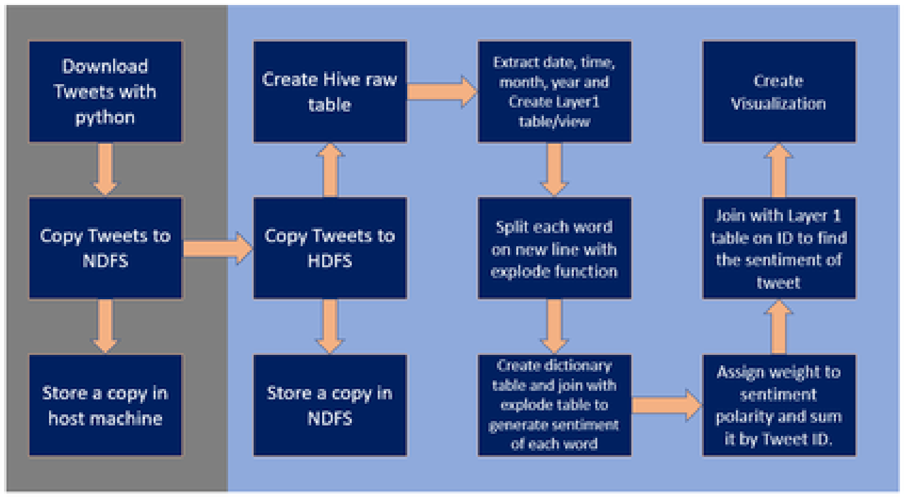
Project 01 - Twitter Sentiment Analysis in Hive

This project will use Apache Hive to analyze the sentiment of incoming twitter data. The pipeline would be as follows.



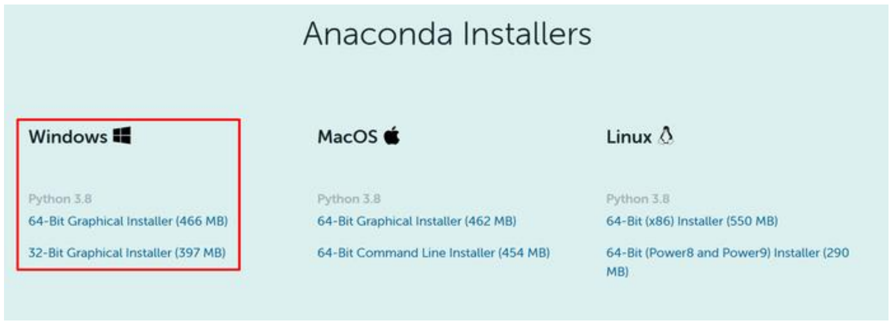
Download Tweets with python
Copy Tweets to NDFS
Store a copy in host machine
Create Hive raw table
Copy Tweets to HDFS
Store a copy in NDFS
Extract date, time, month, year and Create Layer1 table/view
Split each word on new line with explode function
Create dictionary table and join with explode table to generate sentiment of each word
Create Visualization
Join with Layer 1 table on ID to find the sentiment of tweet
Assign weight to sentiment polarity and sum it by Tweet ID.

Getting Data from Twitter API:

The tweet generation code has been written in Python so it is necessary to install it on host machine.

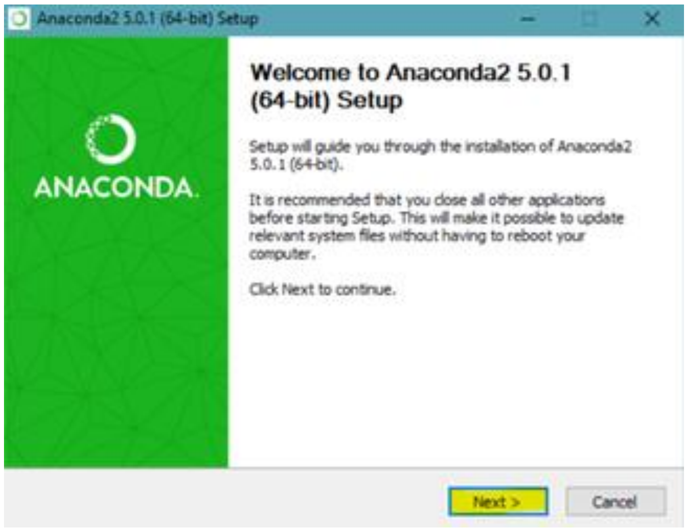
Install Python (Anaconda):

The tweet generation code has been written in Python so it is necessary to install it on host machine. We'll install Anaconda which is a complete package for Python that provides a stable version of Python and a complete IDE and notebook for python coding. In addition, it also provides tools for debugging and testing making it a go-to application for python development. Here's the download link for Anaconda's website. Choose Python 3.8 installer according to your machine specs (32/64 bit).
https://www.anaconda.com/products/individual


Anaconda Installers

Windows	MacOS	Linux
Python 3.8	Python 3.8	Python 3.8
64-Bit Graphical Installer (466 MB)	64-Bit Graphical Installer (462 MB)	64-Bit (x86) Installer (550 MB)
32-Bit Graphical Installer (397 MB)	64-Bit Command Line Installer (454 MB)	64-Bit (Power8 and Power9) Installer (290 MB)

After downloading, open the installer click next and follow the process:

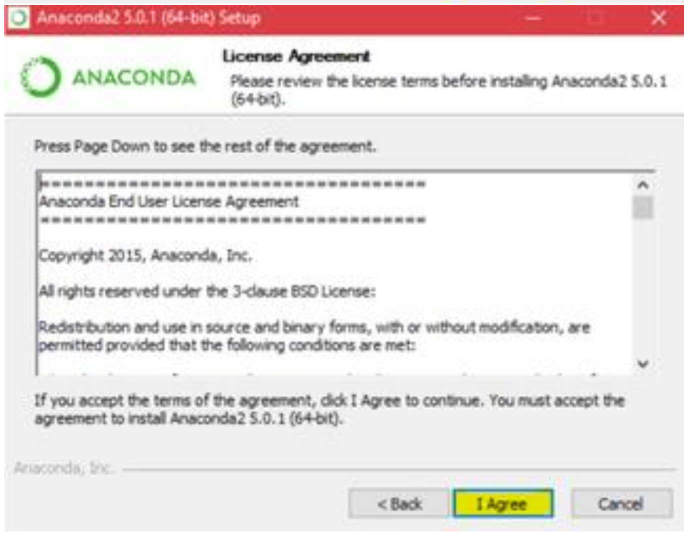

Anaconda2 5.0.1 (64-bit) Setup

Welcome to Anaconda2 5.0.1 (64-bit) Setup

Setup will guide you through the installation of Anaconda2 5.0.1 (64-bit).

It is recommended that you close all other applications before starting Setup. This will make it possible to update relevant system files without having to reboot your computer.

Click Next to continue.


Anaconda2 5.0.1 (64-bit) Setup

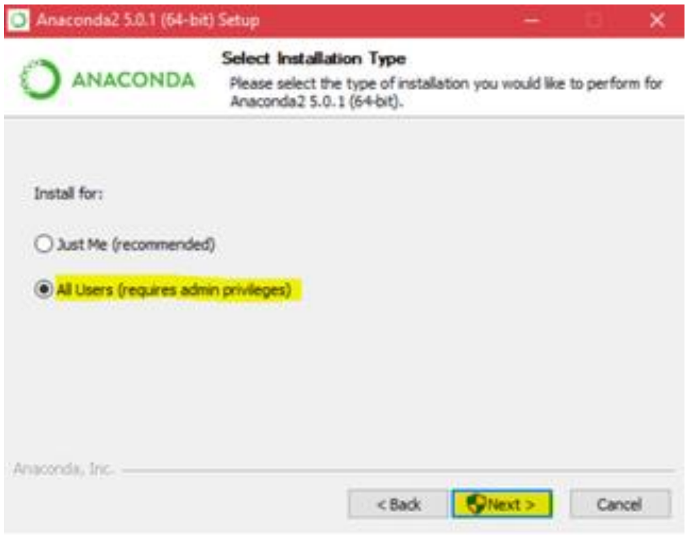
License Agreement

Please review the license terms before installing Anaconda2 5.0.1 (64-bit).

Press Page Down to see the rest of the agreement.

```
===================================
Anaconda End User License Agreement
===================================

Copyright 2015, Anaconda, Inc.

All rights reserved under the 3-clause BSD License:

Redistribution and use in source and binary forms, with or without modification, are
permitted provided that the following conditions are met:
```

If you accept the terms of the agreement, click I Agree to continue. You must accept the agreement to install Anaconda2 5.0.1 (64-bit).


Anaconda2 5.0.1 (64-bit) Setup

Select Installation Type

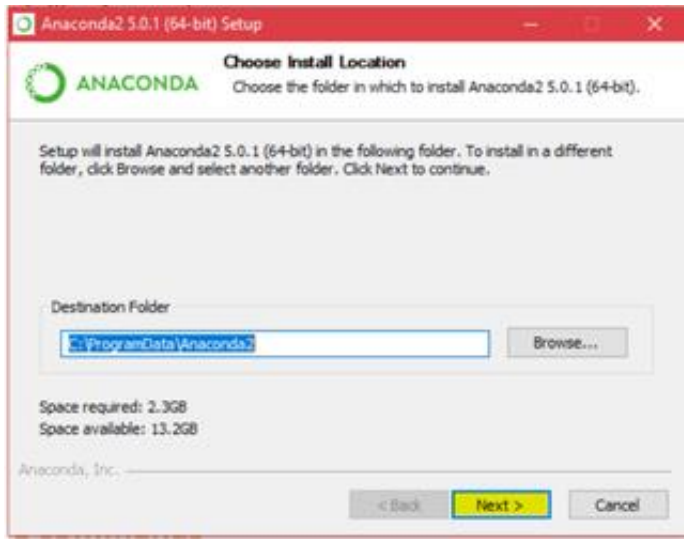
Please select the type of installation you would like to perform for Anaconda2 5.0.1 (64-bit).

Install for:

( ) Just Me (recommended)

(●) All Users (requires admin privileges)

Select the folder you want to install Anaconda in and click next. Take note of the installation path entered in the Destination Folder as we will need it later on. Its best to copy the path in a notepad file for later use.


Anaconda2 5.0.1 (64-bit) Setup

Choose Install Location

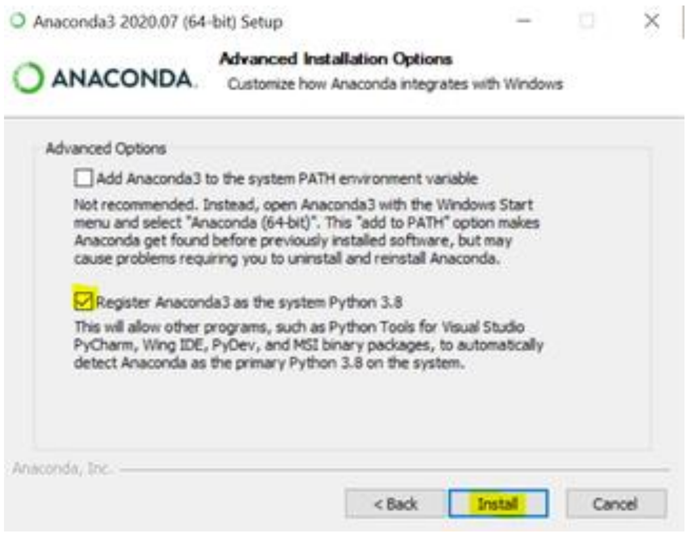
Choose the folder in which to install Anaconda2 5.0.1 (64-bit).

Setup will install Anaconda2 5.0.1 (64-bit) in the following folder. To install in a different folder, click Browse and select another folder. Click Next to continue.

Destination Folder: C:\ProgramData\Anaconda2

Space required: 2.3GB
Space available: 13.2GB

In the Advanced options step for the make sure the Add Anaconda to my PATH environment variable is unchecked as below. This is an option but as the warning below it states, it can cause problems. Check the other option and click install.


Anaconda3 2020.07 (64-bit) Setup

ANACONDA.

Advanced Installation Options

Customize how Anaconda integrates with Windows

Advanced Options

☐ Add Anaconda3 to the system PATH environment variable

Not recommended. Instead, open Anaconda3 with the Windows Start menu and select "Anaconda (64-bit)". This "add to PATH" option makes Anaconda get found before previously installed software, but may cause problems requiring you to uninstall and reinstall Anaconda.

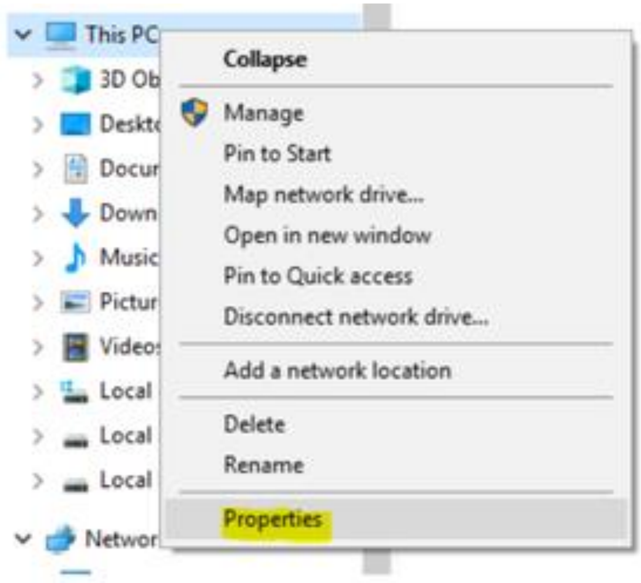
☑ Register Anaconda3 as the system Python 3.8

This will allow other programs, such as Python Tools for Visual Studio PyCharm, Wing IDE, PyDev, and MSI binary packages, to automatically detect Anaconda as the primary Python 3.8 on the system.

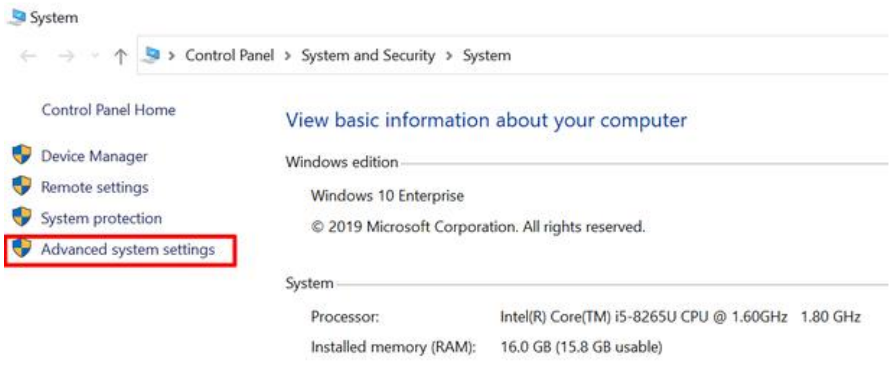
Anaconda, Inc.

< Back | Install | Cancel

The installation will be done after that. Now we need to add the path for Anaconda ourselves as we unchecked the option in the previous point. Open file explorer and on the left side and right click the computer name (usually named This PC but can be different). This will open the system's properties.



Click on Advanced System Settings on the left hand side of the system properties window. A new window will open.



View basic information about your computer

Windows edition

Windows 10 Enterprise

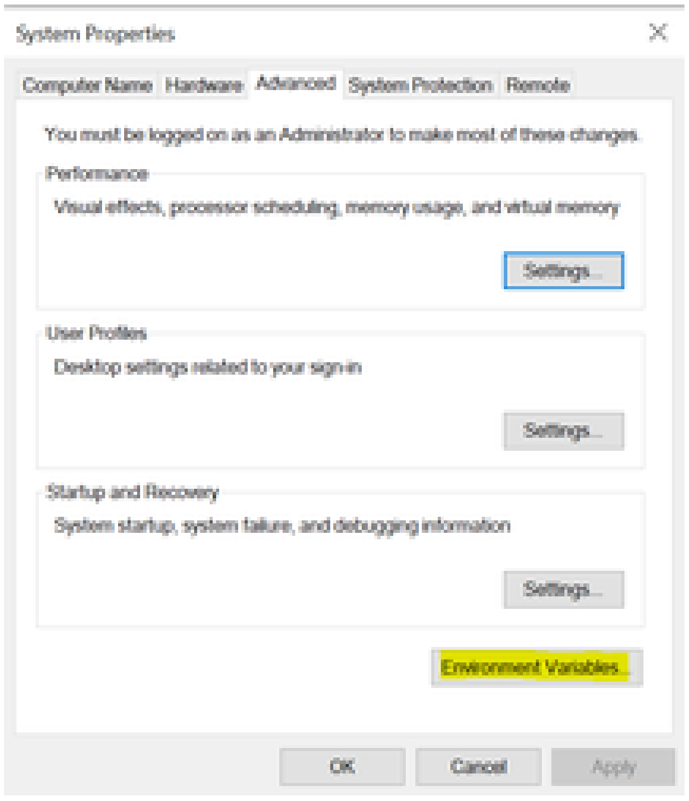
© 2019 Microsoft Corporation. All rights reserved.

System

Processor: Intel(R) Core(TM) i5-8265U CPU @ 1.60GHz 1.80 GHz

Installed memory (RAM): 16.0 GB (15.8 GB usable)

Click on Environment Variables button. A new window will open.



You must be logged on as an Administrator to make most of these changes.

Performance

Visual effects, processor scheduling, memory usage, and virtual memory

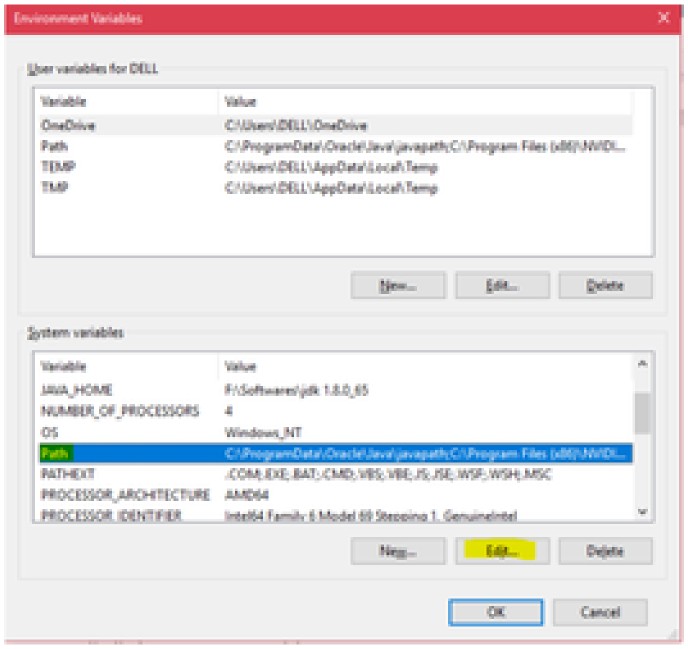
User Profiles

Desktop settings related to your sign-in

Startup and Recovery

System startup, system failure, and debugging information

Find the Path variable in the System variables tab, select it and click Edit. A new window will open.



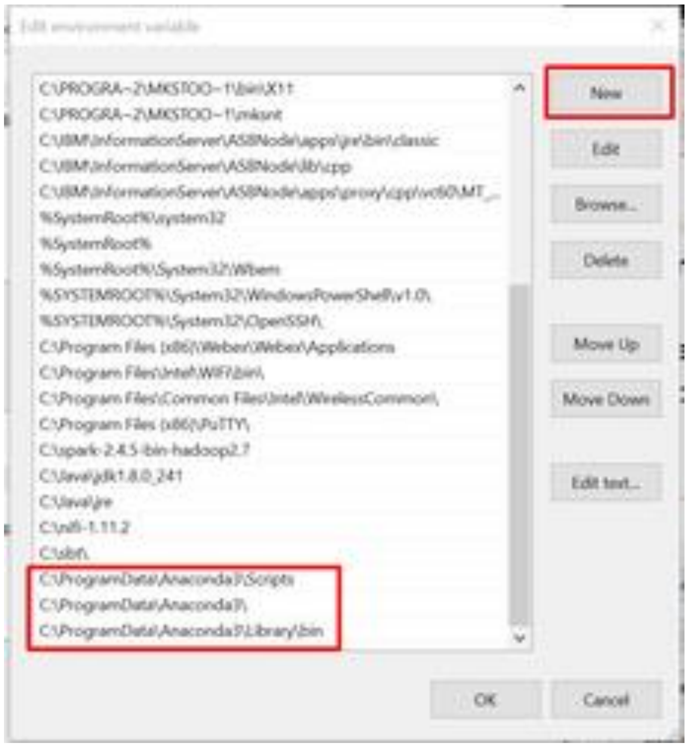
User variables for DELL

Variable	Value
OneDrive	C:\Users\DELL\OneDrive
Path	C:\ProgramData\Oracle\Java\javapath;C:\Program Files (x86)\NVIDI...
TEMP	C:\Users\DELL\AppData\Local\Temp
TMP	C:\Users\DELL\AppData\Local\Temp

System variables

Variable	Value
JAVA_HOME	F:\Softwares\jdk 1.8.0_65
NUMBER_OF_PROCESSORS	4
OS	Windows_NT
Path	C:\ProgramData\Oracle\Java\javapath;C:\Program Files (x86)\NVIDI...
PATHEXT	.COM;.EXE;.BAT;.CMD;.VBS;.VBE;.JS;.JSE;.WSF;.WSH;.MSC
PROCESSOR_ARCHITECTURE	AMD64
PROCESSOR_IDENTIFIER	Intel64 Family 6 Model 69 Stepping 1, GenuineIntel

Click the New button and add these three paths one by one as shown below. Make sure that you DON'T delete previously set environment variables. Then click OK. Click OK again on the Environment Variables window and OK once again on the System Properties window. Now the path is set for Anaconda.



```
C:\PROGRA~2\MKSTOO~1\bin\X11
C:\PROGRA~2\MKSTOO~1\mksnt
C:\IBM\InformationServer\ASBNode\apps\jre\bin\classic
C:\IBM\InformationServer\ASBNode\lib\cpp
C:\IBM\InformationServer\ASBNode\apps\proxy\cpp\vc60\MT_...
%SystemRoot%\system32
%SystemRoot%
%SystemRoot%\System32\Wbem
%SYSTEMROOT%\System32\WindowsPowerShell\v1.0\
%SYSTEMROOT%\System32\OpenSSH\
C:\Program Files (x86)\Webex\Webex\Applications
C:\Program Files\Intel\WiFi\bin\
C:\Program Files\Common Files\Intel\WirelessCommon\
C:\Program Files (x86)\PuTTY\
C:\spark-2.4.5-bin-hadoop2.7
C:\Java\jdk1.8.0_241
C:\Java\jre
C:\nifi-1.11.2
C:\sbt\
C:\ProgramData\Anaconda3\Scripts
C:\ProgramData\Anaconda3\
C:\ProgramData\Anaconda3\Library\bin
```

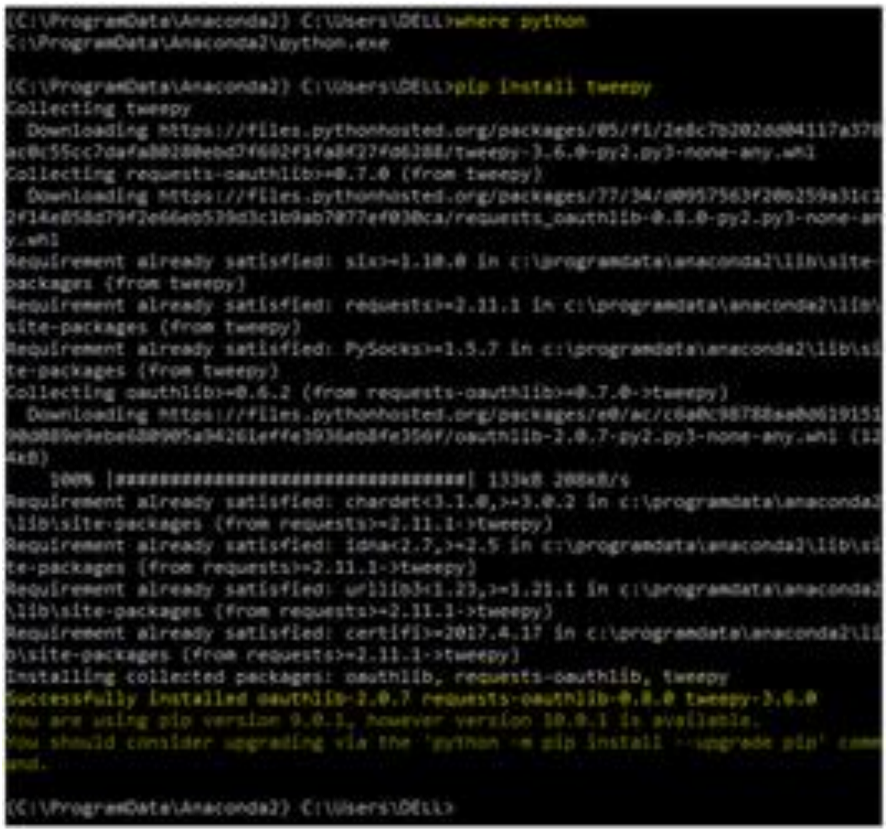
Install Tweepy:

Next we need to install the tweepy library for Python. Tweepy provides a twitter API for Python and we can use it to fetch public tweets from Twitter.
Open the Anaconda Prompt. Enter the command *where python* to make sure python is installed correctly. The command should return the path of the python installation.

```
where python
```

Next enter the command *pip install tweepy* to install tweepy. Pip will automatically find and install tweepy.

```
pip install tweepy
```



```
(C:\ProgramData\Anaconda2) C:\Users\DELL>where python
C:\ProgramData\Anaconda2\python.exe

(C:\ProgramData\Anaconda2) C:\Users\DELL>pip install tweepy
Collecting tweepy
  Downloading https://files.pythonhosted.org/packages/05/f1/2e8c7b202dd04117a378
ac0c55cc7dafa0020ebd7f602f1fa8f27fd6288/tweepy-3.6.0-py2.py3-none-any.whl
Collecting requests-oauthlib>=0.7.0 (from tweepy)
  Downloading https://files.pythonhosted.org/packages/77/34/d0957563f20b259a31c1
2f14e858d79f2e66eb539d3c1b9ab7077ef030ca/requests_oauthlib-0.8.0-py2.py3-none-an
y.whl
Requirement already satisfied: six>=1.10.0 in c:\programdata\anaconda2\lib\site-
packages (from tweepy)
Requirement already satisfied: requests>=2.11.1 in c:\programdata\anaconda2\lib\
site-packages (from tweepy)
Requirement already satisfied: PySocks>=1.5.7 in c:\programdata\anaconda2\lib\si
te-packages (from tweepy)
Collecting oauthlib>=0.6.2 (from requests-oauthlib>=0.7.0->tweepy)
  Downloading https://files.pythonhosted.org/packages/e0/ac/c6a0c98788aa0d619151
00a889e9ebe680905a94261effe3936eb8fe356f/oauthlib-2.0.7-py2.py3-none-any.whl (12
4kB)
    100% |################################| 133kB 288kB/s
Requirement already satisfied: chardet<3.1.0,>=3.0.2 in c:\programdata\anaconda2
\lib\site-packages (from requests>=2.11.1->tweepy)
Requirement already satisfied: idna<2.7,>=2.5 in c:\programdata\anaconda2\lib\si
te-packages (from requests>=2.11.1->tweepy)
Requirement already satisfied: urllib3<1.23,>=1.21.1 in c:\programdata\anaconda2
\lib\site-packages (from requests>=2.11.1->tweepy)
Requirement already satisfied: certifi>=2017.4.17 in c:\programdata\anaconda2\li
b\site-packages (from requests>=2.11.1->tweepy)
Installing collected packages: oauthlib, requests-oauthlib, tweepy
Successfully installed oauthlib-2.0.7 requests-oauthlib-0.8.0 tweepy-3.6.0
You are using pip version 9.0.1, however version 10.0.1 is available.
You should consider upgrading via the 'python -m pip install --upgrade pip' comm
and.

(C:\ProgramData\Anaconda2) C:\Users\DELL>
```

Create .bat file to generate tweets:

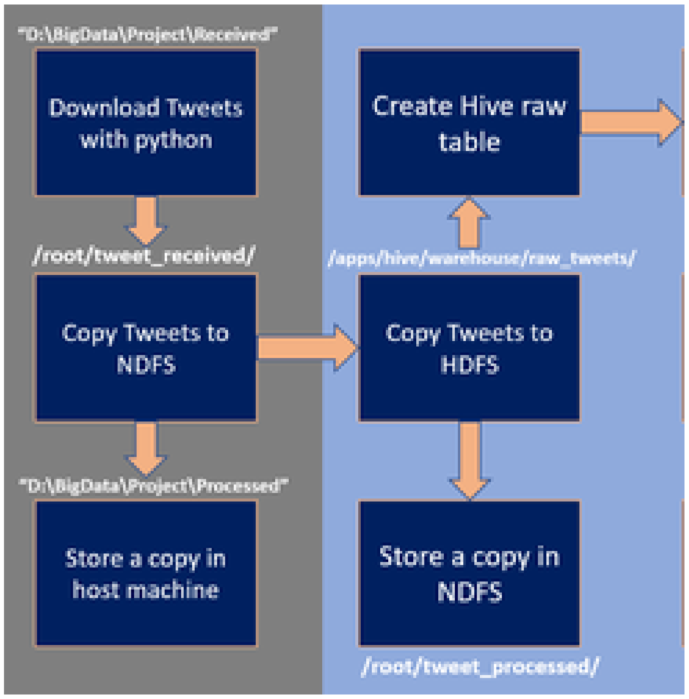
The tweet fetching and transfer has to be done recursively. So, we’ll create a .bat file for this as it will be helpful in automating the process later. This .bat file will be calling the .py file (will be shared with students).

project_py_3.py

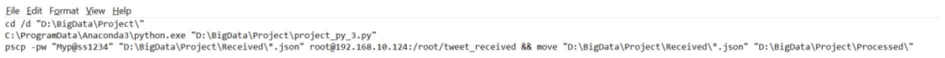
Create a .bat file - just by creating a normal notepad file and saving it with .bat extension - in the folder where you placed your python file. Please make sure not to put them on the desktop. Open the batch file and paste the following code into it:

```
cd /d "path-of-your-python-file"

"path-to-your-python-executable\python.exe" "path-to-pythonfile\name-of-
python-file.py"

pscp -pw "your-vm-root-password" "path-to-where-your-json-filesare-
created\*.json" root@<your-vm-ip-address>:/root/<folder-namewhere-tweets-are
received> && move "path-to-where-your-json-filesare-created\*.json" "path-to-
folder-where-json-files-are-to-bestored-after-processing"
```

So, for example if my directory structure is as below:


"D:\BigData\Project\Received"
Download Tweets with python
/root/tweet_received/
Copy Tweets to NDFS
"D:\BigData\Project\Processed"
Store a copy in host machine
Create Hive raw table
/apps/hive/warehouse/raw_tweets/
Copy Tweets to HDFS
Store a copy in NDFS
/root/tweet_processed/

Then my .bat file would be something like this (assuming I have placed the .py and .bat file in D:\BigData\Project\ )



```
cd /d "D:\BigData\Project\"
C:\ProgramData\Anaconda3\python.exe "D:\BigData\Project\project_py_3.py"
pscp -pw "Myp@ss1234" "D:\BigData\Project\Received\*.json" root@192.168.10.124:/root/tweet_received && move "D:\BigData\Project\Received\*.json" "D:\BigData\Project\Processed\"
```

You can double click the file to run it and it will start downloading tweets.

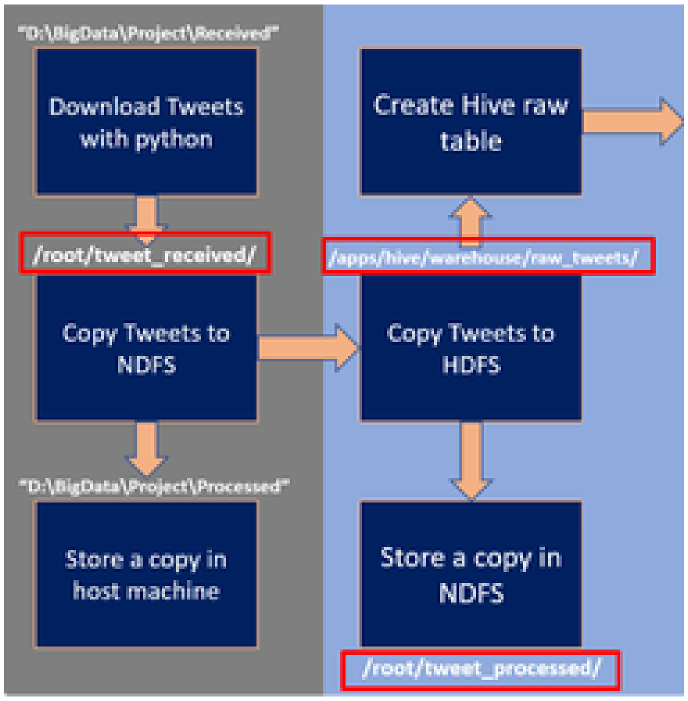
ETL Commands and Hive:

Copy tweets to HDFS :

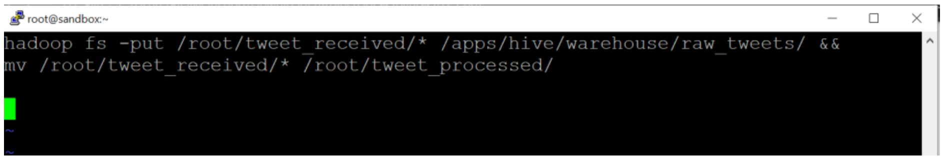
Running the .bat file will copy the tweets to you Linux (NDFS) as well - provided that your VM is running. Next you have to copy it on HDFS to analyze in Hive. For that we will use HDFS put command.

```
hdfs dfs -put ./<folder-where-json-files-are-received-from-laptop>/*
/<destination-folder-in-hdfs> && mv ~<folder-where-json-files-arereceived-
from-laptop>/* ~/<folder-to-put-files-after-processing>
```

for example if my NDFS and HDFS directories are as below:


"D:\BigData\Project\Received"
Download Tweets with python
/root/tweet_received/
Copy Tweets to NDFS
"D:\BigData\Project\Processed"
Store a copy in host machine
Create Hive raw table
/apps/hive/warehouse/raw_tweets/
Copy Tweets to HDFS
Store a copy in NDFS
/root/tweet_processed/

Then my HDFS put command should be like:



```
hadoop fs -put /root/tweet_received/* /apps/hive/warehouse/raw_tweets/ &&
mv /root/tweet_received/* /root/tweet_processed/
```

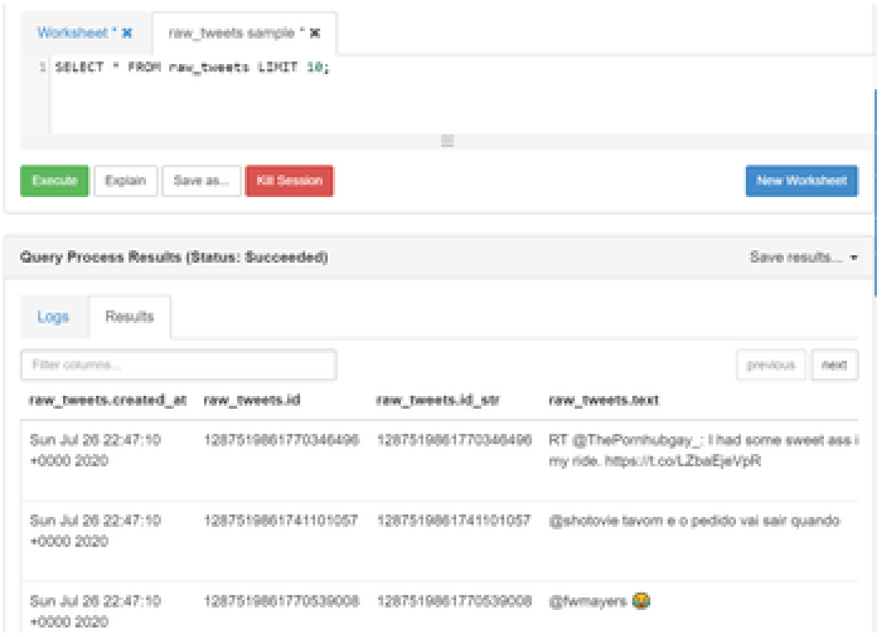
Running the put command will copy all the tweets in tweet_received directory to HDFS' respective directory.
Next, We will create the staging table on top of this HDFS directory.

Creating Hive table:

Now we can create the hive table to load tweets in it. In the query below I have only selected some columns from the .json file to show that we can just load the columns we require later in the process. Use the query or modify it to create the table where tweets will be loaded.

```
CREATE EXTERNAL TABLE IF NOT EXISTS raw_tweets
(
created_at string,
id string,
id_str string,
text string,
source string,
truncated string,
user_tw struct
<
id:string,
id_str:string,
name:string,
screen_name:string,
location:string,
url:string,
description:string,
translator_type:string,
protected:string,
verified:string,
followers_count:string,
friends_count:string,
listed_count:string,
favourites_count:string,
created_at:string,
utc_offset:string,
time_zone:string
>
)
ROW FORMAT SERDE 'org.apache.hive.hcatalog.data.JsonSerDe'
STORED AS TEXTFILE
LOCATION '/apps/hive/warehouse/raw_tweets/' ;
```

Executing this DDL will create the raw table on specified directory. Querying this table should show data as we have already copied it in previous step.



```
SELECT * FROM raw_tweets LIMIT 10;
```

Query Process Results (Status: Succeeded)

raw_tweets.created_at	raw_tweets.id	raw_tweets.id_str	raw_tweets.text
Sun Jul 26 22:47:10 +0000 2020	1287519861770346496	1287519861770346496	RT @ThePornhubgay_: I had some sweet ass i my ride. https://t.co/LZbaEjeVpR
Sun Jul 26 22:47:10 +0000 2020	1287519861741101057	1287519861741101057	@shotovie tavom e o pedido vai sair quando
Sun Jul 26 22:47:10 +0000 2020	1287519861770539008	1287519861770539008	@fwmayers 😭

Create dictionary table :

Next we will create the dictionary table and for that we will use *dictionary.tsv* file. For that copy the dictionary.tsv file to HDFS. I have copied it to /tmp/dictionary/

dictionary.tsv

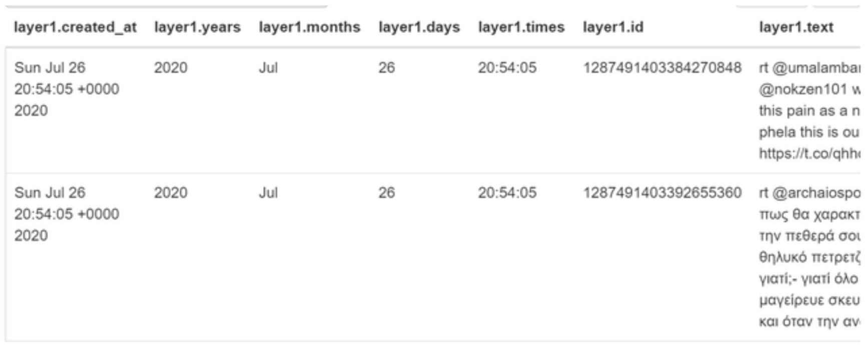
Create the table using the following DDL and load data into this table.

```
CREATE EXTERNAL TABLE my_dictionary (
    type string,
    length int,
    word string,
    pos string,
    stemmed string,
    polarity string
)
ROW FORMAT DELIMITED FIELDS TERMINATED BY '\t'
STORED AS TEXTFILE;

--Following will load data from HDFS directory into Hive table.
LOAD DATA INPATH '/tmp/dictionary/' OVERWRITE INTO TABLE
default.my_dictionary;
```

Create L1, L2 and L3 views :

To analyze the sentiment of a tweet, we need to break it down into words so that we can match the sentiment of words from dictionary table.
First, we will create Layer 1 view which will extract date elements, convert the text to lowercase and remove any new line characters.

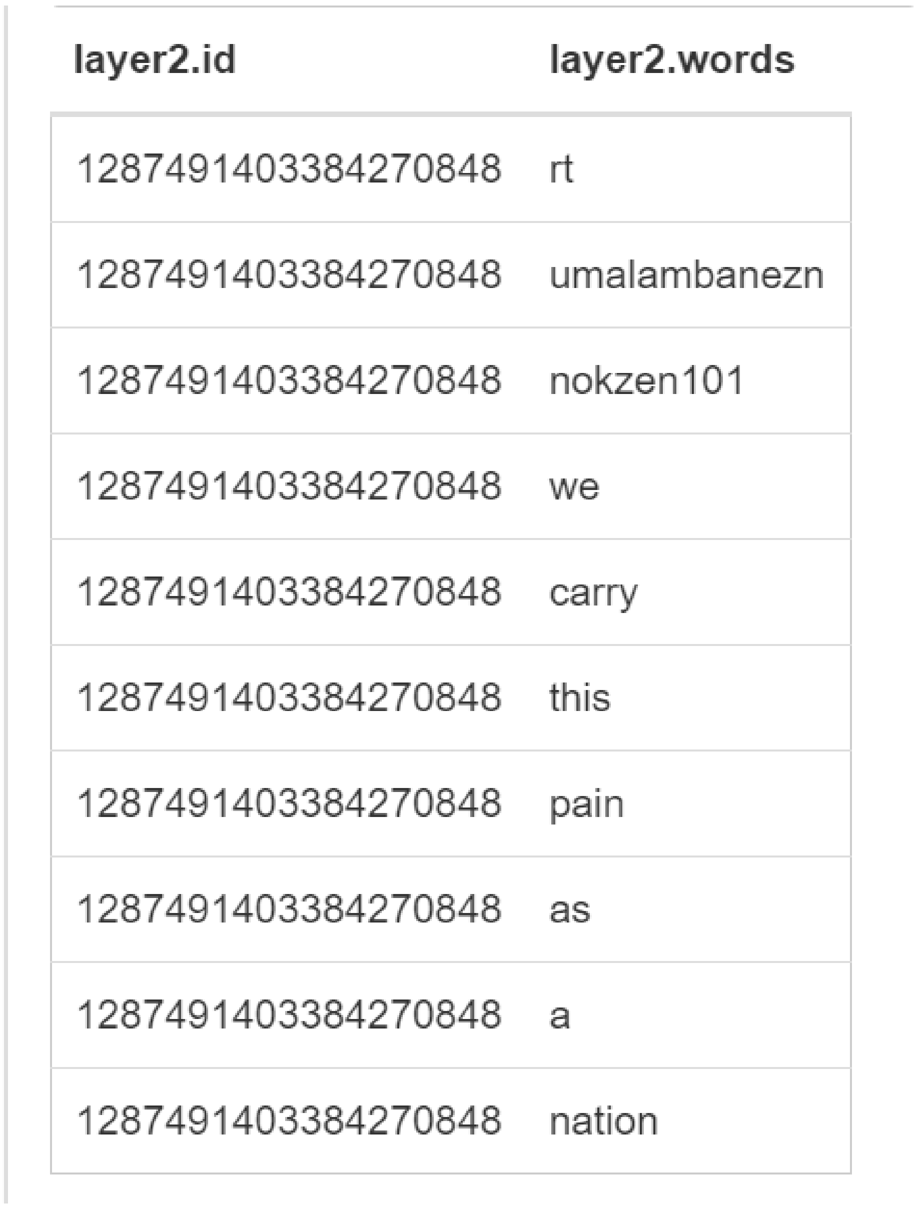
```
create or replace view default.Layer1 AS
select
created_at,
substr(created_at,27,4) as years,
substr(created_at,5,3) as months,
substr(created_at,9,2) as days,
substr(created_at,12,8) as times,
id,
lower(regexp_replace(text,'\n','')) as text
from default.raw_tweets;
```



layer1.created_at	layer1.years	layer1.months	layer1.days	layer1.times	layer1.id	layer1.text
Sun Jul 26 20:54:05 +0000 2020	2020	Jul	26	20:54:05	1287491403384270848	rt @umalambai @nokzen101 w this pain as a n phela this is ou https://t.co/qhh
Sun Jul 26 20:54:05 +0000 2020	2020	Jul	26	20:54:05	1287491403392655360	rt @archaiospo πως θα χαρακτ την πεθερά σοι θηλυκό πετρετζ γιατί;- γιατί όλο μαγείρευε σκευ και όταν την αν

Layer 2 view will explode the tweet by ID and separate each word into new line.

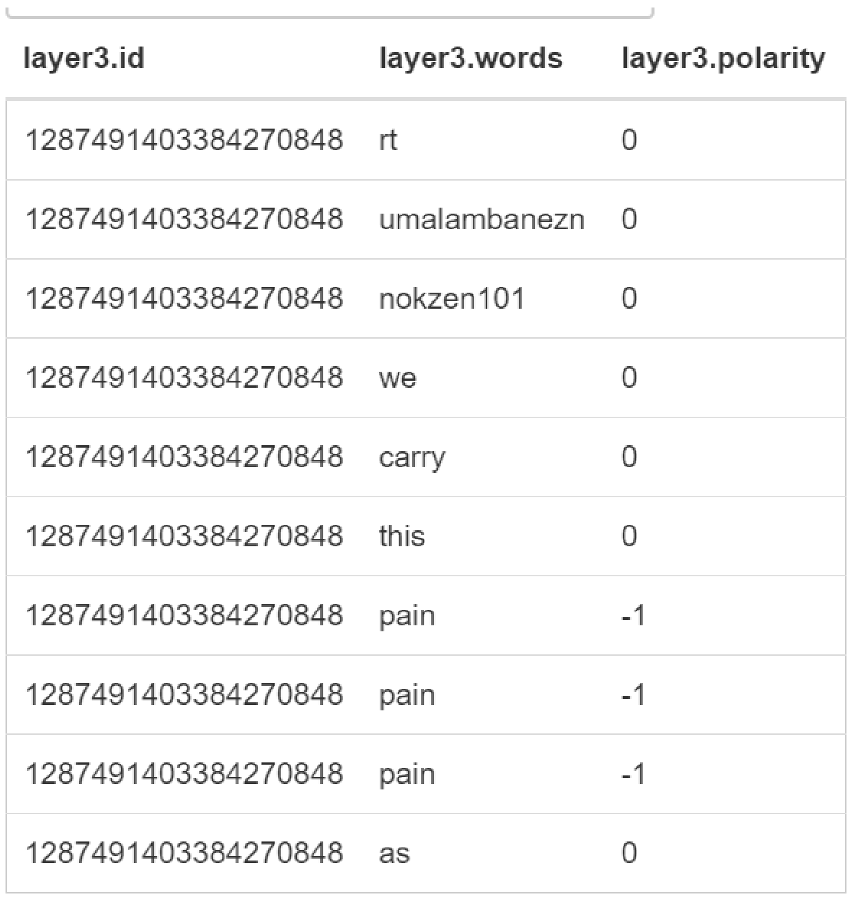
```
create or replace view default.Layer2 AS
select id, words
from default.Layer1
lateral view explode(split(text,'\\W+')) text as words
;
```



layer2.id	layer2.words
1287491403384270848	rt
1287491403384270848	umalambanezn
1287491403384270848	nokzen101
1287491403384270848	we
1287491403384270848	carry
1287491403384270848	this
1287491403384270848	pain
1287491403384270848	as
1287491403384270848	a
1287491403384270848	nation

Layer 3 view matches each word with the dictionary table picks whether polarity is negative or positive and assigns a value -1 or +1 respectively.

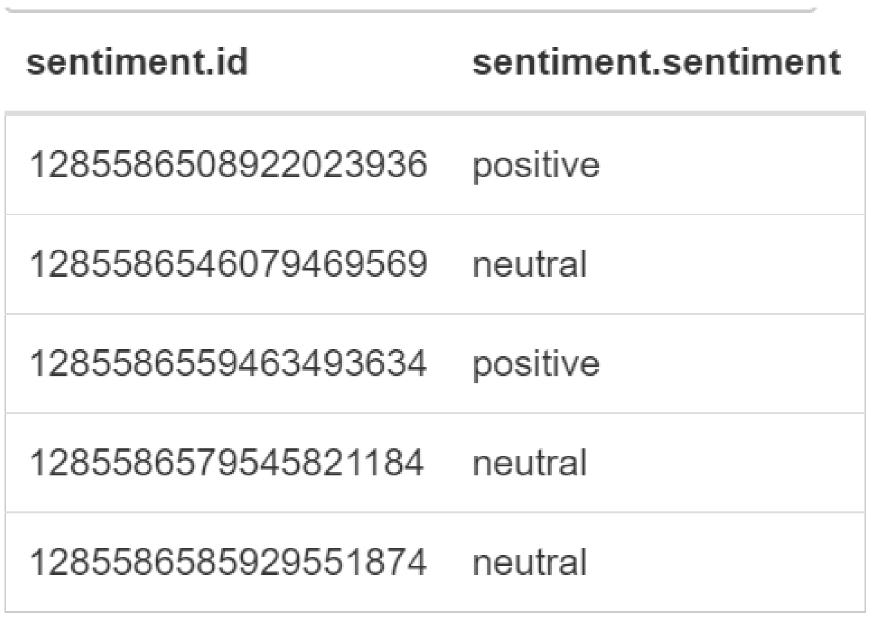
```
Create or Replace view default.Layer3 AS
select
id, L2.words,
case d.polarity
        when 'negative' then -1
        when 'positive' then 1
        else 0 end
as polarity
from Layer2 L2 left outer join my_dictionary d
on L2.words=d.word
;
```



layer3.id	layer3.words	layer3.polarity
1287491403384270848	rt	0
1287491403384270848	umalambanezn	0
1287491403384270848	nokzen101	0
1287491403384270848	we	0
1287491403384270848	carry	0
1287491403384270848	this	0
1287491403384270848	pain	-1
1287491403384270848	pain	-1
1287491403384270848	pain	-1
1287491403384270848	as	0

Finally, Layer 4 will sum up the polarity against a particular ID which generates the collective sentiment of the tweet with that ID. If you are comfortable with writing complex queries, you can also combine Layer 3 and Layer 4 views and just create one view.

```
create or replace view default.sentiment as
select
  id,
  case
    when sum( polarity ) > 0 then 'positive'
    when sum( polarity ) < 0 then 'negative'
    else 'neutral' end as sentiment
 from layer3 l3 group by id;
```



sentiment.id	sentiment.sentiment
1285586508922023936	positive
1285586546079469569	neutral
1285586559463493634	positive
1285586579545821184	neutral
1285586585929551874	neutral

Now, We have the sentiment of each tweet in *default.sentiment* table/view grouped by each ID. If you wish to combine it with the Raw table or Layer 1 view/table to find the actual content of tweet. You can do so using the following query or create a final table using a CTAS statement.

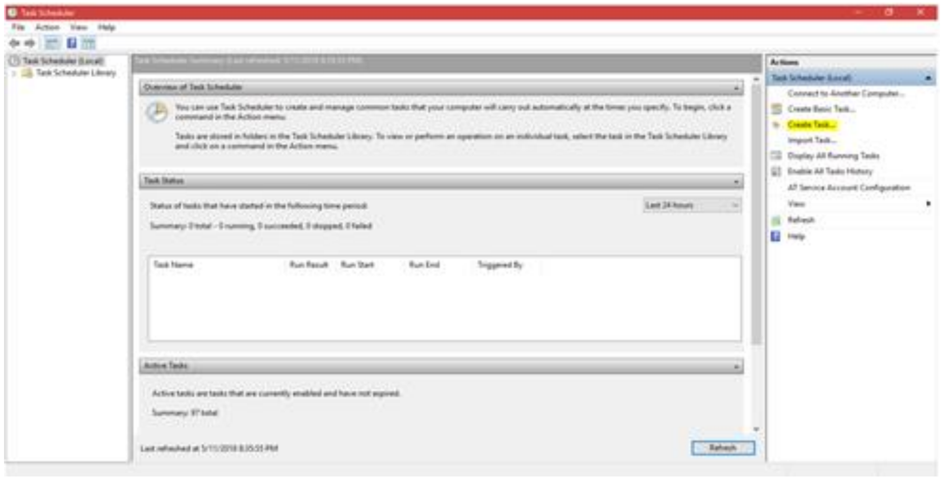
```
SELECT
  L1.*, s.sentiment
FROM layer1 L1 LEFT OUTER JOIN sentiment s on L1.id = s.id
;
```

Automation:

To automate this pipeline you just need to schedule the .bat file and HDFS put command. Rest will be done automatically as views will populate once underlying data does.

Schedule .bat file:

Open up Windows Task Scheduler. It is a default program of Windows and is already installed into your system. Click on Create Task on the right hand side of the window to create a new task:



Task Scheduler

File Action View Help

Task Scheduler (Local)
Task Scheduler Library

Overview of Task Scheduler

You can use Task Scheduler to create and manage common tasks that your computer will carry out automatically at the times you specify. To begin, click a command in the Action menu.

Tasks are stored in folders in the Task Scheduler Library. To view or perform an operation on an individual task, select the task in the Task Scheduler Library and click on a command in the Action menu.

Task Status

Status of tasks that have started in the following time period: Last 24 hours

Summary: 0 total - 0 running, 0 succeeded, 0 stopped, 0 failed

Task Name	Run Result	Run Start	Run End	Triggered By

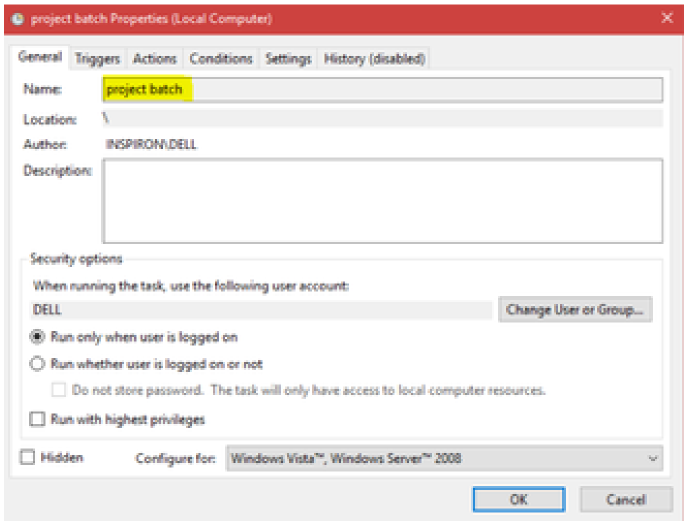
Active Tasks

Active tasks are tasks that are currently enabled and have not expired.

Summary: 97 total

Last refreshed at 5/11/2018 8:35:55 PM

On the General tab enter the name you want to give to the task:



project batch Properties (Local Computer)

Name: project batch

Location: \

Author: INSPIRON\DELL

Description:

Security options

When running the task, use the following user account:

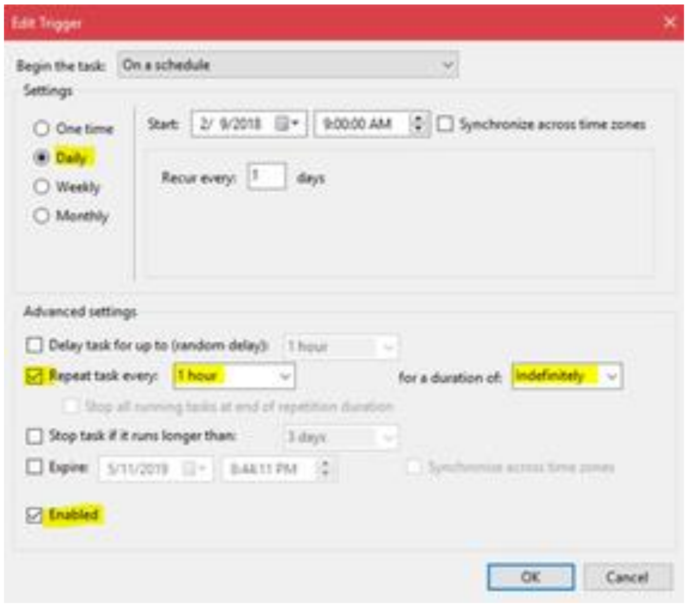
DELL

- [✓] Run only when user is logged on
- Run whether user is logged on or not
- Do not store password. The task will only have access to local computer resources.
- Run with highest privileges

Hidden

Configure for: Windows Vista™, Windows Server™ 2008

On the Triggers tab Click on New to create a new Trigger. Set scheduler to Daily, Repeat to one hour and indefinitely and check enabled as shown in the screenshot below. Click OK and the task will appear in the Triggers tab.



Edit Trigger

Begin the task: On a schedule

Settings

( ) One time
(●) Daily
( ) Weekly
( ) Monthly

Start: 2/ 9/2018 9:00:00 AM ☐ Synchronize across time zones

Recur every: 1 days

Advanced settings

☐ Delay task for up to (random delay): 1 hour

☑ Repeat task every: 1 hour for a duration of: Indefinitely

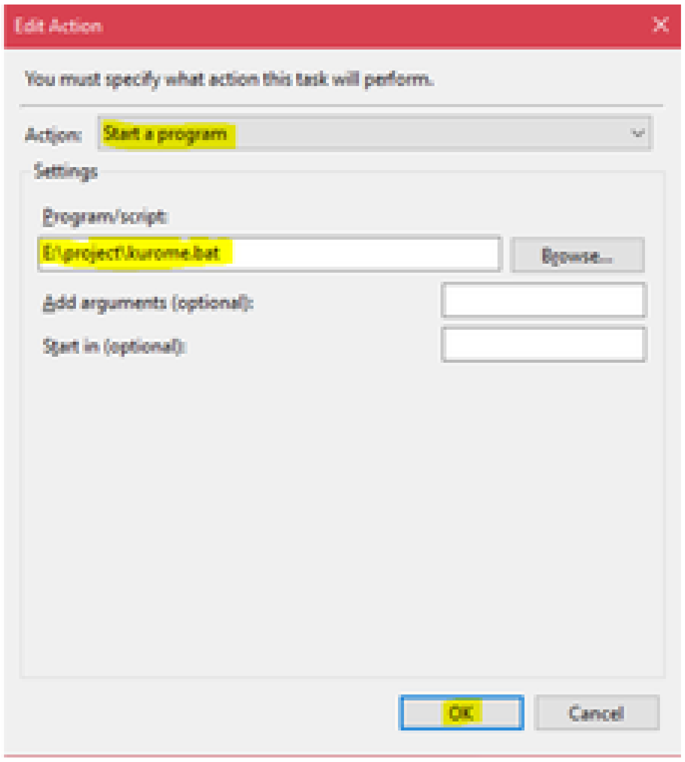
☐ Stop all running tasks at end of repetition duration

☐ Stop task if it runs longer than: 3 days

☐ Expire: 5/11/2019 8:44:11 PM ☐ Synchronize across time zones

☑ Enabled

Go to the Actions tab and Click New. Choose action as Start a Program and give the path to the batch file in the field specified. Click Ok to save it.



Edit Action

You must specify what action this task will perform.

Action: Start a program

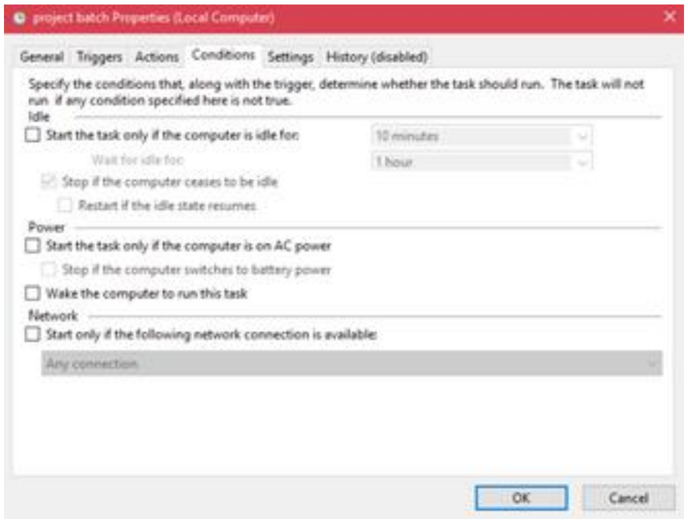
Settings

Program/script: E:\project\kurome.bat

Add arguments (optional):

Start in (optional):

On the Conditions tab uncheck everything. Click OK to create the Task


project batch Properties (Local Computer)

General | Triggers | Actions | Conditions | Settings | History (disabled)

Specify the conditions that, along with the trigger, determine whether the task should run. The task will not run if any condition specified here is not true.

Idle
- [ ] Start the task only if the computer is idle for: 10 minutes
  - Wait for idle for: 1 hour
  - [x] Stop if the computer ceases to be idle
    - [ ] Restart if the idle state resumes

Power
- [ ] Start the task only if the computer is on AC power
  - [ ] Stop if the computer switches to battery power
- [ ] Wake the computer to run this task

Network
- [ ] Start only if the following network connection is available:
  - Any connection

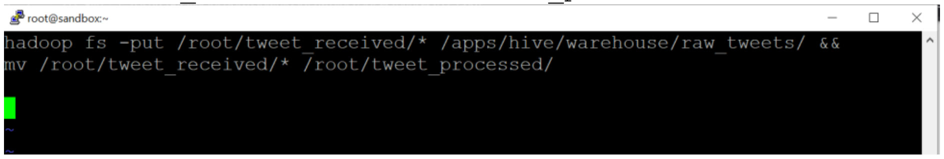
OK | Cancel

According to the time you have set in the Triggers tab the task will activate and run the batch file. Make sure the IP address in the batch file is updated every time the VM is started as the IP address changes every time.

Schedule the HDFS file movement:

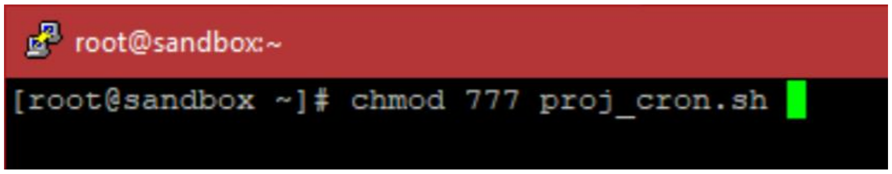
For the HDFS file movement we will use cron scheduler that is a built-in tool for Linux environment just like the windows scheduler.
First of all, create a .sh file using vi editor and paste the command that we used above to copy files from NDFS to HDFS and save this file. Give this files executable permissions as well.

```
hadoop fs -put /root/tweet_received/* /apps/hive/warehouse/raw_tweets/ && mv
/root/tweet_received/* /root/tweet_processed/
```



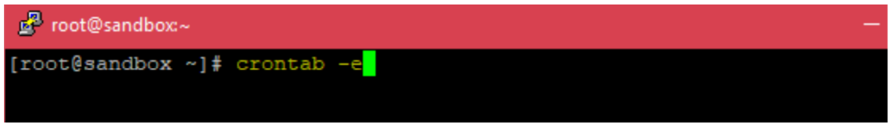
```
root@sandbox:~
hadoop fs -put /root/tweet_received/* /apps/hive/warehouse/raw_tweets/ &&
mv /root/tweet_received/* /root/tweet_processed/

~
~
```

```
chmod 777 <shell-file-name>.sh
```



```
root@sandbox:~
[root@sandbox ~]# chmod 777 proj_cron.sh
```

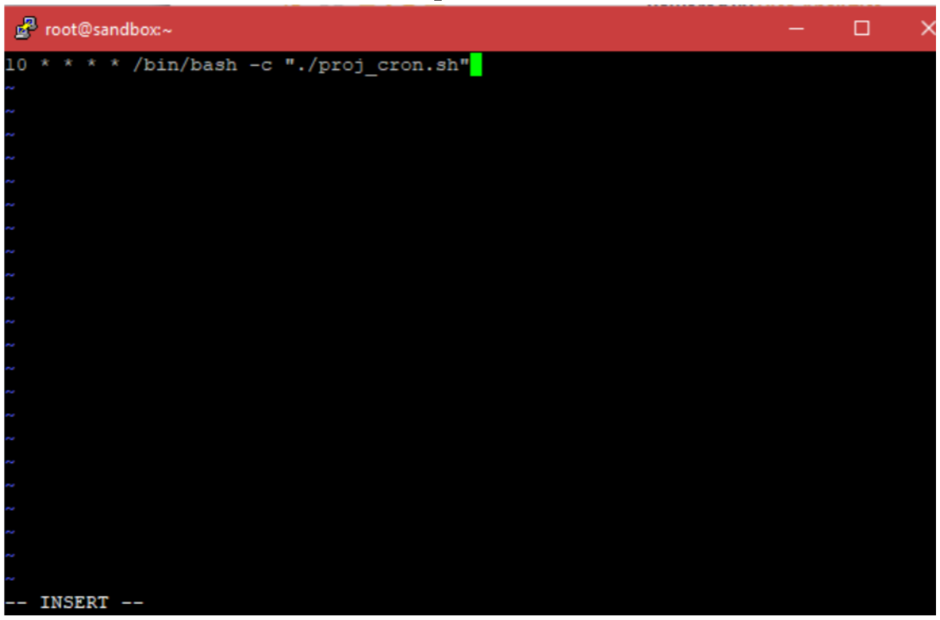
On the shell enter the command *crontab -e* to edit the cron list and add a new cronjob.



```
[root@sandbox ~]# crontab -e
```

The text editor will open up. Enter I to get into insert mode and add the following line:

```
10 * * * * /bin/bash -c "/<your-desired-file-name>.sh"
```



```
10 * * * * /bin/bash -c "./proj_cron.sh"
~
~
~
~
~
~
~
~
~
~
~
~
~
~
~
~
~
~
~
~
~
~
-- INSERT --
```

The above line shows that your scheduled job will run at 10th minute of every hour.
Press Esc and enter :wq to save the cron job.
To view the installed cron job you can enter crontab -l in the shell.

Creating a visualization:

Get creative and create a visualization by connecting Hive with Tableau using the ODBC connector. This task has been left to students so be creative.