Problem 2

Yufei Yin

Question 2

Write a short report (in prose) arguing that the code you've provided works as required and provide a pdf of the report. Include output of R sessions in labelled figures: You could show the translink function being called with different arguments, and show the corresponding elements of the data list proving that correct bus routes are returned. You could also show that for a given element of the data list, parameters to the translink function can be specified in order to return that affected bus route. (5 points)

```
source('translink.R')
```

test 1

```
disruptions = translink(2020, 1, 26, 3)
disruptions$start
```

```
## [1] "406" "401"
```

```
disruptions$stop
```

```
## [1] "23"
```

```
disruptions$df$text
```

```
## [1] "#RiderAlert 23 Beach/Main Street Station detour UPDATE 7:25pm: clear, resuming regular route.^j
## [2] "#RiderAlert 406 Steveston/Brighouse Station detour 6:30 PM. Regular route to No 3 Rd &amp;amp; Saba
## [3] "#RiderAlert 401 One Road/Brighouse Station detour 6:30 PM. Regular route to Westminster &amp;amp; M
```

```
disruptions$df$date
```

```
## [1] "2020-01-26 03:29:02 UTC" "2020-01-26 03:28:41 UTC"
## [3] "2020-01-26 03:23:34 UTC"
```

There are 3 tweets posted at the time we specified (2020-01-26 3 o'clock), from first tweet we noticed that bus 23 is resuming regular route, and from second and third tweets we know that bus 406 and 401 start detouring.

test 2

```
disruptions = translink(2020, 1, 28, 1)
disruptions$start
```

```
## [1] "236" "239" "375" "345" "N24" "335" "230" "229" "228"
```

```
disruptions$stop
```

```
## [1] "335" "345" "375"
```

```
disruptions$df$text
```

```
##  [1] "#RiderAlert 236 Lonsdale Quay detour Mon Jan 27 to Friday Jan 31 nightly from 10:30pm to end o
##  [2] "#RiderAlert 236 Grouse Mountain detour Mon Jan 27 to Friday Jan 31 nightly from 10:30pm to end
##  [3] "#RiderAlert 335 Surrey Central Station / 345 King George / 375 Guildford detours are cleared, s
##  [4] "#RiderAlert 239 Park Royal detour Mon Jan 27 to Friday Jan 31 nightly from 10:30pm to end of se
##  [5] "#RiderAlert 375 Guildford detour 5:12PM. Regular route to 152 St, 76 Ave, 144 St, 84 Ave, 152 S
##  [6] "#RiderAlert 239 Park Royal detour Mon Jan 27 to Friday Jan 31 nightly from 10:30pm to end of se
##  [7] "#RiderAlert 345 King George Station detour 5:12PM. Regular route to 152 St, 76 Ave, 144 St, 84
##  [8] "#RiderAlert N24 Lynn Valley detour Mon Jan 27 to Friday Jan 31 nightly from 10:30pm to end of s
##  [9] "#RiderAlert 335 Surrey Central Station detour 5:12PM. Regular route to 152 St &amp;amp; 76 Ave, 76
## [10] "#RiderAlert 230 Lonsdale detour Mon Jan 27 to Friday Jan 31 nightly from 10:30pm to end of serv
## [11] "#RiderAlert 229 Lynn Valley detour Mon Jan 27 to Friday Jan 31 nightly from 10:30pm to end of s
## [12] "#RiderAlert 228 Lynn Valley detour Mon Jan 27 to Friday Jan 31 nightly from 10:30pm to end of s
```

```
disruptions$df$date
```

```
##  [1] "2020-01-28 01:39:42 UTC" "2020-01-28 01:36:44 UTC"
##  [3] "2020-01-28 01:34:19 UTC" "2020-01-28 01:18:53 UTC"
##  [5] "2020-01-28 01:18:19 UTC" "2020-01-28 01:16:33 UTC"
##  [7] "2020-01-28 01:16:10 UTC" "2020-01-28 01:14:23 UTC"
##  [9] "2020-01-28 01:13:55 UTC" "2020-01-28 01:13:08 UTC"
## [11] "2020-01-28 01:11:10 UTC" "2020-01-28 01:09:36 UTC"
```

From 3rd tweet, bus 335/345/375 detours are cleared, and all the other tweets are bus start detouring. interestingly, bus 335/345/375 start and stop detouring within 1 hour.

According to the test above, i am confidence my function works well, but i can not 100% sure it will extract exactly what we want; however, it will work in most of the case since i have considered lots of situations. For example, i considered lower case upper case, i considered multiple bus number in the same tweet, and special bus number (start with R and N) etc.