DATA MINING

ASSIGNMENT: MILESTONE 4

NORAISHA YUSUF (WQD180008)

UPDATE

Crawl tweets Check sentiment and interpret tweets Merge dataset

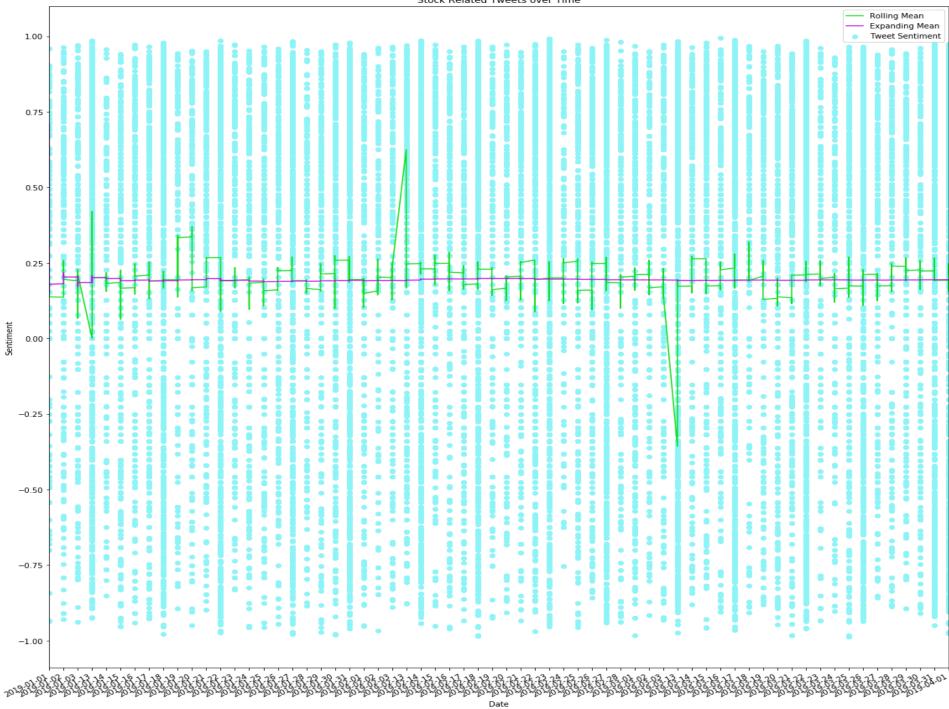


Part 1: Interpreting tweets

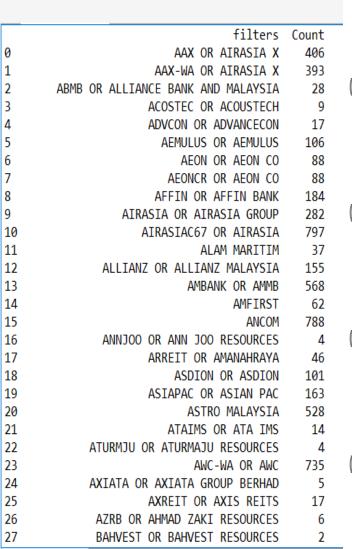
• Sample tweets

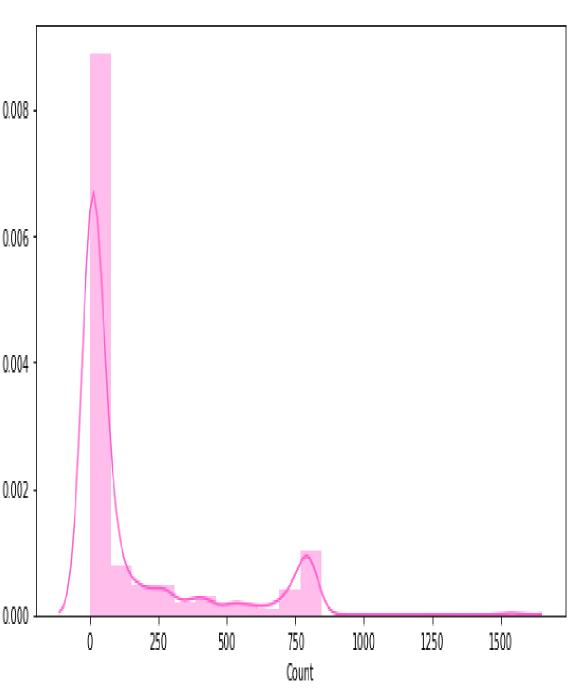
2/1/2019 1:00 POS MALAYSIA Pos Malay 1.09E+18 0 0 0 friyay! (@ DRB-Hicom Defence Services) https://www.swarmapp.com/c/h4lksTjaWLRÂ 2/1/2019 1:00 POS MALAYSIA Pos Malay 1.09E+18 5 8 19 [ANNOUNCEMENT] 2/1/2019 1:01 MALAYSIA AIRPORT Firdaus Ha 2 class=" 1.09E+18 0 0 1 AirAsia ups claim against Malaysia Airports by MYR80 million https://www.flightglobal.com/news/articles/airasia-ups-claim-against-malaysia-airpor 2/1/2019 1:01 AEMULUS OR AEMI emi 2 class=" 1.09E+18 0 0 0 im starting a petition for @JTHS_204 to cancel school tomorrow 2/1/2019 1:01 CRESBLD OR CREST StarBiz 1.09E+18 0 0 0 Crest Builder unit wins RM99.6mil construction job 2/1/2019 1:11 GKENT OR GEORGE Saeger Ryr 1.09E+18 14 0 0 Considering how one-dimensional and boring Superman *could* be, he's been blessed with a lot of outstanding actors. 2/1/2019 1:24 TENAGA NASIONAL EV Club Mi 1.09E+18 1 0 Seems like a no-brainer for @Tenaga_Nasional to offer something similar, or at least a one-time rebate for home EVSE installation. Like the Gillette of the contraction of the contracti																
2/1/2019 0:38 AMBANK OR AMMÉEL I. I. A. Â-ç class=" 1.09E+18	2/1/2019 0:01	PESTECH OR PESTE	StarBiz <	<p< td=""><td>1.09E+18</td><td>1</td><td>0</td><td>2 Pestech secures TNB contracts worth RM280.89mil</td></p<>	1.09E+18	1	0	2 Pestech secures TNB contracts worth RM280.89mil								
2/1/2019 0.49 BIOHLDG OR BIOAL Sable Mug <a 0="" 0.49="" 1="" 2="" 2019="" a="" and="" back="" class="1.09E+18" drb="" drbhcom="" ds="" emulator="" for="" i've="" in="" it="" it's="" just="" louder="" neiver="" on="" or="" people="" played="" ps2="" râ*hâ*h<="" say="" the=""> 0 0 10 friyay! (@ DRBHCOM OR DRB Râ*Hâ*H 0 0 0 friyay! (@ DRBHCOM OR DRB Râ*Hâ*H 1.09E+18 0 0 0 friyay! (@ DRBHCOM OR DRB Râ*Hâ*H 1.09E+18 5 8 19 [ANNOUNCEMENT] 2/1/2019 1:00 FOR MALAYSIA POS Malay 1.09E+18 2 0 1 AirAsia ups claim against Malaysia Airports by MYR80 million https://www.swarmapp.com/c/h4lksTjaWLRÂ 2/1/2019 1:01 REMULUS OR AEMIlemi	2/1/2019 0:34	SAM ENGINEERING	G Samuel.J.E <	<p< td=""><td>1.09E+18</td><td>0</td><td>0</td><td>0 Oje wa ka eng (engineering edition)</td></p<>	1.09E+18	0	0	0 Oje wa ka eng (engineering edition)								
2/1/2019 0:48 PHEFO OR HONG LEC jediMASTE c Joses*** 1.09E+18 0 0 2 And this Datin Sri isriaf*** transmited. And she only sells insurance for Hong Leong! Now wherea£**s Tax Revenue? How did she manage to own shares 2/1/2019 1:00 POS MALAYSIA Pos Malay 1.09E+18 2/1/2019 1:00 MALAYSIA AIRPORT Firidaus Ha 2 p class**** 1.09E+18 0 0 1.09E+18 0 0 1.09E+18 0 <p< td=""><td>2/1/2019 0:38</td><td>AMBANK OR AMM</td><td>IE E. L. I. A. Â <</td><td><p 0="" 1="" 1.09e+18="" 1:01="" 2="" 2019="" aemi="" aemulus="" against="" airasia="" airasia-ups-claim-against-malaysia-airpor="" airports="" articles="" by="" claim="" class="</td><td>1.09E+18</td><td>1</td><td>0</td><td>0 I love you @ammb_nicole pic.twitter.com/bLKtUSrXBd</td></tr><tr><td>2/1/2019 1:00 POS MALAYSIA POS Malay () 1.09E+18 5 8 19 [ANNOUNCEMENT] 2/1/2019 1:00 MALAYSIA AIRPORT Firdaus Ha (> class=" emi<="" https:="" malaysia="" million="" myr80="" news="" or="" td="" ups="" www.flightglobal.com=""><td>2/1/2019 0:39</td><td>BIOHLDG OR BIOA</td><td>L Sable Mug <</td><td><p 0="" 1="" 1.09e+18="" 1:01="" 2="" 2019="" aemilemi<="" against="" airasia="" airasia-ups-claim-against-malaysia-airport="" airports="" articles="" by="" claim="" class="</td><td>1.09E+18</td><td>0</td><td>0</td><td>0 Say it louder for the people in the back and I've never played it on the PS2 just it's emulator and DS</td></tr><tr><td>2/1/2019 1:00 POS MALAYSIA POS Malay < D. 1.09E+18 5 8 19 [ANNOUNCEMENT] 2/1/2019 1:00 MALAYSIA AIRPORT Firdaus Ha < p class=" exenulos="" https:="" malaysia="" million="" myr80="" news="" or="" td="" ups="" www.flightglobal.com=""><td>2/1/2019 0:45</td><td>HLFG OR HONG LE</td><td>(jediMASTE <</td><td><p <="" class="</td><td>1.09E+18</td><td>0</td><td>0</td><td>2 And this Datin Sri isn't married. And she only sells insurance for Hong Leong! Now where's Tax Revenue? How did she manage to own shares</td></tr><tr><td>2/1/2019 1:01 AEMULUS OR AEMIemi</td><td>2/1/2019 0:48</td><td>DRBHCOM OR DRE</td><td>B Râ^†Bâ^†k</td><td><p class=" td=""><td>1.09E+18</td><td>0</td><td>0</td><td>0 friyay! (@ DRB-Hicom Defence Services) https://www.swarmapp.com/c/h4lksTjaWLRÂ</td></p></td></p></td></p></td></p<>	2/1/2019 0:38	AMBANK OR AMM	IE E. L. I. A. Â <	<p 0="" 1="" 1.09e+18="" 1:01="" 2="" 2019="" aemi="" aemulus="" against="" airasia="" airasia-ups-claim-against-malaysia-airpor="" airports="" articles="" by="" claim="" class="</td><td>1.09E+18</td><td>1</td><td>0</td><td>0 I love you @ammb_nicole pic.twitter.com/bLKtUSrXBd</td></tr><tr><td>2/1/2019 1:00 POS MALAYSIA POS Malay () 1.09E+18 5 8 19 [ANNOUNCEMENT] 2/1/2019 1:00 MALAYSIA AIRPORT Firdaus Ha (> class=" emi<="" https:="" malaysia="" million="" myr80="" news="" or="" td="" ups="" www.flightglobal.com=""><td>2/1/2019 0:39</td><td>BIOHLDG OR BIOA</td><td>L Sable Mug <</td><td><p 0="" 1="" 1.09e+18="" 1:01="" 2="" 2019="" aemilemi<="" against="" airasia="" airasia-ups-claim-against-malaysia-airport="" airports="" articles="" by="" claim="" class="</td><td>1.09E+18</td><td>0</td><td>0</td><td>0 Say it louder for the people in the back and I've never played it on the PS2 just it's emulator and DS</td></tr><tr><td>2/1/2019 1:00 POS MALAYSIA POS Malay < D. 1.09E+18 5 8 19 [ANNOUNCEMENT] 2/1/2019 1:00 MALAYSIA AIRPORT Firdaus Ha < p class=" exenulos="" https:="" malaysia="" million="" myr80="" news="" or="" td="" ups="" www.flightglobal.com=""><td>2/1/2019 0:45</td><td>HLFG OR HONG LE</td><td>(jediMASTE <</td><td><p <="" class="</td><td>1.09E+18</td><td>0</td><td>0</td><td>2 And this Datin Sri isn't married. And she only sells insurance for Hong Leong! Now where's Tax Revenue? How did she manage to own shares</td></tr><tr><td>2/1/2019 1:01 AEMULUS OR AEMIemi</td><td>2/1/2019 0:48</td><td>DRBHCOM OR DRE</td><td>B Râ^†Bâ^†k</td><td><p class=" td=""><td>1.09E+18</td><td>0</td><td>0</td><td>0 friyay! (@ DRB-Hicom Defence Services) https://www.swarmapp.com/c/h4lksTjaWLRÂ</td></p></td></p></td></p>	2/1/2019 0:39	BIOHLDG OR BIOA	L Sable Mug <	<p 0="" 1="" 1.09e+18="" 1:01="" 2="" 2019="" aemilemi<="" against="" airasia="" airasia-ups-claim-against-malaysia-airport="" airports="" articles="" by="" claim="" class="</td><td>1.09E+18</td><td>0</td><td>0</td><td>0 Say it louder for the people in the back and I've never played it on the PS2 just it's emulator and DS</td></tr><tr><td>2/1/2019 1:00 POS MALAYSIA POS Malay < D. 1.09E+18 5 8 19 [ANNOUNCEMENT] 2/1/2019 1:00 MALAYSIA AIRPORT Firdaus Ha < p class=" exenulos="" https:="" malaysia="" million="" myr80="" news="" or="" td="" ups="" www.flightglobal.com=""><td>2/1/2019 0:45</td><td>HLFG OR HONG LE</td><td>(jediMASTE <</td><td><p <="" class="</td><td>1.09E+18</td><td>0</td><td>0</td><td>2 And this Datin Sri isn't married. And she only sells insurance for Hong Leong! Now where's Tax Revenue? How did she manage to own shares</td></tr><tr><td>2/1/2019 1:01 AEMULUS OR AEMIemi</td><td>2/1/2019 0:48</td><td>DRBHCOM OR DRE</td><td>B Râ^†Bâ^†k</td><td><p class=" td=""><td>1.09E+18</td><td>0</td><td>0</td><td>0 friyay! (@ DRB-Hicom Defence Services) https://www.swarmapp.com/c/h4lksTjaWLRÂ</td></p></td></p>	2/1/2019 0:45	HLFG OR HONG LE	(jediMASTE <	<p <="" class="</td><td>1.09E+18</td><td>0</td><td>0</td><td>2 And this Datin Sri isn't married. And she only sells insurance for Hong Leong! Now where's Tax Revenue? How did she manage to own shares</td></tr><tr><td>2/1/2019 1:01 AEMULUS OR AEMIemi</td><td>2/1/2019 0:48</td><td>DRBHCOM OR DRE</td><td>B Râ^†Bâ^†k</td><td><p class=" td=""><td>1.09E+18</td><td>0</td><td>0</td><td>0 friyay! (@ DRB-Hicom Defence Services) https://www.swarmapp.com/c/h4lksTjaWLRÂ</td></p>	1.09E+18	0	0	0 friyay! (@ DRB-Hicom Defence Services) https://www.swarmapp.com/c/h4lksTjaWLRÂ
2/1/2019 1:01 CRESBLD OR CREST StarBiz	2/1/2019 1:00	POS MALAYSIA	Pos Malay	<p< td=""><td>1.09E+18</td><td>5</td><td>8</td><td>19 [ANNOUNCEMENT]</td></p<>	1.09E+18	5	8	19 [ANNOUNCEMENT]								
2/1/2019 1:01 CRESBLD OR CREST StarBiz <p *could*="" 0="" 1="" 1.09e+18="" 14="" 1:11="" 1:24="" 2="" 2019="" <p="" @tenaga_nasional="" a="" actors.="" and="" at="" be,="" been="" blessed="" boring="" builder="" club="" considering="" construction="" crest="" ev="" evse="" for="" george="" gillette="" gkent="" gran<="" grant="" he's="" home="" how="" installation.="" job="" least="" like="" lot="" m="" nasional="" no-brainer="" of="" offer="" one-dimensional="" one-time="" or="" outstanding="" rebate="" results="" rm99.6mil="" ryl="" saeger="" seems="" similar,="" something="" superman="" td="" tenaga="" the="" to="" unit="" wins="" with=""><td>2/1/2019 1:00</td><td>MALAYSIA AIRPOR</td><td>T Firdaus Ha <</td><td><p 0="" 1="" 1.09e+18="" 11-15="" 1st="" 2="" 2019="" 3="" 4:47="" 4:48="" 4t="" 5:05="" <="" a="" a!="" advance="" air="" airport="" al="" all="" and="" andersson="" associa="" be="" blueline="" bursa="" calibre="" can="" canâ€"="" class="1.09E+18 8 0 Owe advance to winners match of group A! 1.09E+18 8 0 Owe advance to winners match of group A! 2/1/2019 5:15 MALAYSIA AIRPORT TheMalays < p class=" closed="" closing="" course="" cpo="" day.="" diving="" explore="" factor="" feb="" february="" federal="" for="" friday="" fun="" future="" futures="" group="" hong="" if="" imca="" in="" is="" kb="" kl.="" kylington="" learner="" learning="" leong="" malaysia="" malaysia,="" match="" matter="" mhb="" monday="" more="" necessary="" nhl.="" nhlers="" no="" of="" oil="" on="" one="" or="" owe="" p="" palm="" registration="" resumes="" s="" scheduled="" should="" soon="" supervisor="" td="" territory="" that="" the="" themalays="" then="" there="" to="" today="" too!="" town="" trading="" trading.="" trainee="" valimaki,="" verlook="" wanted="" what="" will="" winners="" with="" you="" you&€"re=""><td>2/1/2019 1:01</td><td>AEMULUS OR AEM</td><td>Nemi <</td><td><p #dataprep<="" #dow="" #sp500="" 0="" 0.06="" 0.86="" 1="" 1"="" 1.09e+18="" 11-15="" 15.19="" 17="" 1st="" 2="" 2,704.10="" 2019="" 23.05="" 24,999.67.="" 3="" 4:47="" 4:48="" 4:50="" 4:59="" 4t="" 5:15="" 7="" <="" @gcpcloud="" a="" advance="" airport="" al="" all="" alr="" and="" andersson="" associa="" average="" be="" blueline="" bursa="" calibre="" can="" can't="" cent,="" class="1" closed="" closing="" course="" cpo="" darren="" datai="" dataprp="" day.="" diving="" explore="" factor="" feb="" february="" federal="" fell="" for="" friday="" fun="" future="" futures="" grouf="" group="" har="" hlfg="" hong="" if="" imca="" impressed="" in="" increased="" industrial="" is="" jones="" kb="" kelington="" kl="" kylington="" langi="" learner="" learning="" lec="" leong="" lynn="" malaysia="" malaysia,="" match="" matter="" me="" mhb="" monday="" more="" necessary="" nhl.="" nhlers="" no="" of="" oil="" on="" one="" or="" overlook="" p="" palm="" per="" points,="" registration="" resumes="" s="" scheduled="" should="" sl="" so="" soon="" sp50(ians="" sp500-hg="" supervisor="" t's="" td="" territory="" that="" the="" themalays="" then="" there="" to="" today="" tool="" trading="" trading.="" trainee="" twee="" valimaki,="" wanted="" what="" will="" winners="" with="" you="" you're="" =""><td>2/1/2019 1:01</td><td>CRESBLD OR CRES</td><td>T StarBiz <</td><td><p< td=""><td>1.09E+18</td><td>0</td><td>0</td><td>0 Crest Builder unit wins RM99.6mil construction job</td></p<></td></p></td></p></td></p>	2/1/2019 1:00	MALAYSIA AIRPOR	T Firdaus Ha <	<p 0="" 1="" 1.09e+18="" 11-15="" 1st="" 2="" 2019="" 3="" 4:47="" 4:48="" 4t="" 5:05="" <="" a="" a!="" advance="" air="" airport="" al="" all="" and="" andersson="" associa="" be="" blueline="" bursa="" calibre="" can="" canâ€"="" class="1.09E+18 8 0 Owe advance to winners match of group A! 1.09E+18 8 0 Owe advance to winners match of group A! 2/1/2019 5:15 MALAYSIA AIRPORT TheMalays < p class=" closed="" closing="" course="" cpo="" day.="" diving="" explore="" factor="" feb="" february="" federal="" for="" friday="" fun="" future="" futures="" group="" hong="" if="" imca="" in="" is="" kb="" kl.="" kylington="" learner="" learning="" leong="" malaysia="" malaysia,="" match="" matter="" mhb="" monday="" more="" necessary="" nhl.="" nhlers="" no="" of="" oil="" on="" one="" or="" owe="" p="" palm="" registration="" resumes="" s="" scheduled="" should="" soon="" supervisor="" td="" territory="" that="" the="" themalays="" then="" there="" to="" today="" too!="" town="" trading="" trading.="" trainee="" valimaki,="" verlook="" wanted="" what="" will="" winners="" with="" you="" you&€"re=""><td>2/1/2019 1:01</td><td>AEMULUS OR AEM</td><td>Nemi <</td><td><p #dataprep<="" #dow="" #sp500="" 0="" 0.06="" 0.86="" 1="" 1"="" 1.09e+18="" 11-15="" 15.19="" 17="" 1st="" 2="" 2,704.10="" 2019="" 23.05="" 24,999.67.="" 3="" 4:47="" 4:48="" 4:50="" 4:59="" 4t="" 5:15="" 7="" <="" @gcpcloud="" a="" advance="" airport="" al="" all="" alr="" and="" andersson="" associa="" average="" be="" blueline="" bursa="" calibre="" can="" can't="" cent,="" class="1" closed="" closing="" course="" cpo="" darren="" datai="" dataprp="" day.="" diving="" explore="" factor="" feb="" february="" federal="" fell="" for="" friday="" fun="" future="" futures="" grouf="" group="" har="" hlfg="" hong="" if="" imca="" impressed="" in="" increased="" industrial="" is="" jones="" kb="" kelington="" kl="" kylington="" langi="" learner="" learning="" lec="" leong="" lynn="" malaysia="" malaysia,="" match="" matter="" me="" mhb="" monday="" more="" necessary="" nhl.="" nhlers="" no="" of="" oil="" on="" one="" or="" overlook="" p="" palm="" per="" points,="" registration="" resumes="" s="" scheduled="" should="" sl="" so="" soon="" sp50(ians="" sp500-hg="" supervisor="" t's="" td="" territory="" that="" the="" themalays="" then="" there="" to="" today="" tool="" trading="" trading.="" trainee="" twee="" valimaki,="" wanted="" what="" will="" winners="" with="" you="" you're="" =""><td>2/1/2019 1:01</td><td>CRESBLD OR CRES</td><td>T StarBiz <</td><td><p< td=""><td>1.09E+18</td><td>0</td><td>0</td><td>0 Crest Builder unit wins RM99.6mil construction job</td></p<></td></p></td></p>	2/1/2019 1:01	AEMULUS OR AEM	Nemi <	<p #dataprep<="" #dow="" #sp500="" 0="" 0.06="" 0.86="" 1="" 1"="" 1.09e+18="" 11-15="" 15.19="" 17="" 1st="" 2="" 2,704.10="" 2019="" 23.05="" 24,999.67.="" 3="" 4:47="" 4:48="" 4:50="" 4:59="" 4t="" 5:15="" 7="" <="" @gcpcloud="" a="" advance="" airport="" al="" all="" alr="" and="" andersson="" associa="" average="" be="" blueline="" bursa="" calibre="" can="" can't="" cent,="" class="1" closed="" closing="" course="" cpo="" darren="" datai="" dataprp="" day.="" diving="" explore="" factor="" feb="" february="" federal="" fell="" for="" friday="" fun="" future="" futures="" grouf="" group="" har="" hlfg="" hong="" if="" imca="" impressed="" in="" increased="" industrial="" is="" jones="" kb="" kelington="" kl="" kylington="" langi="" learner="" learning="" lec="" leong="" lynn="" malaysia="" malaysia,="" match="" matter="" me="" mhb="" monday="" more="" necessary="" nhl.="" nhlers="" no="" of="" oil="" on="" one="" or="" overlook="" p="" palm="" per="" points,="" registration="" resumes="" s="" scheduled="" should="" sl="" so="" soon="" sp50(ians="" sp500-hg="" supervisor="" t's="" td="" territory="" that="" the="" themalays="" then="" there="" to="" today="" tool="" trading="" trading.="" trainee="" twee="" valimaki,="" wanted="" what="" will="" winners="" with="" you="" you're="" =""><td>2/1/2019 1:01</td><td>CRESBLD OR CRES</td><td>T StarBiz <</td><td><p< td=""><td>1.09E+18</td><td>0</td><td>0</td><td>0 Crest Builder unit wins RM99.6mil construction job</td></p<></td></p>	2/1/2019 1:01	CRESBLD OR CRES	T StarBiz <	<p< td=""><td>1.09E+18</td><td>0</td><td>0</td><td>0 Crest Builder unit wins RM99.6mil construction job</td></p<>	1.09E+18	0	0	0 Crest Builder unit wins RM99.6mil construction job
2/1/2019 4:47 FBMKLCI-H6P OR B Palm Oil Al class="1" 1.09E+18 2 1 One factor you canꀙt overlook is the calibre of that blueline with Valimaki, Kylington and Andersson all in the NHL. No future NHLers in the group 2/1/2019 4:47 BURSA MALAYSIA Palm Oil Al class="1" 1.09E+18 3 0 Bursa Malaysia is closed today Friday 1st of February for Federal Territory day. There will be no CPO Futures trading. Trading resumes on Monday 4t 2/1/2019 4:48 MHB OR MALAYSIA KB Associa 1.09E+18 1 0 Registration Closing Soon for IMCA Trainee AIR Diving Supervisor course that is scheduled on the 11-15 Feb 2019 in Malaysia, KL. 2/1/2019 4:50 HLFG OR HONG LEC(T'S SL class="1" 1.09E+18 17 0 7 No matter what learning is necessary and can be fun too! If you're a learner and wanted to explore more then you should explore Hong Leong School For Image and School Fo	2/1/2019 1:11	GKENT OR GEORGI	E Saeger Ryr <	<p< td=""><td>1.09E+18</td><td>14</td><td>0</td><td>O Considering how one-dimensional and boring Superman *could* be, he's been blessed with a lot of outstanding actors.</td></p<>	1.09E+18	14	0	O Considering how one-dimensional and boring Superman *could* be, he's been blessed with a lot of outstanding actors.								
2/1/2019 4:47 KELINGTON GROUF Darren Hal <pre> Class=" 1.09E+18 2 1 0 One factor you can't overlook is the calibre of that blueline with Valimaki, Kylington and Andersson all in the NHL. No future NHLers in the group 2/1/2019 4:47 BURSA MALAYSIA Palm Oil A <pre> Palm Oil A <pre> </pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre>	2/1/2019 1:24	TENAGA NASIONA	L EV Club M	<p< td=""><td>1.09E+18</td><td>1</td><td>1</td><td>0 Seems like a no-brainer for @Tenaga_Nasional to offer something similar, or at least a one-time rebate for home EVSE installation. Like the Gillette r</td></p<>	1.09E+18	1	1	0 Seems like a no-brainer for @Tenaga_Nasional to offer something similar, or at least a one-time rebate for home EVSE installation. Like the Gillette r								
2/1/2019 4:47 KELINGTON GROUF Darren Hal <pre> Class=" 1.09E+18 2 1 0 One factor you can't overlook is the calibre of that blueline with Valimaki, Kylington and Andersson all in the NHL. No future NHLers in the group 2/1/2019 4:47 BURSA MALAYSIA Palm Oil A <pre> Palm Oil A <pre> </pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre>																
2/1/2019 4:47 BURSA MALAYSIA Palm Oil A <pre>pclass=" 1.09E+18</pre>	2/1/2019 4:47	FBMKLCI-H6P OR B	Palm Oil A	p class="	1.09E+18	3	0	0 Bursa Malaysia is closed today Friday 1st of February for Federal Territory day. There will be no CPO Futures trading. Trading resumes on Monday 4th								
2/1/2019 4:48 MHB OR MALAYSIA KB Associa <p #dataprep<="" 0="" 1="" 1.09e+18="" 11-15="" 130011â="" 16%="" 2="" 2019="" 3="" 4:50="" 5:15="" <p="" @gcpcloud="" air="" airports="" class="1.09E+18 17 0 7 No matter what learning is necessary and can be fun too! If you're a learner and wanted to explore more then you should explore Hong Leong St. 2/1/2019 4:59 SP500-HG OR SP500 IANS Twee <p 1.09E+18 2 0 The #Dow Jones Industrial Average fell 15.19 points, or 0.06 per cent, to 24,999.67. The #SP500 increased 23.05 points, or 0.86 per cent, to 2,704.10 2/1/2019 5:08 AIRASIA OR AIRASIA Mineski Pr <p 1.09E+18 8 0 We advance to winners match of group A! 2/1/2019 5:15 MALAYSIA AIRPORT The Malays <p class=" closing="" cny="" course="" datai="" dataprp="" diving="" during="" expects="" feb="" for="" hlfg="" hong="" https:="" imca="" impressed="" in="" is="" kl.="" langi="" lec="" lynn="" malaysia="" malaysia,="" more="" on="" or="" passenger="" registration="" s="" scheduled="" sl="" so="" soon="" supervisor="" t's="" td="" that="" the="" traffic="" trainee="" with="" www.themalaysianinsight.com="" =""><td>2/1/2019 4:47</td><td>KELINGTON GROU</td><td>F Darren Ha</td><td>p class="</td><td>1.09E+18</td><td>2</td><td>1</td><td>0 One factor you can't overlook is the calibre of that blueline with Valimaki, Kylington and Andersson all in the NHL. No future NHLers in the group</td></p>	2/1/2019 4:47	KELINGTON GROU	F Darren Ha	p class="	1.09E+18	2	1	0 One factor you can't overlook is the calibre of that blueline with Valimaki, Kylington and Andersson all in the NHL. No future NHLers in the group								
2/1/2019 4:50 HLFG OR HONG LEC T'S SU < p class=" 1.09E+18 17 0 The #Dow Jones Industrial Average fell 15.19 points, or 0.06 per cent, to 24,999.67. The #SP500 increased 23.05 points, or 0.86 per cent, to 2,704.10 0 We advance to winners match of group A! 2/1/2019 5:15 MALAYSIA AIRPORT The Malays < p class=" 1.09E+18 0 Malaysia Airports expects 16% more passenger traffic during CNY https://www.themalaysianinsight.com/s/130011Â 2/1/2019 5:15 DATAPRP OR DATAL Lynn Langi < p 1.09E+18 3 1 0 so impressed with @GCPcloud #Dataprep	2/1/2019 4:47	BURSA MALAYSIA	Palm Oil A	p class="	1.09E+18	3	0	0 Bursa Malaysia is closed today Friday 1st of February for Federal Territory day. There will be no CPO Futures trading. Trading resumes on Monday 4tl								
2/1/2019 4:59 SP500-HG OR SP50(IANS Twee < p 1.09E+18 2 0 The #Dow Jones Industrial Average fell 15.19 points, or 0.06 per cent, to 24,999.67. The #SP500 increased 23.05 points, or 0.86 per cent, to 2,704.10 0 We advance to winners match of group A! 2/1/2019 5:15 MALAYSIA AIRPORT The Malays < p class=" 1.09E+18 0 1 Malaysia Airports expects 16% more passenger traffic during CNY https://www.themalaysianinsight.com/s/130011Â 2/1/2019 5:15 DATAPRP OR DATAL Lynn Langl < p 1.09E+18 3 1 0 so impressed with @GCPcloud #Dataprep	2/1/2019 4:48	MHB OR MALAYSIA	A KB Associa <	ф	1.09E+18	1	0	0 Registration Closing Soon for IMCA Trainee AIR Diving Supervisor course that is scheduled on the 11-15 Feb 2019 in Malaysia, KL.								
2/1/2019 5:08 AIRASIA OR AIRASI/ Mineski Pr <p 0="" 1="" 1.09e+18="" 2="" 2019="" 5:15="" 8="" <="" <p="" a!="" advance="" airport="" class="1.09E+18 0 1 Malaysia Airports expects 16% more passenger traffic during CNY https://www.themalaysianinsight.com/s/130011Â 2/1/2019 5:15 DATAPRP OR DATA Lynn Langl <p 1.09E+18 3 1 0 so impressed with @GCPcloud #Dataprep</td><td>2/1/2019 4:50</td><td>HLFG OR HONG LE</td><td>(IT'S SU<</td><td>p class=" group="" malays="" malaysia="" match="" of="" td="" the="" to="" we="" winners=""><td>1.09E+18</td><td>17</td><td>0</td><td>7 No matter what learning is necessary and can be fun too! If you're a learner and wanted to explore more then you should explore Hong Leong So</td></p>	1.09E+18	17	0	7 No matter what learning is necessary and can be fun too! If you're a learner and wanted to explore more then you should explore Hong Leong So												
2/1/2019 5:15 MALAYSIA AIRPORT The Malays < p class=" 1.09E+18 0 0 1 Malaysia Airports expects 16% more passenger traffic during CNY https://www.themalaysianinsight.com/s/130011Â 2/1/2019 5:15 DATAPRP OR DATA Lynn Langi < p 1.09E+18 3 1 0 so impressed with @GCPcloud #Dataprep	2/1/2019 4:59	SP500-HG OR SP50	(IANS Twee <	(p	1.09E+18	2	0	0 The #Dow Jones Industrial Average fell 15.19 points, or 0.06 per cent, to 24,999.67. The #SP500 increased 23.05 points, or 0.86 per cent, to 2,704.10								
2/1/2019 5:15 DATAPRP OR DATA Lynn Langi <p #dataprep<="" 0="" 1="" 1.09e+18="" 3="" @gcpcloud="" impressed="" so="" td="" with=""><td>2/1/2019 5:08</td><td>AIRASIA OR AIRASIA</td><td>/Mineski Pr <</td><td>ф</td><td>1.09E+18</td><td>8</td><td>0</td><td>0 We advance to winners match of group A!</td></p>	2/1/2019 5:08	AIRASIA OR AIRASIA	/Mineski Pr <	ф	1.09E+18	8	0	0 We advance to winners match of group A!								
	2/1/2019 5:15	MALAYSIA AIRPOR	T The Malays <	p class="	1.09E+18	0	0	1 Malaysia Airports expects 16% more passenger traffic during CNY https://www.themalaysianinsight.com/s/130011Â								
2/1/2019 5:24 DATAPRP OR DATA Vikram đ <p <="" class="1.09E+18 1 0 Dataprep is awesome! Esp since you can schedule it and forget about it and it keeps working for you :)</td><td>2/1/2019 5:15</td><td>DATAPRP OR DATA</td><td>l Lynn Langi <</td><td>(p</td><td>1.09E+18</td><td>3</td><td>1</td><td>0 so impressed with @GCPcloud #Dataprep</td></tr><tr><td></td><td>2/1/2019 5:24</td><td>DATAPRP OR DATA</td><td>l Vikram ð <</td><td>p class=" td=""><td>1.09E+18</td><td>1</td><td>0</td><td>0 Dataprep is awesome! Esp since you can schedule it and forget about it and it keeps working for you :)</td></p>	1.09E+18	1	0	0 Dataprep is awesome! Esp since you can schedule it and forget about it and it keeps working for you :)												

Stock Related Tweets over Time

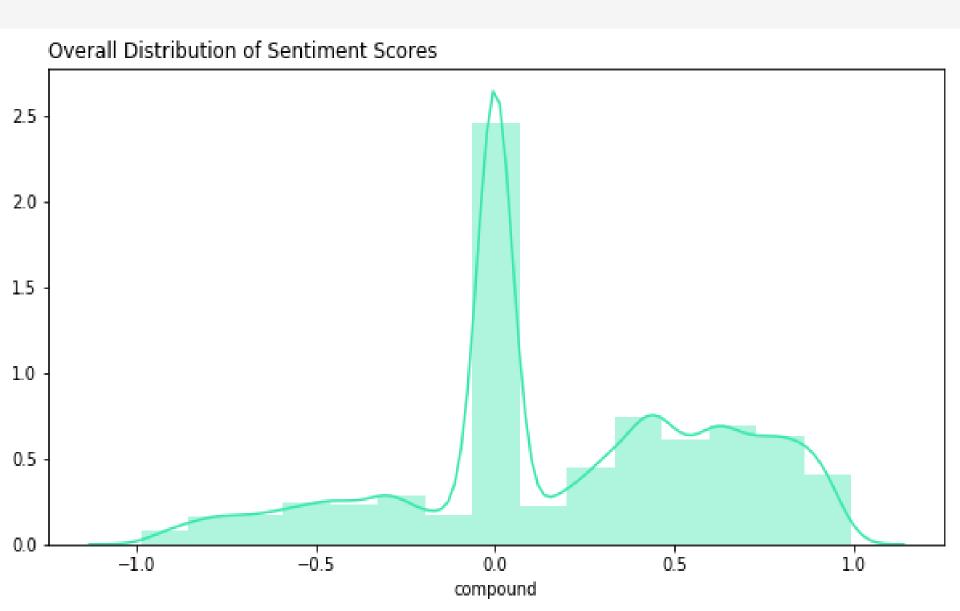


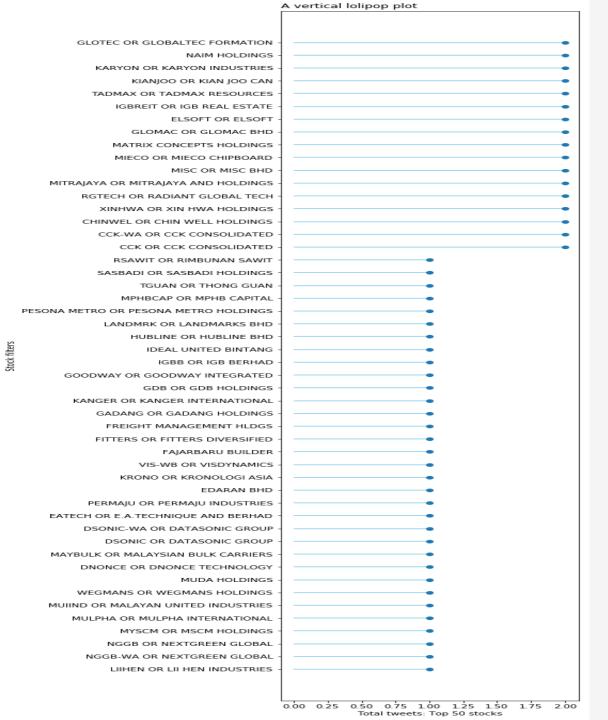
- Distribution of number of tweets by stocks.
- Majority of stocks have lesser than 10 tweets.





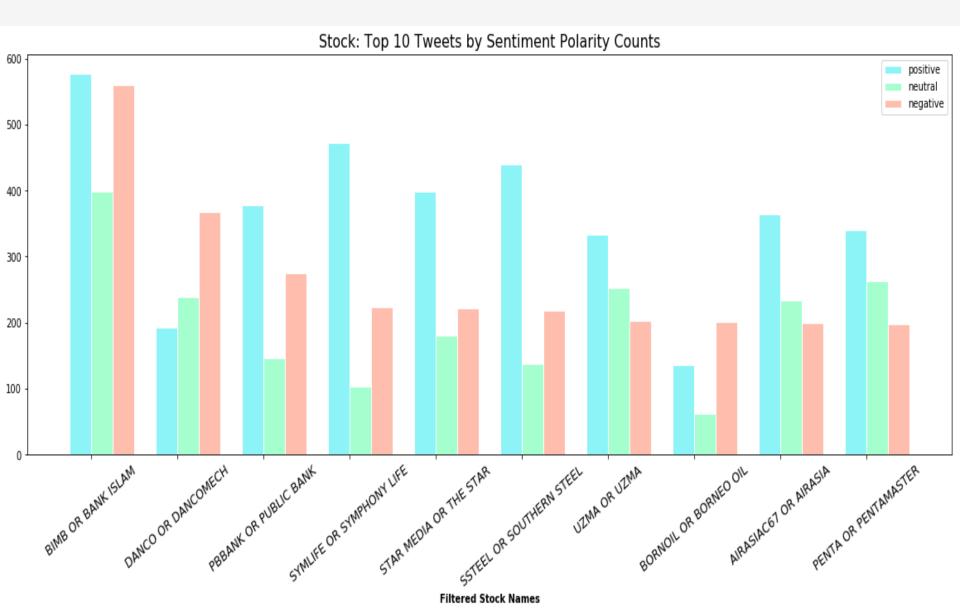
 Majority obtain sentiment scores of near zero, which indicates near-neutral sentiment





Top 50 tweeted stocks (comparison of relative value)

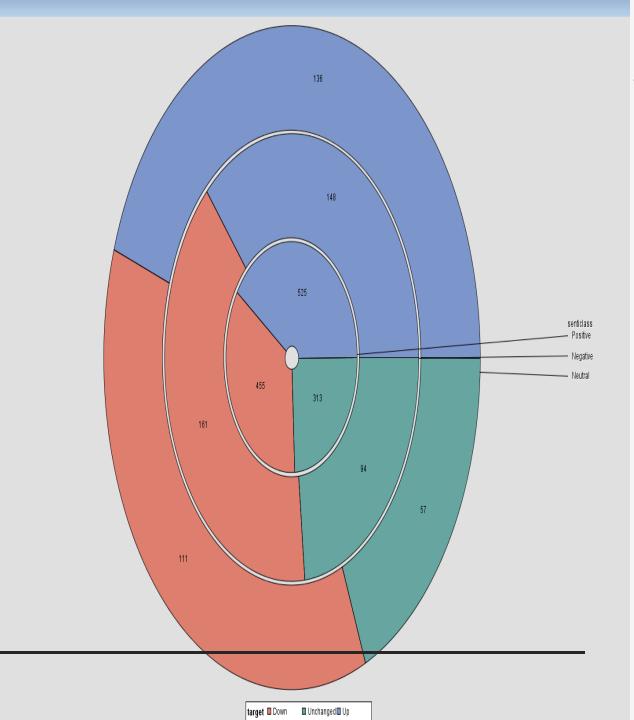
Top 10 tweeted stocks with sentiment polarity labels



Part 2: Interpreting merged dataset

- Target variable for daily changes of closing stock prices
 - ☐ Up price is increasing
 - □ Down price is decreasing
 - ☐ Unchanged no price changes

☐ ☐ ☑ ☐ ☐ ☑													
filters	day	name	board	close	open	high	low	volume	dif	difpct	avg_senti	senticlass	target
AAX OR AIRASIA	2019-01-03/		Main	0.24	0.23	0.24	0.23	8801100	0.009999999999	0.043478260869	0.44853043478	Positive	Up ^
AAX OR AIRASIA	2019-01-04/	AX	Main	0.26	0.24	0.26	0.24	30545800	0.020000000000	0.08333333333333	0.10046666666	Positive	Up
AAX OR AIRASIA	2019-01-07/	AX	Main	0.305	0.26	0.31	0.26	84608800	0.044999999999	0.17307692307	0.08164285714	Positive	Up
AAX OR AIRASIA	2019-01-08/	AX	Main	0.29	0.3	0.305	0.28	69337400	-0.015000000000	-0.049180327868	0.17878666666		Down
AAX OR AIRASIA	2019-01-09/	AX	Wain	0.29	0.29	0.3	0.285	15586200			0.1778	9 Positive	Unchanged
AAX OR AIRASIA	2019-01-10/	AX	Wain	0.29	0.29	0.3	0.285	28305200			-0.067433333333	Negative	Unchanged
AAX OR AIRASIA	2019-01-11/	AX	Wain	0.29	0.295	0.295	0.28	16799600			0.5328	5 Positive	Unchanged
AAX OR AIRASIA	2019-01-14/	AX	Wain	0.285	0.29	0.295	0.28	22089100	-0.005000000000	-0.017241379310	0.39279999999	Positive	Down
AAX OR AIRASIA	2019-01-15/	AX	Main	0.3	0.285	0.3	0.285	32894100	0.015000000000	0.052631578947	0.2056	6 Positive	Up
AAX OR AIRASIA	2019-01-16/	AX	Wain	0.3	0.3	0.305	0.29	25793300				ONeutral	Unchanged
AAX OR AIRASIA	2019-01-17/	AX	Wain	0.295	0.3	0.3	0.29	11649500	-0.005000000000	-0.01666666666666		ONeutral	Down
AAX OR AIRASIA	2019-01-18/	W X	Wain	0.29	0.295	0.295	0.29	5171300	-0.005000000000	-0.016949152542	0.12040000000	Positive	Down
AAX OR AIRASIA	2019-01-22/	W X	Wain	0.29	0.29	0.31	0.28	48519200			0.100478	5Positive	Unchanged
AAX OR AIRASIA	2019-01-23/	W X	Wain	0.285	0.285	0.29	0.28	12168600	-0.005000000000	-0.017241379310	-0.12254000000	.Negative	Down
AAV OD AIDAOIA	0040 04 04	AAV.	п.	0.000	0.00	7.00	0.000	000000	1	^	0.0040000000	D 7	H I I



- As seen in the multi-layer pie chart,
 - ☐ The biggest proportion are neutral tweets (outer pie chart layer), which are dominated by increasing class
 - □ Followed by negative tweets (2nd layer), dominated by decreasing class
 - ☐ Lastly, the positive tweets pie chart (inner most layer) dominated by increasing stock prices

THE END