**TUGAS PRAKTIK BIG DATA**

**TUTORIAL INSTALASI APACHE SPARK**



**DI SUSUN OLEH**

**NAMA : ALEXANDRO ASCHANIO TADA**

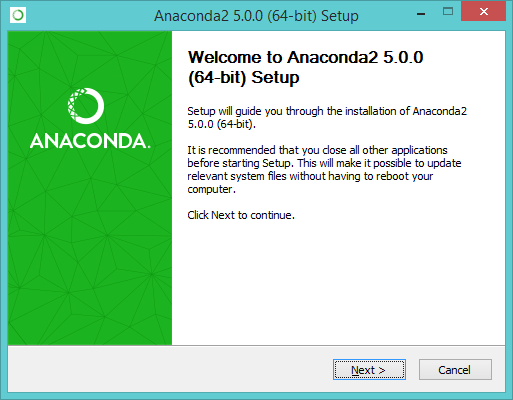
**NIM : 125610009**

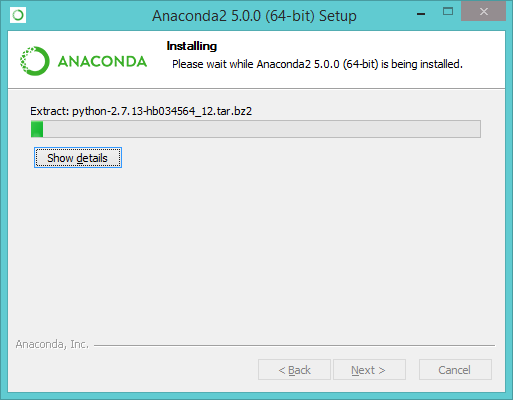
**PRODI : SISTEM INFORMASI**

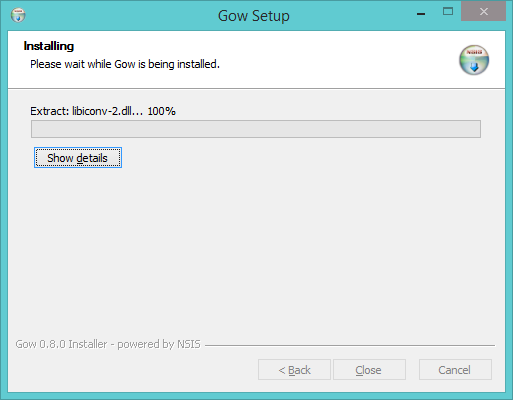
**STMIK AKAKOM YOGYAKARTA**

**2017/2018**

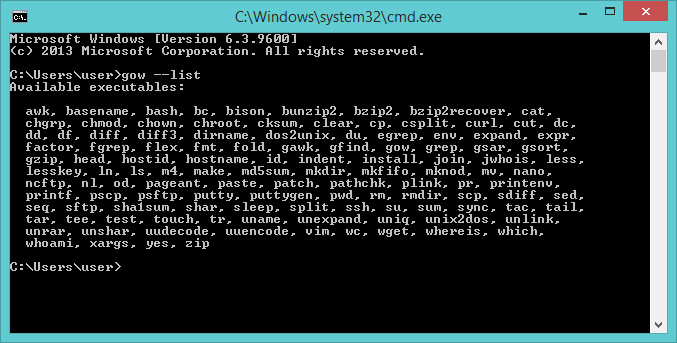
Persyarat: Anaconda dan GOW. Jika Anda sudah memiliki anaconda dan GOW (terinstal), lompat kelangkah 5.



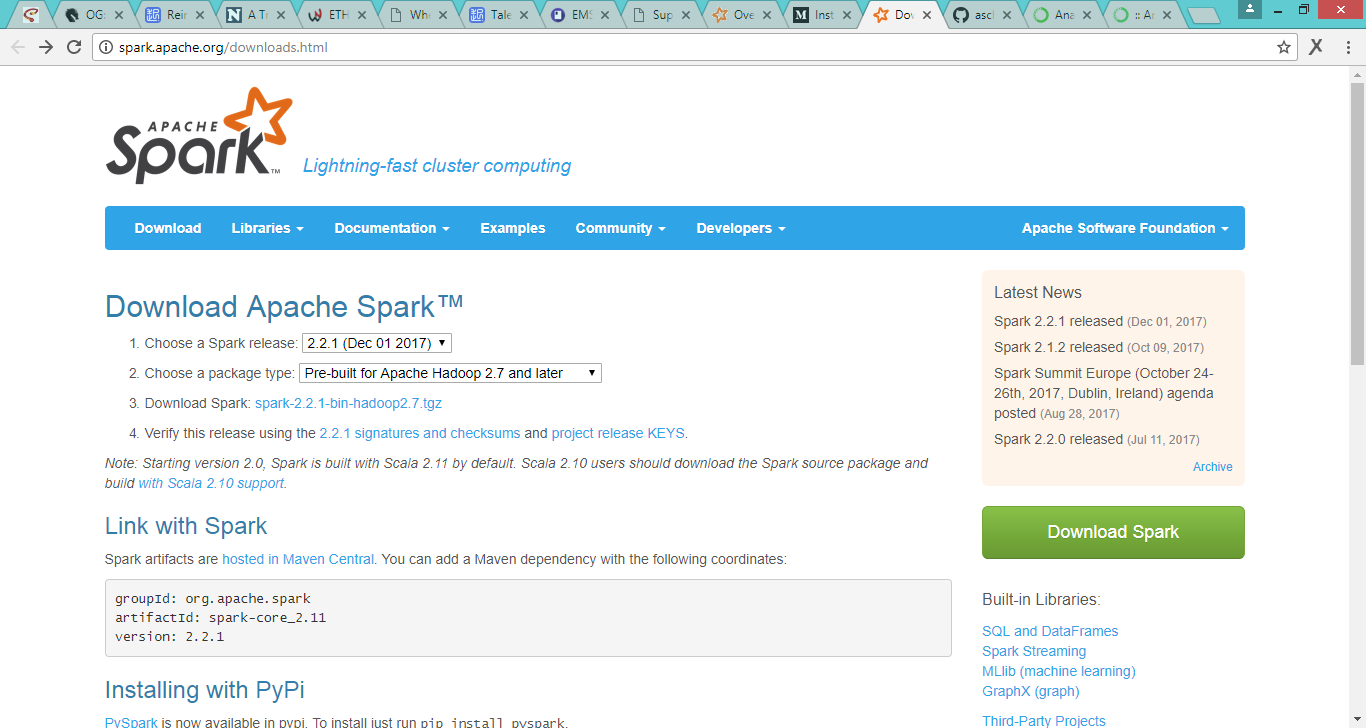




1. Download dan pasang Gnu di windows (GOW). GOW berfungsi sebagai sarana untuk menggunakan perintah linux di windows. Dalam instalasi ini, kita perlu curl, gzip, tar yang disediakan GOW.



1. Download dan instal Anaconda.
2. Tutup dan buka command line (CMD) baru.
3. Pergi ke situs Apache Spark:



a) Pilih rilis Spark

b) Pilih jenis paket

c) Pilih jenis download: (Direct Download)

d) Download Spark

1. Pindahkan file ke tempat yang ingin di unzip:

>mkdir C: \ opt \ spark

>mv C:\Users\user\Downloads\spark-2.2.0-bin-hadoop2.7.tgz C: \ opt \ spark \ spark-2.2.0-bin-hadoop2.7.tgz

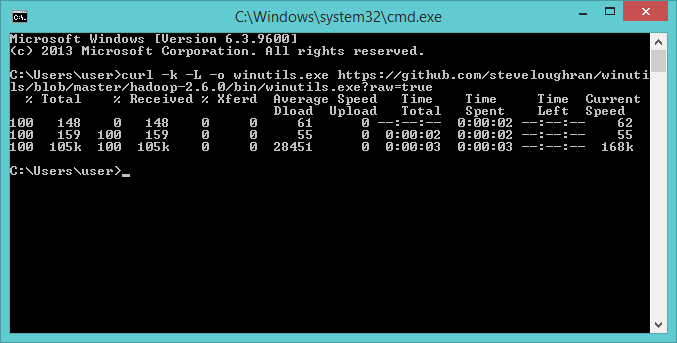
1. Unzip file tersebut. Gunakan perintah di bawah ini:

>gzip -d spark-2.2.0-bin-hadoop2.7.tgz

>tar xvf spark-2.2.0-bin-hadoop2.7.tar

1. Download winutils.exe ke dalam spark-2.1.0-bin-hadoop2.7 \ bin

>curl -k -L -o winutils.exe <https://github.com/steveloughran/winutils/blob/master/hadoop-2.6.0/bin/winutils.exe?raw=true>



1. Pastikan Anda telah menginstal Java 7+ di komputer jika tidak lakukan path di Environment Variables.
2. Selanjutnya, kita akan mengedit variabel lingkungan kita sehingga kita bisa membuka spark notebook di sembarang direktori.

>setx SPARK\_HOME C: \ opt \ spark \ spark-2.2.0-bin-hadoop2.7

>setx HADOOP\_HOME C: \ opt \ spark \ spark-2.2.0-bin-hadoop2.7

>setx PYSPARK\_DRIVER\_PYTHON ipython

>setx PYSPARK\_DRIVER\_PYTHON\_OPTS notebook

>Tambahkan; C: \ opt \ spark \ spark-2.2.0-bin-hadoop2.7 \ bin ke Path.