

Big Data

Assignment 2 / Final Exam
W9-2017

W9 question source is Hadoop

- Map Reduce

- write to HDFS using Command

2 hrs

Date
Set

200 MB data

is Hive table? or not - data

is data DS or not? ~~mean~~, ~~data~~ ~~data~~

min 5 hrs

(Hive) 3
Implication

spark 2

Present
14-15
Nov
Business
Sales

import 1
to Basic → spark 1
to runtime → HIVE / Impala 2

can join / sub-query
is also can

~~trend~~

plot and
SQL

query run
group by
↓

export data
to Hadoop

or

Excel to

Hive / Impala
graph