

Scala: 2.11 Spark: 2.2.1

Task1: Twitter Streaming

1. TrackHashTags

Command:

```
MacBook-Pro-3782:spark-2.2.1-bin-hadoop2.7 watson$ bin/spark-submit --class TrackHashTags Yi_Wei_hw5.jar NW3PXmfvfQYcyTO5mpU79hygB VOuK9zrQY9a4peVb687VALNBNCZzz6E0E2eX1fVwZkkSrDNAJi 835545284-m8v7ga1X0BUDwFtOemz23XXDEui9qy9W9ennxuIJ BBqQ2z2COUCMAlyTWTQrmdUQdZUDDkI4qQDSKHlc7XSnz
```

```
bin/spark-submit --class TrackHashTags Yi_Wei_hw5.jar  
NW3PXmfvfQYcyTO5mpU79hygB  
VOuK9zrQY9a4peVb687VALNBNCZzz6E0E2eX1fVwZkkSrDNAJi  
835545284-m8v7ga1X0BUDwFtOemz23XXDEui9qy9W9ennxuIJ  
BBqQ2z2COUCMAlyTWTQrmdUQdZUDDkI4qQDSKHlc7XSnz
```

The first parameter is consumerKey

The second parameter is consumerSecret

The third parameter is accessToken

The last parameter is accessTokenSecret

2. AveTweetLength

Command:

```
MacBook-Pro-3782:spark-2.2.1-bin-hadoop2.7 watson$ bin/spark-submit --class AveTweetLength Yi_Wei_hw5.jar NW3PXmfvfQYcyTO5mpU79hygB VOuK9zrQY9a4peVb687VALNBNCZzz6E0E2eX1fVwZkkSrDNAJi 835545284-m8v7ga1X0BUDwFtOemz23XXDEui9qy9W9ennxuIJ BBqQ2z2COUCMAlyTWTQrmdUQdZUDDkI4qQDSKHlc7XSnz
```

```
bin/spark-submit --class AveTweetLength Yi_Wei_hw5.jar  
NW3PXmfvfQYcyTO5mpU79hygB  
VOuK9zrQY9a4peVb687VALNBNCZzz6E0E2eX1fVwZkkSrDNAJi  
835545284-m8v7ga1X0BUDwFtOemz23XXDEui9qy9W9ennxuIJ  
BBqQ2z2COUCMAlyTWTQrmdUQdZUDDkI4qQDSKHlc7XSnz
```

The first parameter is consumerKey

The second parameter is consumerSecret

The third parameter is accessToken

The last parameter is accessTokenSecret

Task2: Flajolet-Martin algorithm

Command:

```
n$ bin/spark-submit --class UniqueUserCount Yi_Wei_hw5.jar localhost 9999
```

```
bin/spark-submit --class UniqueUserCount Yi_Wei_hw5.jar localhost  
9999
```

The first parameter is hostname

The second parameter is port number