

# Spark streaming session 29.10.20

## Session 5

Listed the guide on

[Canvas](#)

--

Twitter developer account:

<https://developer.twitter.com>

Install needed tools that you dont allready have:

--

**Java (Java 8 is recommended for Spark)**

<https://www.oracle.com/technetwork/java/javase/downloads/index.html>

--

**Spark**

<https://spark.apache.org/downloads.html>

**Scala IDE for Eclipse**

<http://scala-ide.org/>

Spark MAC guide:

<https://oleweidner.com/blog/2015/getting-started-with-spark-on-osx/>

You can also view the **Install.mp4** video for a 14 minute detailed video on how to install both of them. There is also a PDF with step-by-step guide.

--

**Stream Twitter Data with Spark**

- Follow this guide

<https://acadgild.com/blog/streaming-twitter-data-using-spark>

or

<https://dzone.com/articles/twitter-live-streaming-with-spark-streaming-using>

The video **SparkStreamingOverview.mp4** is a video where he talks about what it is and what it

can be used for. Maybe interesting to watch a a later point.

I made a gif of me running Twitterstreaming from the acadgild.com guide in Scala Eclipse IDE.  
Look at: **RunningSpark.gif**

There is a video with detailed guide on running spark stream for tweets called **SparkTwitter.mp4**

One guide to filter hashtags you want:

<https://mapr.com/blog/spark-streaming-and-twitter-sentiment-analysis/>

My code for filtering specific hashtags:

```
val specificHashtag = stream.map(status => status.getText())
```

```
val t = specificHashtag.filter(status => status contains "#Emergency")
```

Spark streaming **full** doc:

<http://spark.apache.org/docs/latest/streaming-programming-guide.html>