

AN UNEXPLAINED SHIFT IN DISCOURSE

FOR THE LAST THREE MONTHS THE MAINSTREAM MEDIA HAS INDOCTRINATED THE AMERICAN PUBLIC WITH THE IDEA THAT IT IS A MORAL OBLIGATION TO SHELTER IN PLACE DUE TO THE CONOID 19 CRISES. DURING THE MONTHS OF MARCH, APRIL AND MAY OF 2020, "NON-ESSENTIAL" CITIZENS WHO DIDN'T COMPLY WITH THIS DOGMA, WERE REGULARLY SHAMED ON SOCIAL MEDIA AND IN SOME CASES FINED OR EVEN ARRESTED. HOWEVER, SINCE THE GEORGE FLOYD PROTESTS BEGAN IN LATE MAY, THE SAME MAINSTREAM MEDIA WHO SPENT MONTHS STIGMATIZING THE ACT OF GATHERING IN CROWDS HAS ACTIVELY ENCOURAGED THE PROTESTERS TO GATHER IN CROWDS. A LARGE PORTION OF THE AMERICAN PUBLIC HAS INSTANTLY TRANSITIONED FROM ONE RADICAL DICTATE TO THE NEXT DESPITE THE FACT THAT THESE POLICIES OBVIOUSLY CONTRADICT EACH OTHER. THE FACT THAT THE MAINSTREAM MEDIA CAN CHANGE THE NARITIVE SO ABRUPTLY WITHOUT CRITICISM SEEMS TO BEG THE QUESTIONS?

WHAT TYPES OF INDIVIDUALS ARE EASILY INFLUENCED BY THE MAINSTREAM MEDIA?

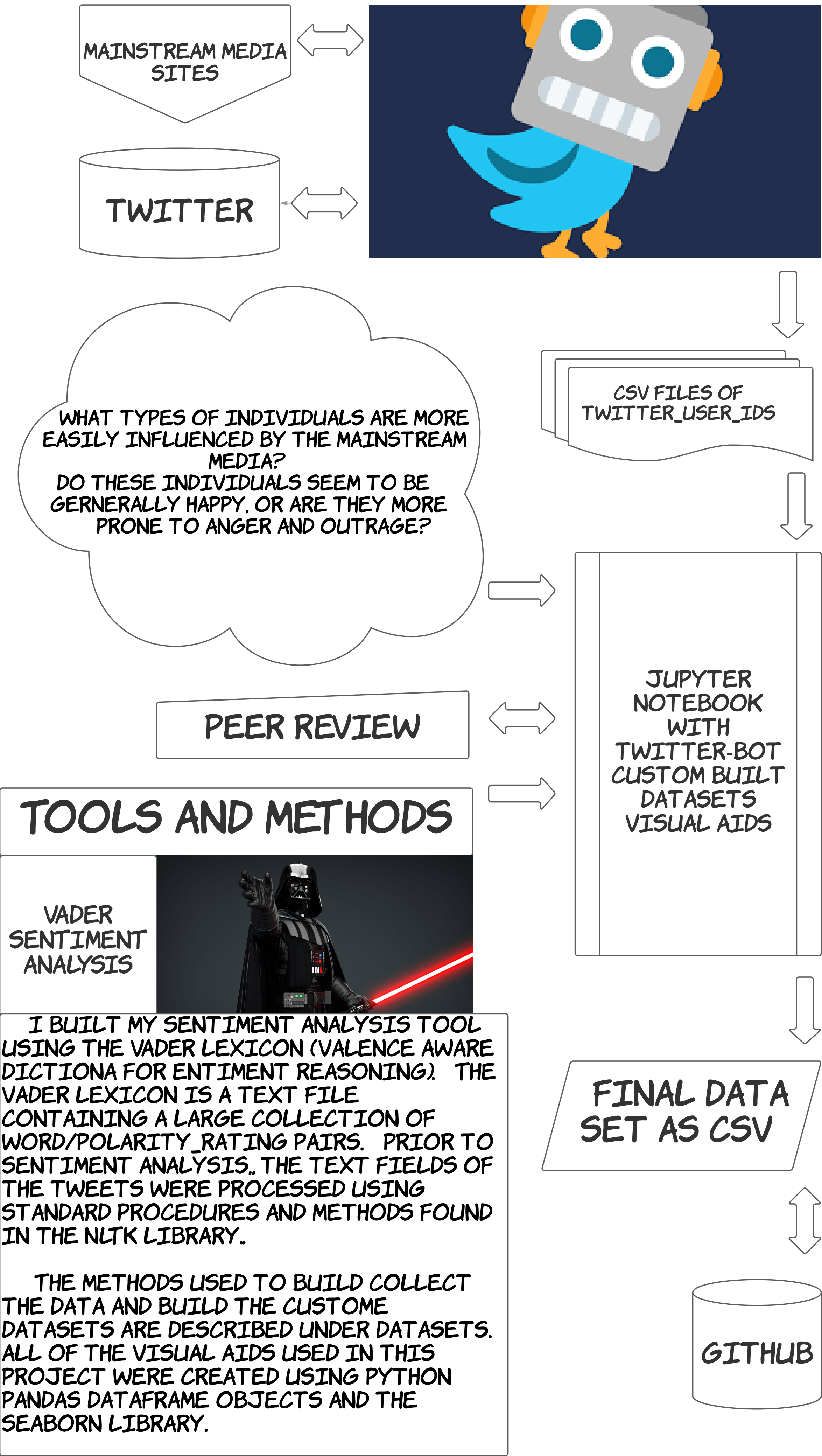
DO THESE INDIVIDUALS SEEM TO BE GENERALLY HAPPY, OR ARE THEY MORE PRONE TO ANGER AND OUTRAGE?

DATASETS

IN ORDER TO BUILD A DATASET CONTAINING INFORMATION ON EASILY INFLUENCED INDIVIDUALS. I NEEDED A WAY TO IDENTIFY THESE SHEEPISH HUMANS. TO ACCOMPLISH THIS, I DECIDED TO USE A JUPYTER NOTEBOOK, AND THE TWEETPY API TO BUILD AN AUTOMATED TWITTER-BOT THAT SCRAPES MAINSTREAM MEDIA WEB SITES FOR IMAGE AND VIDEO LINKS, THEN POSTS THOSE LINKS TO TWITTER. MY THEORY WAS THAT INDIVIDUALS WHO INTERACT WITH A TWITTER-BOT WHO ONLY POSTS MAINSTREAM MEDIA CONTENT, WERE LIKELY TO EASILY CONFORM TO WHATEVER DICTATES THE MAINSTREAM MEDIA SETS OUT. THE OBVIOUS EXAMPLE BEING "SHELTER IN PLACE ON THURSDAY" THEN "PROTEST IN LARGE CROWDS ON FRIDAY."

MY FIRST DATASET, WHICH I AFFECTIONATELY LABELED "SHEEP", WAS A CSV FILE CONTAINING THE TWITTER USER_IDS OF ALL OF THE INDIVIDUALS WHO LIKED, RE-TWEETED OR FOLLOWED MY TWITTER-BOT. I USED THE USER_IDS OF THE SHEEP TO AS INPUTS TO A JUPYTER NOTEBOOK METHOD THAT I WROTE WHICH PULLS THE LAST 25 TWEETS OF EACH SHEEP, THEN ADDS THEM TO MY SECOND DATASET WHICH I CALLED "SHEEP_TWEETS." THE SHEEP_TWEETS DATASET IS A PANDAS DATAFRAME OBJECT CONTAINING ROWS WHICH REPRESENT A SINGLE TWEET OBJECT. A TWEET OBJECTS CONTAINS QUITE A LOT OF INFORMATION BUT, AFTER I FILTERED OUT THE IRRELEVANT FIELDS, I CREATED MY THIRD DATA SET CALLED "IDS_25TWEETS" WHICH WAS ALSO A PANDAS DATAFRAME OBJECT. THE IDS_25TWEETS DATASET CONTAINED ONLY 3 COLUMNS, INDEX, USER_ID OF EACH SHEEP, AND A PYTHON LIST CONTAINING THE TEXT OF A SINGLE TWEET AT EACH INDEX. AT THIS POINT THE IDS_25TWEETS DATASET WAS FULLY CONSTRUCTED AND IN THE APPROPRIATE FORM TO BEGIN THE TEXT PROCESSING THAT IS NECESSARY PRIOR TO SENTIMENT ANALYSIS.

METHODOLGY



INERPRETATIONS OF RESULTS

GRAPHS, CHARTS AND CONCLUSIONS GO HERE.

I DON'T HAVE ANY CONCLUSIONS AT THIS POINT.

VADER LEXICON (VALENCE AWARE DICTIONARY FOR ENTIMENT REASONING).
THE TEXT FILE CAN BE FOUND HERE: ["VADER LEXICON TEXTFILE"](#)

SCRAPED WEB SITE LINKS WILL GO HERE
GITHUB LINK WILL GO HERE

DIGHUM 100
THEORY & METHODS IN THE DIGITAL HUMANITIES
INTRUCTOR: DR. ADAM G. ANDERSON
DATA SCIENTIST: MATHEW BUCK
JUNE 7 2020

AN UNEXPLAINED SHIFT IN DISCOURSE

FOR THE LAST THREE MONTHS THE MAINSTREAM MEDIA HAS INDOCTRINATED THE AMERICAN PUBLIC WITH THE IDEA THAT IT IS A MORAL OBLIGATION TO SHELTER IN PLACE DUE TO THE CONOID 19 CRISES. DURING THE MONTHS OF MARCH, APRIL AND MAY OF 2020, "NON-ESSENTIAL" CITIZENS WHO DIDN'T COMPLY WITH THIS DOGMA, WERE REGULARLY SHAMED ON SOCIAL MEDIA AND IN SOME CASES FINED OR EVEN ARRESTED. HOWEVER, SINCE THE GEORGE FLOYD PROTESTS BEGAN IN LATE MAY, THE SAME MAINSTREAM MEDIA WHO SPENT MONTHS STIGMATIZING THE ACT OF GATHERING IN CROWDS HAS ACTIVELY ENCOURAGED THE PROTESTERS TO GATHER IN CROWDS. A LARGE PORTION OF THE AMERICAN PUBLIC HAS INSTANTLY TRANSITIONED FROM ONE RADICAL DICTATE TO THE NEXT DESPITE THE FACT THAT THESE POLICIES OBVIOUSLY CONTRADICT EACH OTHER. THE FACT THAT THE MAINSTREAM MEDIA CAN CHANGE THE NARITIVE SO ABRUPTLY WITHOUT CRITICISM SEEMS TO BEG THE QUESTIONS?

WHAT TYPES OF INDIVIDUALS ARE EASILY INFLUENCED BY THE MAINSTREAM MEDIA?

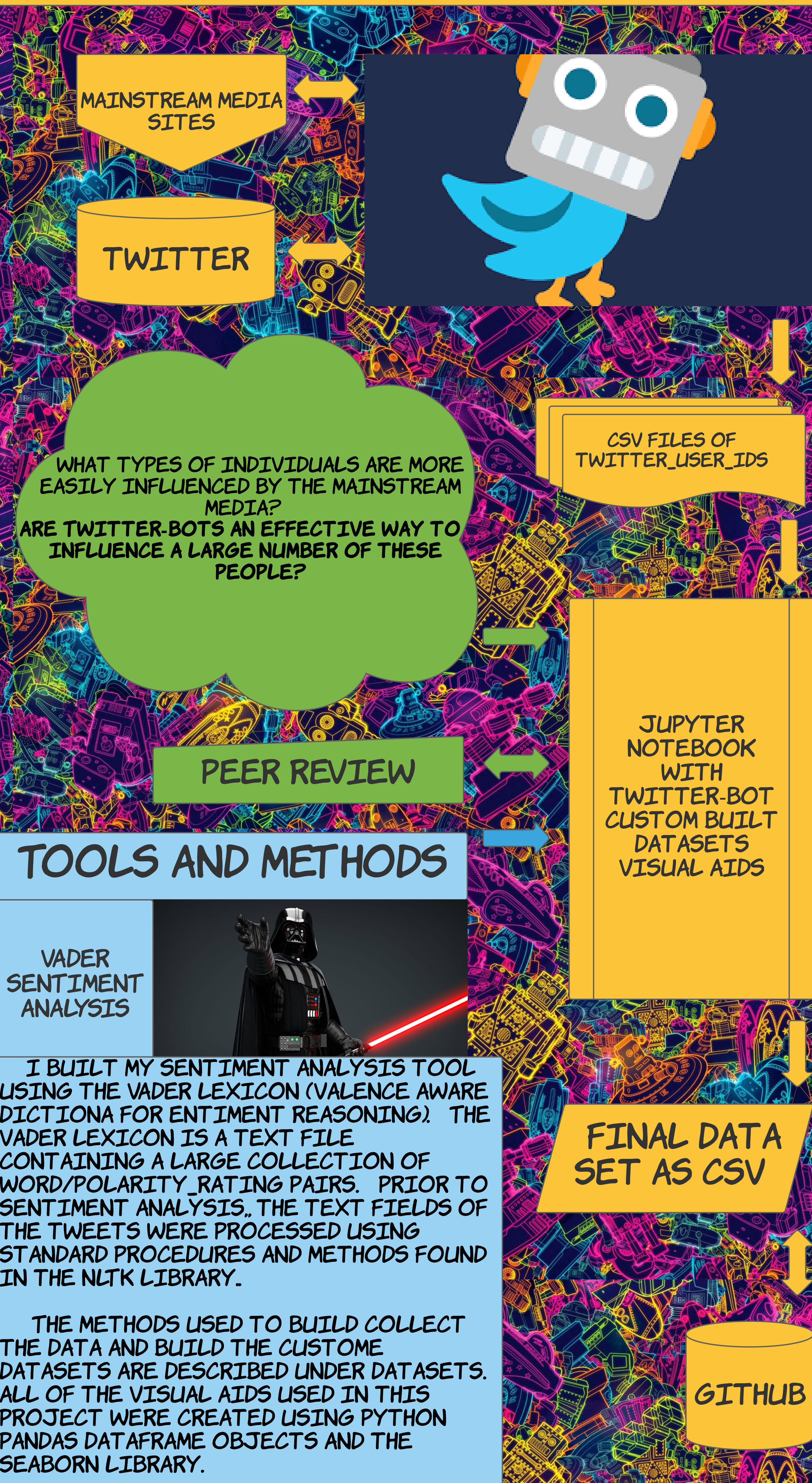
ARE TWITTER-BOTS AN EFFECTIVE WAY TO INFLUENCE A LARGE NUMBER OF THESE PEOPLE?

DATASETS

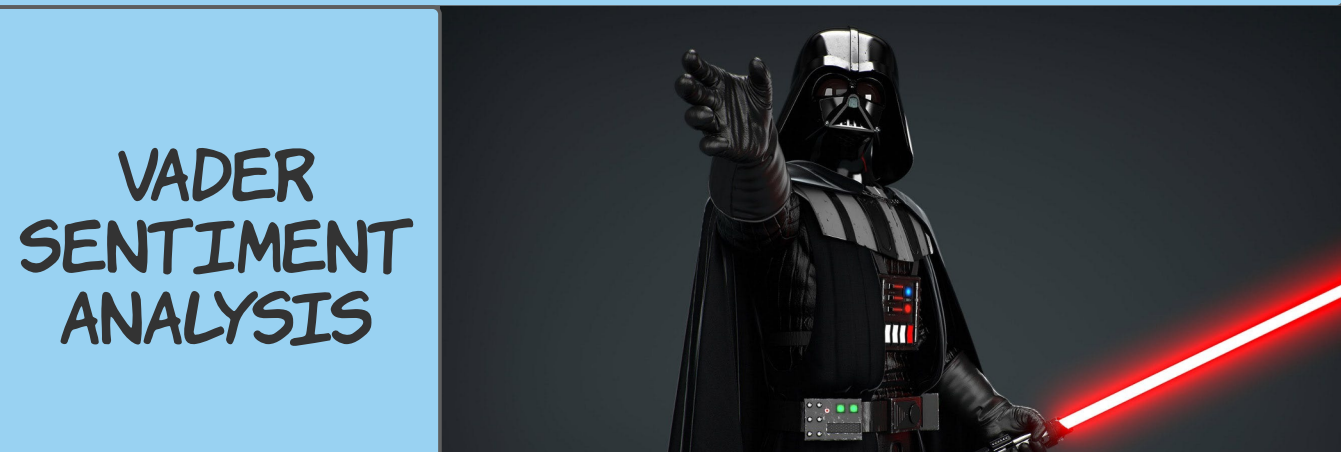
IN ORDER TO BUILD A DATASET CONTAINING INFORMATION ON EASILY INFLUENCED INDIVIDUALS. I NEEDED A WAY TO IDENTIFY THESE SHEEPISH HUMANS. TO ACCOMPLISH THIS, I DECIDED TO USE A JUPYTER NOTEBOOK, AND THE TWEETPY API TO BUILD AN AUTOMATED TWITTER-BOT THAT SCRAPES MAINSTREAM MEDIA WEB SITES FOR IMAGE AND VIDEO LINKS, THEN POSTS THOSE LINKS TO TWITTER. MY THEORY WAS THAT INDIVIDUALS WHO INTERACT WITH A TWITTER-BOT WHO ONLY POSTS MAINSTREAM MEDIA CONTENT, WERE LIKELY TO EASILY CONFORM TO WHATEVER DICTATES THE MAINSTREAM MEDIA SETS OUT. THE OBVIOUS EXAMPLE BEING "SHELTER IN PLACE ON THURSDAY" THEN "PROTEST IN LARGE CROWDS ON FRIDAY."

MY FIRST DATASET, WHICH I AFFECTIONATELY LABELED "SHEEP", WAS A CSV FILE CONTAINING THE TWITTER USER_IDS OF ALL OF THE INDIVIDUALS WHO LIKED, RE-TWEETED OR FOLLOWED MY TWITTER-BOT. I USED THE USER_IDS OF THE SHEEP TO AS INPUTS TO A JUPYTER NOTEBOOK METHOD THAT I WROTE WHICH PULLS THE LAST 25 TWEETS OF EACH SHEEP, THEN ADDS THEM TO MY SECOND DATASET WHICH I CALLED "SHEEP_TWEETS." THE SHEEP_TWEETS DATASET IS A PANDAS DATAFRAME OBJECT CONTAINING ROWS WHICH REPRESENT A SINGLE TWEET OBJECT. A TWEET OBJECTS CONTAINS QUITE A LOT OF INFORMATION BUT, AFTER I FILTERED OUT THE IRRELEVANT FIELDS, I CREATED MY THIRD DATA SET CALLED "IDS_25TWEETS" WHICH WAS ALSO A PANDAS DATAFRAME OBJECT. THE IDS_25TWEETS DATASET CONTAINED ONLY 3 COLUMNS, INDEX, USER_ID OF EACH SHEEP, AND A PYTHON LIST CONTAINING THE TEXT OF A SINGLE TWEET AT EACH INDEX. AT THIS POINT THE IDS_25TWEETS DATASET WAS FULLY CONSTRUCTED AND IN THE APPROPRIATE FORM TO BEGIN THE TEXT PROCESSING THAT IS NECESSARY PRIOR TO SENTIMENT ANALYSIS.

METHODOLGY



TOOLS AND METHODS



I BUILT MY SENTIMENT ANALYSIS TOOL USING THE VADER LEXICON (VALENCE AWARE DICTIONA FOR ENTIMENT REASONING). THE VADER LEXICON IS A TEXT FILE CONTAINING A LARGE COLLECTION OF WORD/POLARITY_RATING PAIRS. PRIOR TO SENTIMENT ANALYSIS, THE TEXT FIELDS OF THE TWEETS WERE PROCESSED USING STANDARD PROCEDURES AND METHODS FOUND IN THE NLTK LIBRARY.

THE METHODS USED TO BUILD COLLECT THE DATA AND BUILD THE CUSTOME DATASETS ARE DESCRIBED UNDER DATASETS. ALL OF THE VISUAL AIDS USED IN THIS PROJECT WERE CREATED USING PYTHON PANDAS DATAFRAME OBJECTS AND THE SEABORN LIBRARY.

INERPRETATIONS OF RESULTS

GRAPHS, CHARTS AND CONCLUSIONS GO HERE.

I DON'T HAVE ANY CONCLUSIONS AT THIS POINT.

VADER LEXICON (VALENCE AWARE DICTIONARY FOR ENTIMENT REASONING). THE TEXT FILE CAN BE FOUND HERE: ["VADER LEXICON TEXTFILE"](#)

SCRAPED WEB SITE LINKS WILL GO HERE
GITHUB LINK WILL GO HERE

DIGHUM 100
THEORY & METHODS IN THE DIGITAL HUMANITIES
INTRUCTOR: DR. ADAM G. ANDERSON
DATA SCIENTIST: MATHEW BUCK
JUNE 7 2020