

EAS 503: PROGRAMMING AND DATABASE FUNDAMENTALS
Preliminary Report: GROUP 12

Rahul Patil (rpatil@buffalo.edu)

Anand (anand6@buffalo.edu)

Supreeth (ssuresh7@buffalo.edu)

Status of the project is as mentioned below:

Tasks	Description	Status	Comments
1	Reading tweets and loading to database using tweet API tweepy directly Currently we are using JSON file loading directly from dataset instead of using API	Completed	04-22-> tried API but it can read max of 200 tweets not more than that. Need to understand API more and work on it, for now we will be using tweet dataset which is already extracted and we will work on direct integration if time permits 04-24-> Finished coding for the preprocessing of tweet data and loading to database checked into Git
2	Reading NYSE stock data and loading to database and querying data from database - table creation, insertion and querying We will use Kaggle dataset for it https://www.kaggle.com/timoboz/tesla-stock-data-from-2010-to-2020	In Progress	
3	Sentiment Analysis - From scratch or using python library for tweet data or use python library	In Progress	04-26-> Analyzed different libraries of python which can be used for sentiment analysis TextBlob Natural Language Toolkit (NLTK) These libraries just give three states of sentiment analysis Positive Negative Neutral But we are looking fine-grained sentiment analysis of tweet.If we don't find we will use one of these libraries to perform sentiment analysis.
4	Correlation of stock table and tweet table data		
5	Interactive plots to represent sentiment analysis and its effects on stock data		
6	Project design document		
7	Project Report		