3-2022

Analysis of the Perspective on Syrian Refugees by Neighboring Countries

Norma Ghanem

ANALYSIS OF THE PERSPECTIVE ON SYRIAN REFUGEES BY NEIGHBORING COUNTRIES

THESIS

Norma Ghanem, 1st Lt, USAF

AFIT-ENS-MS-22-M-130

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT-ENS-MS-22-M-130

ANALYSIS OF THE PERSPECTIVE ON SYRIAN REFUGEES BY NEIGHBORING COUNTRIES

THESIS

Presented to the Faculty

Department of Operations Research

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

in Partial Fulfillment of the Requirements for the

Degree of Master of Science in Operations Research

Norma Ghanem, B.A.

1st Lt, USAF

March 2022

AFIT-ENS-MS-22-M-130

ANALYSIS OF THE PERSPECTIVE ON SYRIAN REFUGEES BY NEIGHBORING COUNTRIES

THESIS

Norma Ghanem, B.A.
1st Lt, USAF

Committee Membership:

Dr. Richard F. Deckro, DBA
Chair

Capt Nicholas T. Boardman, PhD
Member

AFIT-ENS-MS-22-M-130

Abstract

Mass migration destabilizes neighboring states, opening the way for fragile state exploitation by enemies, including those that could undermine U.S. national interests. This study investigates an area of the Levant region, specifically countries neighboring Syria, and analyzes their perspective on Syrian refugees for the time frame of 1 June 2019 - 30 June 2020. This analysis may assist in forming policy and creating strategies to address refugee related issues, both domestic and international. There are three main questions addressed. The first inspects dominant refugee framing, the second explores sentiment (dis)similarity within each country and across countries, and the third investigates drastic sentiment changes and their potential driving factors. Results show that Syrian refugees are dominantly viewed in a political framing. Sentiment similarities are shared across two distinguished groupings, Turkey and Israel in one group and Jordan and Lebanon in the other. External factors that vary in origination are likely to influence high sentiment changes. Future research includes application on South American countries, or a change in topic to analyze perspective towards America post Afghanistan withdrawal.

Acknowledgements

I would like to express my sincere appreciation to my faculty advisor, Dr Deckro, for his guidance and support in completing this thesis. I would also like to thank our reader Dr Boardman, for the great feedback and insights. My family for the unconditional support, and my friends for the motivation. Thank you all.

Norma Ghanem

Table of Contents

List of Figures

List of Tables

Table		Page

List of Acronyms

ADF Average Document Frequency

BoW Bag-of-Words

CSO Bureau of Conflict and Stabilization Operations

CTM Correlated Topic Models

IDF Inverse Document Frequency

LDA Latent Dirichlet Allocation

STM Structured Topic Models

TF Term Frequency

TF-IDF Term Frequency - Inverse Document Frequency

ANALYSIS OF THE PERSPECTIVE ON SYRIAN REFUGEES BY NEIGHBORING COUNTRIES

I. Introduction

Over a decade has passed since the beginning of the Syrian war and Syrian refugees remain the largest group displaced globally. According to a 2022-2023 Resilience and Response Plan by the United Nations High Commissioner for Refugees [1], around 6.7 million are internally displaced within Syria, with an additional 5.6 million displaced in countries neighboring Syria, as displayed in Figure 1.

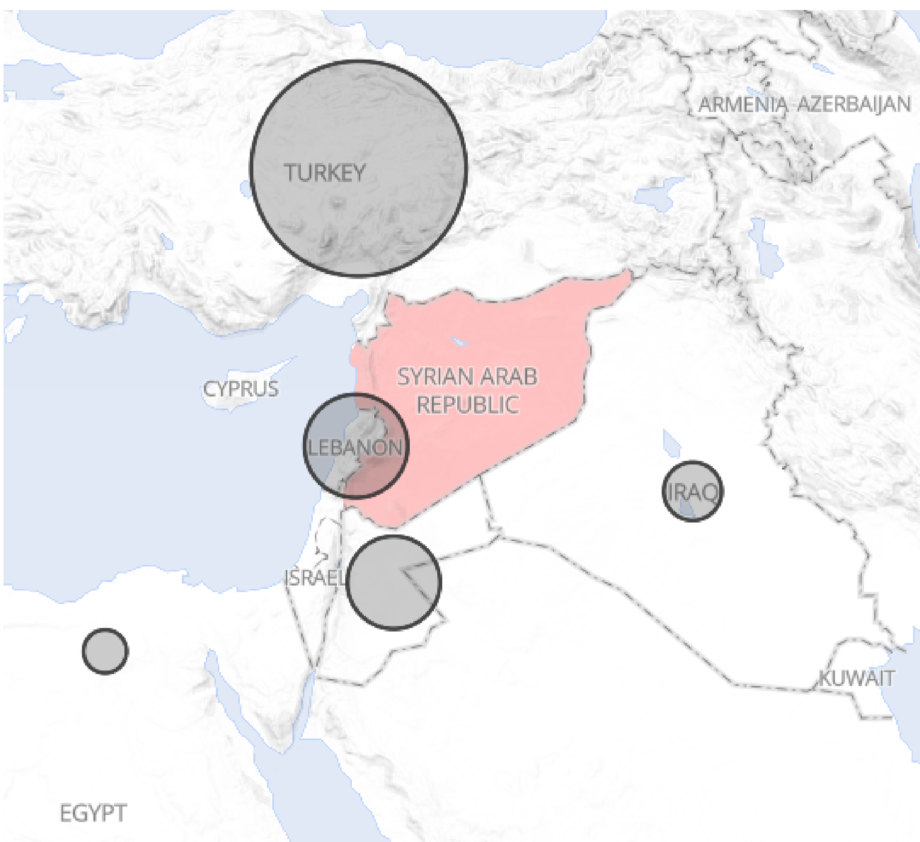


Figure 1. Syrian Refugees in Levant Countries. Source: Adapted from [2]

Syria and its neighboring countries, referred to as the Levant region in this study, have experienced a number of turbulent changes in recent years. Whether it is the severe financial crisis in Lebanon, the turmoil in Jordan's royal family, or the ignited conflict in Israel-Palestine, fissures in stability are becoming strikingly more emergent each day. Being a focal point of the three continents of Asia, Africa, and Europe, this brings concern about the future of the region and the repercussions this future will propagate beyond the Levant.

As stated in a 2020 executive statement by the Bureau of Conflict and Stabilization Operations (CSO) of the U.S. Department of State, mass migration destabilizes neighboring states [3, p.3], opening the way for fragile state exploitation by enemies in that vicinity that directly undermine U.S. national interests. This fragility motivates regional states and their allies to seize on geopolitical interests and establish a projection of power in the global stage [4, p.182], [5, p.403], [6, p.5].

The dynamics at play are not the focus of this study, but it does provide motivation to consider an important aspect in soft power gain, which is winning people over. Justifications and support for future actions become easier when a perspective is shared, and this perspective becomes more influential and effective when it is highlighted in mass media, notably news sources. Studying the perspective of Syria's neighboring countries towards a common subject, the Syrian refugees, is one facet that could provide insight on perspective sharing, and future intention and action indicators of governments in the Levant. Additionally, analyzing the perspectives or reactions of countries towards refugees can assist in forming policy and strategies to address refugee related issues, both domestic and international. The inspiration for this thesis topic comes from a study on the representation of migration in Latvian mass media, where the authors followed a linguistics-based approach to analyze conceptual frames, discourse strategies, and visual representation of migrants [7].

1.1 Problem Statement

This research aims to investigate how news media articles of countries neighboring Syria view Syrian refugees. Specifically,

1. Is there a dominant refugee frame utilized by these countries?
2. Is the sentiment similar or dissimilar within each country, and/or across the countries?
3. If there is a drastic change or intensity in sentiment, when did it happen and are there any factors that could have driven this change?

1.2 Scope, Assumptions, and Limitations

This study was scoped to include readily available digitized news articles in English for the period of 1 June 2019 to 30 June 2020. In addition, articles were collected from at least two different news sources per country based in the Levant region, and specifically those with high sales or popularity. This study was also scoped to focus on text news articles. Analyzing other media formats such as videos, photographs, audios, and so forth could potentially provide more insight in any follow on study.

The data for this study utilized the keyword *refugee* specifically. While the terms *migrant* and *refugee* are used intermittently, a commonly accepted difference between the two terms is choice. Migrants are typically defined as individuals who chose to move, while refugees are usually forced to move due to external factors such as conflict or persecution [8]. The subject of this document is refugee migration.

One limitation is that the target language is in English. A replication of the study with different country-specific languages may provide different insights, and makes a possible future project for comparison. The approach to address the research questions is empirical, utilizing an analysis of news articles (see Section III, Methodology for details). The results of this study can provide information on country dynamics in the Levant, which touches many other aspects, such as foreign and domestic policy, narrative information operations, and a possibly prevalent trend of refugee weaponization.

1.3 Document Outline

The thesis contains five chapters. The first is an introduction to provide background and motivation for studying neighboring countries perspective on Syrian refugees. The problem statement with a breakdown of the research questions, study scope, assumptions, and limitations are included in the first chapter as well. In the second chapter, a literature review of relevant topics is discussed. The third chapter details the methodology utilized to answer the research questions, specifically framing and sentiment analysis. The fourth chapter describes the observations and analyses obtained. The final chapter summarizes the conclusions and identifies possible future work.

II. Background and Related Work

2.1 Overview

This section begins with an introduction of Syrian migration and media studies. A brief background on text analysis is then provided since the media in this study is text articles. It ends with a discussion of framing and sentiment techniques, which are migration representation specific techniques.

2.2 Introduction

Migration related studies have typically branched from two main domains: internal migration research and international migration research, with some studies covering both [9, p.290]. According to an empirical study by Asya *et al.* [10, p.477], the volume of migration related studies have expanded significantly since the mid 1990's, but remained somewhat consistent regarding the range of topics within the field.

In the past decade, one prominent topic of interest has been media representation of refugees. The 2015-2016 period reportedly had the highest number of refugee arrivals in Europe, most of which were Syrian refugees [11]. This caused an increase in popularity of Syrian refugee representation studies, largely in Europe. While some included Levant countries, most of the Syrian refugee representation studies were based on Turkish media near the 2015-2016 time frame, such as [12], [13], and [14]. Even the few studies that included Syria's other neighboring countries utilized data from that time frame as well. For example, a recent (2021) study on media coverage of female refugees in Jordan and Lebanon was based on 2012-2016 data [15]. In late 2019, the European Commission declared that the migratory situation had "returned to pre-crisis levels" [16, p.2], and it seems that since then there has been less insight into the news media perspective on Syrian refugees past that year, and more so encapsulating the Levant region as a whole.

As social media became more prevalent in recent years, many researchers shifted to leverage platforms such as Twitter, Facebook, and others as data sources for migration representation studies instead of news media. However, using social media as a source relies on two important assumptions [17]. First, is that the account user has the correct country or location specified, if the location is enabled at all. For international migration representation, this is not a likely issue. However, for country specific representation, it is an assumption, and in some cases could require legal and monetary commitments. Second, is that social media users share the representation, and provide their views honestly. Some people tend to avoid certain topics, especially on social media where work colleagues and friends or family can access. Therefore, it is sometimes questionable if social media data can accurately represent the target country population.

Despite these assumptions, numerous studies have been conducted using social media to gauge refugee representation, and in different niches. For example, some look at representation from a gender perspective [18], or target texts in a specific language [19], while others focus on some type of European Union perspective of Syrian refugees, which was especially common after the 2015-2016 period [20], [21].

2.3 Context Presentation and News Media

When analyzing news media for migration representation, researchers often address the idea of framing the migration. Many utilize Entman's description of framing as a process of selection and salience in communication [22, p.52]. The general idea is that when submitting a piece of communication, the presenter specifies the context and highlights a key piece of the information, providing it in a frame. News media, whether intentional or not, will typically deploy some form of framing whilst communicating information. The research on media's framing of refugees is extensive ([23], [24], [25], [26], and others), but a common feature is that media has a role in influencing policy or public opinion.

In the same token though, many governments and influential personnel in the Levant own media outlets or are affiliated in some fashion. For example, about 78% of media outlets in Lebanon are politically affiliated [27]. In Jordan, all publications require licenses from the government to operate, and news reporting and commentary is influenced by the government according to a report by the U.S. Department of State [28, p.12]. Many pro-government businessmen in Turkey have become media owners as well, and Coşkun characterizes this as part of a government strategy to capture mass media in Turkey [29, p.651]. It would be reasonable then to infer that the framing likely reflects the affiliated entity's viewpoint, or propagated viewpoint as well. Collecting successive framing instances over time with a fixed target subject can demonstrate a change in viewpoint. Depending upon how detailed the frame is, this can also indicate a change in sentiment, if the framing changes considerably. While some researchers may analyze such frames and sentiment in a qualitative manner, this can be accomplished quite efficiently in a quantitative manner via text analysis.

2.4 Relevant Background for Text Analysis

The main challenge with analyzing text data, such as news articles, is the inability to perform mathematical computations on language directly. Features from text data must be extracted in some form of numerical representation to perform computations and analyses. A classical and simple approach to do so is to use a Bag-of-Words (BoW) method [30, p.414]. In this approach, a text data sample, such as a sentence, is simplified into a set of words with typically a frequency count of each word. While the order of the words is irrelevant for the set, an index can be used to aggregate and compare the two samples. In this manner, a text can be simplified into numbers as frequencies across sentences for comparison.

Another similar, and more general approach to the classical BoW that does take into account the sequence of the words is the n-gram method [31, p.867]. For example, in a bi-gram (where $n = 2$), the elements of the set of words would contain pairs. For any given word in the text, there would be an element in the set containing the target word with a word prior, and the target word with a word after, effectively capturing the adjacent information in a pair. However, when analyzing a large number of documents, referred to as a corpus, the relative importance of a word cannot be captured by a simple frequency. For instance, in English text, articles and pronouns are highly frequent but provide little to no informational value, while proper nouns are less frequent but have the most meaningful impact [32].

There are a number of ways to determine the relative importance of a word, or term, compared to others in a corpus. One of the most notable term weighing schemes is the Term Frequency - Inverse Document Frequency (TF-IDF) weight [33, p.2], as displayed in Equation 1.

$$W_{x,y} = t_{x,y} \cdot \ln \frac{n}{d_x} \tag{1}$$

where:

$W_{x,y}$ is the weight for term x in document y

$t_{x,y}$ is the frequency of term x in document y

n is the number of documents in the corpus

d_x is the number of documents containing x

For a given term, the Term Frequency (TF) is the count of instances of that term divided by the total number of terms in the document [33]. The Inverse Document Frequency (IDF) is the natural logarithm of the total number of documents divided by the number of documents with the given term [33]. The product of these two gives the TF-IDF weight. This is one of the most basic and simple term weighing schemes, but variations and novel schemes continue being developed.

As Jiang *et al.* note, typically the schemes can be generalized to the product of some term frequency factor multiplied by a collection frequency factor, and a summary table of the most conventional examples can be found in [33, p.3]. The same authors also proposed a novel (2021) term weighing scheme based on Improved TF-IDF utilizing an Average Document Frequency (ADF), which they define as "the variance between the [Document Frequency] DF value of a specific term and the average of all DF values in a corpus" [33, p.5]. The conclusions of their study indicated that term weighing schemes with ADF perform better than TF-IDF by weighing words more reasonably [33, p.28].

Approaches such as the TF-IDF and ADF are frequency-based approaches of a concept known as word embedding. Simplified, word embeddings are representations of words in vector form, and this involves the transformation of text data to numerical representation for analysis. Word embedding based on deep learning have been a main focus of researchers, and a recent survey of these methods with detailed descriptions can be found in [33].

2.5 Techniques for Framing and Sentiment Analyses

Given that framing and sentiment are integral to our study, a relevant background of techniques specific to these two topics is discussed. For framing analysis, an important aspect to consider is the construction of the defining frames. Some researchers prefer to use predetermined frames and then label each sample based upon that predetermined set; this process is referred to as topic classification. Others group common sample topics to create the frames after the entire corpus has been processed, which is referred to as topic modeling [34, p.351]. For predetermined frames, or topic classification, the approach typically encompasses supervised machine learning techniques, such as Support Vector Machines or Naive Bayes [35]. Support Vector Machines is a non-probabilistic model that attempts to classify data points by identifying a hyperplane that separates data in feature space. Naive Bayes is a simple probabilistic classification algorithm derived from Bayes theorem. There are numerous other supervised methods that researches can apply, see [35] for further details on supervised methods. On the other hand, topic modeling is an unsupervised machine learning approach that leverages groups of words together to determine the broader context compared to stand alone word meanings. For example, the context in which the word *book* is utilized in the following two sentences differs:

- He read a book at the library.
- Remember to book a room at the hotel.

Typically the word *book* as a stand alone is referenced in the context of the first sentence, as a noun or object. However, topic modeling takes into account the grouping of words to determine a broader context, such as hotel booking, which is treated as a verb as displayed in the second sentence. Topic modeling involves statistical modeling used to discover the underlying topics in a corpus [36, 37]. Three well known and largely utilized topic modeling methods are Latent Dirichlet Allocation (LDA), Correlated Topic Models (CTM), and Structured Topic Models (STM) [36].

LDA is a generative probabilistic model. It was built by Blei, Ng, and Jordan as an extension of the probabilistic Latent Semantic Indexing model created by Hofmann in 1999 [38]. LDA has three main assumptions. The first is that it is essentially a BoW method, where the sequence of the terms in a document are irrelevant. The second is that the sequence of the documents themselves are irrelevant as well and can be interchanged. The third is that the topics of the corpus are uncorrelated. The goal of the LDA model is to form clusters of co-occurring terms, which identifies the group of topics that each document represents. There are several algorithms that can be utilized in this method, such as Collapsed Gibbs Sampling, Mean Field Variational Methods, Expectation Propagation, and others. See [38] for more details on the LDA method.

One limitation of the LDA method is that it does not address topic correlation. If a document is comprised of 30% topic A and 70% topic B, LDA does not account for topics A and B being correlated. Blei and Lafferty improved the LDA method by taking this factor into account in the CTM method [39]. Later in 2013, Roberts *et al.* presented the STM framework, which is built from a combination of methods, to include CTM [40]. This framework is particularly useful since it leverages meta data of the corpus, such as document time or source. For further details and a survey of other topic modeling methods that can be used for framing analysis, see [40] and [36].

Another aspect to consider in framing analysis is whether a sample, such as a news article, contains multiple frames or is only fit into a single frame [34, p.350]. In a quantitative frame analysis study by Carver *et al.*, the authors allowed a sample to contain multiple frames, which were ranked by order of relevance. They then assigned a weight to each frame based on the number of frames present, and performed summary statistic analysis [41, p.458]. While the method requires a human coder to label the frames, the concept is straightforward and provides a simple format to account for multiple frames.

In addition to the frame of an article in which refugees are represented, there is also the sentiment of the language used. This sentiment can be thought of as an underlying tone in which the information is communicated. While the frame itself is like a picture the news source is displaying, the sentiment is likened to a filter for that picture, for example grey-scaled, colored, or faded picture filters. Sentiment analysis is a typical method that researchers have used to gauge the representation of migrants or refugees based on a "sentiment score." A sentiment score is essentially a label for a text sample, and can be considered a numerical representation of sentiment intensity. This specifies the tone of the language used for the text, ranging from negative to positive. The news source's perspective and demeanor towards Syrian refugees is inferred from the sentiment label for a given text article, and an aggregation of a country's local news sources provides a view of the country's perspective, or propagated perspective.

Sentiment analysis can be accomplished using a variety of polarity scales, some as simple as a binary 0 or 1 representing negative and positive, while others have larger ranges, such as -5 to 5 [42, 43, 44]. This spectrum includes more levels for the sentiment classification, such as extremely negative, very negative, slightly negative, neutral, *etc.* The general process of sentiment analysis begins with pre-processing the text. Pre-processing the text is vital and can have significant affects on the results, as it is a method of reducing noise in the text [45]. The pre-processing can be summarized in three steps. The first involves data formatting, such as converting all text files to an appropriate, consistent encoding. The second involves data transformation, such as removal of white space, punctuation, numbers, and stop words, expanding abbreviations, lower case conversion, stemming, lemmatization, tokenization, and others [45, 46].

Stop words are words that do not carry important information or content meaning, such as articles, conjunctions, adpositions, and other short function words [46].

Stemming and lemmatization are both processes to reduce a word to a base form. The difference is that stemming is typically a heuristic process, while lemmatization involves converting to a proper morphological root of a word [47]. While lemmitization might be more accurate, stemming is typically used since it is an easier process to implement, has faster processing speed, and the difference in accuracy is normally insignificant. Tokenization is breaking down text into minimal components, such as sentences, phrases, words, or some form of "token." This key concept also highlights the complexity of natural language processing, especially in terms of syntax, semantics, and even pragmatics. Sentiment at the sentence level might differ from the word level due to these complexities. Certain linguistics features, such as valence shifters, are a prime example [48]. Consider the phrase: *He rarely likes running.* From a word tokenization approach, the sample text might be positively polarized due to the term *like.* However, from a sentence approach, the valence shifter *rarely* transfers the sentiment to a true negative polarity. Such challenges in sentiment analysis are observed and continuously being improved upon by researchers. See [49] for a more detailed survey of recent challenges.

The third and final step for text pre-processing involves data filtering, which is extracting text patterns by applying specific functions [46], such as TF-IDF described earlier. The text is then processed through a lexicon classifier that returns a response on the sentiment spectrum (positive, negative, *etc.*).

2.6 Summary

This section provided a recent outlook on Syrian migration and refugee studies and the context of data used for these studies. A summary of relevant text analysis approaches is also provided. Finally, descriptions of framing and sentiment analysis are discussed as these two techniques are the main methods for use in this study.

III. Methodology

3.1 Overview

This section provides the detailed methodology used for this study. First, the data collection method is explained. Second, the data-processing applied to the text articles is summarized. Third, the reasoning and approaches for framing analysis are presented. Lastly, the sentiment analysis methods are described, including the statistical tests utilized. Refer to Appendix A for an overview diagram of the process.

3.2 Data Collection

Data was collected for 1 June 2019 to 30 June 2020, a recent period that contains two important timeline events. First is the European Commission's declaration of the end of the what is known as the "migratory crisis" in October 2019, which falls in the selected time period [16]. Second is a conflict escalation in Idlib that started in December 2019 and continued into 2020, causing many Syrians to flee to Turkey [50]. The articles collected included the keywords "Syria" and "refugee," from fifteen news sources. At least two different news sources were assigned per country, to include either the most influential or the highest sales revenue for that news source [51], [52], [53], [54], [55], [56], [57] and are displayed in Table 1. These newspapers were selected since they were local or based in the target country, and theoretically are a better proxy to the country's perspective than external or non-regional newspapers.

A few remarks regarding the selection of news articles as a medium are presented. Since this study requires observations that are country-specific, it was a main reason for avoiding social media and selecting official news source articles. Social media does not always have a geo-location option feature, and even if it does, many users turn off location options for safety.

Table 1. Sources per Country.

Country	Sources
Iraq	National Iraqi News Agency (NINA), Shafaq News
Israel	Haaretz, Jerusalem Post
Jordan	Amman Net, Ammon News, Jordan News Agency (Petra), Jordan Times, Roya News
Lebanon	Al Manar, Al Nahar, The Daily Star, National News Agency (NNA)
Turkey	Daily Sabah, Hurriyet

In addition, even if users had a location specified, many immigrants of Levant countries would keep their family's country as their country of origin despite only experiencing the country through vacations or short visits; this can provide a skewed perspective on local issues compared to the citizens. This also brings forth the concern of the accuracy of the information in social media. Many might not wish to present their honest views on issues online, especially if there are negative consequences attached. Lastly, the people's perspective in some of these countries might be outweighed as the government would proceed with their intentions regardless. Considering all the uncertainties and multitude of factors that come with social media information, it was deemed more appropriate to utilized official news articles text for the purpose of this study.

3.3 Data Pre-processing

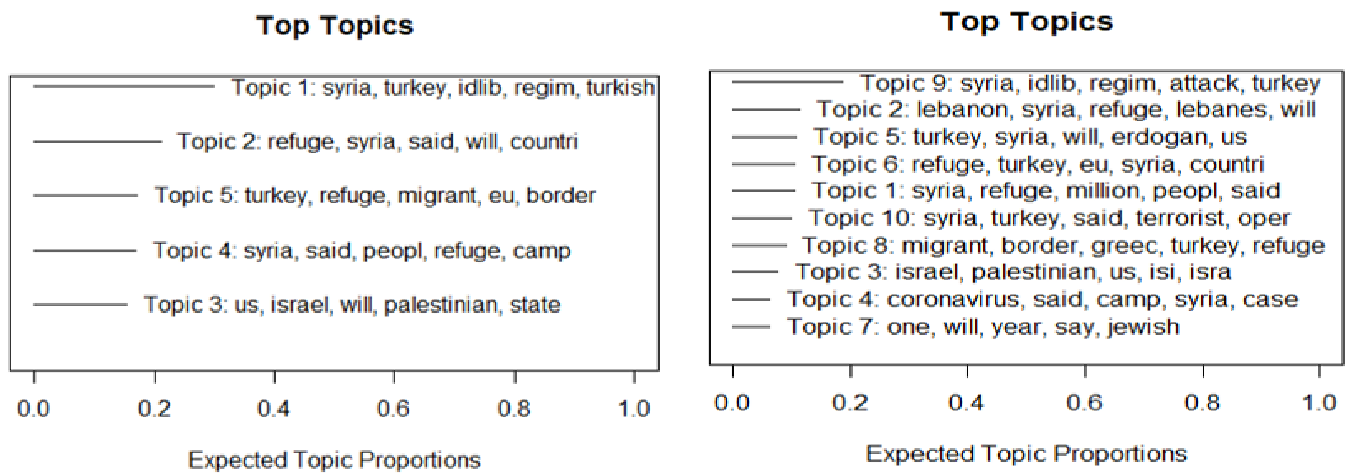
Every article in the selected time frame from the new sources in Table 1 was collected. There were too few Iraqi articles within the criteria period for a representative sample, and therefore Iraq was removed from the countries studied. The final number of article samples after removing the Iraqi observations was 1397 articles. The article samples are as follows: 312 from Israel, 43 from Jordan, 216 from Lebanon, and 826 from Turkey. Refer to Table 8 in Appendix B for further breakdown details. For BoW methods and exploratory analysis, the text articles were pre-processed. All text files were converted to UTF-8 encoding (Unicode Transformation Format-8 bits). This encoding is preferred since it is backwards compatible with other text file formats. The corpus is created of single term tokens, each term composed of two or more characters. Numbers, symbols, punctuation, and URLs were removed. All characters were converted to lower case. All stop words were removed, and all the tokens were stemmed. Additional filler words from split hyphens were removed as well, such as *al, el, il*, which translate to article terms in Arabic. The minimum allotted term frequency was 50, with a minimum document frequency of 10.

3.4 Framing Analysis

The initial attempt was to utilize STM to get a general idea of the possible frames. Topic modeling via the *stm* package [58] was performed with the set of topics (K) ranging from five to ten, based on the size of the corpus. However, the resulting topics from STM are too refined (see Figure 2 for examples), either indicating specific events or not quite indicating any framing theme at all. Thus, for the purpose of this study and for a more rigid structure for analysis, five abstract frames were created and are displayed in Table 2. The basis for selecting these categories is a survey of migration representation articles, then generalized into overarching topics.

Table 2. Frame Categories and Description

Category	**Description**
Political	Involves government or policy discussion
Economic	Relating to country funds or any economic impact
Crime	Any criminal aspect either by or against refugees
Documentary	Current conditions of refugee or process of fleeing
Other	Not related to any of the above



(a) K=5 (b) K=10

Figure 2. STM with K Topics

To account for multiple frames in a given article, each article can be assigned up to two frames. If an article has two frames, the more prevalent frame (primary frame) is weighted at 2/3, and the other (alternate frame) at 1/3. Absence of a frame is indicated by a zero. This provides a matrix of samples with frame features in the range of $[0, 1]$. Percentages of frame compositions are then calculated while accounting for the mixture of frames per article. Additionally, the *crime* frame has a label to indicate the victim or offender in the article context; this distinction is based on the article's wording as of the date of the article. All labels were assigned by the same individual. Frames were then analyzed by country as well as time period.

3.5 Sentiment Analysis

The first approach was to break down an article into tokens of content words, determine the sentiment of the words individually, and then aggregate the individual word sentiments to assign a sentiment score to the article as a whole. There are multiple lexicons (dictionaries) available through the *R* software [59], and in this study, three are utilized. The first is identified as the Bing lexicon, and is a database of English words categorized as either positive or negative [43]. The second is the National Research Council Canada's Sentiment and Emotion Lexicons, referred to as the NRC lexicon, and it differs from the Bing binary categorization in that it parses the words further into categories of emotion, such as anger, anticipation, disgust, and so forth [44]. The third is the Afinn lexicon; it differs from the other two lexicons in that it utilizes a spectrum for word sentiment scoring, ranging from -5 to 5 [42]. The creator for this lexicon selected the scale range from *very negative* as -5 to *very positive* as 5, which allows for intensity levels in sentiment analysis, i.e., a word can be more negative than another in a measurable score. For a simple demonstration, consider for instance the following two terms: *unpleasant* and *horrendous*. The Oxford dictionary defines horrendous as *extremely* unpleasant, indicating a greater intensity from one word to the other. For more information on these lexicons, see [43], [44], and [42]. The general procedure for sentiment analysis is summarized as follows:

1. Load and transform articles text into data frame with meta data.
2. Pre-process and clean text articles (see Data Pre-Processing section).
3. Parse each text article into tokens per observation (*e.g.* words per article).
4. Load target sentiment lexicon.
5. Match by left-joining lexicon to tokens and address non-matching tokens.
6. If non-numeric lexicon, assign numeric value to categories (negative as -1, positive as 1, neutral as 0).
7. Aggregate by observation to determine net numeric value of sentiment for each article, labeled as *sentiment score*. These scores are unbounded. Standardize scores.
8. Bind sentiment scores to meta data for analysis.

The lexicons are not all-inclusive of English content words, but instead focus on sentiment words. There are also tokens that have no predefined matching sentiment value, such as proper names or locations. For consistency, attaching sentiment to names or locations was omitted as this sentiment can be subjective and vary depending upon the country, news source, or time period. Instead, all the non-matching tokens were addressed by assigning a neutral value to the token. Since the text articles were formal communication, colloquialism or double-meaning slang terms are not deemed a concern. Once the sentiment scores were aggregated and attached to the meta data, they can then be analyzed by available prominent features.

A second approach was to determine the sentiment of each sentence, and then aggregate the sentiment score by sentences per article rather than words per article. The package *sentimentr* [60] is deemed best for this approach since it takes into account valence shifters. Using the Bing dictionary, a comparison of the sentiment score via word and sentence tokens indicated that $\approx 8\%$ of the data had a one-unit difference, and $\approx 2\%$ had greater than a one-unit difference. Overall, both approaches provide relatively similar sentiment scores in this case. The advantage of the word token approach though is that it is faster in terms of processing time. Hence, the word token approach was adopted and deemed sufficient for this study application.

Comparison of the lexicons was maintained throughout the analysis in the case of a major difference between the sentiment scores. For the NRC lexicon, while there are multiple emotion classifications, the four relevant to the context of the study were explored: joy, anger, fear, and sadness. This distinction of the type of sentiment is more informative than positive or negative sentiment classification. The sentiment score results are added to the meta data for analysis. Finally, non-parametric statistical tests with an alpha level of 0.05 are applied for insights based on meta data feature groupings such as country, news source, or time period.

The reasoning for the use of non-parametric statistical tests is as follows: Nayak and Hazra [61] mention that the first step to determining which statistical test is best utilized for numerical data is to identify if the data follows the parameters of a known distribution curve, typically the normal distribution. Since the sentiment score sets failed the Shapiro-Wilk test, did not fit known distributions, and are essentially an ordinal scale, non-parametric statistical tests were deemed appropriate. Specifically, the Mann-Whitney U test for two group analyses and the Kruskal-Wallis H test for groups of larger than two.

Nayak and Hazra state that utilizing a two-group test, such as the Mann-Whitney U test "to a multiple group situation increases the possibility of incorrectly rejecting the null hypothesis" [61]. Essentially, the Kruskal-Wallis H test is an extension of the Mann-Whitney U test that accounts for larger groups. These tests only determine if there is a difference in groups, but cannot determine where the difference lies. The post hoc Dunn's test serves this purpose and identifies specifically which two data sets differ [61]. Refer to Table 3 for the null and alternate hypotheses of these tests. Nayak and Hazra provide a compact summary of the statistical test selection process, as displayed in Figure 3.

Table 3. Statistical Tests Null and Alternate Hypotheses

Test	**Hypotheses**
Kruskal-Wallis	H_0 : The distribution of sentiment scores for all groups are the same. H_A : At least one is different.
Dunn's Test	H_0 : There is no difference between the groups. H_A : There is a difference between the groups.

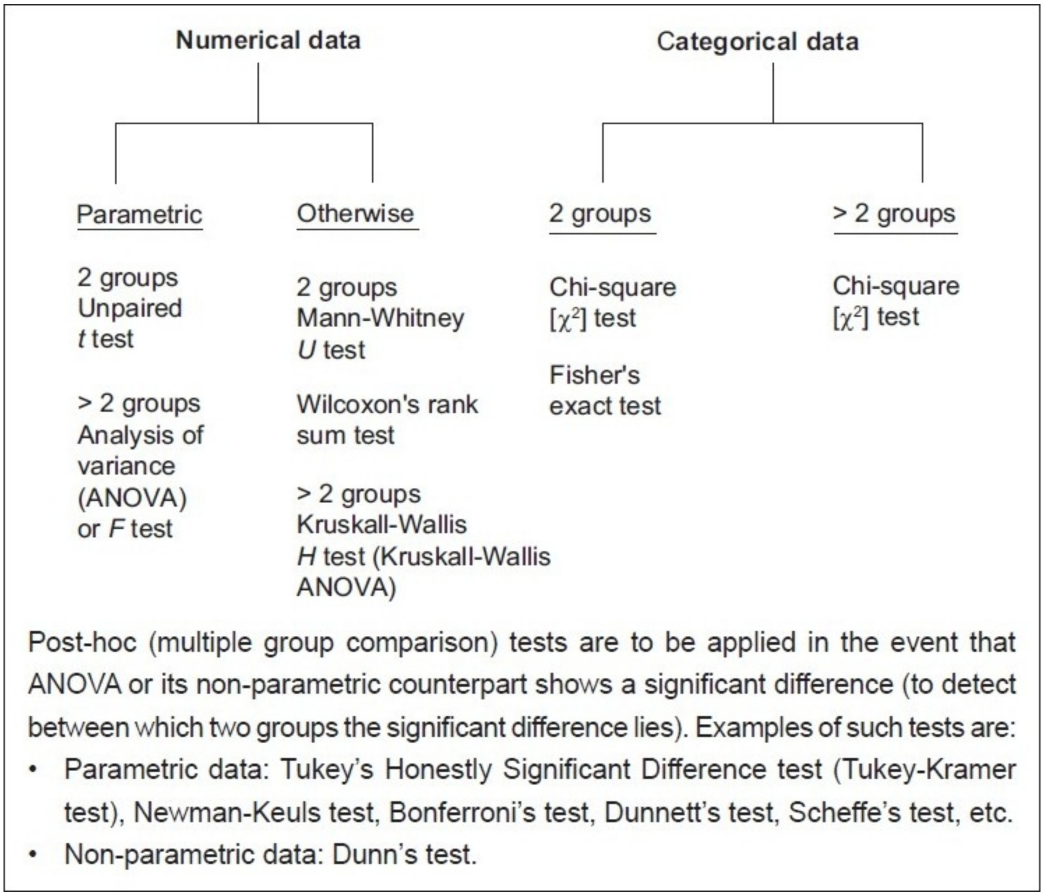


Figure 3. Nayak and Hazra's Test Selection Process. Source: Adapted from [61, p.1].

3.6 Summary

This section discussed the detailed steps of the methodology used to collect, process, and then analyze the data. Next, the selected framing analysis approach is provided and described. Finally, the steps for performing sentiment analysis are presented with an explanation of the selection of the statistical tests for analysis.

IV. Observations and Analysis

4.1 Overview

This section explores the data and discusses the results of the frame and sentiment analyses. All results and observations are in the context of the articles with the keywords "Syria" and "refugee," from the time period of 1 June 2019 - 30 June 2020.

During the selected period, and for the identified news sources, there were too few Iraqi article samples for a conclusive analysis. Therefore, Iraq was removed from the country samples, alongside the associated observations. While Arabic or Kurdish written articles might provide additional observations for a more representative sample, these were beyond the scope of this study. The majority of locally based Iraqi news sources did not have an English equivalent, and this could be a reason for the lack of samples. Another reasoning could be that Iraq had many internal protests and domestic grievances that might have taken priority over Syrian refugee coverage during this time. Regardless, the analysis was completed excluding Iraqi observations due to insufficient samples. All analysis was completed with *R* software [59] (see Appendix C for packages list), and all statistical tests had an alpha of 0.05 level of significance.

4.2 Exploratory Analysis

Preliminary analysis on the corpus is conducted via frequency diagnostics. The top ten terms filtered by a minimum term frequency of 50 and a minimum document frequency of ten are displayed in Figure 4. These top ten term results have highlighted the prevalence of themes such as religion and terrorist groups in the refugee articles corpus. These themes and the association of major powers in the region was explored. Each theme in Table 4 was formulated from a list of associated words. For example, the theme *United States* included terms such as *Trump*, *America*, *Washington*, and others that indicated an association to the U.S. The presence of a theme relative to the number of Syrian refugee articles by country was computed to gauge association in an exploratory manner. The results are summarized in Table 4 and briefly discussed. From all of the Israeli news samples mentioning Syrian refugees, approximately 62% also included a term associated with religion, which is relatively high compared to the remaining countries. Terrorism associated terms were largely present in all the country samples, the highest being in Israel and Turkey article groups. The association of such themes in refugee articles is concerning, as it shifts focus from the universal human aspect association to an individual or group preference bias. Finally, some insight on the association of major powers revealed the United States as the largest percentage mentioned across all the sample countries. It is reasonable to assume that the U.S. has a large association to refugee news in the Levant area, and possibly due to foreign policy or United Nations initiatives and financial contributions.

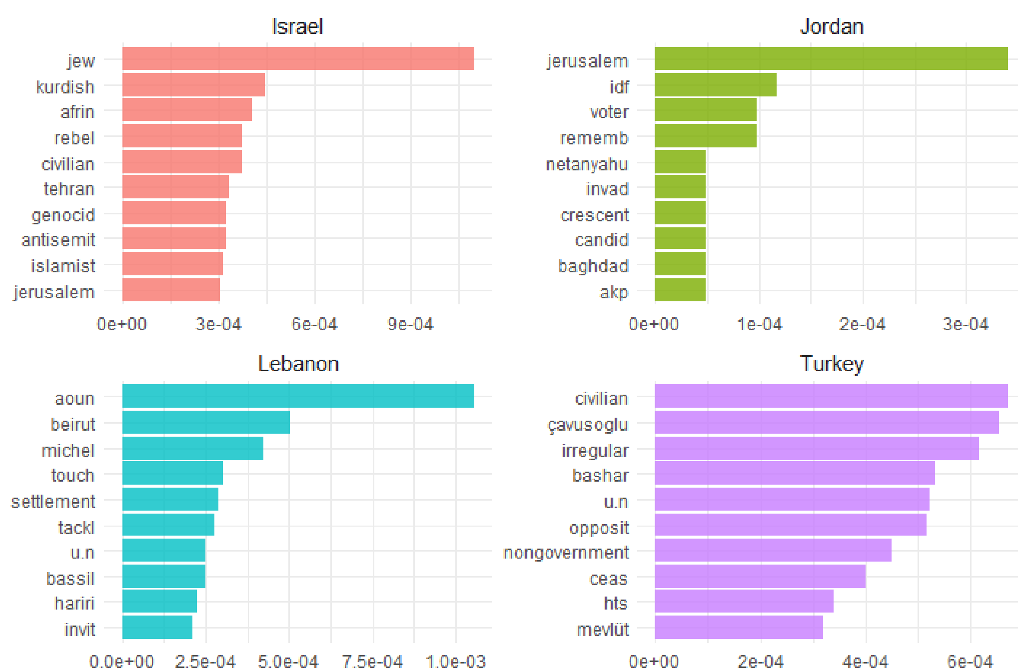


Figure 4. Top Ten TF-IDF Words by Country

Table 4. Topic Frequency Percentages per Country

Associated Topic	Israel	Jordan	Lebanon	Turkey
Religion	61.86	9.30	13.89	11.50
Terrorism	74.36	60.47	48.15	73.12
United States	96.15	90.70	81.02	96.49
China	8.65	4.65	1.39	5.21
Russia	49.36	4.70	19.44	51.82

4.3 Frame Analysis Results

To address the first research question, the selected countries from the Levant region as a whole have a dominantly political frame of Syrian refugees for the sample period. As displayed in Figure 5, over half of the frames composition is political. Documentary and economic frames were both approximately 7% of the total frame compositions, and the *other* category was approximately 14%. In this *other* category, approximately 35% were of a type of humanitarian charity or mission, 25% were health-related (coronavirus), 21% were religious-related, and the remaining were miscellaneous topics, such as entertainment, local disasters, marriage rates, *etc.* The crime frame was the least utilized, forming around 5.5% of the total observations. These crime frames were further separated into categories to distinguish between the refugee representation as either a victim or offender. Approximately 70% of the crime frames were acts of crime against refugees by non-terrorist groups, 21% were of terrorist crimes against refugees, and the remaining 9% were of crimes committed by refugees. There were two abnormal instances of a crime frame by an individual impersonating a refugee.

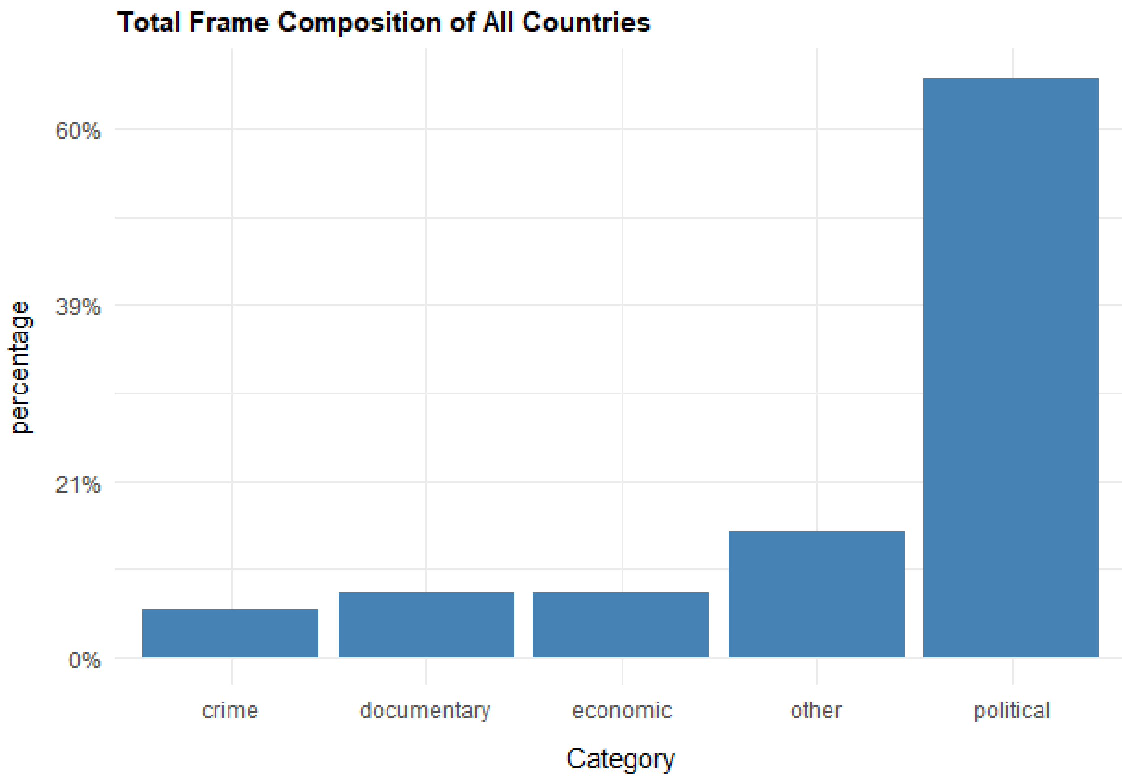


Figure 5. Frame Category Composition of All Samples

Further partitioning of the frames by country is shown in Figure 6. While Israel, Lebanon, and Turkey have a clear distinction as to the political frame being the largest utilized compared to the remaining frames, it is not the case for Jordan. Jordan's economic, political, and "other" frame compositions are roughly equivalent, each accounting for almost a third of Jordan's total frame compositions. This is a potential indication of a more varied framing of Syrian refugees in Jordan, while the remaining countries lean more heavily towards political frames. It could also be due to Jordan's press censorship, specifically an offense under their Anti-Terrorism Law that can charge individuals or entities with "disturbing relations with a foreign country" [62], and thus discourage political contexts.

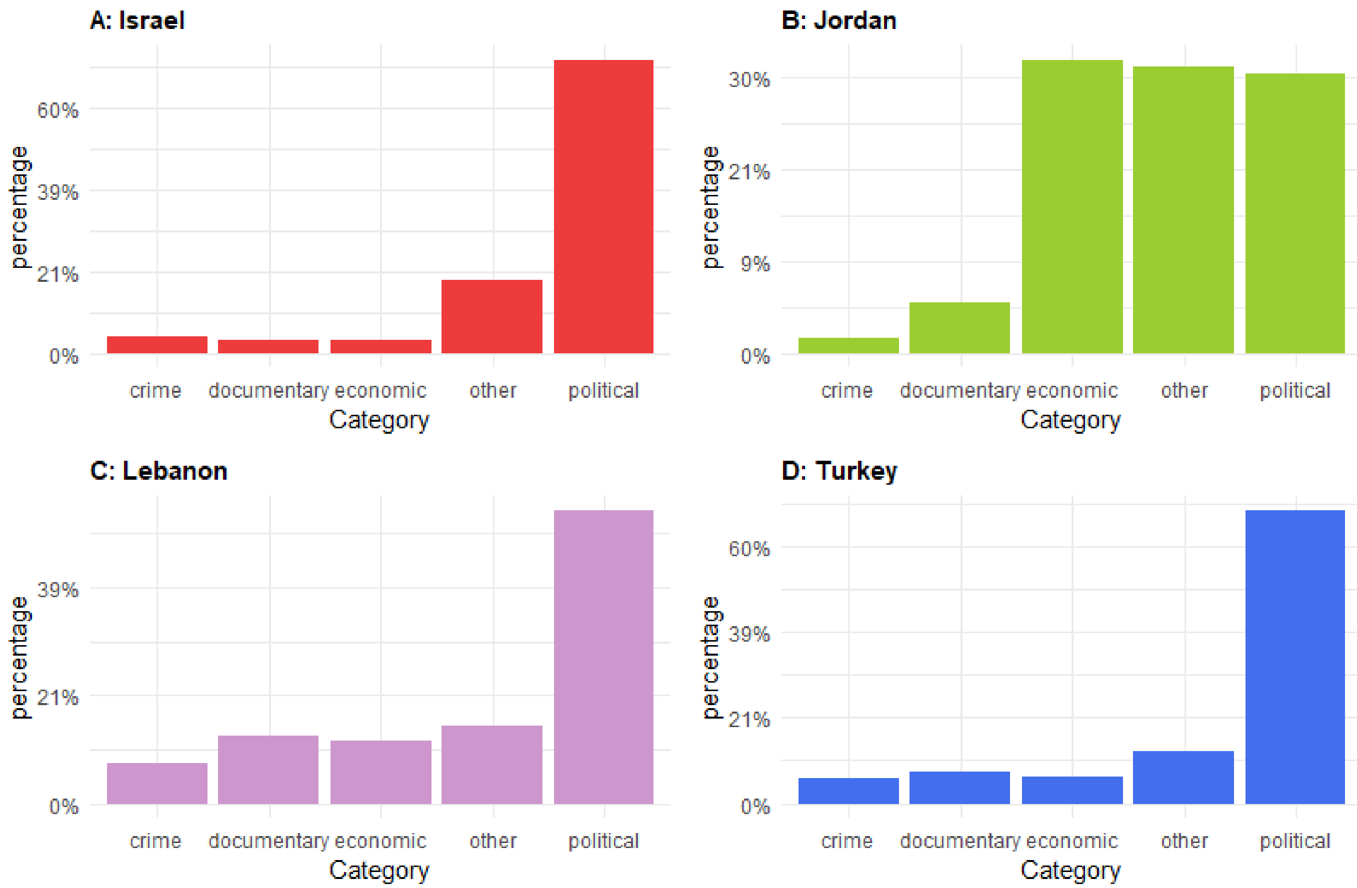


Figure 6. Frame Compositions Partitioned by Country

Another aspect to consider is that these frames are from an aggregate period. Mapping the frame compositions over smaller units of time confirms the dominance of the political frame throughout the entire selected period with the exception of the month of April 2020, as displayed in Figure 7. The dominant frame of this month is "other," and the most likely event to correspond with this gradual shift from March to April is the coronavirus outbreak in the region. The World Health Organization began publishing situational reports for the Eastern Mediterranean during this time, and encouraged states to provide coronavirus prevention guidance for the Easter and Ramadan holidays in April [63]. Thus, it is not unexpected for an emphasis of the outbreak to be included in all news, including Syrian refugee articles to cover the status of the outbreak in camps, as well as guidelines and testing plans for all. As for the peak intensity of political frames building from December 2019 to February 2020, it is potentially due to a chain reaction of compounding events from the reignited conflict in Idlib that eventually subsided with a ceasefire agreement between Russia and Turkey in March 2020 [64].

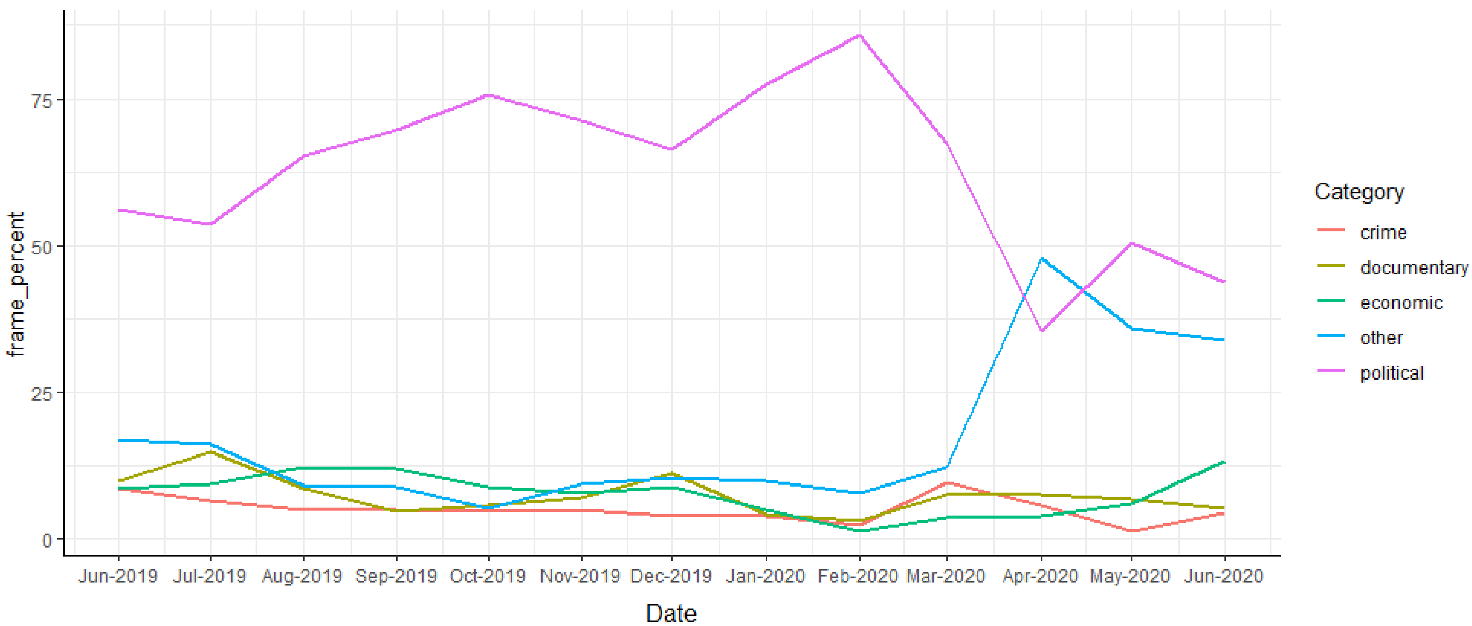


Figure 7. Frame Compositions Over Time

Overall, it is reasonable to assume that Syrian refugees were largely portrayed in a political frame for the selected period. In the past few years, one of the most alarming concerns where politics and refugees collide is the concept of refugee weaponization. This concept involves the exploitation of human migration, both voluntary and involuntary, to achieve some objective, such as political, militarized, economic, or other [65]. In 2016, during a hearing of the U.S. Senate Armed Services Committee, General Breedlove stated that Russia and the Assad Regime were "deliberately weaponizing migration in an attempt to overwhelm European structures and break European resolve" [66]. Whether that was the true intent of Russia or not is irrelevant. The perception of this concept ultimately led to materialization, and within the past five years other states have utilized this strategy for political leverage. Prominent examples include Belarus and Turkey's refugee weaponization against the European Union [67, 68].

With the large association of refugees with politics, the issue of refugee weaponization is likely to become even more normalized. Moreover, the shared perspective of these countries in framing Syrian refugees politically can gradually reinforce acceptance of the concept and consideration as a viable option for power. Logically, refugee weaponization is an easy and inexpensive option for less powerful states to gain leverage in a cost imposing strategy against a more powerful state. Plans and policy to properly address and prepare for such an issue, in both foreign and domestic instances, can help maintain leverage for the U.S. and allies. Observation of the situation in the Levant region and establishing lesson learned and mitigations for potentially similar scenarios can prove useful for the U.S. and allies.

4.4 Sentiment Analysis Results

The second research question is addressed via sentiment analysis by country, as well as news source. The third research question is then addressed by mapping sentiment over time. Shared sentiment towards Syrian refugees can indicate shared perspective for a grouping of states. Exploring the sentiment, especially as a reaction to possible events, can clarify some of the dynamics of the complex relationships these countries have, as well as their posture and intentions towards future policies or conflict. This information can prove useful for negotiations or leverage in the global political stage.

4.4.1 By Country and News Sources

Analysis from the standardized sentiment scores using Bing's lexicon suggest that some countries share similar sentiment towards Syrian refugees in their articles for the selected time period. Jordan and Lebanon seem to fall in one group, and Turkey and Israel in another, as displayed in Figure 8.

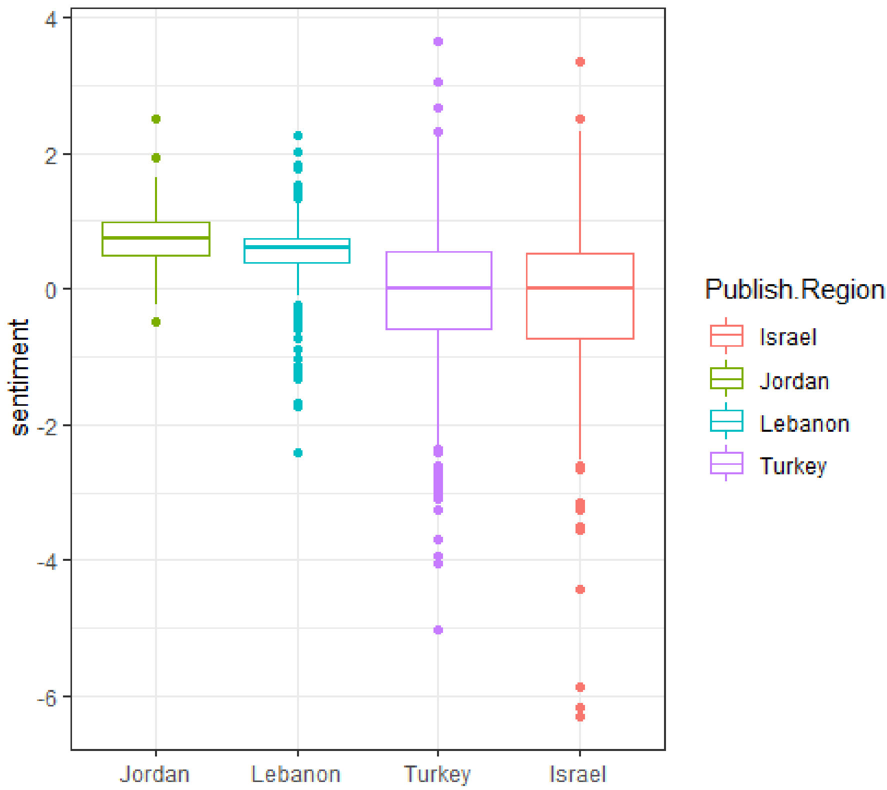


Figure 8. Standardized Sentiment Scores by Country – Bing Dictionary

For statistical significance, a Kruskal-Wallis test was performed, resulting with a p-value less than 0.05. The null hypothesis is that the distribution of sentiment scores for all groups are the same, and the alternate is that at least one group is different. The result indicates that there is likely statistically significant differences between sentiment scores of the countries.

To find the pair differences between countries, a Dunn's test was performed, adjusted with a Bonferroni correction. The null hypothesis for Dunn's test is that there is no difference between each pair of countries, while the alternate is that there is a difference. The results are displayed in Table 5. There is statistically significant difference between all comparison pairs except the Jordan - Lebanon, and Israel - Turkey pairs. This suggests that Jordan and Lebanon are similar in their sentiment perspective towards Syrian refugees, and this perspective differs from the one shared by Israel and Turkey. This divide could naturally be due to Jordan and Lebanon being Arab countries, with historical similarities and traditions shared with Syrians and not with Turks or Israelis. The division could also be an indication of current affairs, whereas Turkey and Israel have larger global diplomatic indexes, economies, and military spending budgets compared to economically struggling Lebanon and Jordan that are possibly breaking under refugee strain [69], [70], [71].

Table 5. Bing Lexicon: Dunn's Test with Bonferroni Correction via *FSA* package [72]

Comparison Group	**Adjusted p-value**
Israel - Jordan	< 0.05
Israel - Lebanon	< 0.05
Jordan - Lebanon	> 0.05
Israel - Turkey	> 0.05
Jordan - Turkey	< 0.05
Lebanon - Turkey	< 0.05

The same analysis was then performed using Afinn's lexicon for comparison. The standardized scores portrayed a similar grouping as the Bing lexicon, as shown in Figure 9. The same tests were performed for confirming statistical significance. The Kruskal-Wallis test resulted with a p-value less than 0.05, where the null hypothesis is that the distribution of sentiment scores for all groups are equivalent, and the alternate is that at least one group is different. Dunn's test also showed consistent results as with the sentiment scores from the Bing lexicon (see Table 6).

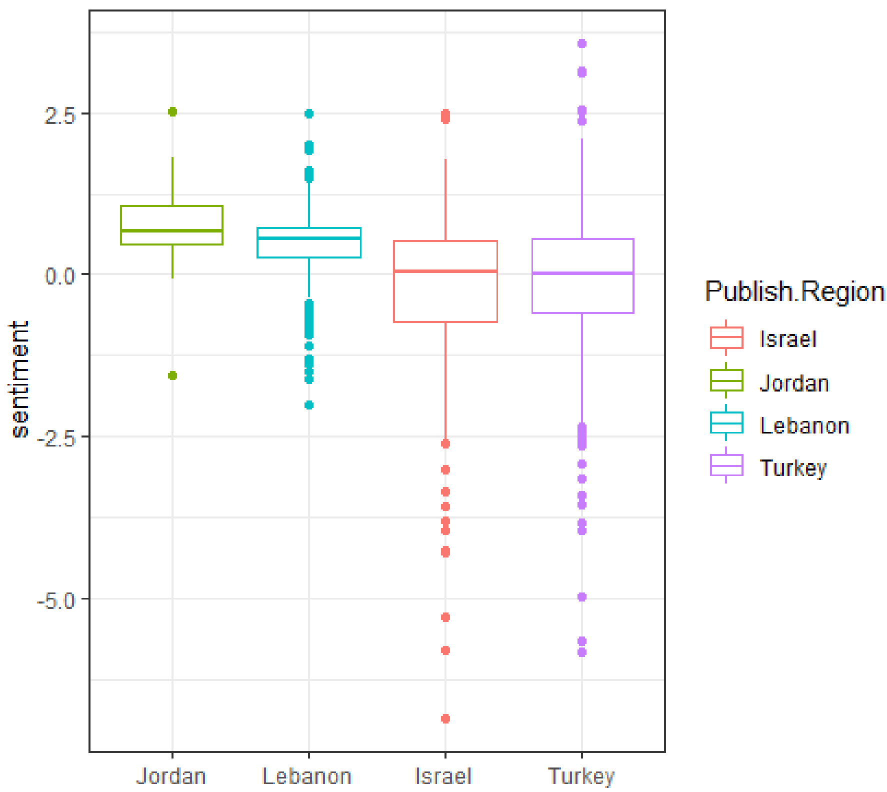


Figure 9. Standardized Sentiment Scores by Country – Afinn Dictionary

Both lexicons confirm that there is a difference in sentiment intensity between the Jordan-Lebanon group, and the Israel-Turkey group for the selected time period observations. The underlying deduction is that Israel and Turkey tend to have more negative sentiment intensity in their articles of Syrian refugees than Jordan and Lebanon. Comparison of the frequency of the raw sentiment scores shows that the samples from Israel and Turkey had the largest percentage of negative sentiment articles. Approximately 80% of Turkey's articles had an overall negative sentiment with the Bing lexicon, and roughly 76% with the Afinn lexicon. For Israel, negative articles were approximately 81% and 80% of the observations via the Bing and Afinn lexicons respectively. Lebanon hovered in the mid-range with 56% via the Bing lexicon, and 52% via the Afinn lexicon. Jordan had the least percentage of negative sentiment articles; 39% for the Bing lexicon and 33% for the Afinn lexicon. Table 7 provides further details.

Table 6. Afinn Lexicon: Dunn's Test with Bonferroni Correction via *FSA* package [72]

Comparison Group	**Adjusted *p*-value**
Israel - Jordan	< 0.05
Israel - Lebanon	< 0.05
Jordan - Lebanon	> 0.05
Israel - Turkey	> 0.05
Jordan - Turkey	< 0.05
Lebanon - Turkey	< 0.05

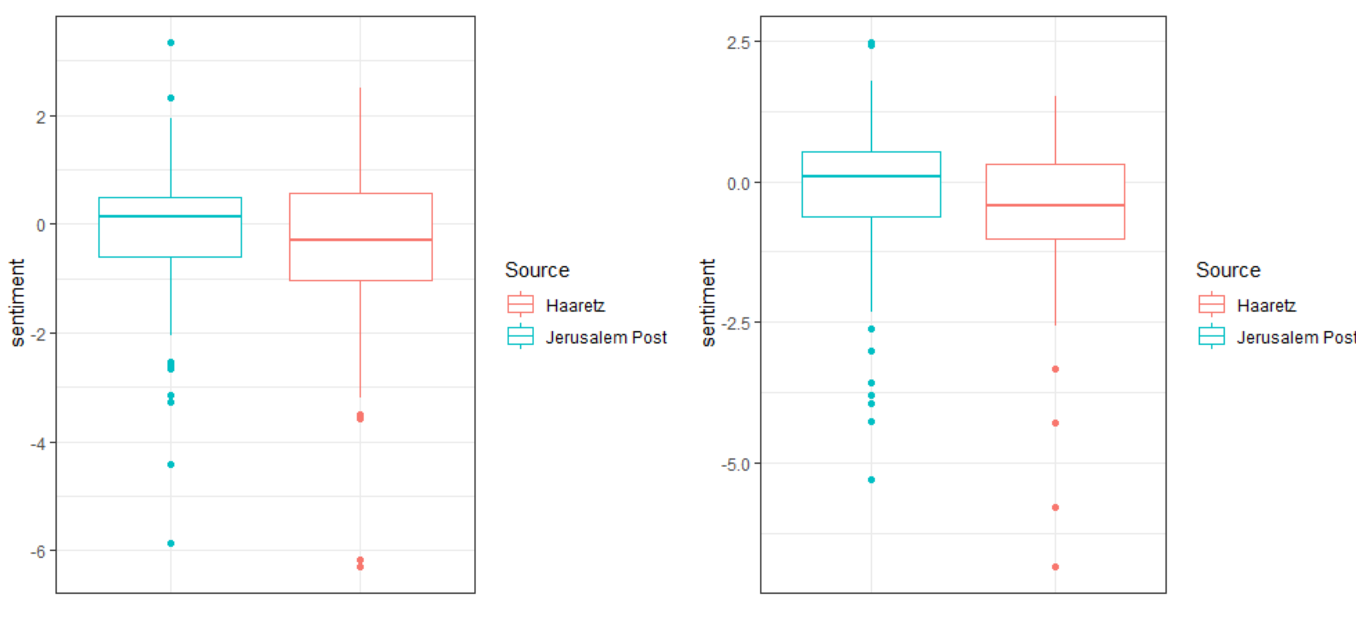
Another aspect for analysis is the breakdown of each country's sentiment by news source. These news sources are popular and influential in projecting the country's perspective, albeit propagated or not. They play an important role in shaping the concept of the Syrian refugee in the target country, and sometimes can be significant enough to evoke change in the public's sentiment towards the Syrian refugees. A

Table 7. Words Sentiment Percentage by Country

	Bing Lexicon			Afinn Lexicon		
	Positive	Negative	Neutral	Positive	Negative	Neutral
Israel	16.03	81.41	2.56	18.59	79.49	1.92
Jordan	51.16	39.53	9.31	62.79	32.56	4.65
Lebanon	30.56	56.48	12.96	38.43	52.78	8.79
Turkey	16.95	79.66	3.39	21.19	76.27	2.54

variation between the news sources' sentiment is a strong indicator for variation in inter-state perspectives, which is a sign of differentiation within the nation, and less of group think or dictated news reporting. The sentiment variation can also hint to flexibility in postures of a nation in regards to a target topic, and in this case specifically, Syrian refugees.

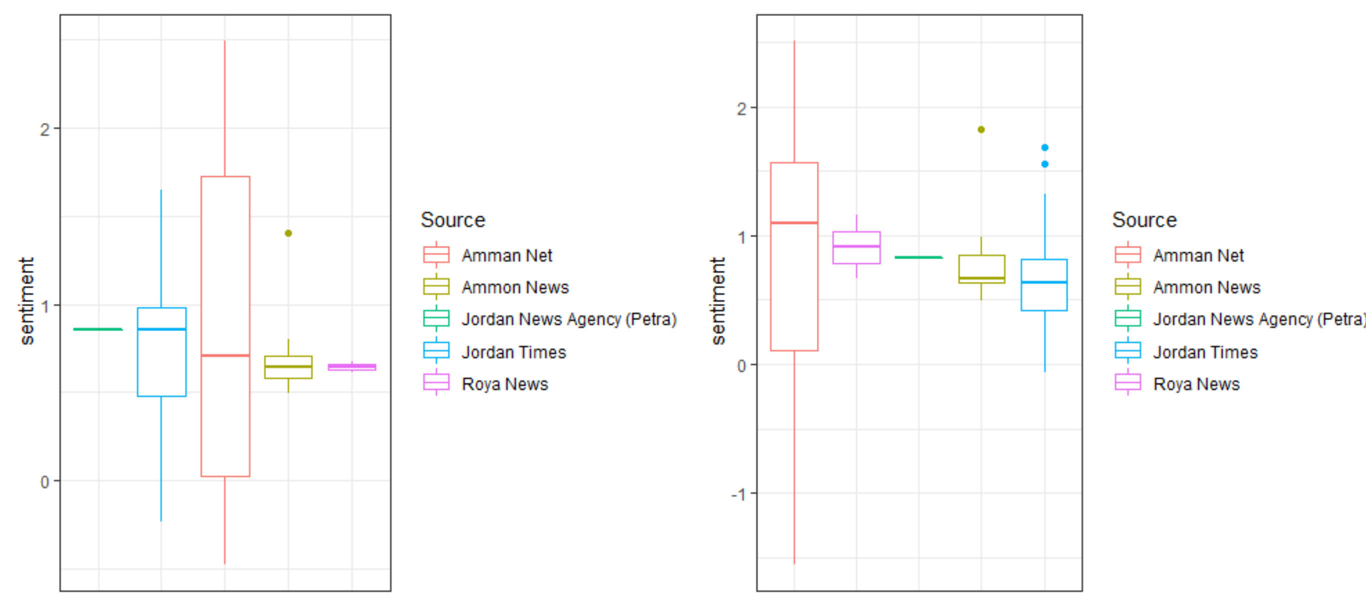
The breakdown of the news sources sentiment for Israel is displayed in Figure 10. The Jerusalem Post scores are slightly left skewed with the Bing lexicon. Since there are two groups, a Mann-Whitney U test is performed, with a null hypothesis that the two sources come from the same population (i.e. the two news sources can be treated as the same news source), and the alternate hypothesis is that they do not. The test results in a p-value less than 0.05 for both Bing and Afinn lexicons. Therefore there is likely a significant difference in sentiment towards Syrian refugees between the two news sources. The main difference between the two sources is possibly that Haaretz is more left-centered and Jerusalem Post is right-centered.



(a) Bing Lexicon (b) Afinn Lexicon

Figure 10. Israel - Standardized Sentiment Scores by News Source

For Jordan's news sources, the sentiment breakdown is displayed in Figure 11. Jordan News Agency (Petra) has a single observation as shown. Amman Net has the largest dispersion in both lexicons. A possible reason for the distinction is that it is the only identified non-profit news source from the group. However, to determine if there is a statistical significant change in sentiment between the news sources, a Kruskal-Wallis test was performed. The resulting p-value was greater than 0.05, thus failing to reject the null hypothesis that the distribution of sentiment scores for all news sources are the same; i.e., essentially they can be treated as the same news source due to lack of evidence indicating otherwise. It is not surprising that Jordan's news sources are similar in their sentiment, as news censorship is quite well known to be present in Jordan.



(a) Bing Lexicon

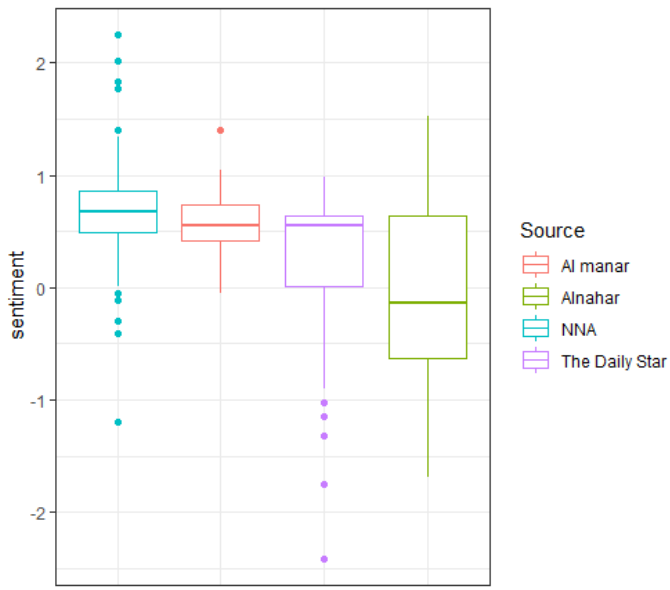
(b) Afinn Lexicon

Figure 11. Jordan - Standardized Sentiment Scores by News Source

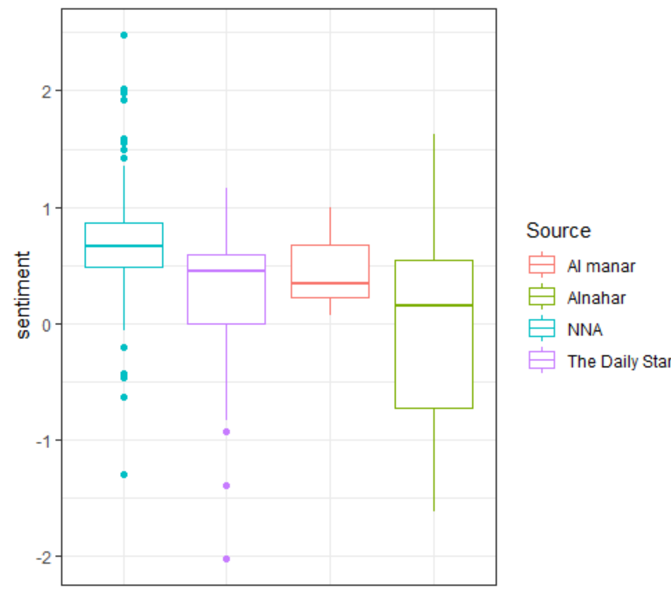
The Jordanian government is very cautious of the country's relations with foreign countries, and this is written in law and terms for media licenses. An example of this extreme caution played out in 2019, when an Iraqi TV satellite that is based in Amman was suspended for covering protests in Iraq [62]. It could very well be that Jordan has a restricted and possibly censored sentiment towards Syrian refugees as a precaution for foreign relations with Syria. The relations between the two neighbors had declined during the Syrian civil war, with the Jordanian monarch being one of the first Arab leaders to call for Bashar Al-Assad to step down. However, the stance seems to have changed over the years, especially after the international crossing between the two countries, the Nasseb-Jabeer border, was reopened in 2018. This was most likely in an effort to restore trade for a weakening economy in Jordan [73]. While the border was later closed in August 2020 due to coronavirus, it does reinforce the intended stance of Jordan towards Syria, and their reluctance in projecting negative sentiment towards Syrian refugees as a precaution for country relations.

As for Lebanon's news sources, the sentiment is displayed in Figure 12 for both Bing and Afinn lexicons. Al-nahar has the largest dispersion between samples among the sources. The Lebanese National News Agency (NNA) and Al-manar are the least dispersed. The Kruskal-Wallis test had a p-value less than 0.05, rejecting the null hypothesis that the distribution of sentiment scores for all news sources are the same, and indicating a significant difference among them. A Dunn's test was performed to determine pair differences. The null hypothesis is that there is no difference the pair, and the alternate is that there is a likely difference. The results indicated that the NNA was dissimilar to the other news sources. This difference could be due to the NNA being the only state-owned news agency from the group.

Overall, the NNA had a less negative tone than the three other news sources, which are all owned by prevalent political entities in Lebanon. Al-manar is operated by the terrorist organization Hezbollah. Al-nahar is 98% owned by two of the largest politically involved families in Lebanon, the Harriri family and the Tueni family. The



(a) Bing Lexicon



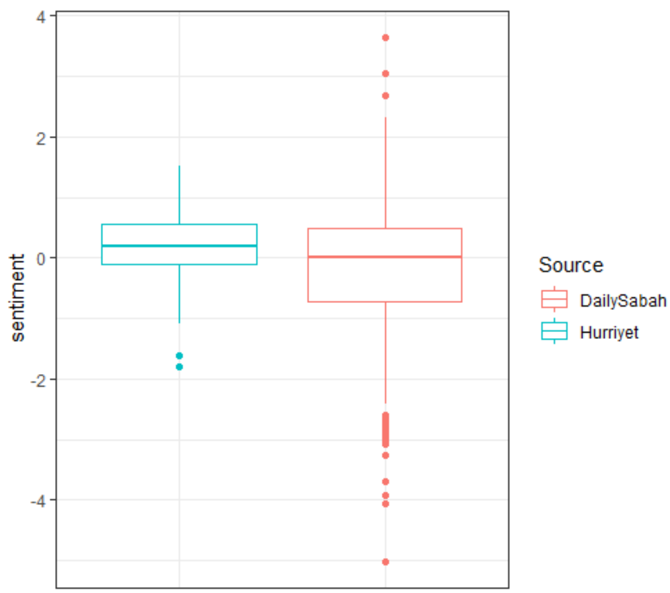
(b) Afinn Lexicon

Figure 12. Lebanon - Standardized Sentiment Scores by News Source

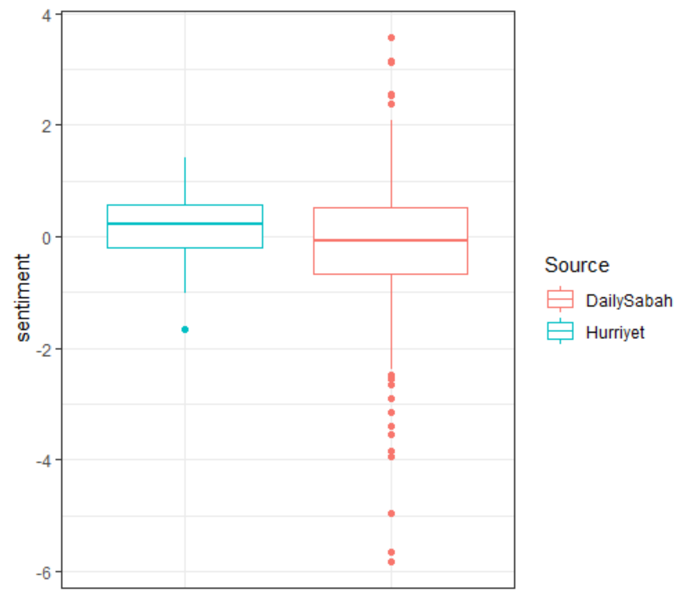
Daily Star is 100% owned by the Harriri family as well. As for the NNA, the director has remained the same, but the minister of information has changed alongside the new formed governments in 2019 and 2020. In 2019 the minister was Jamal Jarrah, a member of the Future Movement political organization founded by Saad Harriri from the Harriri family, and in 2020 the minister was Manal Abdel Samad, a member of the Lebanese Democratic Party. The more neutral tone of the NNA could be due to less interference resulting from the rapid turnover in ministers. As for the sentiment from the other news sources, it most likely has a political touch based on the owners.

Finally, the breakdown of news sources sentiment for Turkey is displayed in Figure 13 for both Bing and Afinn lexicons. The Daily Sabah has a slightly wider dispersion than Hurriyet. To test significant difference, a Mann Whitney U test was performed, resulting with a p-value less than 0.05 for both lexicons. Thus, the null hypothesis that the two news sources come from the same population is rejected. From the samples collected, Hurriyet had a more neutral tone overall than Daily

Sabah. While both news sources are considered pro-government, Hurriyet is typically seen as left-wing and Daily Sabah right-wing, which could be a possible reason for the sentiment distinction.



(a) Bing Lexicon



(b) Afinn Lexicon

Figure 13. Turkey - Standardized Sentiment Scores by News Source

In general, analysis for the second research question has deduced two key items thus far. The first is that there are two similar groups of countries; Israel and Turkey form one group, and Jordan and Lebanon form the other. Moreover, the first group of Israel and Turkey tends to have more negative polarity in the news articles than Jordan and Lebanon. The second item is that all the sample countries have variation in their news sources' sentiment with the exception of Jordan. The following subsection addresses the third research question by analyzing change in sentiment over the selected time period.

4.4.2 By Sentiment Responses Over Time

In this subsection, drastic sentiment changes are highlighted with feasible factors that could have driven the sentiment change. While sentiment analysis with a positive to negative spectrum is informative, a deeper dissection based upon emotions delivers better sentiment clarity. The NRC lexicon provided this extra layer with multiple emotion classification, and in this case four relevant emotions were investigated: joy, anger, fear, and sadness. Potential patterns of reaction behavior can assist policy makers and analysts in understanding the narrative of these countries towards Syrian refugees, the dynamics of the complex relationships among the host countries and Syria, as well as future stances or intentions that could alter foreign policy decisions.

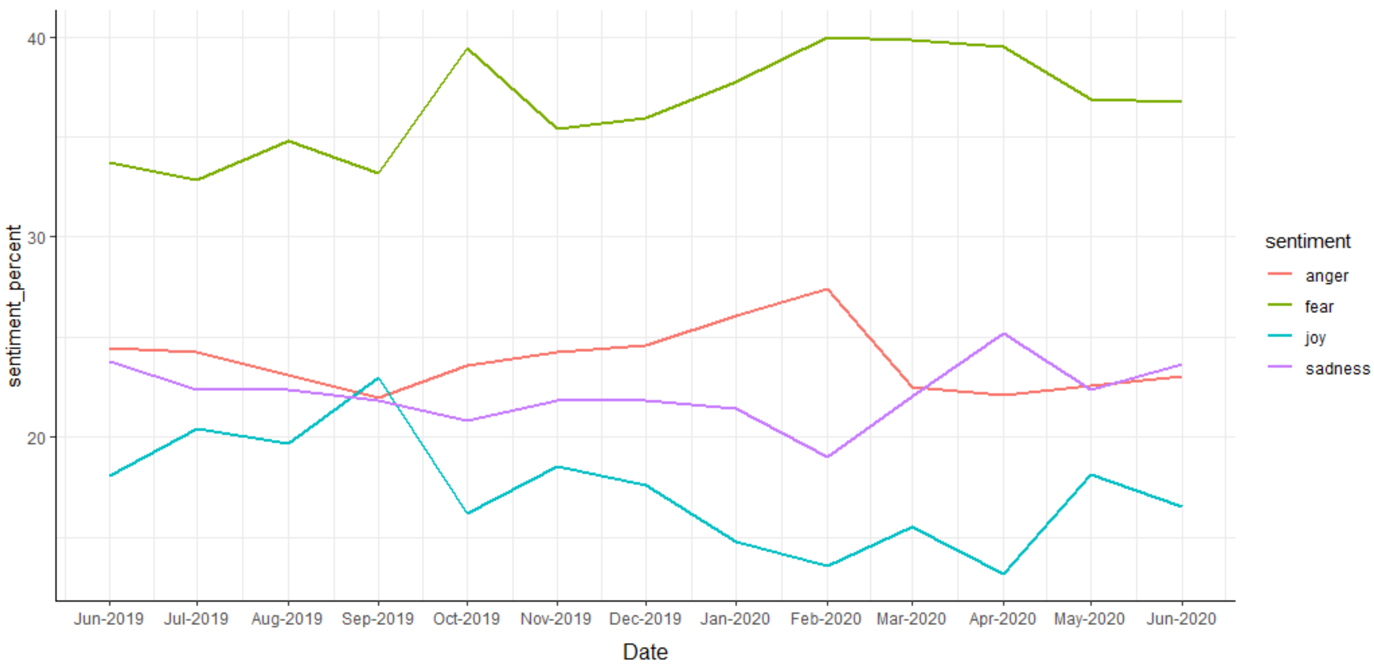


Figure 14. Neighboring Countries' Sentiment, June 2019 - June 2020

Events from the Syrian side are compared to the sentiment percentages displayed in Figure 14, to gauge reactions as a whole of the neighboring countries. For the duration of the selected time period, the fear sentiment was the most prevalent throughout the duration of the sample articles associated with Syrian refugees. The fear sentiment percentage peaked during the months of October 2019, and February-April 2020. In October 2019, two critical events were announced that could have contributed to a change in sentiment. The first was the withdrawal of U.S. troops in northern Syria, a move that represented a "significant shift in U.S. foreign policy" [74]. The second is the death of the Islamic State of Iraq and the Levant (ISIL) leader Abu Bakr al-Baghdadi [75]. The implications of the first event could be the fear due to uncertainty that comes with the loss of a major power's involvement, while the second is possibly the concerns of revenge or increased conflict. Both events entail follow-on consequences through the region, specifically more forced migration from Syria. Thus it is plausible that the neighboring countries perceive an association between the ramifications of these events and Syrian refugees, as is disseminated by the sentiment tone in the news articles.

As for the February-April 2020 time frame, multiple significant developments occurred. The Balyun airstrikes in February, for example, caused a chain reaction of events, to include Operation Spring Shield by Turkey, protests against Russia in Istanbul, as well as the opening of Turkey's western border as retaliation for not receiving more support from the European Union for intervention in Syria [76, 77]. All these events are again indicators of conflict continuation and more reasons for refugee displacements. Another relevant factor was the spread of coronavirus in the region starting in February, with the first case confirmed by the Syrian government in March 2020 [78]. It is interesting how the beginning of this time frame had the highest anger percentage (February 2020) and then at the end had shifted entirely to the highest sadness percentage (April 2020). Over a span of three months, the viewpoint on Syrian refugees by its neighbours changed, possibly due to coronavirus, displacements and effects of conflicts, or a mixture of both. If this relatively quick shift is a reaction to the threat of coronavirus, it does bring forth the notion that the Levant region can consolidate in empathetic sentiment towards a greater outside threat. One implication of this notion is that when the region shares a perspective on an all-inclusive threat, it is possible for joint posture, and then potentially, cooperation.

The other intriguing time stamp is September 2019, the only point during the selected period where the joy sentiment overtook any other emotion. A major highlight of this month could be the proposal announcement by Turkey at the United Nations general assembly to create a safe zone and resettle millions of Syrians [79]. The proposal included the idea of providing financial assistance for aid in returning Syrian refugees from the neighboring countries to the safe zone, especially from Jordan and Lebanon. The sentiment could be a result of the prospect of the financial assistance and lessening the refugee strain on those countries, or simply due to the direction of proposed stability in northern Syria which would infer less forced migrations. In either case, a prospective economic gain or stability is the result and is possibly a factor for the joy sentiment.

The sentiments discussed thus far are of the selected countries as a whole in alignment with the events from the Syrian side. Comparisons with the individual countries sentiment and their internal events can point to any additional potential sentiment change indicators, or stark differences in sentiment from the majority. Checking for these differences can indicate which, if any, internal events might have also been a contributing factor to drastic change in sentiment towards Syrian refugees. Highlights of sentiments for each country and the internal events during the corresponding month are discussed in the subsequent paragraphs.

First, the major points are discussed for Israel. The joy sentiment percentage peaked in September 2019 for Israel. The main internal event from this month is the Israeli elections, the second round after a tie in April of the same year [80]. The following month in October 2019, this switched to a peak sentiment percentage of fear. Highlighted events include the failure of forming a stable governing body again, and the signs of a deepening political crisis post elections [81]. The fear sentiment percentage peaked again in March 2020, when the elections were held again for a third time [82]. During this month, there were nationwide protests against Benjamin Netanyahu demanding his resignation, mainly due to corruption scandals [83]. The anger sentiment peaked during the month of February 2020. The first case of coronavirus was confirmed in February 2020 in Israel [84]. While the sadness sentiment was relatively consistent throughout the entire period, it reached the lowest level during the month of February 2020, when anger peaked. Overall, the fear percentage sentiment superseded all the other sentiments for the entire period. There are no obvious major clashes in sentiment compared to the majority.

Second, the main points for Lebanon are summarized. In the samples from Lebanon, the fear sentiment peaked during the months of October 2019, December 2019, and February-March 2020. October 2019 witnessed the largest spontaneous national protests across Lebanon in years, and was the start of a political crisis in which its effects still manifest today [85, 86]. February 2020 was when the first coronavirus case was confirmed in Lebanon, and March is when the first death was recorded [84, 87]. There were no major or relevant internal events in December 2019, but fear and sadness sentiments peaked during this month, while anger and joy were at the lowest levels. A speculation could be that this is the sentiment of the end of the year with dire political and economic conditions causing fear and sadness for a bleak outlook on the state of the nation. Anger peaked during October 2019 and January 2020, the main internal events as a cause are possibly the national protests [85, 86, 88]. Fear was generally the largest sentiment percentage throughout the selected time period, and there were no evident disconnects from the majority region sentiment. The month of December 2019 provides an interesting insight. This month is when the conflict in Idlib escalated, leading to thousands of more Syrian refugees fleeing the violence [89]. This instance supports the assumption that sentiment change towards Syrian refugees is more likely driven by external factors or events rather than internal.

Third, the key internal aspects for Turkey are considered. The fear sentiment was high during the months of October 2019, and then February-May 2020. October 2019 was when Operation Peach Spring began, a Turkish offensive into north-eastern Syria [90]. It is also during this time that American troops were withdrawn from north-eastern Syria that had been supporting Kurdish allies [91]. As for February 2020, two major earthquakes in Turkey, as well as a deadly airstrike in Balyun that was possibly assisted by the Russian Air Force took place [92, 93]. Russia denied involvement with the incident, but protests against Russia still occurred the day following the strike at the Russian consulate in Istanbul [94]. The highest level of anger percentage sentiment occurred in February as well. In March, the first coronavirus case was confirmed in Turkey, bringing a series of governmental measures that were not eased until the end of May 2020 [95, 96]. Sadness was prevalent in June 2020. This month is when military operations against Iraq began (Operation Claw-Eagle and Operation Claw-Tiger) [97]. The joy sentiment only peaked during September 2019, which is when Erdogan's plan was announced at the UN General Assembly to settle Syrian refugees in north-eastern Syria [79]. Overall, there are no distinct contradictions in sentiment compared to the majority.

Lastly, for Jordan, there were no consistently discerning sentiment trends as with the other countries. Fear was high during the month of December 2019, but sadness was at the lowest level. The highest level of anger sentiment was in October 2019. Protests starting with a teacher's union demanding better pay had been ongoing during this month [98]. The highest fear percentage was in March 2020. In this month, Jordan had declared a state of emergency due to coronavirus [99]. Overall, there are no stark or evident instances of contradiction in sentiment compared to the majority.

While internal events might have some contribution to a drastic change in sentiment towards Syrian refugees, it is more likely that externally-related factors are the main drivers to producing inflection points for sentiment change. A general proposition is that the initial framing is likely a by-product of internal factors of the host nation, but then drastic sentiment changes after a solidified frame are caused by external factors. These external factors can vary in origination, such as foreign nations' policy changes and operations, pandemics, or the typical conflict clashes continued from the Syrian war.

4.5 Summary

In this chapter, news article samples from Israel, Jordan, Lebanon and Turkey were analyzed. The section began with results of an exploratory analysis, then frames analysis, and finally sentiment analysis. The insights obtained answered three research questions. The first addressed the existence of a dominant refugee frame utilized by the sample countries, the second investigated the sentiment of these countries, and the third examined potential factors behind drastic sentiment changes.

V. Conclusion

In this case study, the perspective of host countries neighboring Syria towards Syrian refugees was analyzed through news media articles via two components: framing and sentiment. The proof of concept is established in that perspective on a fixed target, in this case, Syrian refugees, can be analyzed through framing and sentiment. Framing is the context for which information is portrayed, and sentiment is the underlying tone relaying the information. From the sample of countries selected, Iraq was removed due to insufficient news article observations.

Exploratory analysis provided insights prior to addressing the research questions. The key items are the prevalence of themes such as religion or terrorism in the refugee articles' corpus based on country, as well as the large association of the United States to refugee news in the Levant area, possibly due to foreign policy or United Nations initiatives and financial contributions.

The frame and sentiment analyses addressed three specific questions. The first is to determine if there was a dominant refugee frame utilized by the sample countries. Based on the observations collected, it is indicative of a dominant frame, and that Syrian refugees were largely portrayed in a political frame for the June 2019 - June 2020 period.

The second is to determine sentiment similarity or dissimilarity within each country and across the countries. The analysis displayed similarity of sentiment across two groups of countries. Israel and Turkey formed one similar grouping, while Jordan and Lebanon formed the other. Moreover, the first group of Israel and Turkey tended to have more negative polarity in the news articles than Jordan and Lebanon. In addition, all the sample countries have variation in their news sources' sentiment with the exception of Jordan. This result is unsurprising, as news and media censorship is well known to be in Jordan.

The third is to determine periods of drastic change or intensity in sentiment, and potential factors that could have driven the change. For the Levant samples as a whole, fear was the prevalent sentiment, with the largest peaks at two time frames. The first is October 2019, and the second is February-April 2020. Possible factors for the first period include the withdrawal of U.S. troops in northern Syria, an indicator of a shift in U.S. foreign policy, as well as the death of ISIL leader Abu Bakr al-Baghdadi. As for the February-April 2020 period, there were two routes of events that could be potential factors. The first series of events are possibly caused by the Balyun airstrikes and are all indicators of conflict continuation and refugee displacements. The second route is the emergence of the coronavirus pandemic. A key observation is that the beginning of this time frame had the highest anger percentage (February 2020) and by the end had shifted entirely to the highest sadness percentage (April 2020). If this shift in sentiment is a reaction to the threat of coronavirus, it does bring forth the notion that the Levant region can consolidate in empathetic sentiment towards a greater outside threat. The implications of this notion is that when the region shares a perspective on an all-inclusive threat, it might be possible for joint posture, and then potentially, cooperation. Country-specific details on sentiment change over time can be found in chapter four.

A general proposition from this study is that the initial framing is likely a by-product of internal factors of the host nation, but then drastic sentiment changes after a solidified frame are caused by external factors. These external factors can vary in origination, such as foreign nations' policy changes and operations, pandemics, or the typical conflict clashes continued from the Syrian war.

As for potential applications and implications of this case study, they mainly involve policy, planning, and posture. Determining the framing of refugees for a country is a potential indicator for the possibility of refugee weaponization, especially for highly political framing. While Turkey has already conveyed actions of refugee weaponization on a higher level than the remaining Levant countries, the U.S. and allies cannot eliminate the high potential for these remaining countries to follow suite. Moreover, classifying countries and grouping them into sets of similarity may highlight intended postures or perspectives on fixed target issues, such as refugee weaponization. For a larger sample of countries, such as the entire Asian continent, these sets of similarity can provide a form of simplification by grouping to assist policy makers. Additionally, one of the region's peak sentiment changes is directly linked to U.S. operations and policy, indicating likely influence towards perspectives in the region prior to the withdrawal from Afghanistan.

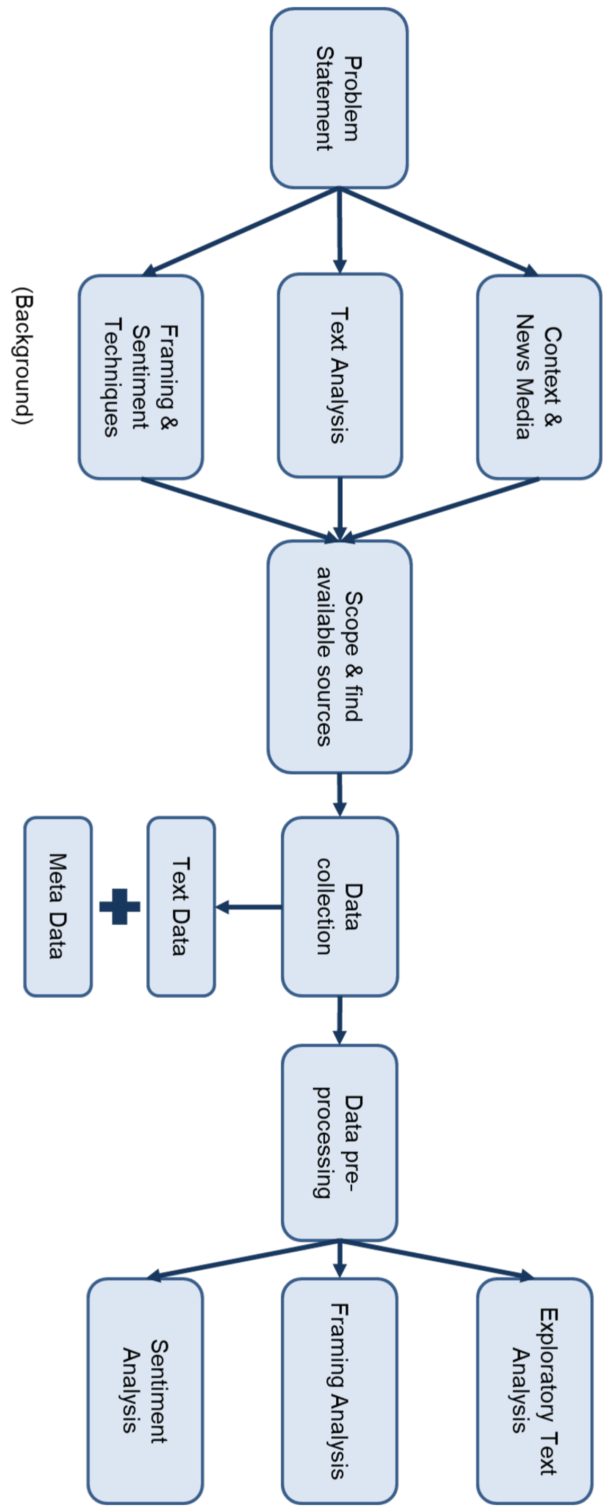
5.1 Potential Future Applications and Research

While the U.S. had large association to refugee news in the Levant during the selected time period of June 2019-June 2020, as well as likely influence via operations and policy, this might not be the case after the withdrawal from Afghanistan in August 2021. A potential study on the perspective of countries towards America specifically can be analyzed with the same methodology, and changing the fixed target to the U.S. rather than Syrian refugees. The route for potentially measuring indications of influence through framing and sentiment change over time can apply to a number of longitudinal study applications in the information operations field. It may also be a proxy measure for information campaigns.

Another possibility is to apply this methodology, with proper approvals, on a domestic front, specifically for South American countries to analyze the possibility of refugee weaponization or simply current foreign perspectives on migration. Such insights may assist the U.S. in addressing potential migration issues that affect the southern border. Moreover, as displayed in this study, external factors are also likely to change sentiment towards refugees in a host country. Identifying external factors that influence the host countries' perspectives can gauge external involvement, and in turn, potential affects on migration flows.

There are also many potential refinements to this study. For example, including all the languages of the Levant sample countries, such as Kurdish or Turkic languages, Hebrew, and others. In addition, including non-regional news sources might provide a basis for comparison with the local sources. Additionally, expanding to include the entire Middle East, or potentially Europe or Asia can provide a larger picture on the relations of countries to one another regarding refugees in general.

Appendix A. Overview Diagram of Study Process



Appendix B. Additional Details

This appendix provides a) the breakdown of the article samples by news source and b) additional details to include the lower/upper fences, bottom/top quartiles, as well as the median for some of the figures in the document.

Table 8. Breakdown of Article Samples

	News Source	Israel	Jordan	Lebanon	Turkey
1	Haaretz	68	-	-	-
2	Jerusalem Post	244	-	-	-
3	Amman Net	-	8	-	-
4	Ammon News	-	8	-	-
5	Petra	-	1	-	-
6	Jordan Times	-	24	-	-
7	Roya News	-	2	-	-
8	Al Manar	-	-	16	-
9	Al Nahar	-	-	20	-
10	NNA	-	-	120	-
11	The Daily Star	-	-	60	-
12	Daily Sabah	-	-	-	663
13	Hurriyet	-	-	-	163
	Total	312	43	216	826

Table 9. Details for Figure 7: Standardized Sentiment Scores by Country – Bing Dictionary

	Entity	LowerFence	BottomQuartile	Median	TopQuartile	UpperFence
1	Jordan	-0.23	0.49	0.74	0.98	1.65
2	Lebanon	-0.11	0.37	0.62	0.74	1.28
3	Turkey	-2.29	-0.60	0.01	0.56	2.25
4	Israel	-2.54	-0.72	0.01	0.51	2.31

Table 10. Details for Figure 8: Standardized Sentiment Scores by Country – Afinn Dictionary

	Entity	LowerFence	BottomQuartile	Median	TopQuartile	UpperFence
1	Jordan	-0.07	0.46	0.66	1.06	1.82
2	Lebanon	-0.36	0.26	0.56	0.73	1.42
3	Israel	-2.58	-0.73	0.03	0.53	1.79
4	Turkey	-2.32	-0.60	0.02	0.56	2.09

Table 11. Details for Figure 9: Israel - Standardized Sentiment Scores by News Source (Afinn)

	Entity	LowerFence	BottomQuartile	Median	TopQuartile	UpperFence
1	Jerusalem Post	-2.32	-0.63	0.10	0.53	1.79
2	Haaretz	-2.58	-1.02	-0.43	0.31	1.52

Table 12. Details for Figure 9: Israel - Standardized Sentiment Scores by News Source (Bing)

	Entity	LowerFence	BottomQuartile	Median	TopQuartile	UpperFence
1	Jerusalem Post	-2.05	-0.61	0.13	0.49	1.95
2	Haaretz	-3.20	-1.04	-0.29	0.56	2.49

Table 13. Details for Figure 10: Jordan - Standardized Sentiment Scores by News Source (Afinn)

	Entity	LowerFence	BottomQuartile	Median	TopQuartile	UpperFence
1	Amman Net	-1.56	0.11	1.09	1.56	2.52
2	Roya News	0.66	0.79	0.91	1.03	1.16
3	Petra	0.83	0.83	0.83	0.83	0.83
4	Ammon News	0.50	0.63	0.66	0.84	0.99
5	Jordan Times	-0.07	0.42	0.63	0.82	1.32

Table 14. Details for Figure 10: Jordan - Standardized Sentiment Scores by News Source (Bing)

	Entity	LowerFence	BottomQuartile	Median	TopQuartile	UpperFence
1	Petra	0.86	0.86	0.86	0.86	0.86
2	Jordan Times	-0.23	0.48	0.86	0.98	1.65
3	Amman Net	-0.48	0.02	0.71	1.72	2.49
4	Ammon News	0.49	0.59	0.65	0.71	0.80
5	Roya News	0.62	0.63	0.65	0.66	0.68

Table 15. Details for Figure 11: Lebanon - Standardized Sentiment Scores by News Source (Afinn)

	Entity	LowerFence	BottomQuartile	Median	TopQuartile	UpperFence
1	NNA	-0.07	0.49	0.66	0.86	1.36
2	Daily Star	-0.83	0.00	0.45	0.60	1.16
3	Al Manar	0.07	0.22	0.35	0.68	0.99
4	Al Nahar	-1.62	-0.73	0.15	0.55	1.62

Table 16. Details for Figure 11: Lebanon - Standardized Sentiment Scores by News Source (Bing)

	Entity	LowerFence	BottomQuartile	Median	TopQuartile	UpperFence
1	NNA	0.01	0.49	0.68	0.86	1.34
2	Al Manar	-0.05	0.42	0.56	0.74	1.04
3	Daily Star	-0.90	0.01	0.56	0.63	0.98
4	Al Nahar	-1.69	-0.63	-0.14	0.63	1.52

Table 17. Details for Figure 12: Turkey - Standardized Sentiment Scores by News Source (Afinn)

	Entity	LowerFence	BottomQuartile	Median	TopQuartile	UpperFence
1	Hurriyet	-1.03	-0.20	0.23	0.58	1.42
2	Daily Sabah	-2.38	-0.66	-0.07	0.53	2.09

Table 18. Details for Figure 12: Turkey - Standardized Sentiment Scores by News Source (Bing)

	Entity	LowerFence	BottomQuartile	Median	TopQuartile	UpperFence
1	Hurriyet	-1.08	-0.11	0.19	0.56	1.52
2	Daily Sabah	-2.41	-0.72	0.01	0.49	2.31

Appendix C. List of Software Packages

This appendix provides a list of packages utilized. All packages were set at default settings unless otherwise specified.

3.1 Data Manipulation and Pre-processing Packages

- base R [59]
- readtext [100]
- stringr [101]
- tidyverse (includes ggplot2) [102]
- reshape2 [103]
- lubridate [104]
- quanteda [105]
  - Minimum term frequency of 50, minimum document frequency of 10.
  - See *Methodology* chapter, *Data Pre-processing* section for further details on corpus cleaning

3.2 Framing and Sentiment Techniques, Statistical Testing Packages

- stm [58] for topic modeling
  - Parameter of topics K tested in range from 5 to 40
- Multiple packages for sentiment analysis:
  - sentimentr [60]

  - Bing Lexicon [43]
  - Afinn Lexicon [42]
  - NRC Lexicon [44]
- FSA [72] for statistical testing
  - α level of 0.05

3.3 Document and Formatting Packages

- xtable [106]
- cowplot [107]

Bibliography

1. "Syria Emergency," United Nations High Commissioner for Refugees. [Online]. Available: https://www.unhcr.org/en-us/syria-emergency.html [Accessed: Feb. 17, 2022]

2. The United Nations High Commissioner for Refugees Operational Data Portal, "Syria Regional Refugee Response." [Online]. Available: https://data2.unhcr.org/en/situations/syria [Accessed: Feb. 17, 2022]

3. Department of State, "Functional Bureau Strategy: Bureau of Conflict and Stabilization Operations (CSO)," Under Secretary for Civilian Security, Democracy, and Human Rights, Washington, DC, USA, 2020. [Online]. Available: https://www.state.gov/wp-content/uploads/2020/05/CSO-UNCLASS-FBS-20200504-508.pdf [Accessed: Apr. 20, 2021]

4. H. A. Sevilla Jr, "Middle east geopolitics and china-india strategic interaction in the new era," *Asian Journal of Middle Eastern and Islamic Studies*, vol. 14, no. 2, pp. 179–193, 2020. [Online]. doi: 10.1080/25765949.2020.1760541.

5. S. Colombo and E. S. i Lecha, "Europe and the 'new' middle east," *Journal of Balkan and Near Eastern Studies*, vol. 23, no. 3, pp. 403–422, 2021. [Online]. doi: 10.1080/19448953.2021.1888246.

6. D. D. Kaye, L. Robinson, J. Martini, N. Vest, and A. L. Rhoades, "Reimagining U.S. Strategy in the Middle East: Sustainable Partnerships, Strategic Investments (Executive Summary)," RAND Corp., Santa Monica, CA, USA, RR-A958-2, 2021. [Online]. doi: 10.7249/PE265.

7. A. Rožukalne, S. Kruks, A. Stakle, and I. Skulte, “Representation of migration in Latvian mass media (2015-2016): Deny voice to the voiceless,” *Informacijos Mokslai*, vol. 87, pp. 13–35, 2020.

8. United Nations High Commimissioner for Refugees (UNHCR), “UNHCR viewpoint: ‘Refugee’ or ‘migrant’ – Which is right?” [Online]. Available: https://www.unhcr.org/news/latest/2016/7/55df0e556/unhcr-viewpoint-refugee-migrant-right.html [Accessed: May 2, 2021]

9. J. Nestorowicz and M. Anacka, “Mind the Gap? Quantifying Interlinkages between Two Traditions in Migration Literature,” *International Migration Review*, vol. 53, no. 1, pp. 283–307, Mar. 2019. [Online]. doi: 10.1177/0197918318768557.

10. A. Pisarevskaya, N. Levy, P. Scholten, and J. Jansen, “Mapping migration studies: An empirical analysis of the coming of age of a research field,” *Migration Studies*, vol. 8, no. 3, pp. 455–481, 2020. [Online]. doi: 10.1093/migration/mnz031.

11. Operational Data Portal, “Mediterranean Situation,” Operational Data Portal: Refugee Situations. [Online]. Available: http://data2.unhcr.org/en/situations/mediterranean?page=1&view=grid&Type%255B%255D=3&Search=%2523monthly%2523%7cjournal=Refugees/Migrants [Accessed: May 27, 2021]

12. U. Sunata and E. Yıldız, “Representation of syrian refugees in the turkish media,” *Journal of Applied Journalism & Media Studies*, vol. 7, no. 1, pp. 129–151, 2018. [Online]. doi: 10.1386/ajms.7.1.129_1.

13. N. Narlı and M. Özaşçılar, “Representation of syrian women and children refugees’ health in turkish daily newspapers,” *Journal of International Migration and Integration*, pp. 1–15, 2019. [Online]. doi: 10.1007/s12134-019-00732-6.

14. D. Onay-Coker, “The representation of syrian refugees in turkey: a critical discourse analysis of three newspapers,” *Continuum*, vol. 33, no. 3, pp. 369–385, 2019. [Online]. doi: 10.1080/10304312.2019.1587740.

15. A. S. Haider, S. S. Olimy, and L. S. Al-Abbas, “Media coverage of syrian female refugees in jordan and lebanon,” *SAGE Open*, vol. 11, no. 1, pp. 1–21, 2021. [Online]. doi: 10.1177/2158244021994811.

16. European Commission, “Communication from the Commission to the European Parliament, the European Council and the Council: Progress Report on the Implementation of the European Agenda on Migration,” European Commission, Brussels, Belgium, 2019. [Online]. Available: https://ec.europa.eu/home-affairs/sites/default/files/what-we-do/policies/european-agenda-migration/20191016_com-2019-481-report_en.pdf [Accessed: Apr. 12, 2021]

17. A. Righi, “Assessing migration through social media: a review,” *Mathematical Population Studies*, vol. 26, no. 2, pp. 80–91, 2019.

18. M. Gallego, E. Gualda, and C. Rebollo, “Women and refugees in twitter: Rhetorics on abuse, vulnerability and violence from a gender perspective,” *Journal of Mediterranean knowledge*, vol. 2, no. 1, pp. 37–58, 2017. [Online]. doi: 10.26409/2017JMK2.1.03.

19. N. Öztürk and S. Ayvaz, "Sentiment analysis on twitter: A text mining approach to the syrian refugee crisis," *Telematics and Informatics*, vol. 35, no. 1, pp. 136–147, 2018. [Online]. doi: 10.1016/j.tele.2017.10.006.

20. R. Kreis, "#refugeesnotwelcome: Anti-refugee discourse on twitter," *Discourse & Communication*, vol. 11, no. 5, pp. 498–514, 2017. [Online]. doi: 10.1177/1750481317714121.

21. C. Arcila-Calderón, D. Blanco-Herrero, M. Frías-Vázquez, and F. Seoane, "Refugees welcome? online hate speech and sentiments in twitter in spain during the reception of the boat aquarius," *Sustainability*, vol. 13, no. 5, p. 2728, 2021. [Online]. doi: 10.3390/su13052728.

22. R. M. Entman, "Framing: Toward clarification of a fractured paradigm," *Journal of Communication*, vol. 43, no. 4, pp. 51–58, Dec. 1993. [Online]. doi: 10.1111/j.1460-2466.1993.tb01304.

23. T. Heidenreich, F. Lind, J.-M. Eberl, and H. G. Boomgaarden, "Media Framing Dynamics of the 'European Refugee Crisis': A Comparative Topic Modelling Approach," *Journal of Refugee Studies*, vol. 32, no. 1, pp. i172–i182, 12 2019.

24. X. Zhang and L. Hellmueller, "Visual framing of the european refugee crisis in der spiegel and cnn international: Global journalism in news photographs," *International Communication Gazette*, vol. 79, no. 5, pp. 483–510, 2017.

25. A. Lawlor and E. Tolley, "Deciding who's legitimate: News media framing of immigrants and refugees," *International Journal of Communication*, vol. 11, p. 25, 2017.

26. N. Corbu, R. Buturoiu, and F. Durach, "Framing the refugee crisis in online media: A romanian perspective," *Romanian Journal of Communication and Public Relations*, vol. 19, no. 2, pp. 5–18, 2017.

27. "Media Ownership Monitor: Lebanon," Samir Kassir Foundation. [Online]. Available: https://lebanon.mom-rsf.org/en/media/ [Accessed: May 25, 2021]

28. Department of State, "2019 Country Reports on Human Rights Practices: Jordan," Bureau of Democracy, Human Rights, and Labor, Washington, DC, USA, 2019. [Online]. Available: https://www.state.gov/wp-content/uploads/2020/02/JORDAN-2019-HUMAN-RIGHTS-REPORT.pdf [Accessed: May 22, 2021]

29. G. B. Coşkun, "Media capture strategies in new authoritarian states: the case of turkey," *Publizistik*, vol. 65, no. 4, pp. 637–654, Sep. 2020. [Online]. doi: 10.1007/s11616-020-00600-9.

30. E. Isaeva and D. Aldarova, "Text-mining in terms of methodology and development," in *2021 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus)*, 2021. [Online]. doi: 10.1109/ElConRus51938.2021.9396437, pp. 413–416.

31. M. Schonlau, N. Guenther, and I. Sucholutsky, "Text mining with n-gram variables," *The Stata Journal*, vol. 17, no. 4, pp. 866–881, 2017. [Online]. doi: 10.1177/1536867X1801700406.

32. V. Fromkin, B. Hayes, S. Curtiss, A. Szabolcsi, T. Stowell, E. Stabler, D. Sportiche, H. Koopman, P. Keating, P. Munro *et al.*, *Linguistics: An Introduction to Linguistic Theory.* Wiley, 2001.

33. Z. Jiang, B. Gao, Y. He, Y. Han, P. Doyle, and Q. Zhu, "Text classification using novel term weighting scheme-based improved tf-idf for internet media reports." *Mathematical Problems in Engineering*, pp. 1 – 30, 2021. [Online]. doi: 10.1155/2021/6619088.

34. J. Matthes, "What's in a frame? a content analysis of media framing studies in the world's leading communication journals, 1990-2005," *Journalism & Mass Communication Quarterly*, vol. 86, no. 2, pp. 349–367, 2009. [Online]. doi: 10.1177/107769900908600206.

35. V. Nasteski, "An overview of the supervised machine learning methods," *Horizons*, vol. 4, pp. 51–62, 2017. [Online]. doi: 10.20544/HORIZONS.B.04.1.17.P05.

36. I. Vayansky and S. A. Kumar, "A review of topic modeling methods," *Information Systems*, vol. 94, p. 101582, 2020. [Online]. doi: 10.1016/j.is.2020.101582.

37. D. M. Blei, "Probabilistic topic models," *Communications of the ACM*, vol. 55, no. 4, p. 77–84, 2012. [Online]. doi: 10.1145/2133806.2133826.

38. D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet Allocation," *The Journal of Machine Learning Research*, vol. 3, pp. 993–1022, 2003. [Online]. doi: 10.5555/944919.944937.

39. D. M. Blei and J. Lafferty, "Correlated topic models," *Advances in neural information processing systems*, vol. 18, pp. 147–155, 2006.

40. M. E. Roberts, B. M. Stewart, and D. Tingley, "stm: An R package for structural topic models," *Journal of Statistical Software*, vol. 91, no. 2, pp. 1–40, 2019.

41. R. Carver, E. Rødland, and J. Breivik, "Quantitative frame analysis of how the gene concept is presented in tabloid and elite newspapers," *Science Communication*, vol. 35, pp. 449–475, 2013.

42. F. Äruprup Nielsen, “A new anew: Evaluation of a word list for sentiment analysis in microblogs,” *CoRR*, vol. abs/1103.2903, 2011. [Online]. Available: http://arxiv.org/abs/1103.2903 [Accessed: Sep. 1, 2021]

43. M. Hu and B. Liu, “Mining and summarizing customer reviews,” in *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD ’04. New York, NY, USA: ACM, 2004, p. 168–177. [Online]. Available: http://doi.acm.org/10.1145/1014052.1014073 [Accessed: Sep. 1, 2021]

44. S. M. Mohammad and P. D. Turney, “Crowdsourcing a word-emotion association lexicon,” *Computational Intelligence*, vol. 29, no. 3, pp. 436–465, 2013. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-8640.2012.00460.x [Accessed: Sep. 1, 2021]

45. E. Haddi, X. Liu, and Y. Shi, “The role of text pre-processing in sentiment analysis,” *Procedia Computer Science*, vol. 17, pp. 26–32, 2013. [Online]. doi: 10.1016/j.procs.2013.05.005, first International Conference on Information Technology and Quantitative Management.

46. I. Feinerer, K. Hornik, and D. Meyer, “Text mining infrastructure in r,” *Journal of Statistical Software, Articles*, vol. 25, no. 5, pp. 1–54, 2008. [Online]. doi: 10.18637/jss.v025.i05.

47. C. Manning, P. Raghavan, and H. Schütze, *Introduction to Information Retrieval*. Cambridge University Press, 2008.

48. L. Polanyi and A. Zaenen, *Contextual Valence Shifters*. Dordrecht: Springer Netherlands, 2006, pp. 1–10.

49. D. M. E.-D. M. Hussein, "A survey on sentiment analysis challenges," *Journal of King Saud University - Engineering Sciences*, vol. 30, no. 4, pp. 330–338, 2018. [Online]. doi: 10.1016/j.jksues.2016.04.002.

50. Foreign Policy Research Institute, "TIMELINE OF THE SYRIAN WAR," FPRI, Philadelphia, PA, USA, 2020. [Online]. Available: https://www.fpri.org/wp-content/uploads/2020/09/timeline.pdf [Accessed: Mar. 22, 2021]

51. "Daily Sabah," European Press Roundup. [Online]. Available: https://www.eurotopics.net/en/156478/daily-sabah [Accessed: Apr. 28, 2021]

52. "17 Mayıs-23 Mayıs haftası Tiraj Raporu," Medya Radar. [Online]. Available: https://www.medyaradar.com/tirajlar [Accessed: May 24, 2021]

53. R. Tal, "The Israeli Press," Israel Ministry of Foreign Affairs. [Online]. Available: https://mfa.gov.il/MFA/MFA-Archive/1998/Pages/The%20Israeli%20Press.aspx [Accessed: Apr. 4, 2021]

54. "Lebanon profile - Media," BBC News. [Online]. Available: https://www.bbc.com/news/world-middle-east-14648683 [Accessed: Apr. 5, 2021]

55. "Iraq profile - Media," BBC News. [Online]. Available: https://www.bbc.com/news/world-middle-east-14546541 [Accessed: Apr. 5, 2021]

56. "Jordan profile - Media," BBC News. [Online]. Available: https://www.bbc.com/news/world-middle-east-14636310 [Accessed: Apr. 5, 2021]

57. "World Newspapers and Magazines: Jordan," Worldpress.org. [Online]. Available: https://www.worldpress.org/newspapers/MIDEAST/Jordan.cfm [Accessed: Apr. 5, 2021]

58. M. E. Roberts, B. M. Stewart, D. Tingley, E. M. Airoldi *et al.*, "The structural topic model and applied social science," in *Advances in neural information processing systems workshop on topic models: computation, application, and evaluation*, vol. 4. Harrahs and Harveys, Lake Tahoe, 2013, pp. 1–20.

59. R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2021. [Online]. Available: https://www.R-project.org/ [Accessed: Mar. 2, 2021]

60. T. W. Rinker, *sentimentr: Calculate Text Polarity Sentiment*, Buffalo, New York, 2019, version 2.7.1. [Online]. Available: http://github.com/trinker/sentimentr [Accessed: Sep. 1, 2021]

61. B. K. Nayak and A. Hazra, "How to choose the right statistical test?" *Indian journal of ophthalmology*, vol. 59, no. 2, p. 85, 2011.

62. Department of State, "2020 Country Reports on Human Rights Practices: Jordan," Bureau of Democracy, Human Rights, and Labor, Washington, DC, USA, 2020. [Online]. Available: https://www.state.gov/wp-content/uploads/2021/10/JORDAN-2020-HUMAN-RIGHTS-REPORT.pdf [Accessed: May 22, 2021]

63. World Health Organization, "Coronavirus disease 2019 (COVID-19) Weekly Situation Report 05," Regional Office for the Eastern Mediterranean, Cairo, Egypt, Tech. Rep. 5, 2020. [Online]. Available: http://www.emro.who.int/images/stories/coronavirus/covid-19-sitrep-5.pdf?ua=1 [Accessed: Oct. 1, 2021]

64. BBC News. (2020, March) Syria war: Russia and Turkey agree Idlib ceasefire. [Online]. Available: https://www.bbc.com/news/world-middle-east-51747592 [Accessed: Jan. 6, 2021]

65. N. D. Steger, "The weaponization of migration: examining migration as a 21st century tool of political warfare," NAVAL POSTGRADUATE SCHOOL MONTEREY CA MONTEREY, United States, Tech. Rep., 2017.

66. G. P. Breedlove, "United States Senate Committee on Armed Services: HEARING TO RECEIVE TESTIMONY ON UNITED STATES EUROPEAN COMMAND," Washington, D.C., March 2016. [Online]. Available: https://www.armed-services.senate.gov/imo/media/doc/16-20_03-01-16.pdf [Accessed: Oct. 1, 2021]

67. Washington Post, "The E.U. is furious that Belarus allowed more than 4,000 migrants to cross into Europe." [Online]. Available: https://www.washingtonpost.com/politics/2021/08/13/eu-is-furious-that-belarus-allowed-more-than-4000-migrants-cross-into-europe/ [Accessed: Oct. 14, 2021]

68. Reuters, "Turkey's Erdogan threatened to flood Europe with migrants." [Online]. Available: https://www.reuters.com/article/us-europe-migrants-eu-turkey/turkeys-erdogan-threatened-to-flood-europe-with-migrants-greek-website-idUSKCN0VH1R0 [Accessed: Oct. 14, 2021]

69. Lowy Institute, "Global Diplomacy Index: 2019 Country Ranking," 2019. [Online]. Available: https://globaldiplomacyindex.lowyinstitute.org/country_rank.html [Accessed: Oct. 14, 2021]

70. The World Bank, "Data for Turkey, Lebanon, Jordan, Israel: GDP." [Online]. Available: https://data.worldbank.org/?locations=TR-LB-JO-IL [Accessed: Oct. 14, 2021]

71. The World Bank, "Military expenditure (current USD) - Turkey, Lebanon, Jordan, Israel." [Online]. Available: https://data.worldbank.org/indicator/MS.MIL.XPND.CD?end=2020&locations=TR-LB-JO-IL&start=2019 [Accessed: Oct. 14, 2021]

72. D. H. Ogle, J. C. Doll, P. Wheeler, and A. Dinno, *FSA: Fisheries Stock Analysis*, 2021, R Package version 0.9.1 Non-Parametric Testing. [Online]. Available: https://github.com/droglenc/FSA [Accessed: Sep. 12, 2021]

73. A. AJI and O. AKOUR, "Syria reopens vital crossing with jordan, un post with golan," Oct 2018. [Online]. Available: https://apnews.com/article/5a3b44a592f94bfdb9700fc810044a06 [Accessed: Oct. 14, 2021]

74. BBC News. Turkey-Syria border: Kurds bitter as US troops withdraw. [Online]. Available: https://www.bbc.com/news/world-middle-east-49960973 [Accessed: Oct. 18, 2021]

75. BBC News. Abu Bakr al-Baghdadi: IS leader 'dead after US raid' in Syria. [Online]. Available: https://www.bbc.com/news/world-us-canada-50200339 [Accessed: Oct. 18, 2021]

76. BBC News. Syria war: Alarm after 33 Turkish soldiers killed in attack in Idlib. [Online]. Available: https://www.bbc.com/news/world-middle-east-51667717 [Accessed: Oct. 18, 2021]

77. Daniel Boffey. Clashes as thousands gather at Turkish border to enter Greece. The Guardian. [Online]. Available: https://www.theguardian.com/

world/2020/mar/01/thousands-gather-at-turkish-border-to-cross-into-greece [Accessed: Oct. 18, 2021]

78. Suleiman Al-Khalidi. Syria confirms first coronavirus case as fears grow it could spread. Reuters. [Online]. Available: https://www.reuters.com/article/us-health-coronavirus-syria/syria-confirms-first-coronavirus-case-as-fears-grow-it-could-spread-idUSKBN21912A [Accessed: Oct. 18, 2021]

79. Patrick Wintour. Recep Tayyip Erdoğan proposes 'safe zone' for refugees in Syria. The Guardian. [Online]. Available: https://www.theguardian.com/world/2019/sep/24/erdogan-proposes-plan-for-refugee-safe-zone-in-syria [Accessed: Oct. 18, 2021]

80. Daniel Estrin and Orr Hirschauge. (2019, September) Netanyahu Fights To Hang On In Another Israeli Election. Here's What To Know. NPR. [Online]. Available: https://www.npr.org/2019/09/16/760464318/netanyahu-fights-to-hang-on-in-another-israeli-election-heres-what-to-know [Accessed: Nov. 17, 2021]

81. BBC News. (2019, October) Israel PM Netanyahu fails to form government ahead of deadline. [Online]. Available: https://www.bbc.com/news/world-middle-east-50132760 [Accessed: Nov. 17, 2021]

82. BBC News. (2020, March) Israelis vote in unprecedented third general election in a year. [Online]. Available: https://www.bbc.com/news/world-middle-east-51612360 [Accessed: Nov. 17, 2021]

83. Reuters. (2020, March) Thousands of Israelis protest against Netanyahu ahead of election. [Online]. Available: https://www.reuters.com/article/uk-israel-

election-netanyahu-protests/thousands-of-israelis-protest-against-netanyahu-ahead-of-election-idUSKBN2BC0PJ [Accessed: Nov. 17, 2021]

84. Virginia Pietromarchi. (2020, February) Coronavirus in Middle East: What you need to know. Al-Jazeera. [Online]. Available: https://www.aljazeera.com/news/2020/2/25/coronavirus-in-middle-east-what-you-need-to-know [Accessed: Nov. 17, 2021]

85. Vivian Yee. (2019, October) Lebanon Protests Unite Sects in Demanding New Government. The New York Times. [Online]. Available: https://www.nytimes.com/2019/10/23/world/middleeast/lebanon-protests.html [Accessed: Nov. 17, 2021]

86. Al-Jazeera. (2019, October) Lebanon protests: Thousands demand 'fall of the regime' in Beirut. [Online]. Available: https://www.aljazeera.com/economy/2019/10/18/lebanon-protests-thousands-demand-fall-of-the-regime-in-beirut [Accessed: Nov. 17, 2021]

87. Reuters. (2020, March) Lebanon records first death from coronavirus. [Online]. Available: https://www.reuters.com/article/us-health-coronavirus-lebanon/lebanon-records-first-death-from-coronavirus-idUSKBN20X16R [Accessed: Nov. 17, 2021]

88. Reuters. (2020, January) Lebanese security forces, protesters clash for second night. [Online]. Available: https://www.reuters.com/article/us-lebanon-crisis-protests/lebanese-security-forces-protesters-clash-for-second-night-idUSKBN1ZI0DS [Accessed: Nov. 17, 2021]

89. Vivian Yee and Hwaida Saad. (2019, December) Syrian Offensive Sends Tens of Thousands Fleeing. The New York Times. [Online]. Available-

able: https://www.nytimes.com/2019/12/23/world/middleeast/syria-idlib-russia-aid-refugees.html [Accessed: Nov. 17, 2021]

90. BBC News. (2019, October) Turkey launches ground offensive in northern Syria. [Online]. Available: https://www.bbc.com/news/world-middle-east-49983357 [Accessed: Nov. 17, 2021]

91. BBC News. (2019, October) Turkey-Syria border: Kurds bitter as US troops withdraw. [Online]. Available: https://www.bbc.com/news/world-middle-east-49960973 [Accessed: Nov. 17, 2021]

92. BBC News. (2020, February) Syria war: Alarm after 33 Turkish soldiers killed in attack in Idlib. [Online]. Available: https://www.bbc.com/news/world-middle-east-51667717 [Accessed: Nov. 17, 2021]

93. BBC News. (2020, February) Earthquake kills at least 9 in Turkey, injures many in Iran. [Online]. Available: https://www.bbc.com/news/world-europe-51603400 [Accessed: Nov. 17, 2021]

94. Carlotta Gall. (2020, February) Protesters in Turkey denounce Russia over Idlib assault. The New York Times. [Online]. Available: https://www.nytimes.com/2020/02/27/world/middleeast/russia-turkey-syria-war-strikes.html [Accessed: Nov. 17, 2021]

95. Reuters. (2020, April) Turkey has most coronavirus cases outside Europe and U.S. [Online]. Available: https://www.reuters.com/article/us-health-coronavirus-turkey/turkey-has-most-coronavirus-cases-outside-europe-and-u-s-idUSKBN2210P0 [Accessed: Nov. 17, 2021]

96. Reuters. (2020, May) Turkey COVID cases below 20,000 for first time since mid-March. [Online]. Available: https://www.reuters.com/world/middle-

east/turkey-covid-cases-below-20000-first-time-since-mid-march-2021-05-08/ [Accessed: Nov. 17, 2021]

97. Al-Jazeera. (2020, June) Turkey sends special forces into northern Iraq. [Online]. Available: https://www.aljazeera.com/news/2020/6/17/turkey-sends-special-forces-into-northern-iraq [Accessed: Nov. 17, 2021]

98. Al-Jazeera. (2019, October) Jordan teachers end four-week strike in pay deal with government. [Online]. Available: https://www.aljazeera.com/news/2019/10/6/jordan-teachers-end-four-week-strike-in-pay-deal-with-government [Accessed: Nov. 17, 2021]

99. Ali Younes. (2020, March) Jordan imposes state of emergency to curb coronavirus pandemic. Al-Jazeera. [Online]. Available: https://www.aljazeera.com/news/2020/3/17/jordan-imposes-state-of-emergency-to-curb-coronavirus-pandemic [Accessed: Nov. 17, 2021]

100. K. Benoit and A. Obeng, *readtext: Import and Handling for Plain and Formatted Text Files*, 2021, r package version 0.81. [Online]. Available: https://CRAN.R-project.org/package=readtext [Accessed: Jan. 17, 2022]

101. H. Wickham, *stringr: Simple, Consistent Wrappers for Common String Operations*, 2019, r package version 1.4.0. [Online]. Available: https://CRAN.R-project.org/package=stringr [Accessed: Jan. 17, 2022]

102. H. Wickham, M. Averick, J. Bryan, W. Chang, L. D. McGowan, R. François, G. Grolemund, A. Hayes, L. Henry, J. Hester, M. Kuhn, T. L. Pedersen, E. Miller, S. M. Bache, K. Müller, J. Ooms, D. Robinson, D. P. Seidel, V. Spinu, K. Takahashi, D. Vaughan, C. Wilke, K. Woo, and H. Yutani, "Welcome to the tidyverse," *Journal of Open Source Software*, vol. 4, no. 43, p. 1686, 2019.

103. H. Wickham, “Reshaping data with the reshape package,” *Journal of Statistical Software*, vol. 21, no. 12, pp. 1–20, 2007. [Online]. Available: http://www.jstatsoft.org/v21/i12/ [Accessed: Jan. 17, 2022]

104. G. Grolemund and H. Wickham, “Dates and times made easy with lubridate,” *Journal of Statistical Software*, vol. 40, no. 3, pp. 1–25, 2011. [Online]. Available: https://www.jstatsoft.org/v40/i03/ [Accessed: Jan. 17, 2022]

105. K. Benoit, K. Watanabe, H. Wang, P. Nulty, A. Obeng, S. Müller, and A. Matsuo, “quanteda: An r package for the quantitative analysis of textual data,” *Journal of Open Source Software*, vol. 3, no. 30, p. 774, 2018. [Online]. Available: https://quanteda.io [Accessed: Jan. 17, 2022]

106. D. B. Dahl, D. Scott, C. Roosen, A. Magnusson, and J. Swinton, *xtable: Export Tables to LaTeX or HTML*, 2019, r package version 1.8-4. [Online]. Available: https://CRAN.R-project.org/package=xtable [Accessed: Jan. 17, 2022]

107. C. O. Wilke, *cowplot: Streamlined Plot Theme and Plot Annotations for 'ggplot2'*, 2020, r package version 1.1.1. [Online]. Available: https://CRAN.R-project.org/package=cowplot [Accessed: Jan. 17, 2022]

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704–0188

1. REPORT DATE *(DD–MM–YYYY)*
24–03–2022

2. REPORT TYPE
Master's Thesis

3. DATES COVERED *(From — To)*
Sept 2020 — Mar 2022

4. TITLE AND SUBTITLE
Analysis of the Perspective on Syrian Refugees by Neighboring Countries

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Ghanem, Norma, 1st Lt, USAF

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Air Force Institute of Technology
Graduate School of Engineering and Management (AFIT/EN)
2950 Hobson Way
Wright-Patterson AFB OH 45433-7765

8. PERFORMING ORGANIZATION REPORT NUMBER
AFIT-ENS-MS-22-M-130

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Intentionally Left Blank

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Distribution Statement A. Approved for Public Release; Distribution Unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Mass migration destabilizes neighboring states, opening the way for fragile state exploitation by enemies, including those that could undermine U.S. national interests. This study investigates an area of the Levant region, specifically countries neighboring Syria, and analyzes their perspective on Syrian refugees for the time frame of 1 June 2019 - 30 June 2020. This analysis may assist in forming policy and creating strategies to address refugee related issues, both domestic and international. There are three main questions addressed. The first inspects dominant refugee framing, the second explores sentiment (dis)similarity within each country and across countries, and the third investigates drastic sentiment changes and their potential driving factors. Results show that Syrian refugees are dominantly viewed in a political framing. Sentiment similarities are shared across two distinguished groupings, Turkey and Israel in one group and Jordan and Lebanon in the other. External factors that vary in origination are likely to influence high sentiment changes. Future research includes application on South American countries, or a change in topic to analyze perspective towards America post Afghanistan withdrawal.

15. SUBJECT TERMS
Information operations, refugee migration, sentiment analysis, framing analysis, perspective analysis

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON / 19b. TELEPHONE NUMBER *(include area code)*
U	U	U	UU	84	Dr. Richard F. Deckro, AFIT/ENS; (937) 255-3636 x4325; richard.deckro@afit.edu