neng Huang, and Hsinchun Chen. 2012. "Evaluating Sentiment in Financial News Articles." Decision Support Systems 53 (3): 458–64. https://doi.org/10.1016/j.dss.2012.03.001.

S. Baur, S. Schickram, A. Homulenko, N. Martinez, and A. Dyskin, "Urban air mobility: The rise of a new mode of transportation," Roland Berger Focus, November, 2018. https://doi.org/10.1016/j.procs.2021.03.049.

Sul, Hong Kee, Alan R. Dennis, and Lingyao Ivy Yuan. 2017. "Trading on Twitter: Using Social Media Sentiment to Predict Stock Returns." Decision Sciences 48 (3): 454–88. https://doi.org/10.1111/deci.12229.

Thipphavong, D. P., Apaza, R., Barmore, B., Battiste, V., Burian, B., Dao, Q., ... & Verma, S. A. (2018). Urban air mobility airspace integration concepts and considerations. In 2018 Aviation Technology, Integration, and Operations Conference (p. 3676). https://doi.org/10.2514/6.2018-3676.

Vascik, P. D., and Hansman, R. J. 2017. Evaluation of key operational constraints affecting on-demand mobility for aviation in the Los Angeles basin: ground infrastructure, air traffic control and noise. In 17th AIAA Aviation Technology, Integration, and Operations Conference (p. 3084). https://doi.org/10.2514/6.2017-3084.

Williams, Kate, and Joan Durrance. 2008. "Social Networks and Social Capital: Rethinking Theory in Community Informatics." The Journal of Community Informatics 4 (3): 1–20. http://ci-journal.net/index.php/ciej/article/view/465/430.

Witken, R., 1979. New York Airways acts to file for bankruptcy. The New York Times, May 16. https://www.nytimes.com/1979/05/16/archives/new-york-airways-acts-to-file-for-bankruptcy-suing-sikorsky.html.

Yedavalli, P., and J. Mooberry. 2019 "An Assessment of Public Perception of Urban Air Mobility," Airbus UTM. https://storage.googleapis.com/blueprint/AirbusUTM_Full_Community_PerceptionStudy.pdf.