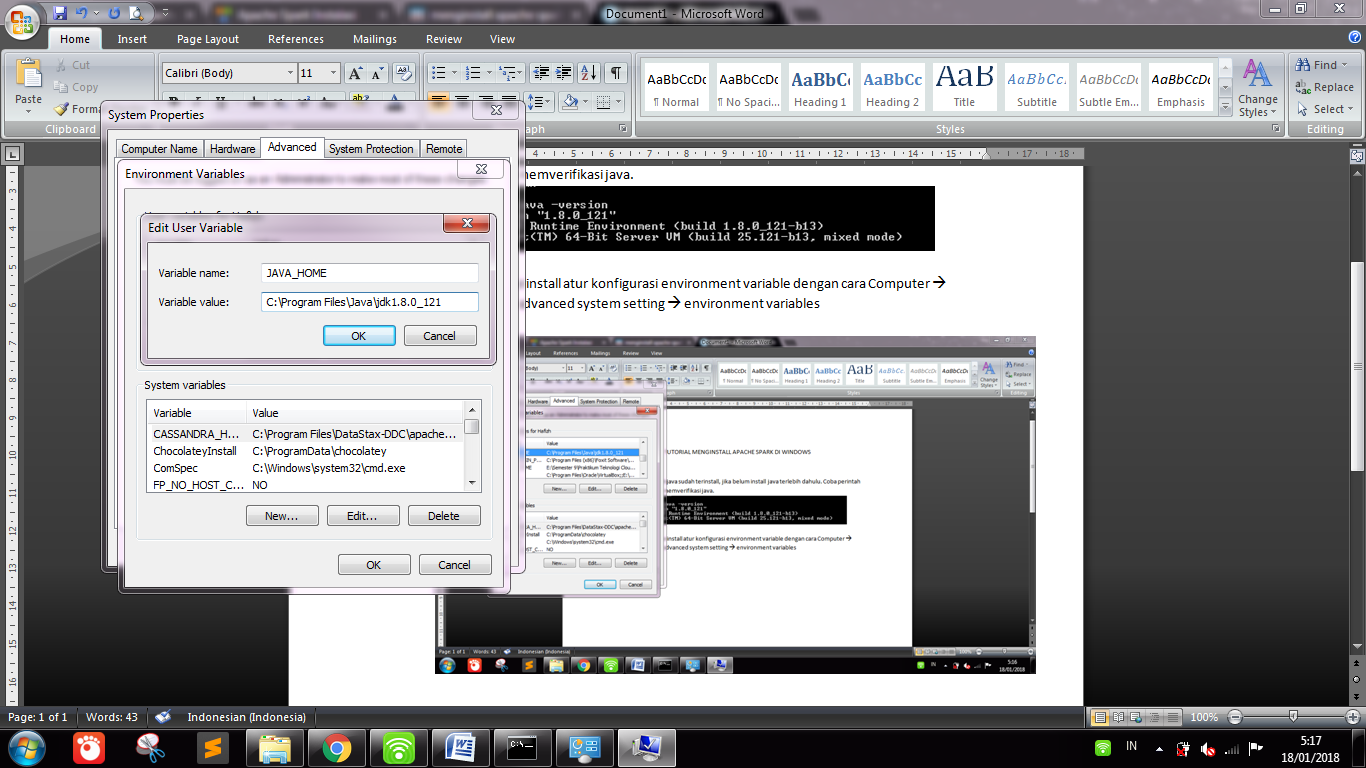
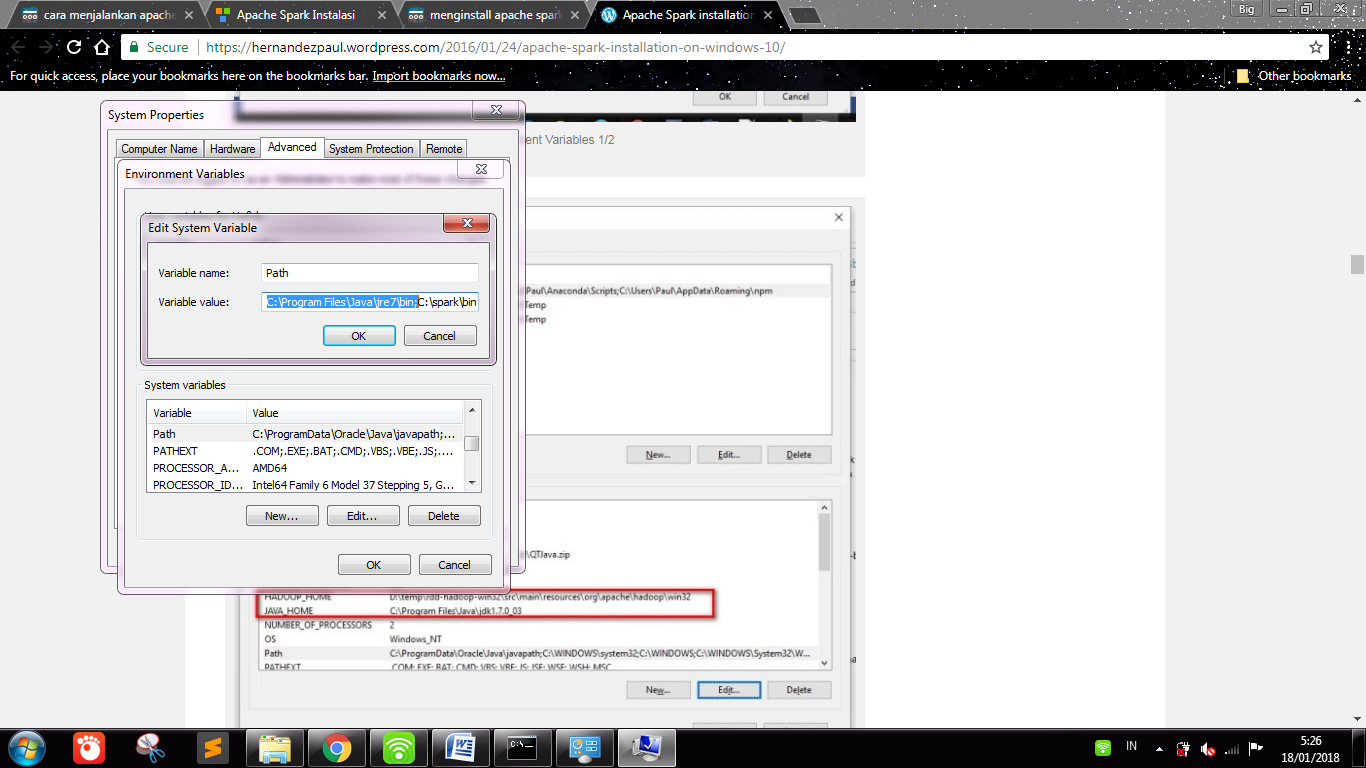
TUTORIAL MENGINSTALL APACHE SPARK DI WINDOWS

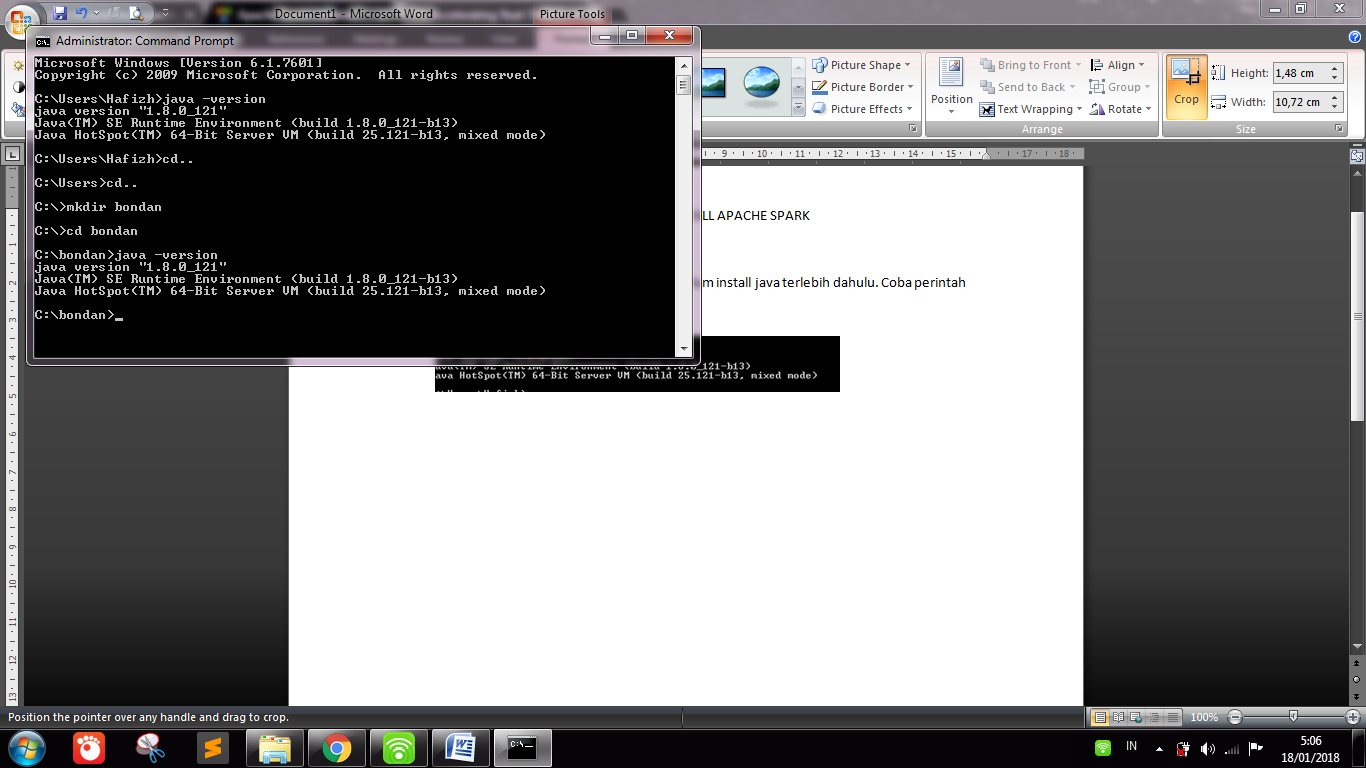
1. Periksa apakah java sudah terinstall, jika belum install java terlebih dahulu.

Atur konfigurasi environment variable dengan cara Computer 🡪 properties 🡪 advanced system setting 🡪 environment variables. Atur seperti berikut.



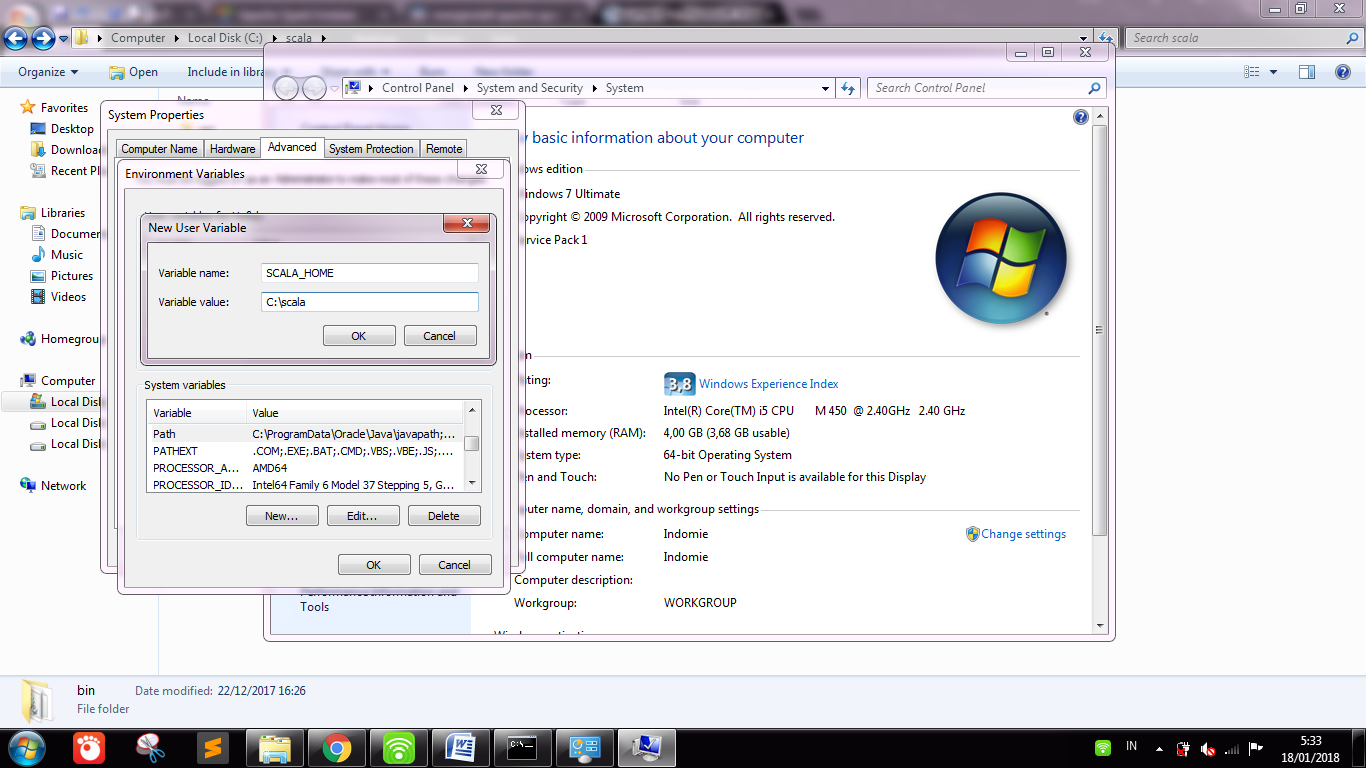


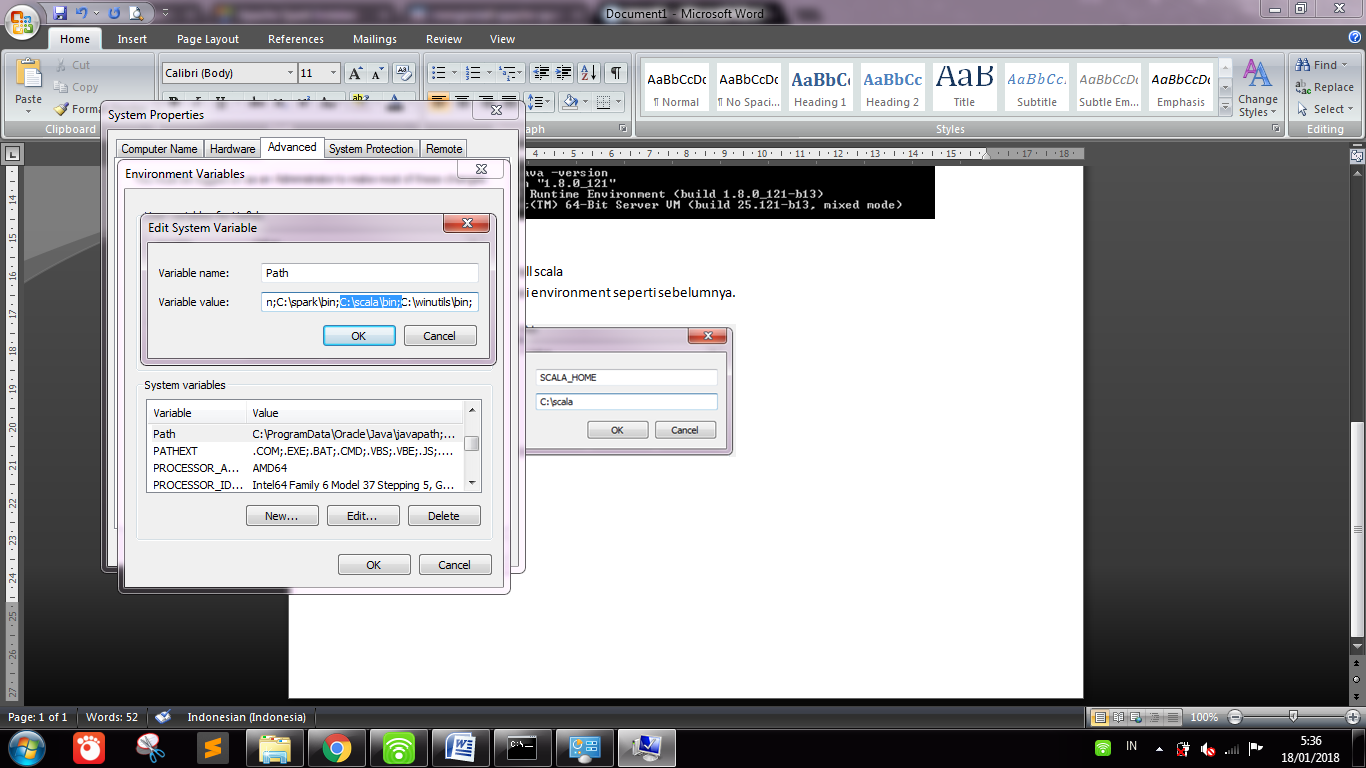
Coba perintah berikut untuk mengeceki java.



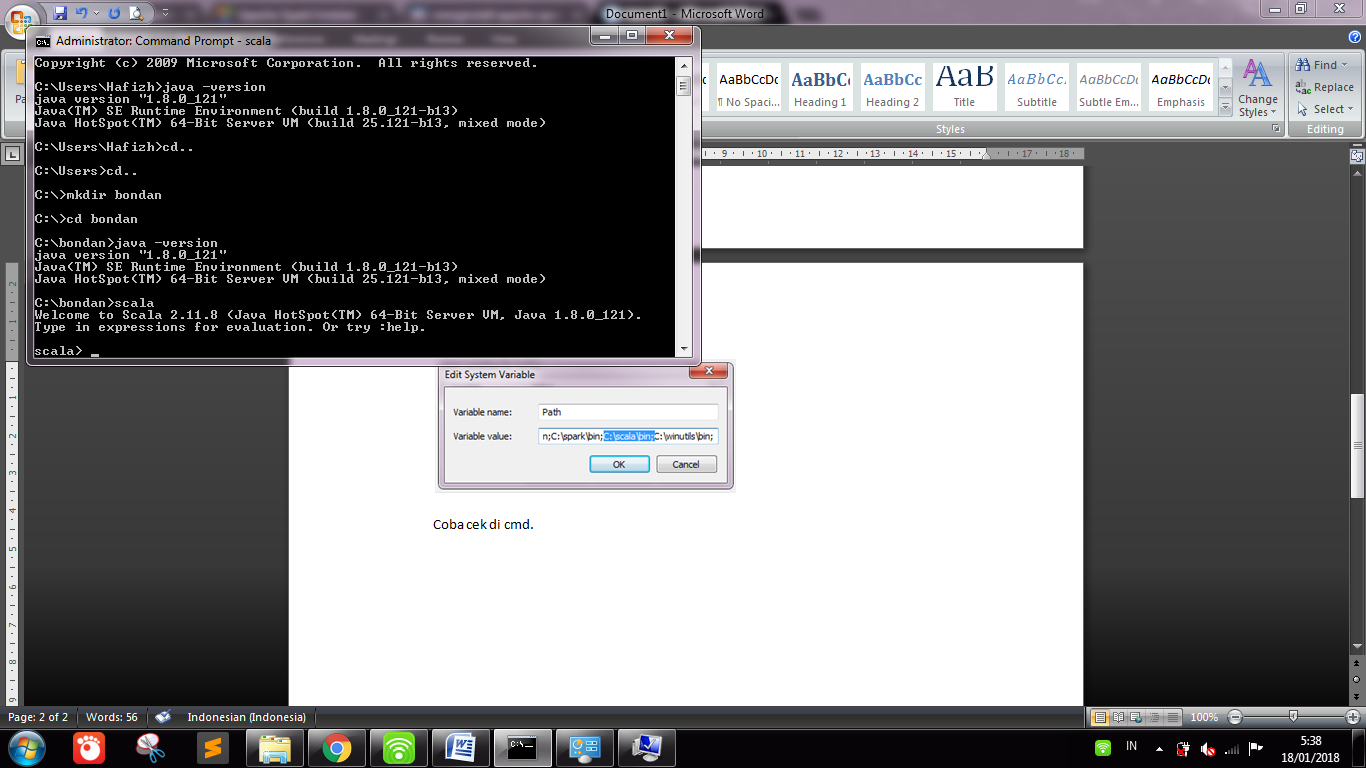
1. Kemudian install scala

Atur konfigurasi environment seperti sebelumnya.

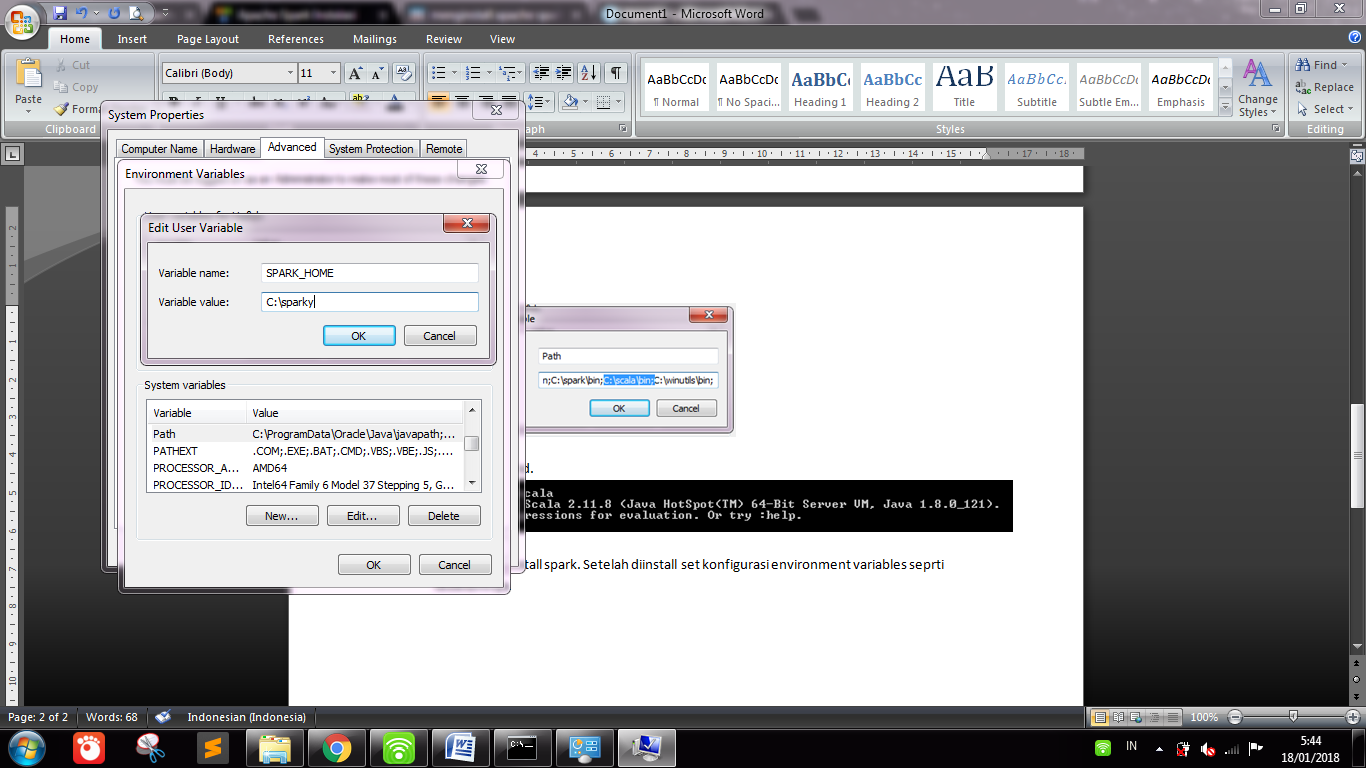


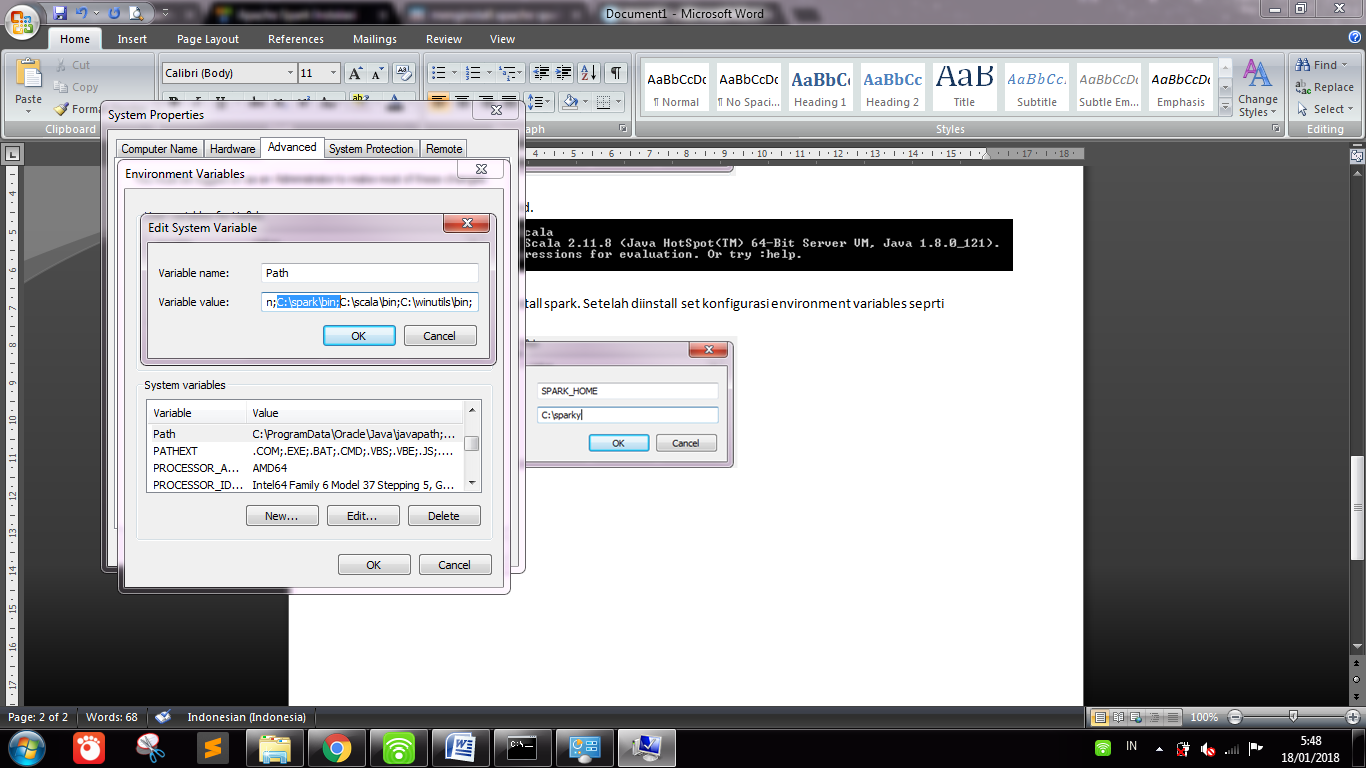


Coba cek di cmd.

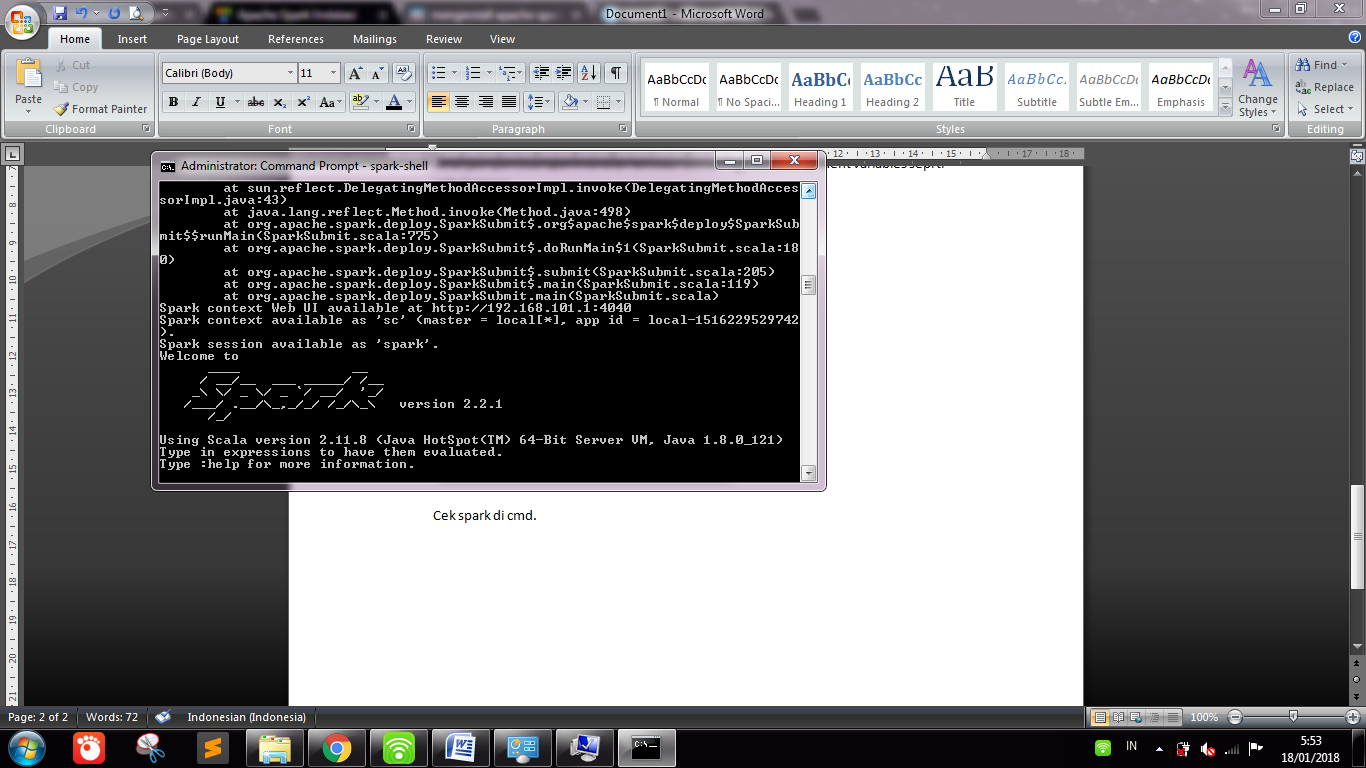


1. Selanjutnya install spark. Setelah diinstall set konfigurasi environment variables seprti sebelumnya.

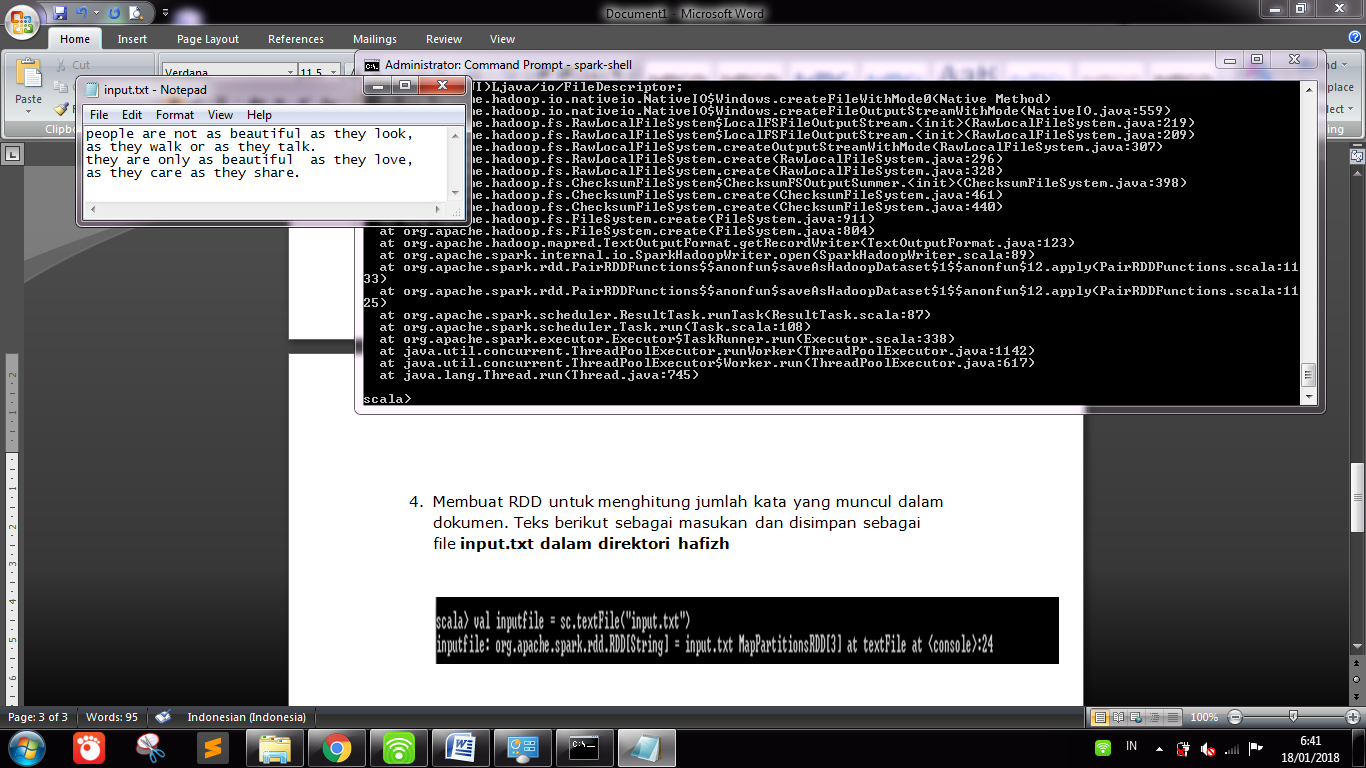




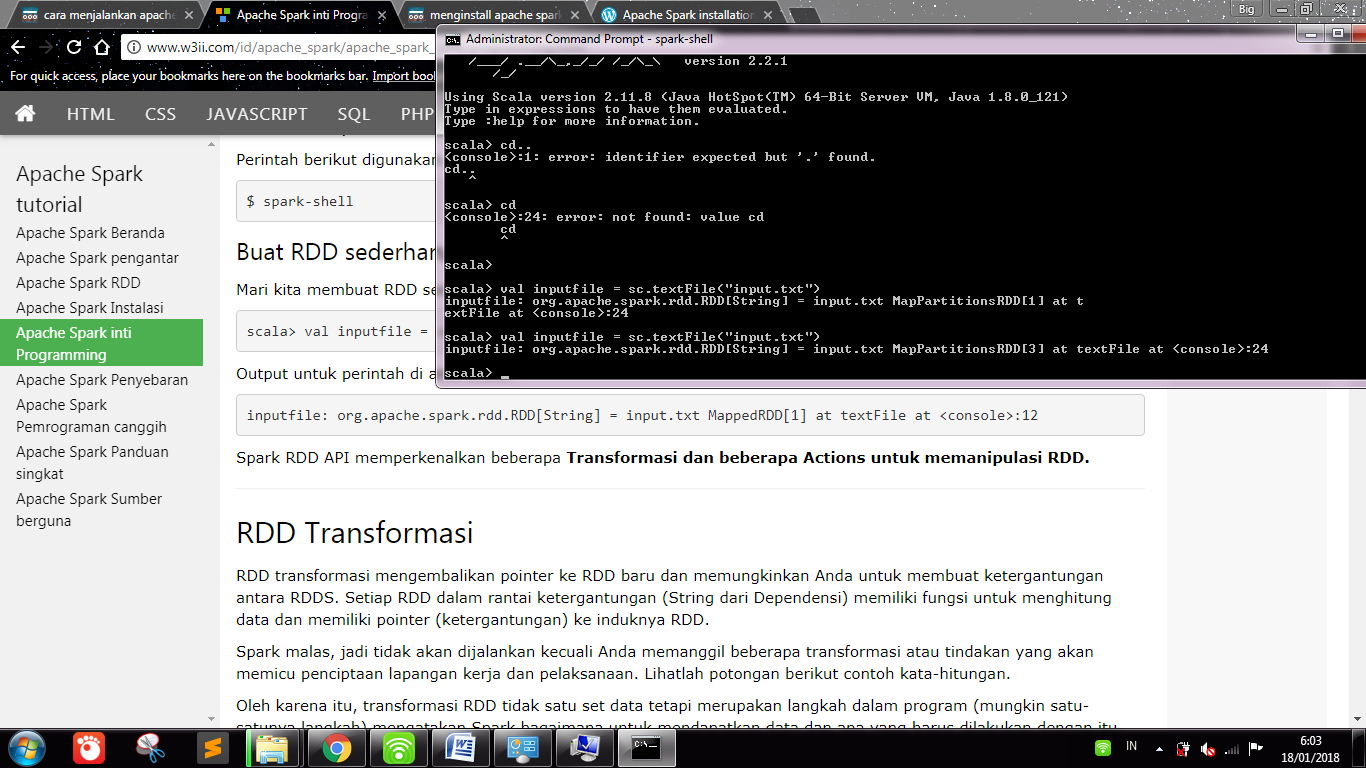
Cek spark di cmd.



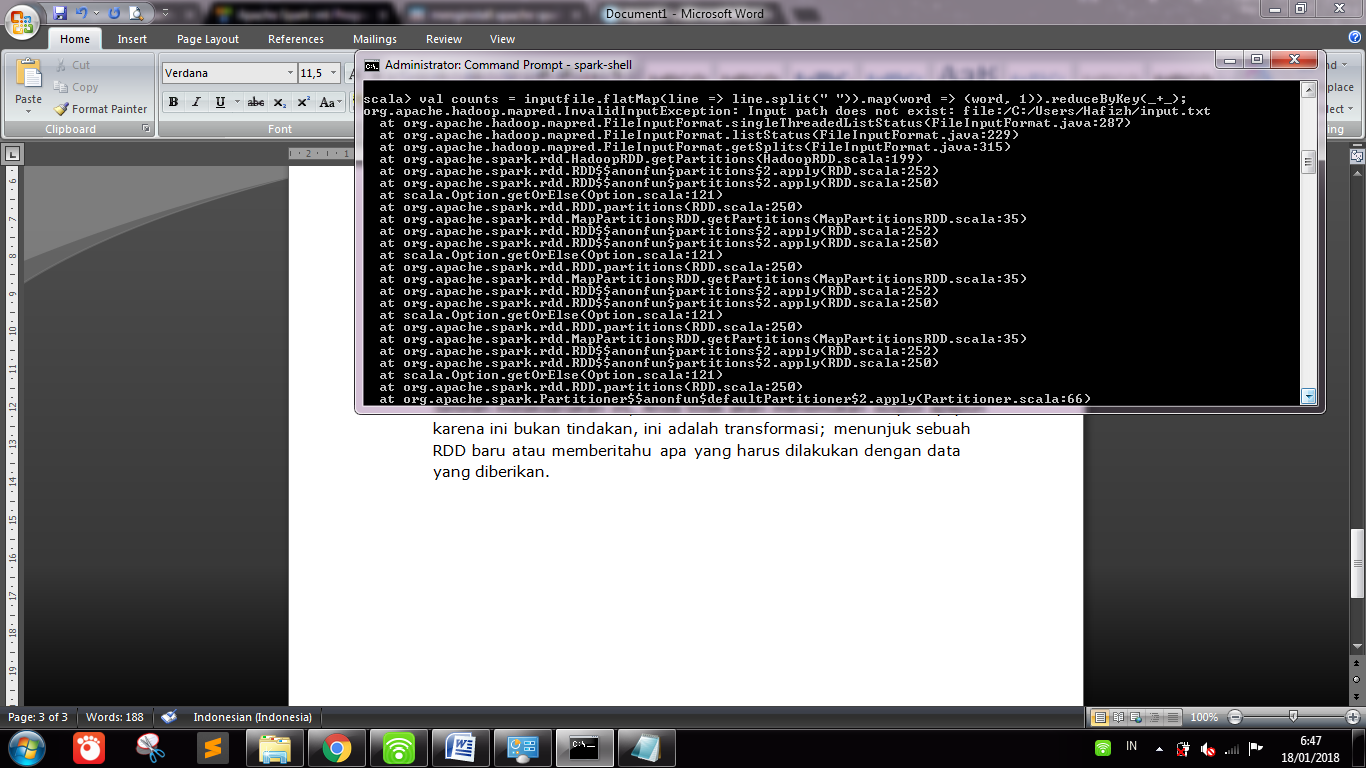
1. Membuat RDD untuk menghitung jumlah kata yang muncul dalam dokumen. Teks berikut sebagai masukan dan disimpan sebagai file **input.txt dalam direktori hafizh**

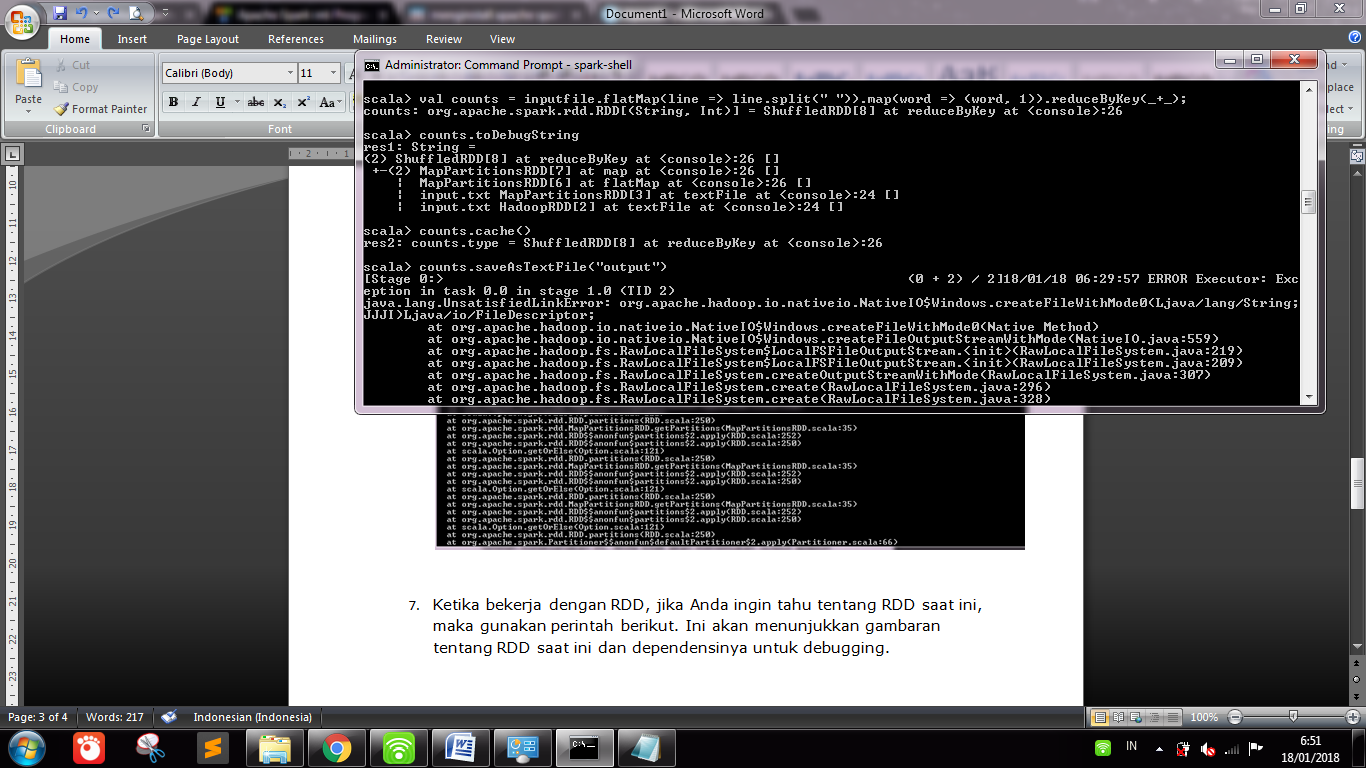


1. Perintah berikut digunakan untuk membaca file dari lokasi tertentu. Di sini, RDD baru dibuat dengan nama inputfile. String yang diberikan sebagai argumen dalam textfile ( "") metode adalah path absolut untuk nama file input. Namun, jika hanya nama file yang diberikan, maka itu berarti bahwa file input adalah di lokasi saat ini.

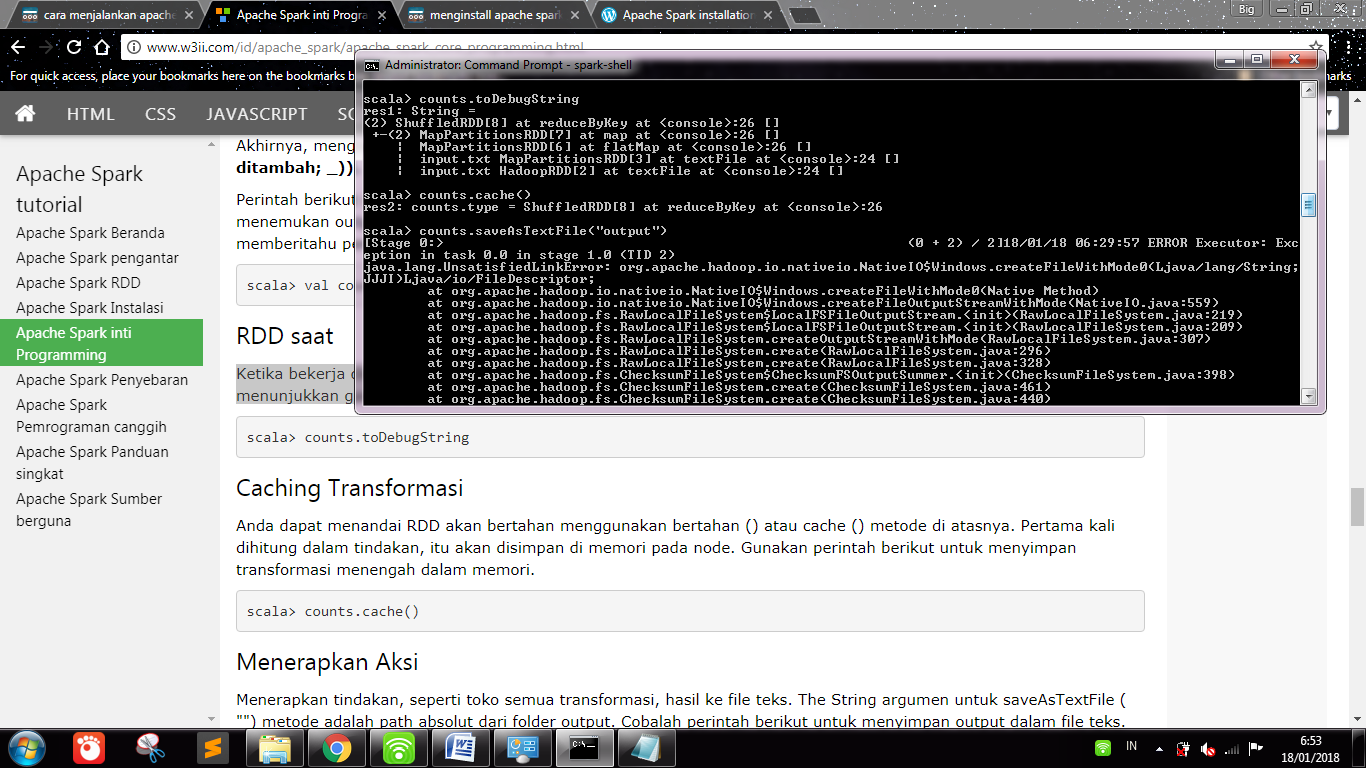


1. Perintah berikut ini digunakan untuk mengeksekusi logika jumlah kata. Setelah melaksanakan ini, Anda tidak akan menemukan output apapun karena ini bukan tindakan, ini adalah transformasi; menunjuk sebuah RDD baru atau memberitahu apa yang harus dilakukan dengan data yang diberikan.

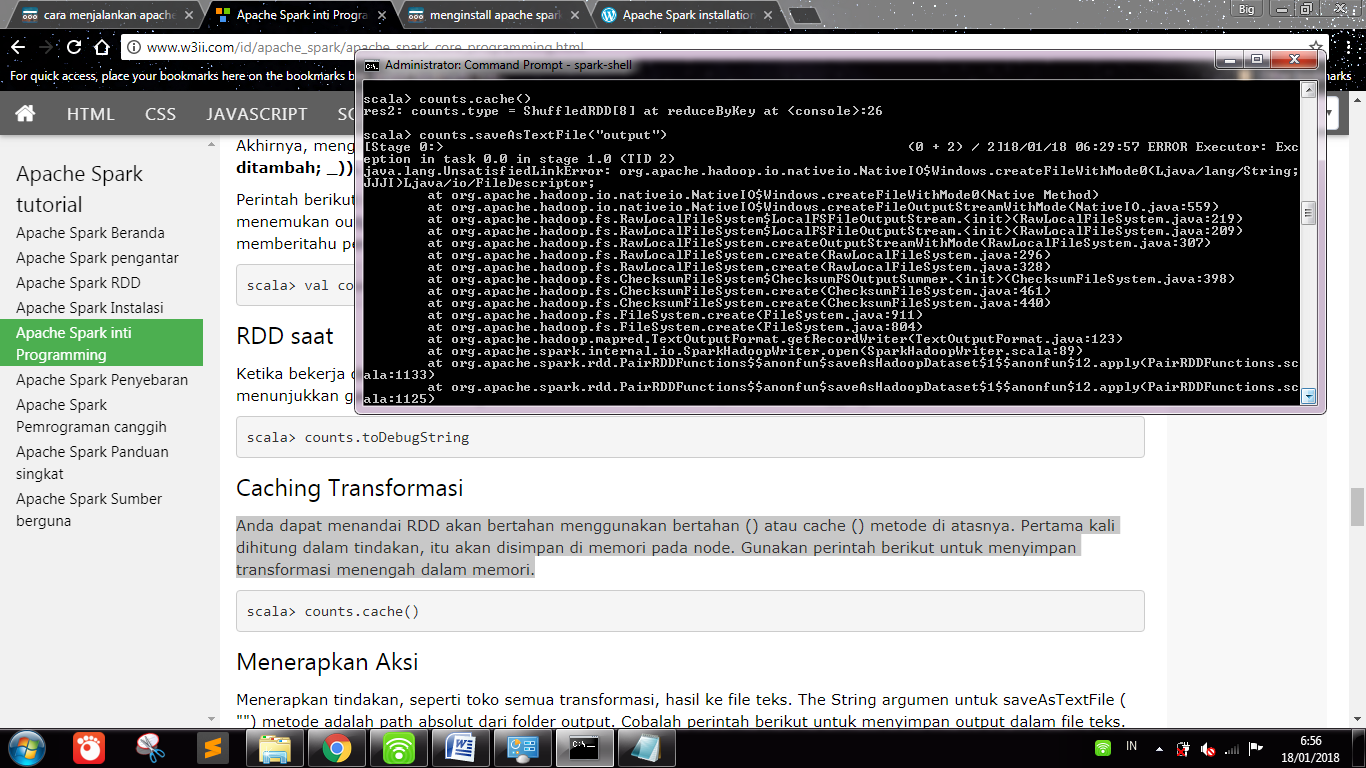




1. Perintah berikut digunakan untuk menunjukkan gambaran tentang RDD saat ini dan dependensinya untuk debugging.



1. Perintah berikut digunakan untuk menghitung tranformasi hasil pemghitungan.



1. perintah berikut untuk menyimpan output dalam file teks.

