



大數據資料處理 – Hadoop MapReduce Programming Windoop (MS-Window+Hadoop)

授課教授：王經篤 (Jing-Doo Wang)

Email: jdwang@asia.edu.tw

亞洲大學 資訊工程系 (CSIE, Asia University)

學習目標

- Windoop安裝 (含Eclipse4.5 Mar)
- 觀察網頁介面Name Node & Resource Manager
- 執行一範例程式PresentElection (Eclipse 環境)

Windoop

- 掌握Hadoop翱翔雲端：Windoop應用實作指南(附光碟)
- 作者：[許清榮、林奇暲、買大誠](#)
- 出版社：[博碩](#)
- 出版日期：2012/12/24



Windoop Download (Hadoop V1)

安全 | <https://code.google.com/archive/p/windoop/downloads>

Google Code

Search this site

Projects Search All

Project

Source

Issues

Wikis

Downloads

windoop

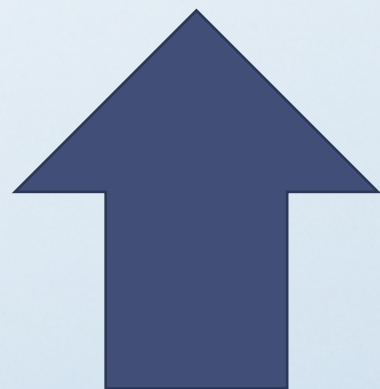
	Uploaded	Size
windoop_1.2.1_Mahout_Demo.7z	Oct 20, 2013	57.56MB
windoop_1.2.1_app_jre7_x64_zh_TW.7z	Oct 12, 2013	60.44MB
windoop_1.2.1_app_jre7_x64_zh_TW.7z	Sep 27, 2013	187.8MB
windoop_1.2.1_app_jre7_x64_zh_TW.7z	Sep 27, 2013	156.47MB
windoop_1.2.1_app_jre7_x64_en.7z	Sep 27, 2013	155.69MB

<https://code.google.com/archive/p/windoop/downloads>

Windoop Download (Hadoop V2)

- [\(Hadoop 2.0\) windoop_2.7.1_with_HBase_jre8_x64_zh_TW.7z](#) (304M)(新版)(參考)
- 感謝[林奇暲](#) 先生提供

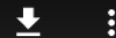
<https://drive.google.com/file/d/0BznAn3K15kCWS092R0ZGM2lIUG8/view>



請點擊網址

windoop_2.7.1_with_HBase_jre8_x64_zh_TW.7z

選擇開啟工具 ▾



※進入之後會出現下面狀況是正常的

糟糕！
顯示預覽畫面時發生問題。
正在重試...

↓ 下載

點擊下載即可



Google 雲端硬碟無法為這個檔案掃描病毒。

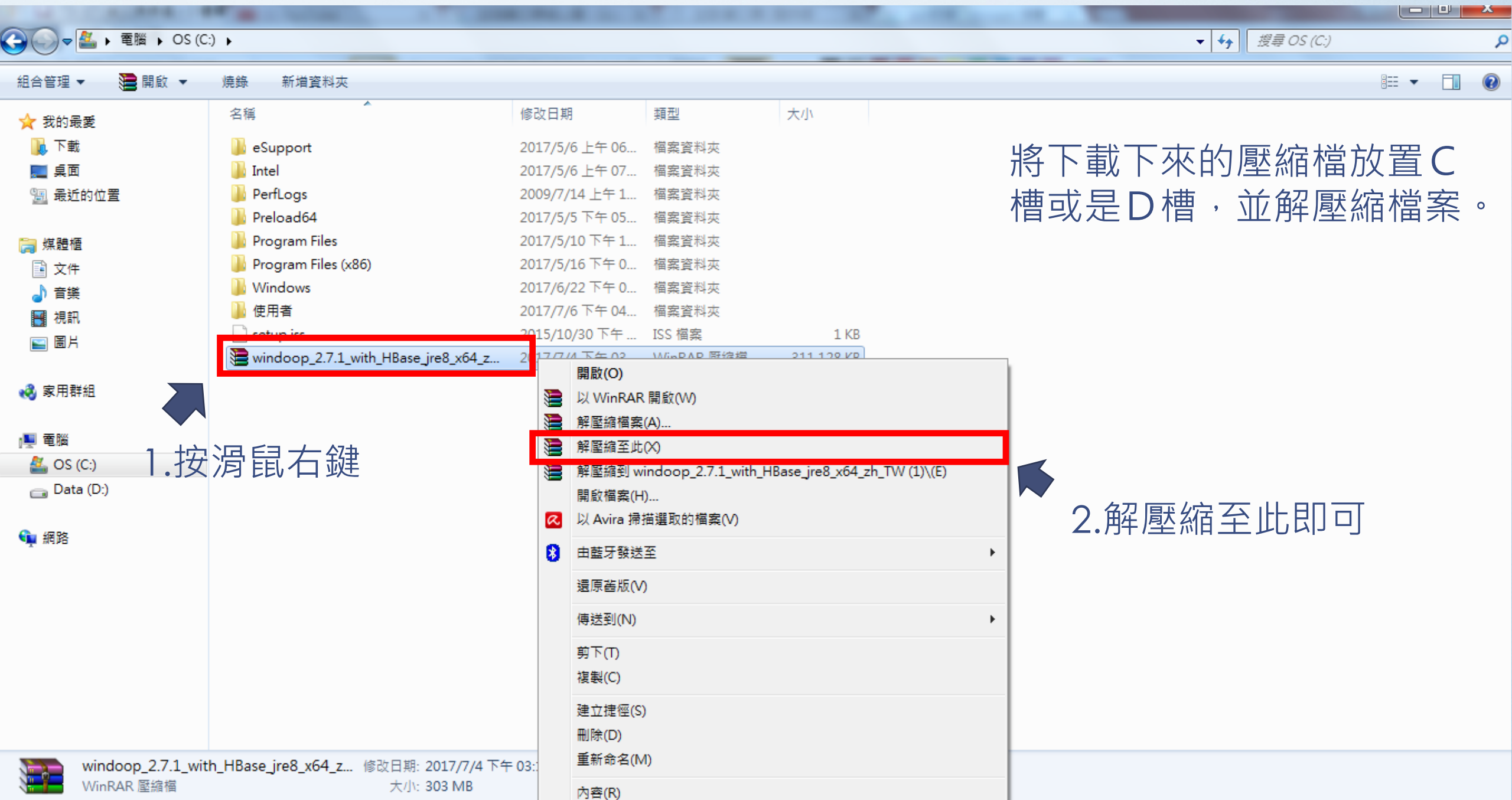
[windoop_2.7.1_with_HBase_jre8_x64_zh_TW.7z](#) (304M) 的大小超過掃描上限，您仍要下載這個檔案嗎？

仍要下載



© 2017 Google - [說明](#) - [隱私權與條款](#)

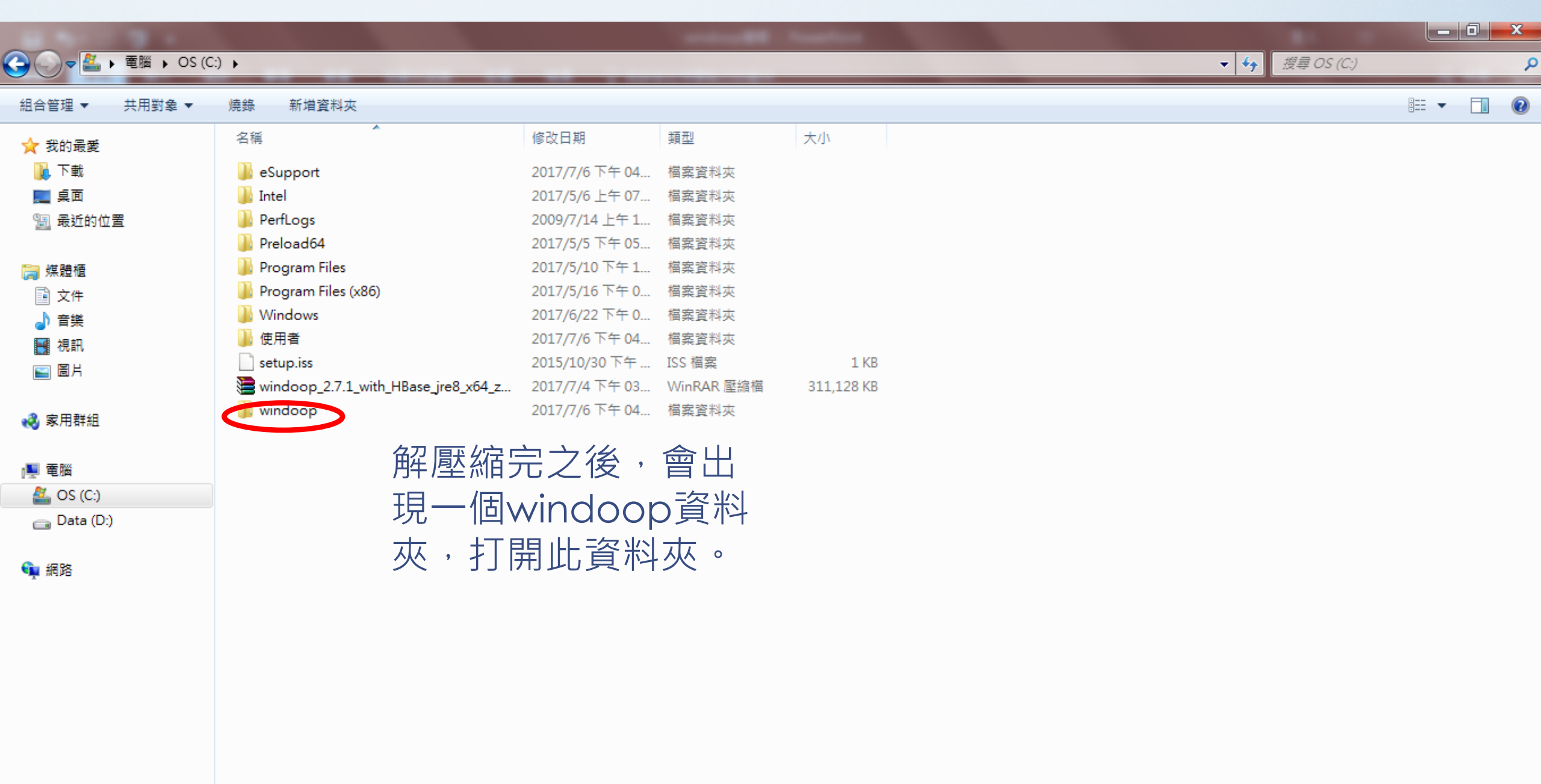
繼續下載



將下載下來的壓縮檔放置 C 槽或是 D 槽，並解壓縮檔案。

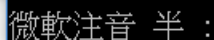
1.按滑鼠右鍵

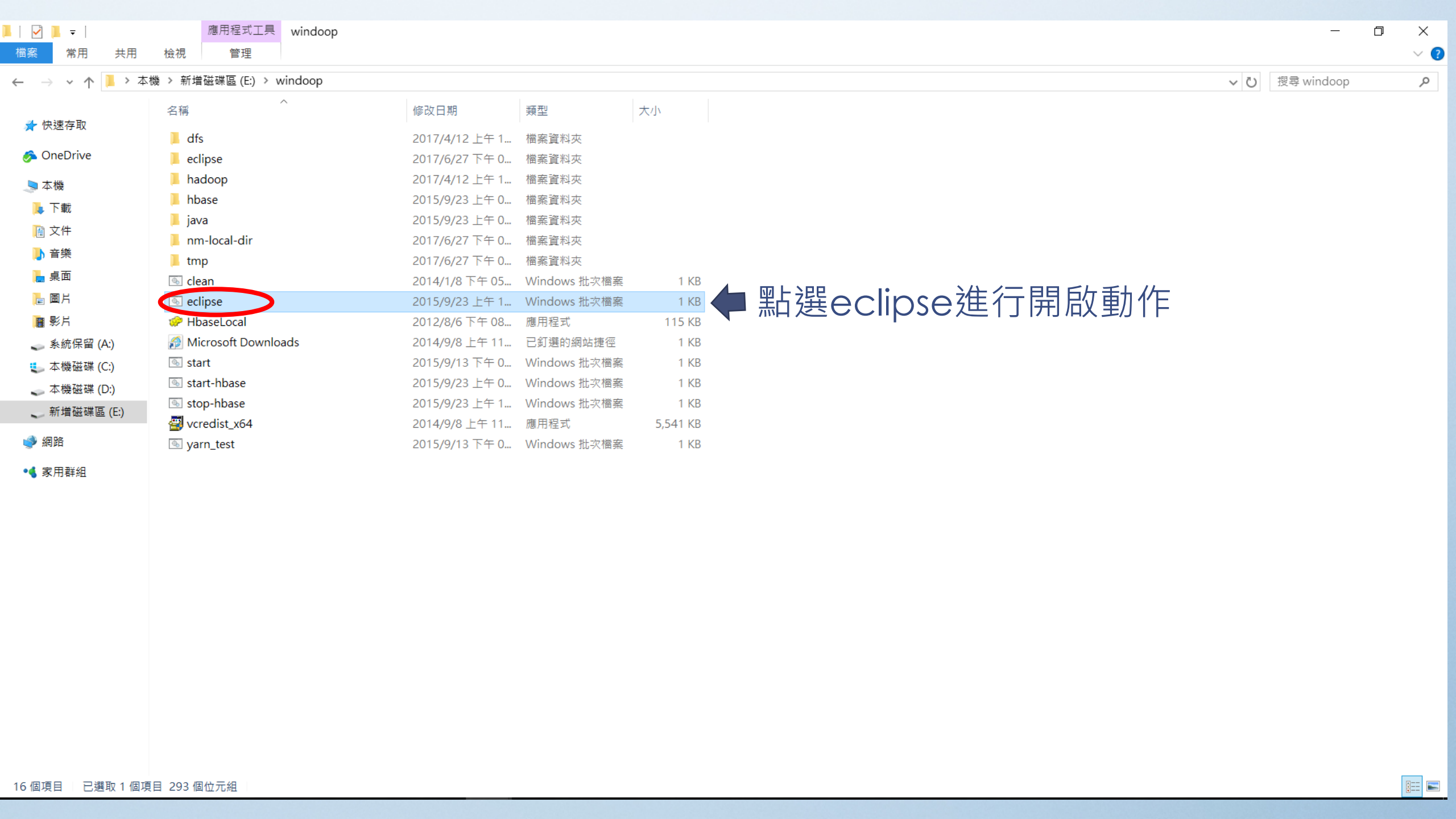
2.解壓縮至此即可



解壓縮完之後，會出現一個windoop資料夾，打開此資料夾。

搜尋 window





名稱	修改日期	類型	大小
dfs	2017/4/12 上午 1...	檔案資料夾	
eclipse	2017/6/27 下午 0...	檔案資料夾	
hadoop	2017/4/12 上午 1...	檔案資料夾	
hbase	2015/9/23 上午 0...	檔案資料夾	
java	2015/9/23 上午 0...	檔案資料夾	
nm-local-dir	2017/6/27 下午 0...	檔案資料夾	
tmp	2017/6/27 下午 0...	檔案資料夾	
clean	2014/1/8 下午 05...	Windows 批次檔案	1 KB
eclipse	2015/9/23 上午 1...	Windows 批次檔案	1 KB
HbaseLocal	2012/8/6 下午 08...	應用程式	115 KB
Microsoft Downloads	2014/9/8 上午 11...	已釘選的網站捷徑	1 KB
start	2015/9/13 下午 0...	Windows 批次檔案	1 KB
start-hbase	2015/9/23 上午 0...	Windows 批次檔案	1 KB
stop-hbase	2015/9/23 上午 1...	Windows 批次檔案	1 KB
vcredist_x64	2014/9/8 上午 11...	應用程式	5,541 KB
yarn_test	2015/9/13 下午 0...	Windows 批次檔案	1 KB

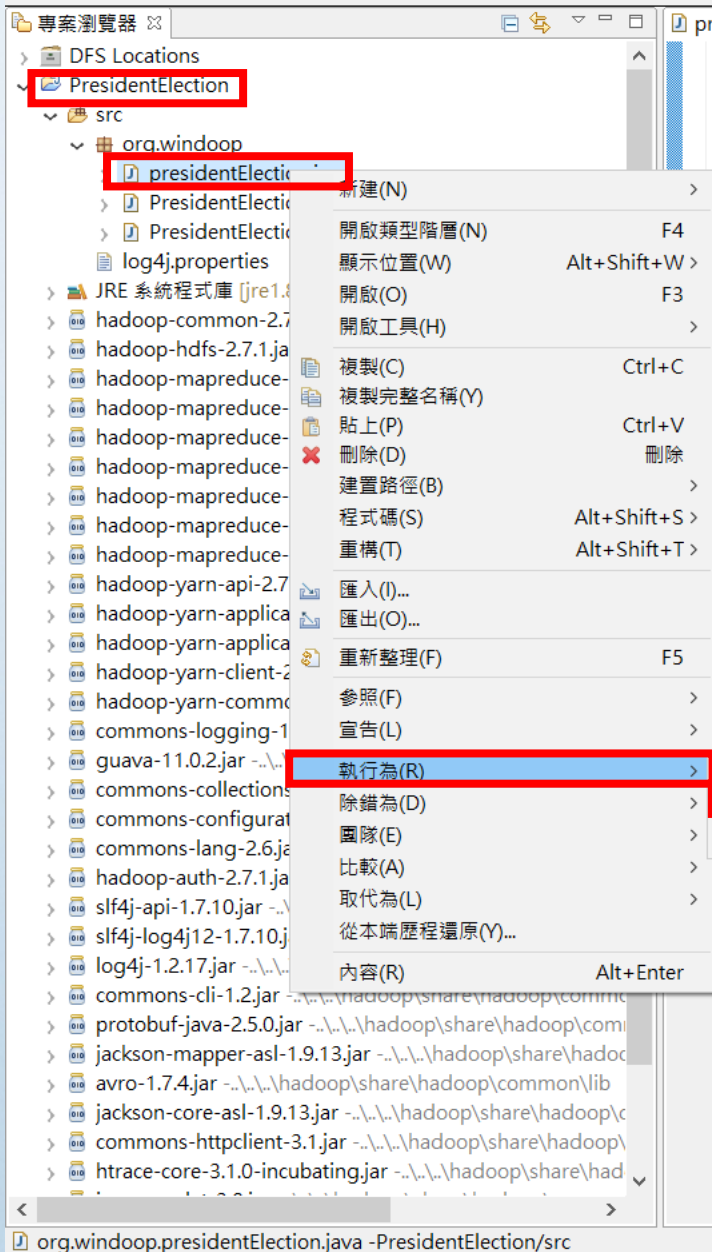
點選eclipse進行開啟動作

Summary

Security is off.
Safemode is off.
281 files and directories, 242 blocks = 523 total filesystem object(s).
Heap Memory used 73.07 MB of 145.5 MB Heap Memory. Max Heap Memory is 910.5 MB.
Non Heap Memory used 45.09 MB of 46.06 MB Committed Non Heap Memory. Max Non Heap Memory is -1 B.

也可在localhost:50070中，觀察除了live node以外的其他詳細訊息

Configured Capacity:	431.01 GB
DFS Used:	6.37 GB (1.48%)
Non DFS Used:	83.05 GB
DFS Remaining:	341.59 GB (79.25%)
Block Pool Used:	6.37 GB (1.48%)
DataNodes usages% (Min/Median/Max/stdDev):	1.48% / 1.48% / 1.48% / 0.00%
Live Nodes	1 (Decommissioned: 0)
Dead Nodes	0 (Decommissioned: 0)
Decommissioning Nodes	0
Total Datanode Volume Failures	0 (0 B)
Number of Under-Replicated Blocks	242
Number of Blocks Pending Deletion	0
Block Deletion Start Time	2017-6-27 14:08:54



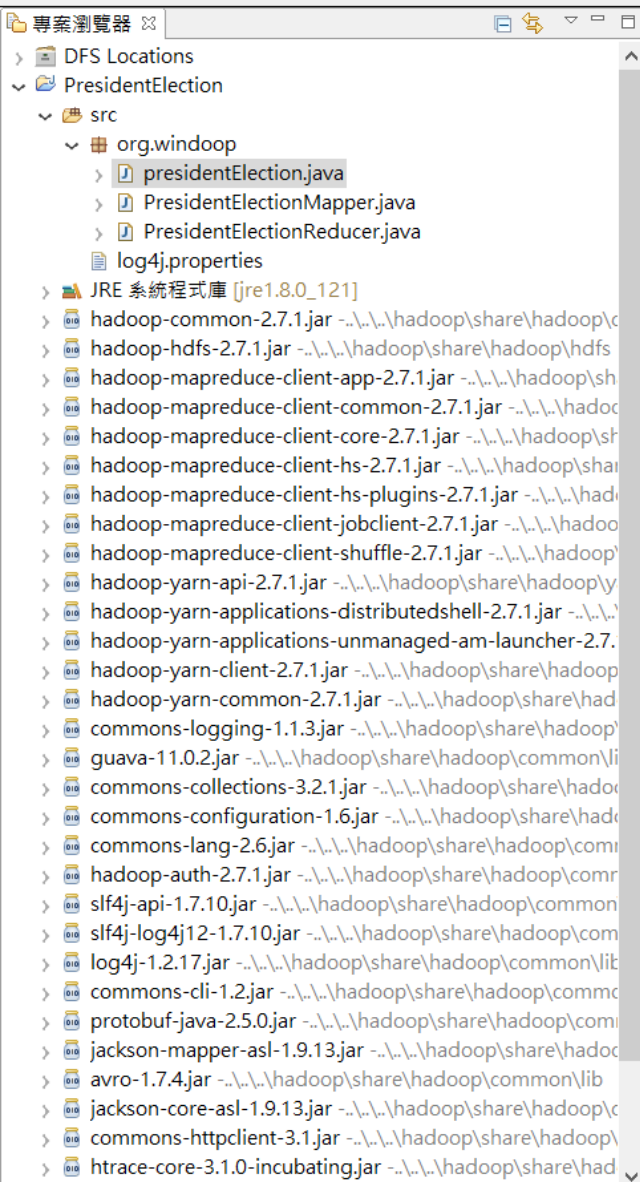
```
String[] candidate = {"A","B","C"};
String[] location = {"Taichung", "Taipei", "Kaohsiung"};
Random r;
boolean useResourceManager = true;
//boolean useJobTracker = true; <-- for hadoop 1.0

//Step 0. Set configuration
Configuration conf = new Configuration();

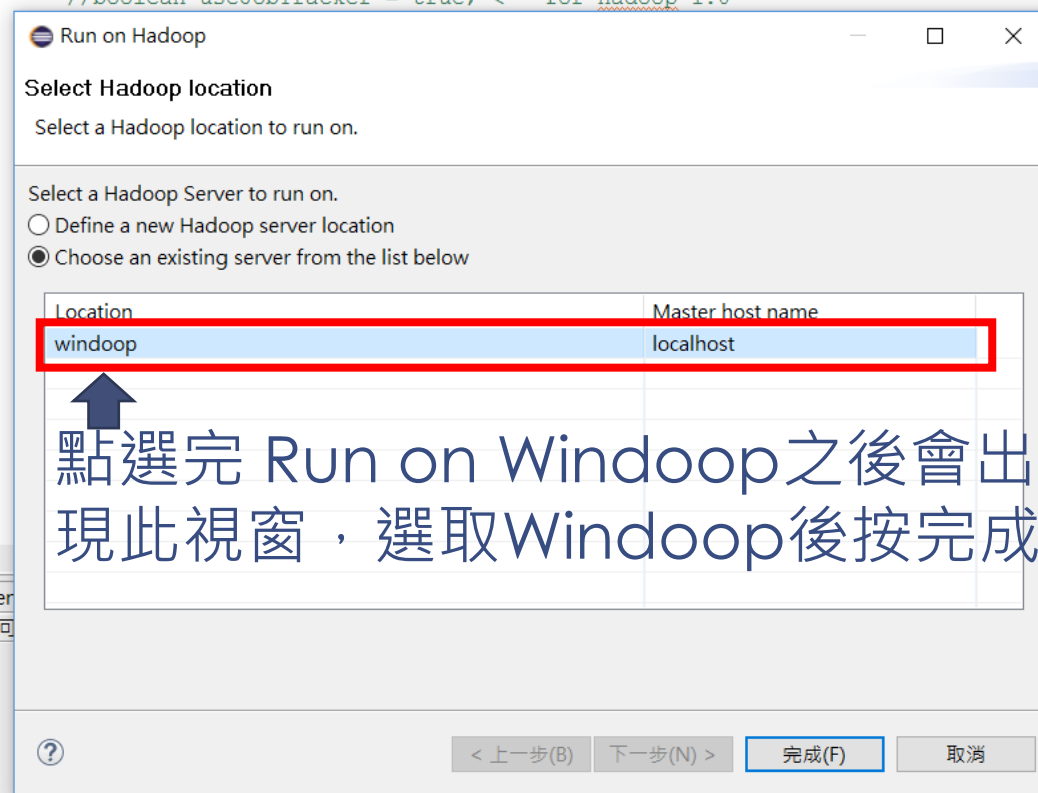
/* for hadoop 1.0
conf.set("fs.default.name", "hdfs://localhost:9000");
if(useJobTracker)
    conf.set("mapred.job.tracker", "localhost:9001");
else
    conf.set("mapred.job.tracker", "local");
*/

if(useResourceManager)
{
    conf.set("fs.defaultFS", "hdfs://0.0.0.0:9000");
    conf.set("mapreduce.framework.name", "yarn");
    //conf.set("yarn.nodemanager.aux-services", "mapreduce_shuffle");
    //conf.set("yarn.resourcemanager.hostname", "0.0.0.0");
    //conf.set("yarn.nodemanager.hostname", "0.0.0.0");
}
```

點選PresidentElection底下src，
接著PresidentElection.java右鍵，
選執行為 Run on windoop



```
String[] candidate = {"A", "B", "C"};
String[] location = {"Taichung", "Taipei", "Kaohsiung"};
Random r;
boolean useResourceManager = true;
//boolean useJobTracker = true; <-- for hadoop 1.0
```



點選完 Run on Windoop之後會出現此視窗，選取Windoop後按完成



DFS Locations
PresidentElection
src
org.windoop
presidentElection.java
PresidentElectionMapper.java
PresidentElectionReducer.java
log4j.properties
JRE 系統程式庫 [jre1.8.0_121]
hadoop-common-2.7.1.jar
hadoop-hdfs-2.7.1.jar
hadoop-mapreduce-client-app-2.7.1.jar
hadoop-mapreduce-client-common-2.7.1.jar
hadoop-mapreduce-client-core-2.7.1.jar
hadoop-mapreduce-client-hs-2.7.1.jar
hadoop-mapreduce-client-hs-plugins-2.7.1.jar
hadoop-mapreduce-client-jobclient-2.7.1.jar
hadoop-mapreduce-client-shuffle-2.7.1.jar
hadoop-yarn-api-2.7.1.jar
hadoop-yarn-applications-distributedshell-2.7.1.jar
hadoop-yarn-applications-unmanaged-am-launcher-2.7.1.jar
hadoop-yarn-client-2.7.1.jar
hadoop-yarn-common-2.7.1.jar
commons-logging-1.1.3.jar
guava-11.0.2.jar
commons-collections-3.2.1.jar
commons-configuration-1.6.jar
commons-lang-2.6.jar
hadoop-auth-2.7.1.jar
slf4j-api-1.7.10.jar
slf4j-log4j12-1.7.10.jar
log4j-1.2.17.jar
commons-cli-1.2.jar
protobuf-java-2.5.0.jar
jackson-mapper-asl-1.9.13.jar
avro-1.7.4.jar
jackson-core-asl-1.9.13.jar
commons-httpclient-3.1.jar
htrace-core-3.1.0-incubating.jar

```
String[] candidate = {"A", "B", "C"};
String[] location = {"Taichung", "Taipei", "Kaohsiung"};
Random r;
boolean useResourceManager = true;
//boolean useJobTracker = true; <-- for hadoop 1.0

//Step 0. Set configuration
Configuration conf = new Configuration();

/* for hadoop 1.0
conf.set("fs.default.name", "hdfs://localhost:9000");
if(useJobTracker)
    conf.set("mapred.job.tracker", "localhost:9001");
else
    conf.set("mapred.job.tracker", "local");
*/

if(useResourceManager)
{
    conf.set("fs.defaultFS", "hdfs://0.0.0.0:9000");
    conf.set("mapreduce.framework.name", "yarn");
    //conf.set("yarn.nodemanager.aux-services", "mapreduce_shuffle");
    //conf.set("yarn.resourcemanager.hostname", "0.0.0.0");
    //conf.set("yarn.nodemanager.hostname", "0.0.0.0");
}
```

Problems Tasks Javadoc Map/Reduce Locations 主控台
<已终止> presidentElection [Java 應用程式] C:\Program Files\Java\jre1.8.0_121\bin\javaw.exe (2017年6月27日 下午2:20:19 直接存取)

```
Total committed heap usage (bytes)=737148928
Shuffle Errors
BAD_ID=0
CONNECTION=0
IO_ERROR=0
WRONG_LENGTH=0
WRONG_MAP=0
WRONG_REDUCE=0
File Input Format Counters
Bytes Read=9000000
File Output Format Counters
Bytes Written=28
```

Job Done!

※如果出現 Job Failed代表在 start 執行
的時候未使用系統管理員身分執行

執行成功會出現 Job Done的字樣

執行成功後，點擊 DFS Locations 的 Windoop 下有一個 user 的資料夾，裡面會有一個 input 跟 output 資料夾

Map/Reduce - part-r-00000 (28.0 b, r3) - Eclipse平台

檔案(F) 編輯(E) 程式碼(S) 重構(T) 導覽(N) 搜尋(A) 專案(P) 執行(R) 視窗(W) 說明(H)

專案瀏覽器

- DFS Locations
 - windoop
 - (2)
 - tmp (1)
 - user (1)
 - (17)
 - input (3)
 - Kaohsiung.txt (2.9 Mb, r3)
 - Taichung.txt (2.9 Mb, r3)
 - Taipei.txt (2.9 Mb, r3)
- PresidentElection
 - src
 - org.windoop
 - presidentElection.java
 - PresidentElectionMapper.java
 - PresidentElectionReducer.java
 - log4j.properties
 - JRE 系統程式庫 [jre1.8.0_121]
 - hadoop-common-2.7.1.jar
 - hadoop-hdfs-2.7.1.jar
 - hadoop-mapreduce-client-app-2.7.1.jar

快速存取

Java Map/Reduce

概要

無法取得概要。

選票結果

A	999045
B	999474
C	1001481

點選 output 的 part-00000 (28.0b,r3)
可觀看選票結果

已選取 1 個項目

Problems Tasks Javadoc Map/Reduce Locations 主控台

<已終止> presidentElection [Java 應用程式] C:\Program Files\Java\jre1.8.0_121\bin\javaw.exe (2017年6月27日 下午2:20:19 直接存取)

```
Total committed heap usage (bytes)=737148928
Shuffle Errors
BAD_ID=0
CONNECTION=0
IO_ERROR=0
WRONG_LENGTH=0
WRONG_MAP=0
WRONG_REDUCE=0
File Input Format Counters
Bytes Read=9000000
File Output Format Counters
Bytes Written=28
Job Done!
```