Analysis of Over 1,600 Public Tweets
from February and March 2019
on the Military Housing Problem

by Dr. Lynn Marie Houston

Introduction:

Issues with Army housing gained national attention on December 27, 2018 with the Reuter's publication of an investigative report revealing unsafe and unsanitary living conditions of families living on military installations in the United States. Most of these apartments and houses were owned and managed by Corvias, a housing company owned by John Picerne, a resident of Rhode Island.

The reporters who broke the story—which reported specifically on housing conditions at Fort Bragg, North Carolina, and Fort Meade, Maryland—were initially alerted to the situation in October 2018 because of a widely circulated online petition begun by a military wife and mother. Her petition began in August of that year (posted at change.org), and followed numerous complaints by her and others in online forums, including Twitter.

Because users originally shared information about poor housing via Twitter, and because the issue gained attention at the highest levels of the Department of Defense, the issue of military housing seemed like a natural one to investigate using data science practices such as text mining and natural language processing. The results of this study were shared with the senior leaders of Aberdeen Proving Ground, an active U.S. Army installation located in Maryland.

Data:

A sampling of over 1600 public tweets using the keywords "Corvias" (the name of the privately-owned company that holds the contract for DOD base housing) and/or "Military Housing" were gathered from early February to mid-March (figures US1500a and US1500b). During this time, the U.S. Army was conducting housing inspections and town halls to address the reports of poor quality housing on military installations. The data set was limited by search parameter keywords, geographic limiters (the U.S.), and by Twitter policies, as well as daily/monthly data rate limits. Very few of the tweets were tagged with latitude and longitude coordinates, making a more geographically-focused analysis (such as by city) impossible; however, 141 unique tweets (figures MD141a and MD 141b) were associated with users that listed "Maryland" as a location somewhere in their profile information (as either the place where they were posting or their hometown location).

The full set of tweets spanned at least three categories of social media posts: 1) complaints (and kudos) by residents, 2) official postings from Corvias and military installations, and 3) news articles. Retweets were deleted so as to get a more comprehensive picture of the issue's complexity at all levels rather than results skewed in favor of the popular news stories that were frequently shared, reducing the size of the dataset before analysis.

Project Goals:

- To better understand the problems residents of Aberdeen Proving Ground (APG) faced with their on-post housing and in their interactions with Corvias staff

- To determine whether the Army's initiatives were working to address the problems
- To learn more about the evolution of the military housing problem

Methods:

Two keyword searches were used to collect public posts on Twitter dated from early February to mid-March 2019 that used "Corvias" and/or "military housing." Another search was run to generate an additional 141 tweets about "Corvias" from users located in the state of Maryland.

Data was then cleaned to isolate text and to remove URLs, "stop words" (articles, prepositions, etc.), and special characters.

```
natlCorpus <- tm_map(natlCorpus, FUN = content_transformer(tolower)
natlCorpus <- tm_map(natlCorpus, removePunctuation)
natlCorpus <- tm_map(natlCorpus, removeNumbers)
removeURL <- function(x) gsub("http[[:alnum:]]*", '', x)
natlCorpus <- tm_map(natlCorpus, content_transformer(removeURL))

natlClean <- tm_map(natlCorpus, stripWhitespace)
```

Word frequency was counted and plotted on a graph. Finally, natural language processing and machine learning were run in the form of a sentiment analysis. The plots included in this report were generated using the NRC Word-Emotion Association Lexicon. This lexicon associates individual words with eight different sentiments and was created by Saif Mohammed and Peter Turney using crowdsourcing techniques approved by the Canadian NRC Research Ethics Board (#2009-94).

Conclusions/Results:

<u>Word Frequency Results (Appendices A and B)</u>

Aberdeen Proving Ground (APG) was not mentioned by name in any public tweet in the data set. However, Fort Bragg, Fort Meade, Fort Riley, and Fort Polk were specifically called out (MD141a), as is consistent with national news reports.

From a visual scan of the tweets, it seems that those using a variation of "thanks" (MD141a) were mainly being sarcastic, calling out bad behavior by Corvias in the guise of "thanking" them. However, a few instances of "thanks" (and related synonyms) suggest that the joy reflected in the sentiment analysis results comes from residents expressing happiness to have their Corvias concerns heard and validated by reporters and Army leadership.

While there were still posts being made during this timeframe that revealed additional or ongoing struggles to remedy substandard housing conditions (see frequent words "mold," "air," and "termites"), it would seem that the problem already peaked in terms of public debate. The most recent data (from March 2019) indicates that the direction of the discussion has shifted to campaign donations made by John Picerne (Corvias CEO) to Senator Jack Reed (Democrat—Rhode Island). The frequency of words like "senjackreed," "senblumenthal," "senthomtillis," "close ties," "ceo," "johnpicerne," "democrat," "district," and "election" demonstrates that the military housing issue has now morphed into a political battle in Washington, involving hot-button topics such as campaign funding practices and money for the border wall (versus money for better military housing).

Recommendation #1: Army leadership may have an opportunity here to model best practices for effecting positive change by remaining on-task, maintaining a results-oriented posture, and shifting the conversation back to the families who are most affected by substandard housing conditions.

Recommendation #2: Educate the families living on Army installations to help them distinguish between issues that can be handled locally (and what the proper process is) versus issues that need to be lobbied at the level of their political representation.

Sentiment Analysis Results (Appendices C and D)

The larger score for positivity (over negativity) reflects the efforts by the DOD and Corvias to counter the problem. This has also resulted in a high score for trust. The contrast between a firm base of trust and elevated marks for fear and anticipation (or "nervousness") indicate that the situation remains somewhat dynamic.

Descriptions of ill health were unique to many Twitter users (some residents bemoaned mold-induced "asthma," while others referenced the same illness by talking about frequent their "inhaler" use). The lack of common vocabulary to discuss these health matters (issues of poor health are often more culturally acceptable when spoken about in euphemisms) accounts for why health issue complaints did not show up any visible way in the word frequency graph, even though a visual scan of the tweets suggests that this is a major factor contributing to the sentiment score for fear.

In conclusion, the language of public tweets suggests that, while they are still fearful and nervous about the details of their individual housing problems and possible side effects to their health, residents have regained trust in Army leadership as a result of the outreach that has happened since the problem was brought to national attention.

Recommendation #3: Army installations would do well to harness data analytics to monitor public opinion of the Army brand, especially during a time of lower recruitment rates. This particular housing issue would have been visible through simple key word searches of public fora on the internet as early as August 2018, 4 months before the Reuter's report came out that attained such wide publicity.

Maryland versus National

Maryland users expressed much less fear than users at the national level. Coupled with the greater expression of trust by Maryland users, this is most likely reflective either of positive past or current crisis management on the part of Army leadership, as well as a history of strong community relations.

Although the US1500a word frequency list (Appendix B) included words associated with anger, the anger score for sentiment was about the same, proportionally, on both the national and State of Maryland lists, and occurred with a below average frequency. Nationwide, the public were more surprised by the military housing conditions than the state population was. This might be attributed to differences in users' familiarity with military installations, suggesting that tweets in the State of Maryland came from users more familiar with conditions on military bases. This means that analyses such as MD141a and MD141b (Appendices A and C), while a smaller data set, are probably more reflective of public opinion at and around APG.

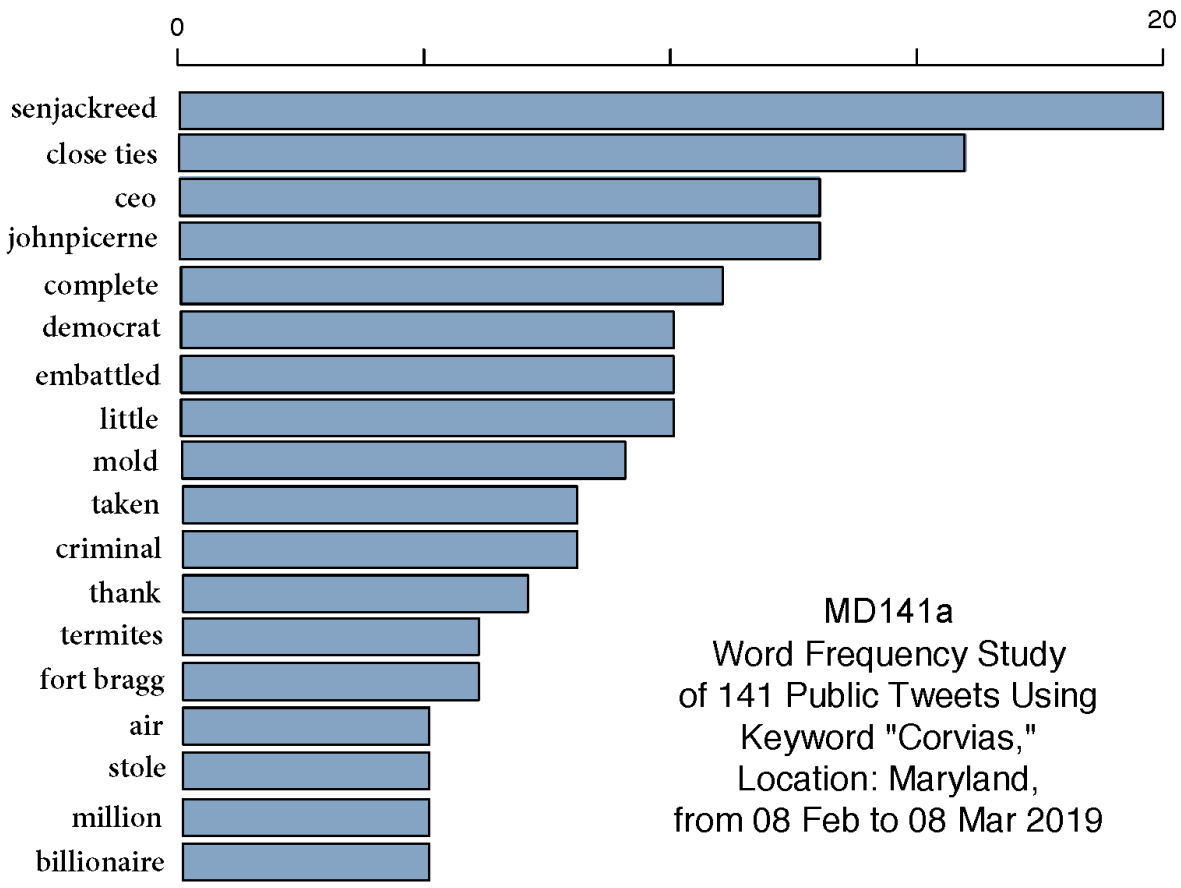
Important Note:

This study was done independently. The researcher received no funding from the U.S. Government, Department of Defense, or the U.S. Army. No Twitter user information was shared with the government.

Researcher's Credentials:

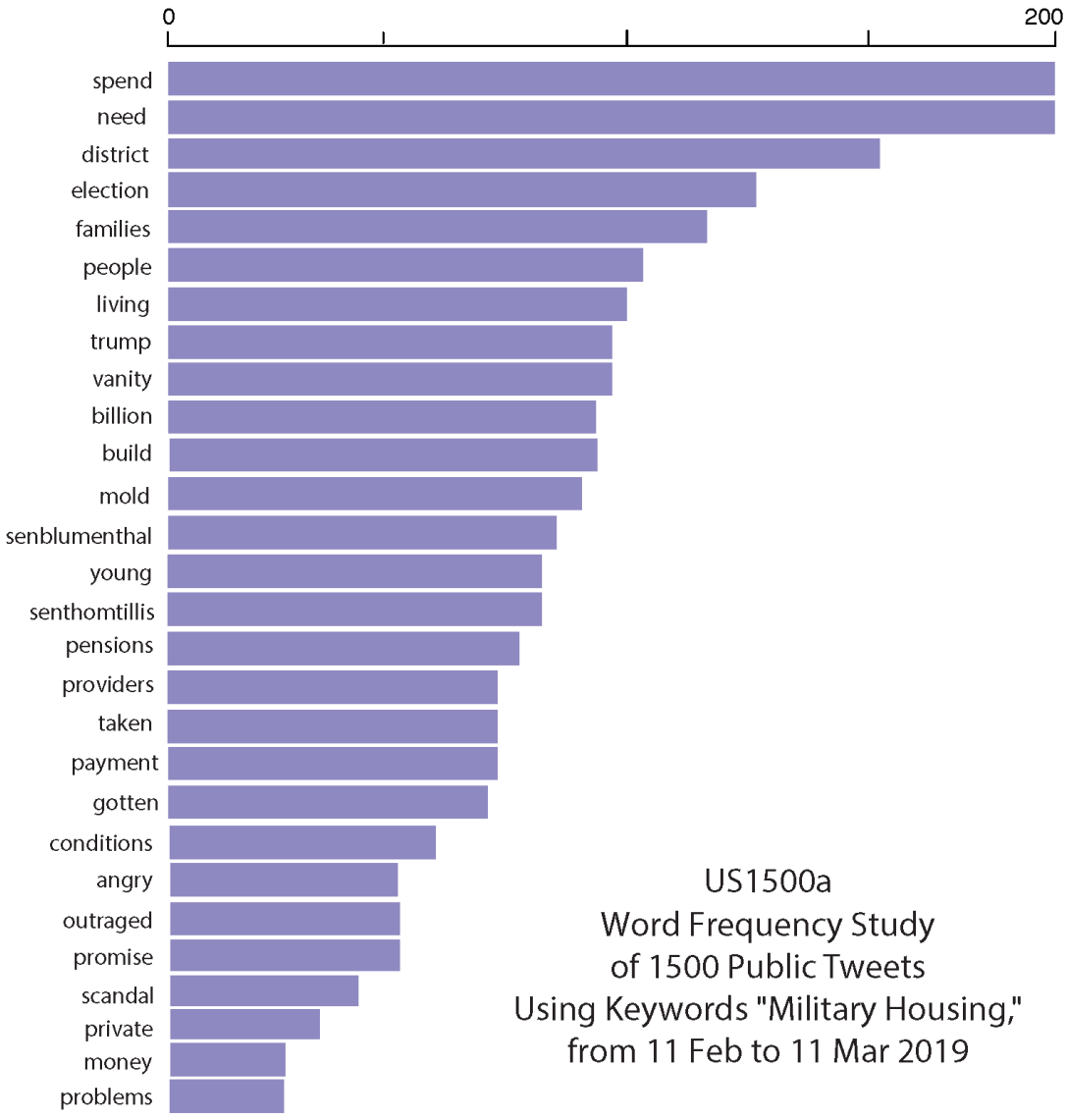
Lynn Marie Houston holds a Ph.D. in English (Cultural Studies) from Arizona State University. For her dissertation, she investigated health crises in the U.S. meat industry in conjunction with the rhetoric of health in 20th century American literature. She used interdisciplinary research methodologies, including close readings of memoirs, ethnographic fieldwork, and qualitative analysis of data from Internet sources. She is a technical writer-editor for the U.S. Army's Communications and Electronics Command as well as the author of several books and numerous peer-reviewed articles.

APPENDIX A



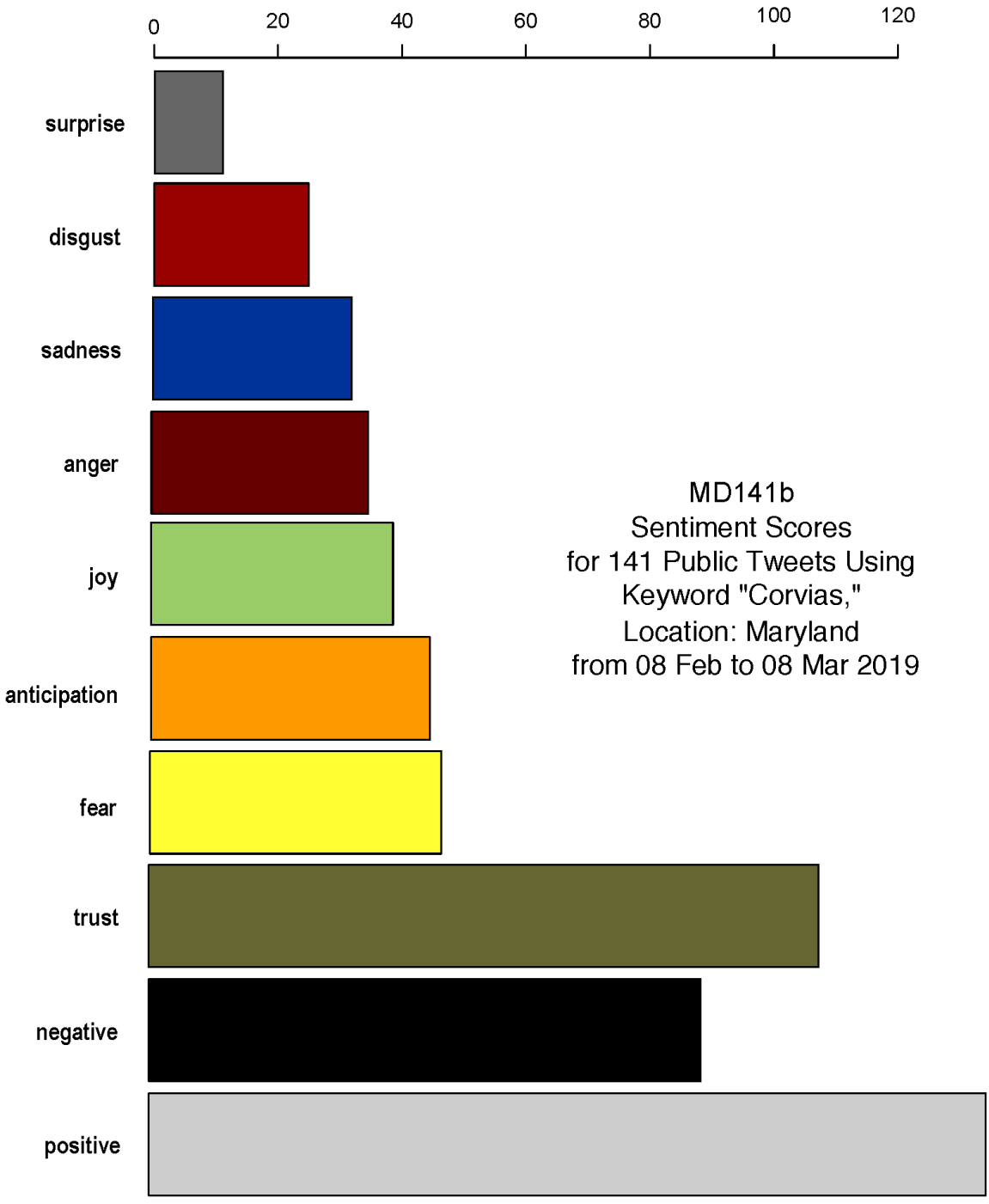
MD141a
Word Frequency Study
of 141 Public Tweets Using
Keyword "Corvias,"
Location: Maryland,
from 08 Feb to 08 Mar 2019

APPENDIX B

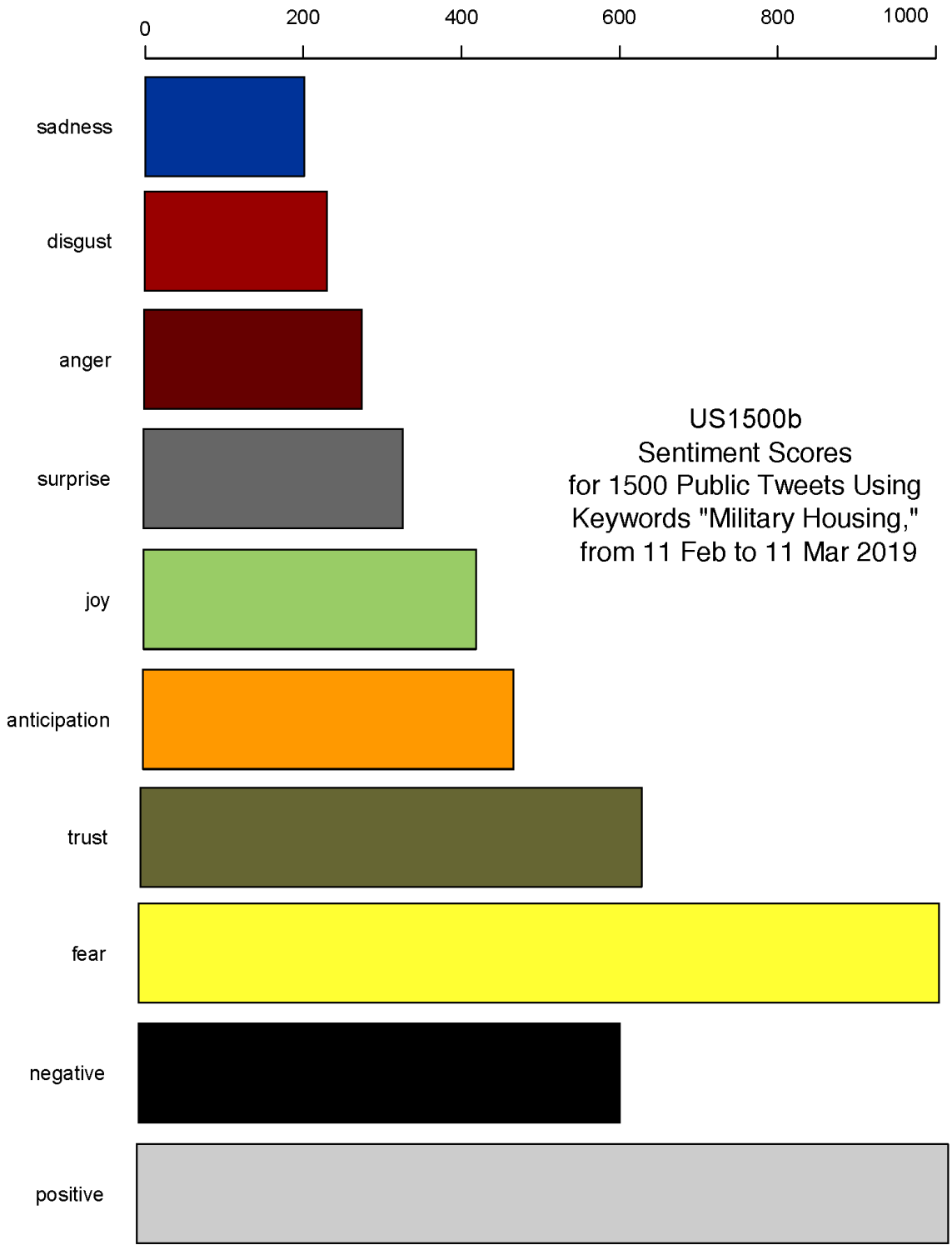


US1500a
Word Frequency Study
of 1500 Public Tweets
Using Keywords "Military Housing,"
from 11 Feb to 11 Mar 2019

APPENDIX C


0
20
40
60
80
100
120
surprise
disgust
sadness
anger
joy
anticipation
fear
trust
negative
positive
MD141b
Sentiment Scores
for 141 Public Tweets Using
Keyword "Corvias,"
Location: Maryland
from 08 Feb to 08 Mar 2019

Appendix D


0
200
400
600
800
1000
sadness
disgust
anger
surprise
joy
anticipation
trust
fear
negative
positive
US1500b
Sentiment Scores
for 1500 Public Tweets Using
Keywords "Military Housing,"
from 11 Feb to 11 Mar 2019