

大數據資料處理 – Hadoop MapReduce Programming

Hadoop Java 程式 載入 修改與執行

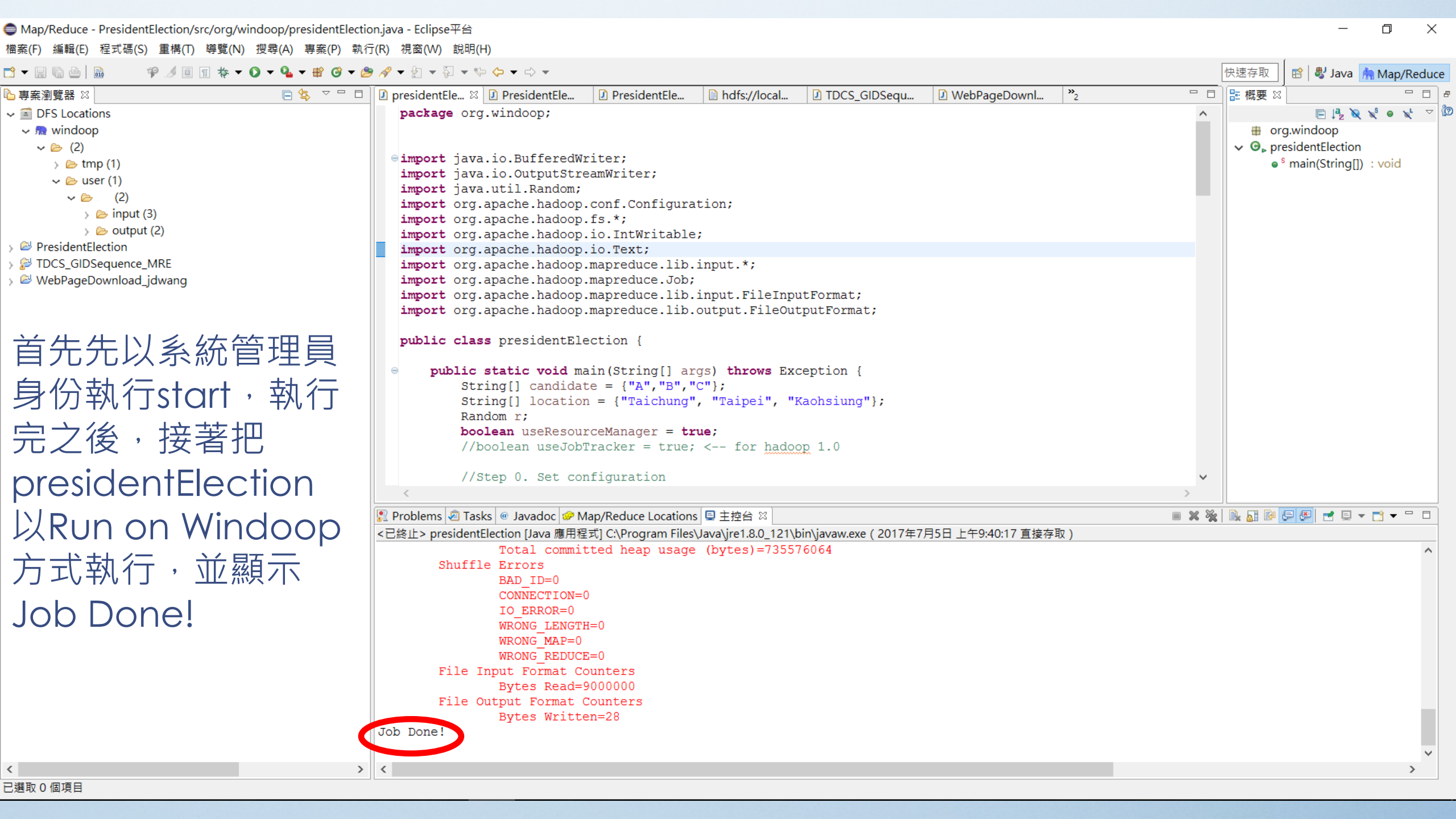
授課教授：王經篤 (Jing-Doo Wang)

Email: jdwang@asia.edu.tw

亞洲大學 資訊工程系 （CSIE, Asia University）

學習目標

- 下載 Java MapReduce 程式
- 匯入 Java MapReduce 程式
- 修改外部 JAR 路徑
- 設定程式環境變數
- 了解 Mapper 與 Reduce 的關係



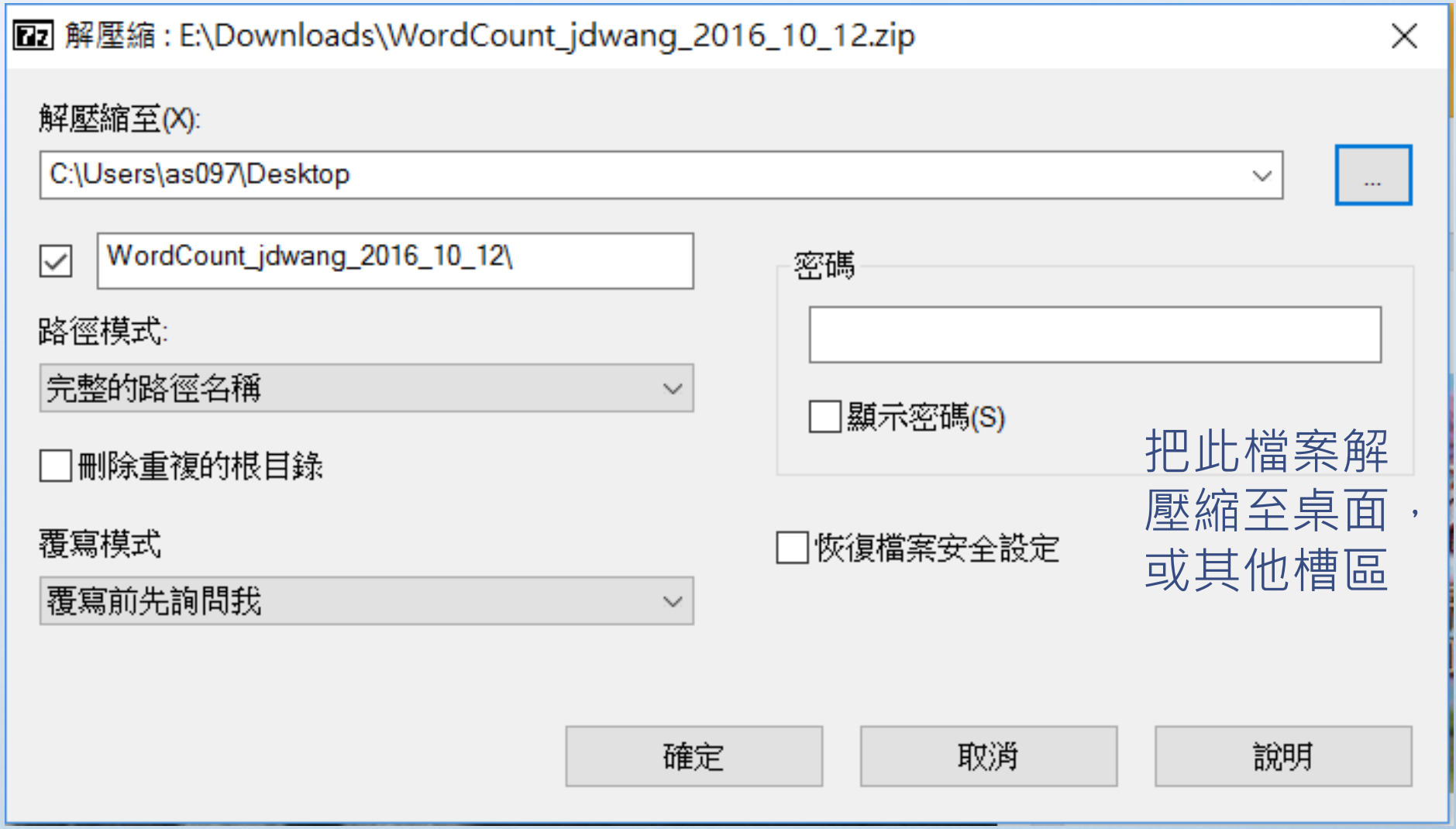
首先先以系統管理員身份執行start，執行完之後，接著把presidentElection以Run on Windoop方式執行，並顯示Job Done!

WordCount_jdwang_2016_10_12.zip



點選此段網址，下載壓縮檔案

http://dns2.asia.edu.tw/~jdwang/TeachingCourses/Hadoop/WordCount_jdwang_2016_10_12.zip



把此檔案解壓縮至桌面，或其他槽區

Map/Reduce - PresidentElection/src/org/windoop/presidentElection.java - 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)

新建(N) Alt+Shift+N >
開啟檔案(.)...
關閉(C) Ctrl+W
全部關閉(L) Ctrl+Shift+W
儲存(S) Ctrl+S
另存新檔(A)...
全部儲存(E) Ctrl+Shift+S
回復(T)
移動(V)...
重新命名(M)... F2
重新整理(F) F5
將行定界字元轉換成(V) >
列印(P)... Ctrl+P
切換工作區(W) >
Restart
匯入(I)...
匯出(O)...
內容(R) Alt+Enter
1 presidentElection.java [PresidentEl...]
2 TDCS_GIDSequence_MRE.java [TDCS_GID...]
3 WebPageDownload_jdwang_1.java [WebP...]
結束(X)

選取檔案底下的匯入

```java
package org.windoop;


import java.io.BufferedWriter;
import java.io.OutputStreamWriter;
import java.util.Random;
import org.apache.hadoop.conf.Configuration;
import org.apache.hadoop.fs.*;
import org.apache.hadoop.io.IntWritable;
import org.apache.hadoop.io.Text;
import org.apache.hadoop.mapreduce.lib.input.*;
import org.apache.hadoop.mapreduce.Job;
import org.apache.hadoop.mapreduce.lib.input.FileInputFormat;
import org.apache.hadoop.mapreduce.lib.output.FileOutputFormat;

public class presidentElection {

    public static void main(String[] args) throws Exception {
        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
```

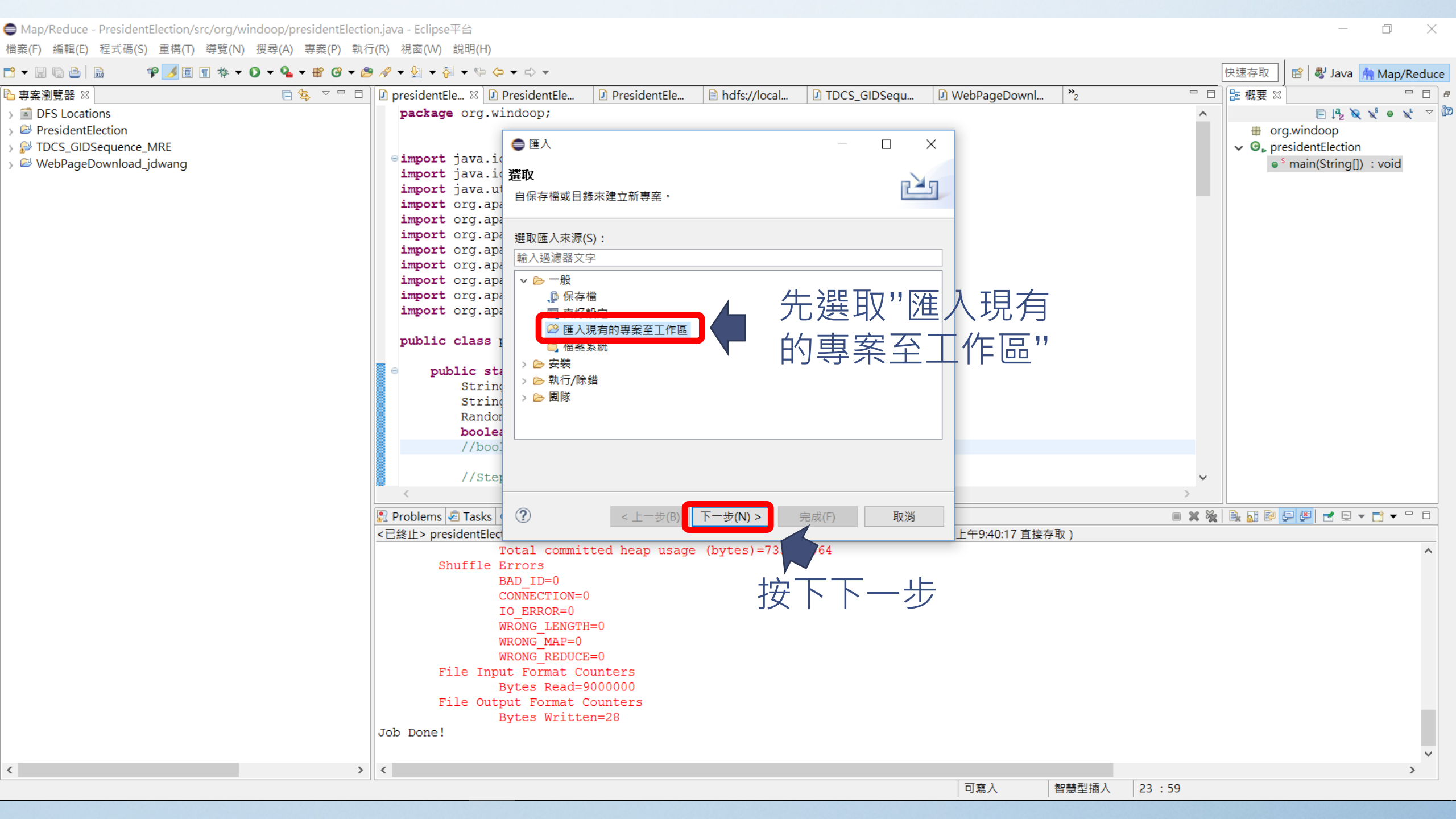
概要
org.windoop
presidentElection
main(String[]) : void

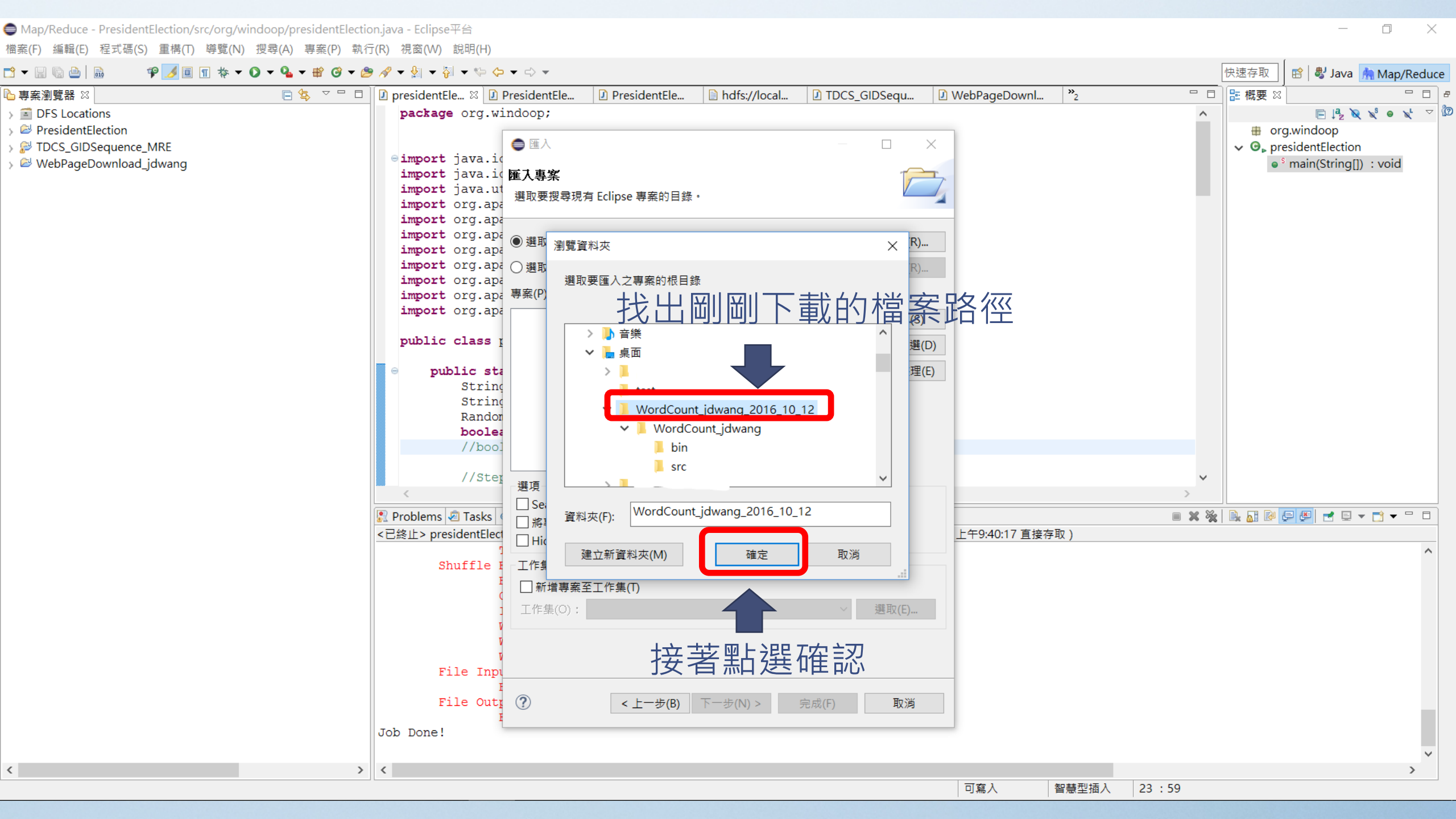
Problems Tasks Javadoc Map/Reduce Locations 主控台

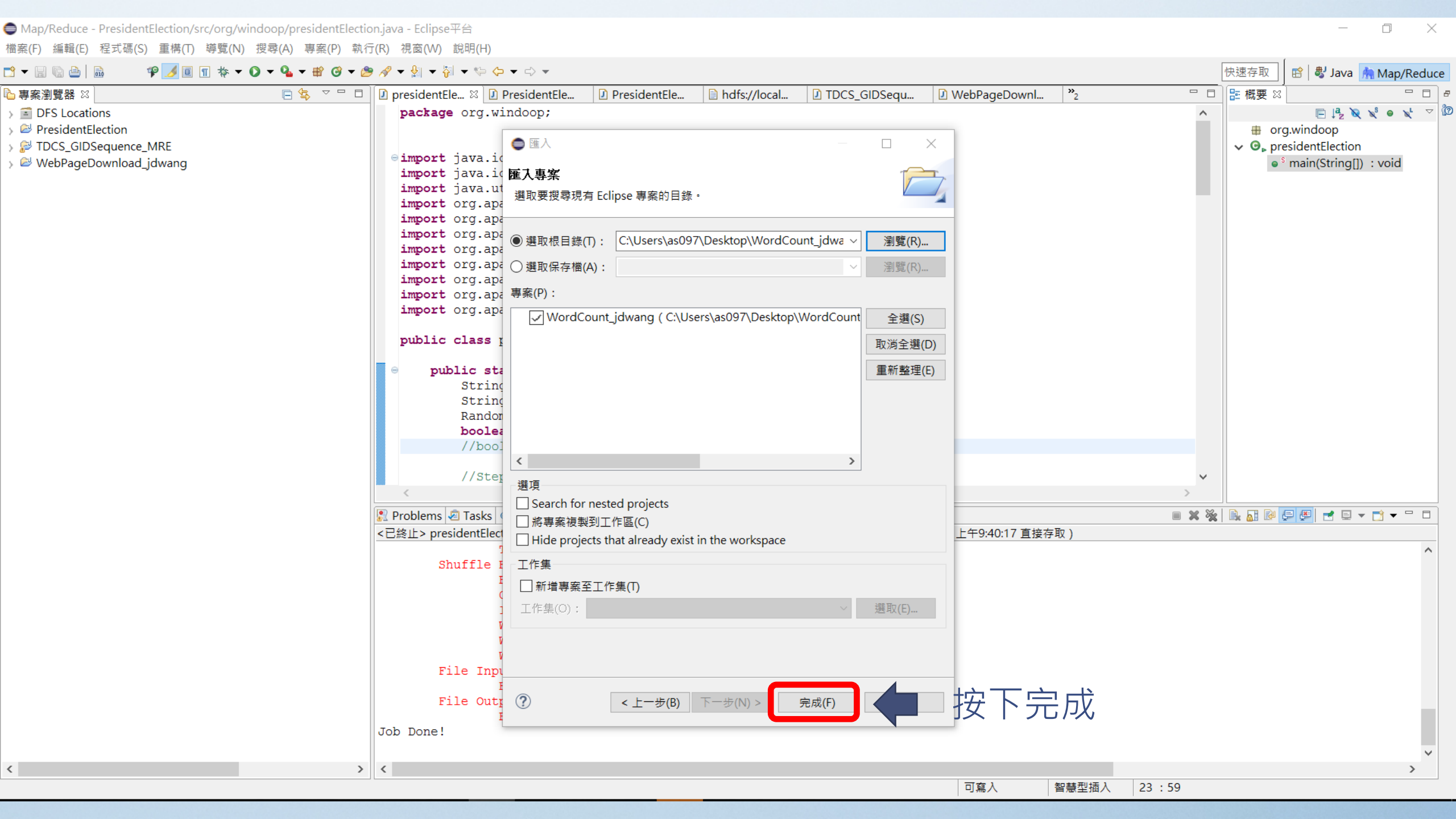
<已終止> presidentElection [Java 應用程式] C:\Program Files\Java\jre1.8.0_121\bin\javaw.exe (2017年7月5日 上午9:40:17 直接存取)

```
                Total committed heap usage (bytes)=735576064
        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!
```

可寫入 智慧型插入 23 : 59


Map/Reduce - PresidentElection/src/org/windoop/presidentElection.java - 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
PresidentElection
TDCS_GIDSequence_MRE
WebPageDownload_jdwang
匯入
選取
自保存檔或目錄來建立新專案。
選取匯入來源(S)：
輸入過濾器文字
一般
保存檔
匯入現有的專案至工作區
檔案系統
安裝
執行/除錯
團隊
< 上一步(B)
下一步(N) >
完成(F)
取消
先選取"匯入現有的專案至工作區"
按下下一步
概要
org.windoop
presidentElection
main(String[]) : void
Problems
Tasks
Total committed heap usage (bytes)=73
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!


找出剛剛下載的檔案路徑
接著點選確認
匯入專案
選取要搜尋現有 Eclipse 專案的目錄。
瀏覽資料夾
選取要匯入之專案的根目錄
WordCount_jdwang_2016_10_12
WordCount_jdwang
bin
src
資料夾(F):
WordCount_jdwang_2016_10_12
建立新資料夾(M)
確定
取消
新增專案至工作集(T)
< 上一步(B)
下一步(N) >
完成(F)
取消
Job Done!


Map/Reduce - PresidentElection/src/org/windoop/presidentElection.java - 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
PresidentElection
TDCS_GIDSequence_MRE
WebPageDownload_jdwang
匯入
匯入專案
選取要搜尋現有 Eclipse 專案的目錄。
選取根目錄(T)：
C:\Users\as097\Desktop\WordCount_jdwa
瀏覽(R)...
選取保存檔(A)：
專案(P)：
WordCount_jdwang (C:\Users\as097\Desktop\WordCount
全選(S)
取消全選(D)
重新整理(E)
選項
Search for nested projects
將專案複製到工作區(C)
Hide projects that already exist in the workspace
工作集
新增專案至工作集(T)
工作集(O)：
選取(E)...
< 上一步(B)
下一步(N) >
完成(F)
按下完成
Problems
Tasks
Job Done!
概要
org.windoop
presidentElection
main(String[]) : void

Map/Reduce - PresidentElection/src/org/windoop/presidentElection.java - 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
- PresidentElection
- TDCS_GIDSequence_MRE
- WebPageDownload_jdwang
- WordCount_jdwang
  - src
    - (預設套件)
      - Mapper_jdwang.java
      - Reducer_jdwang.java
      - WordCount_jdwang.java
  - JRE 系統程式庫 [jre1.8.0_121]

完成之後
WordCount_jdwang
旁邊會出現一個驚嘆號

presidentEle... | PresidentEle... | PresidentEle... | hdfs://local... | TDCS_GIDSe

```
package org.windoop;


import java.io.BufferedWriter;
import java.io.OutputStreamWriter;
import java.util.Random;
import org.apache.hadoop.conf.Configuration;
import org.apache.hadoop.fs.*;
import org.apache.hadoop.io.IntWritable;
import org.apache.hadoop.io.Text;
import org.apache.hadoop.mapreduce.lib.input.*;
import org.apache.hadoop.mapreduce.Job;
import org.apache.hadoop.mapreduce.lib.input.FileInputFormat;
import org.apache.hadoop.mapreduce.lib.output.FileOutputFormat;

public class presidentElection {

    public static void main(String[] args) throws Exception {
        String[] candidate = {"A","B","C"};
        String[] location = {"Taichung", "Taipei", "Kaohsiung"};
        Random r;
        boolean useResourceManager = true;
        //boolean useJobTracker = true; <-- for hadoop 1.0

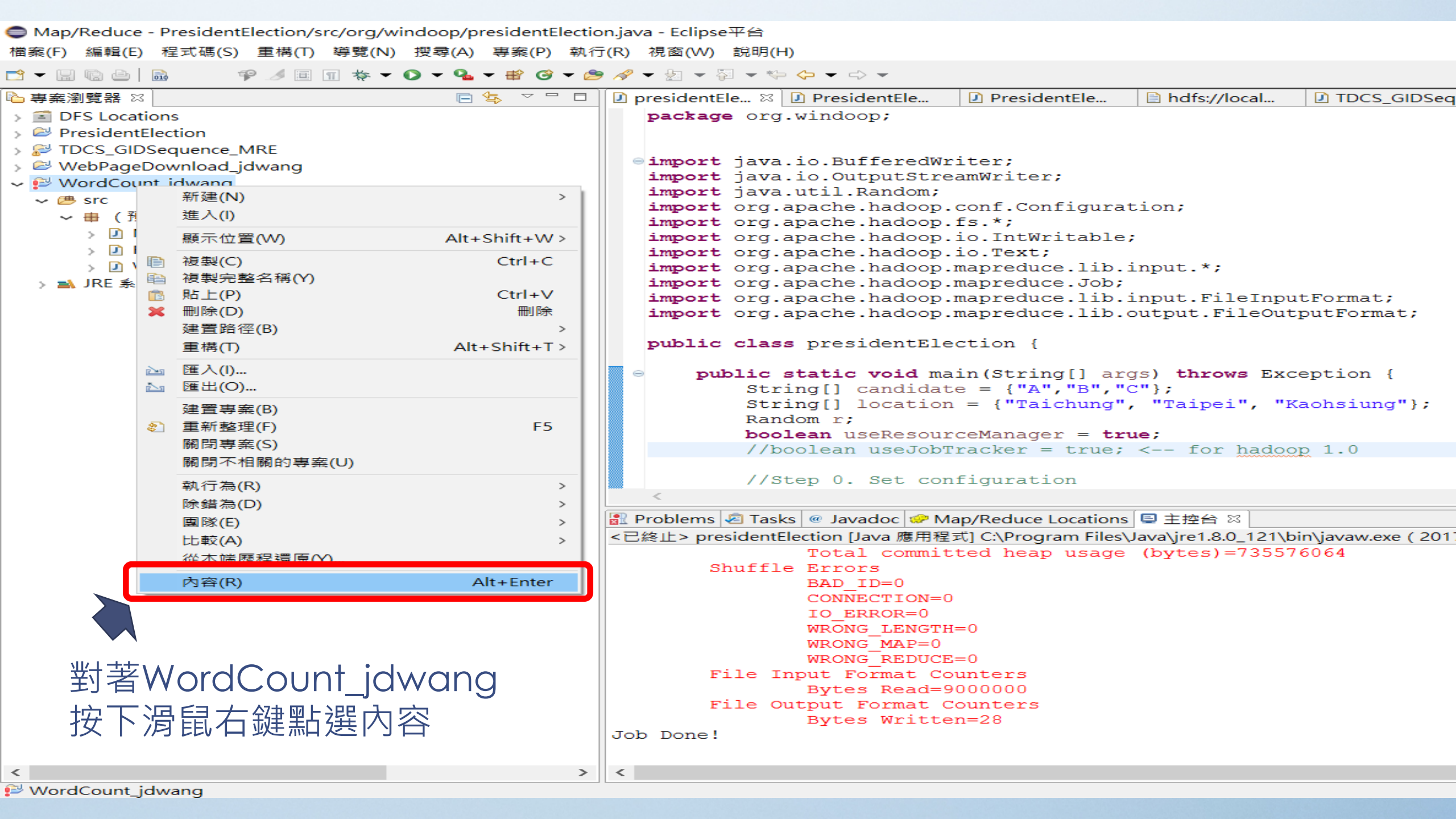
        //Step 0. Set configuration
```

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 (20

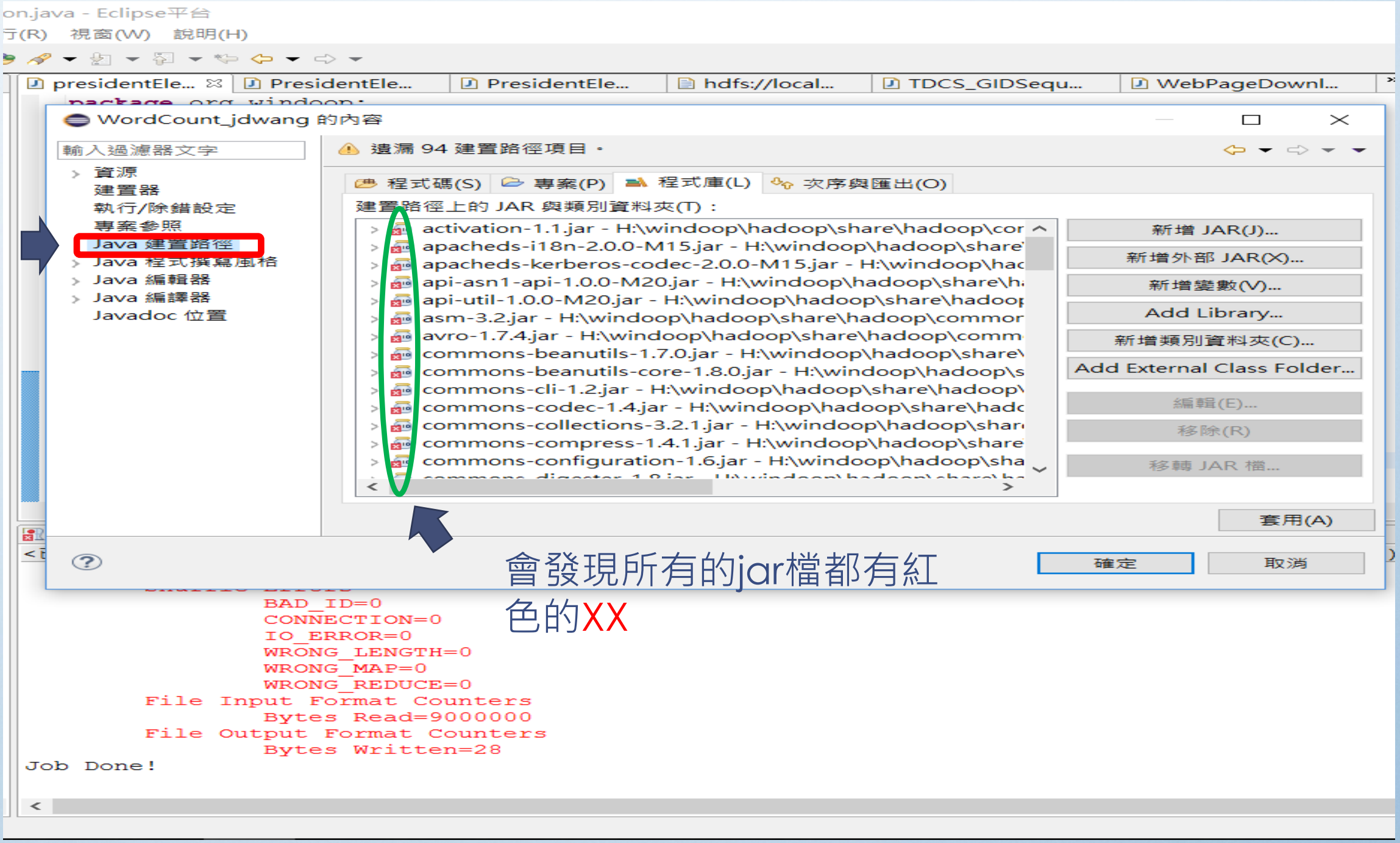
```
                Total committed heap usage (bytes)=735576064
        Shuffle Errors
                BAD_ID=0
                CONNECTION=0
                IO_ERROR=0
                WRONG_LENGTH=0
                WRONG_MAP=0
                WRONG_REDUCE=0
        File Input Format Counters
                Bytes Read=9000000
        File Output Format Counters
                Bytes Written=28
Job Done!
```

已選取 0 個項目

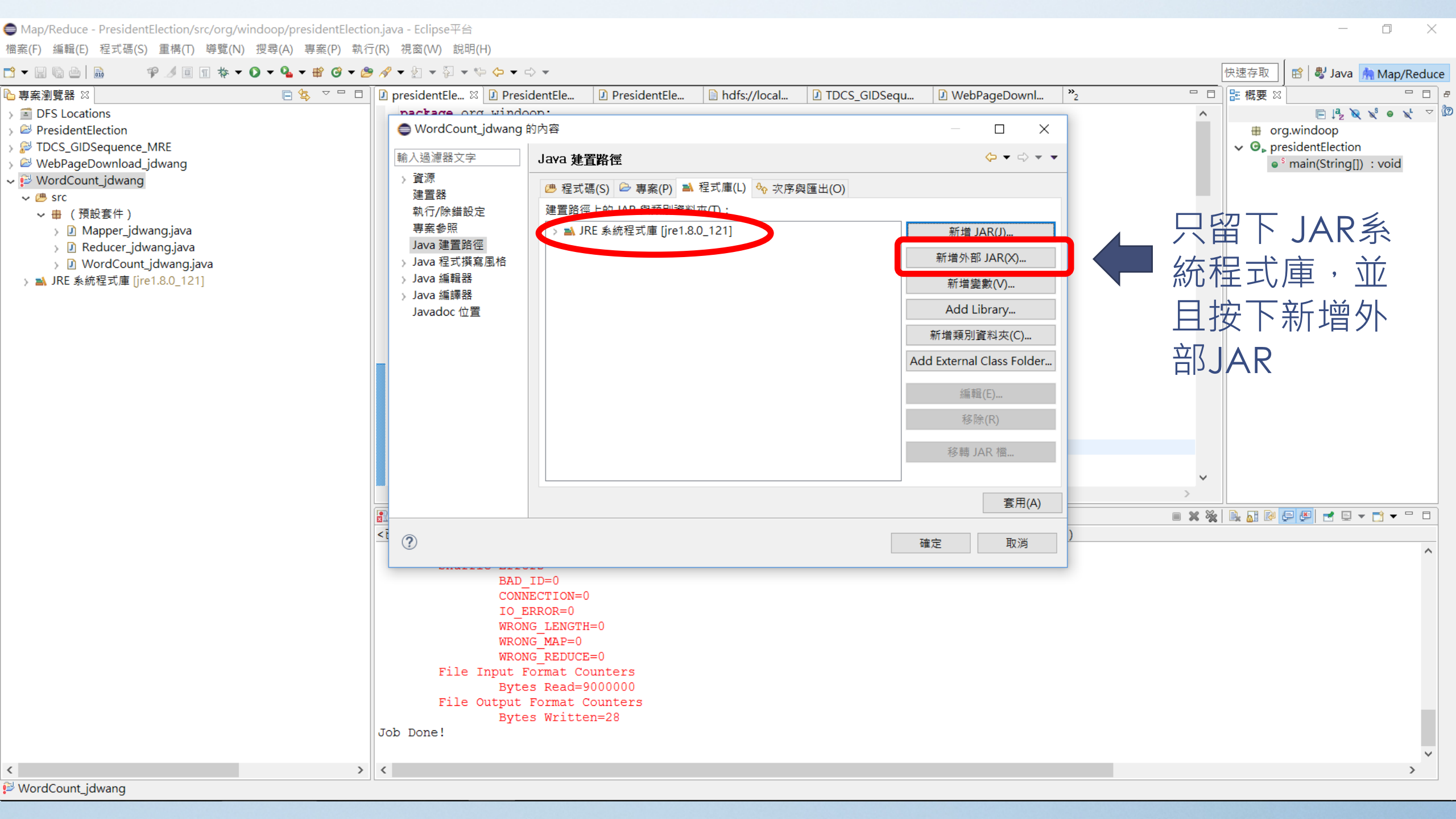


對著WordCount_jdwang
按下滑鼠右鍵點選內容

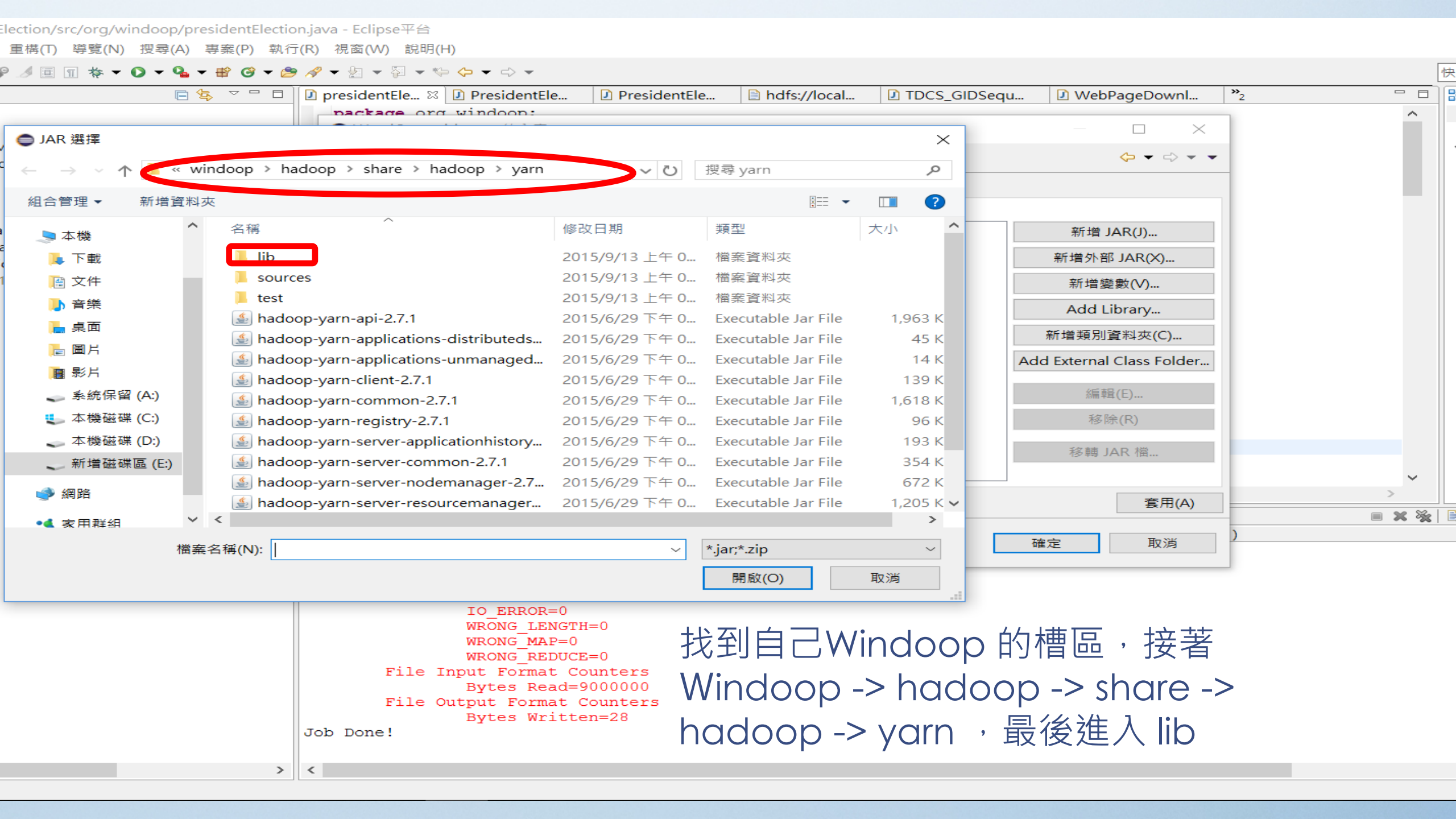
先點選Java建置路徑



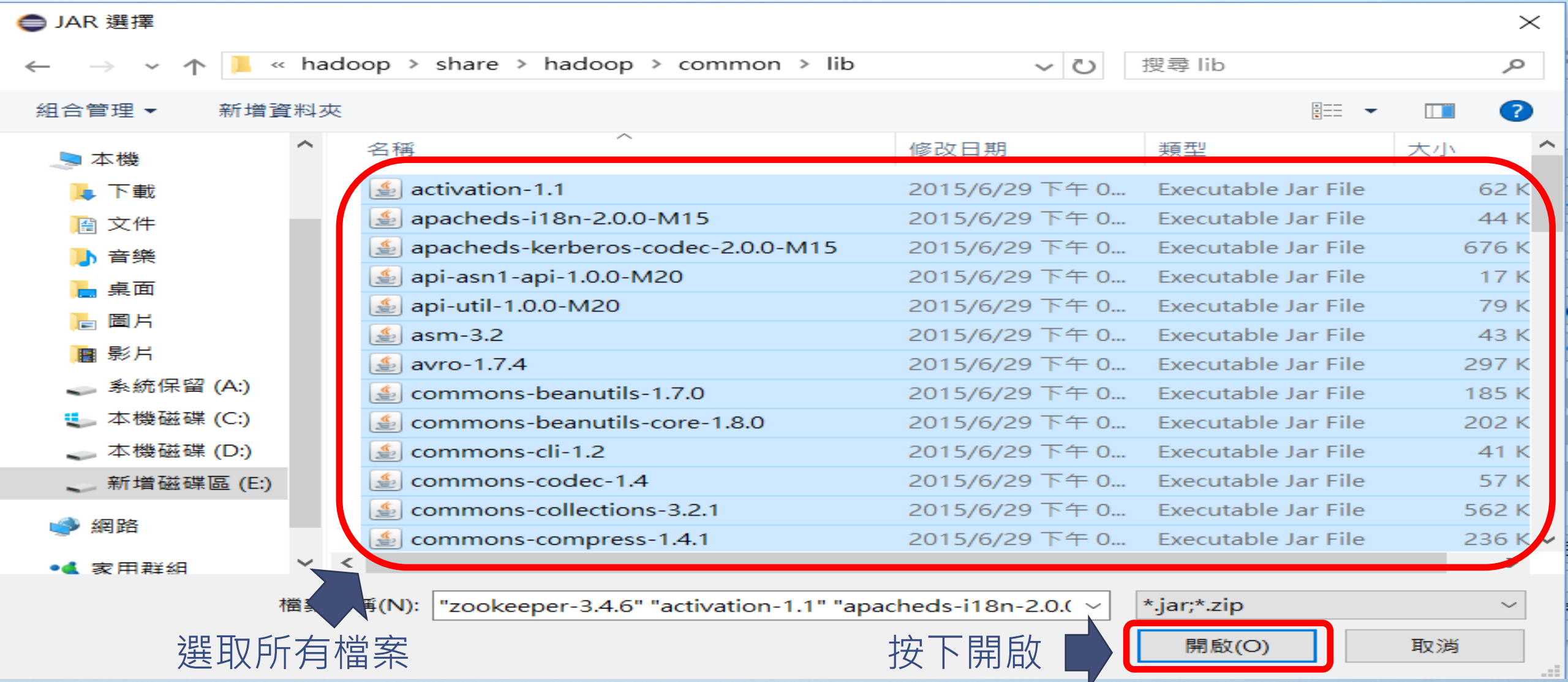
會發現所有的jar檔都有紅色的XX

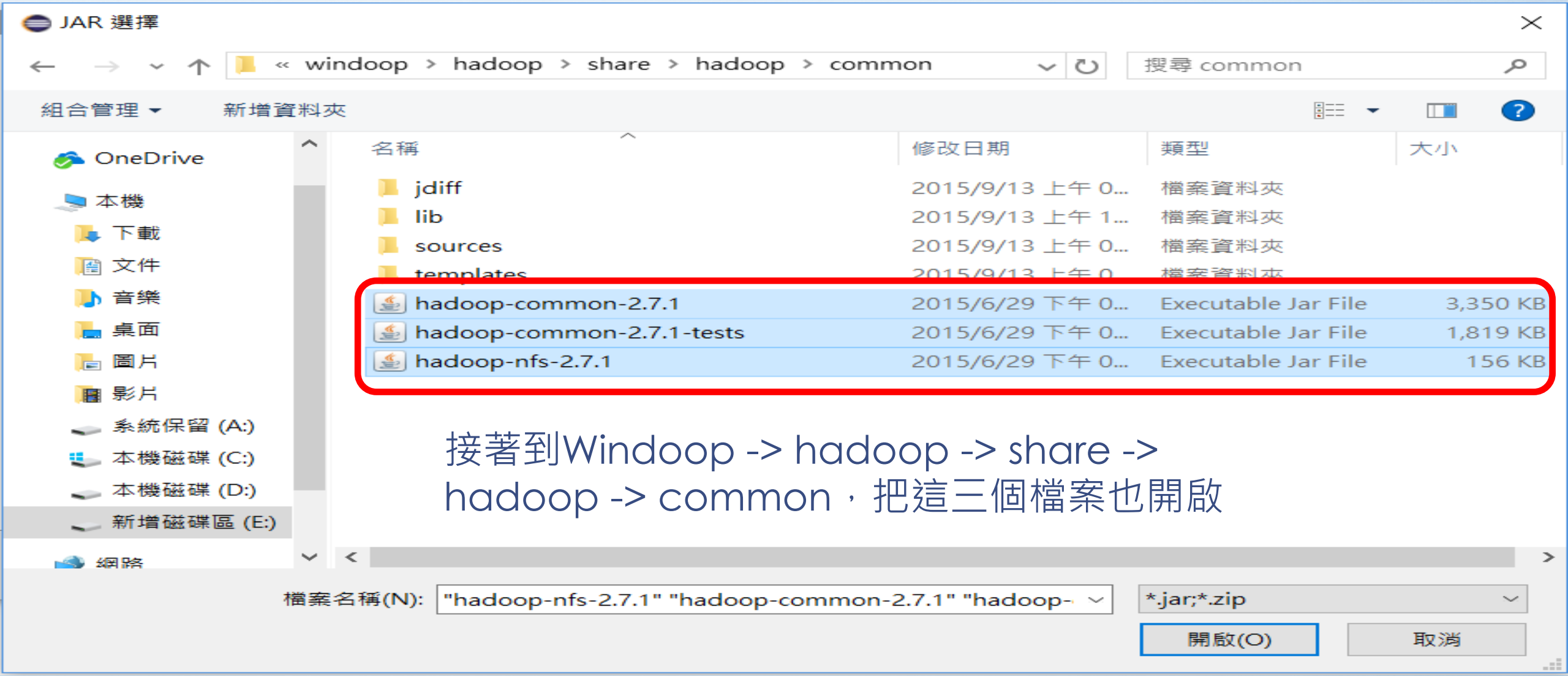


只留下 JAR系統程式庫，並且按下新增外部JAR

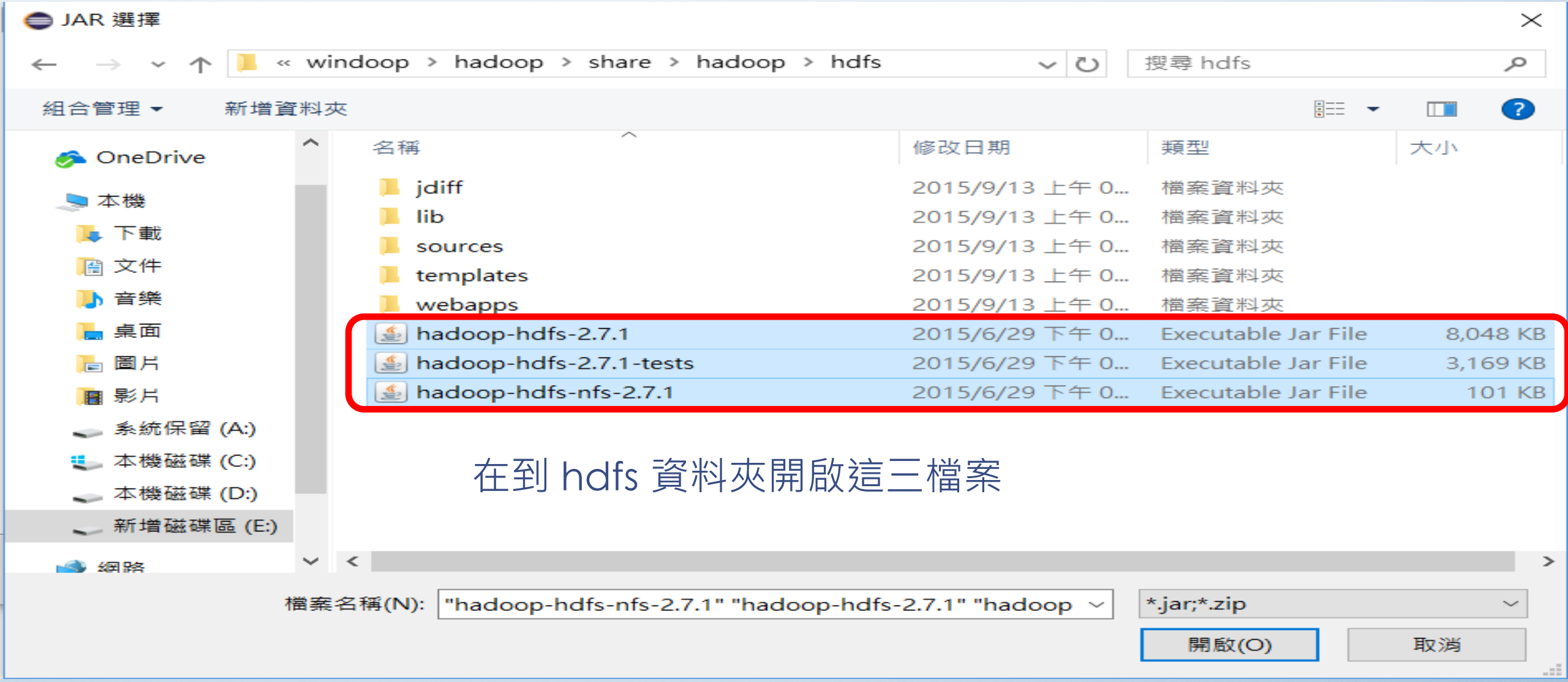


找到自己Windoop 的槽區，接著 Windoop -> hadoop -> share -> hadoop -> yarn ，最後進入 lib

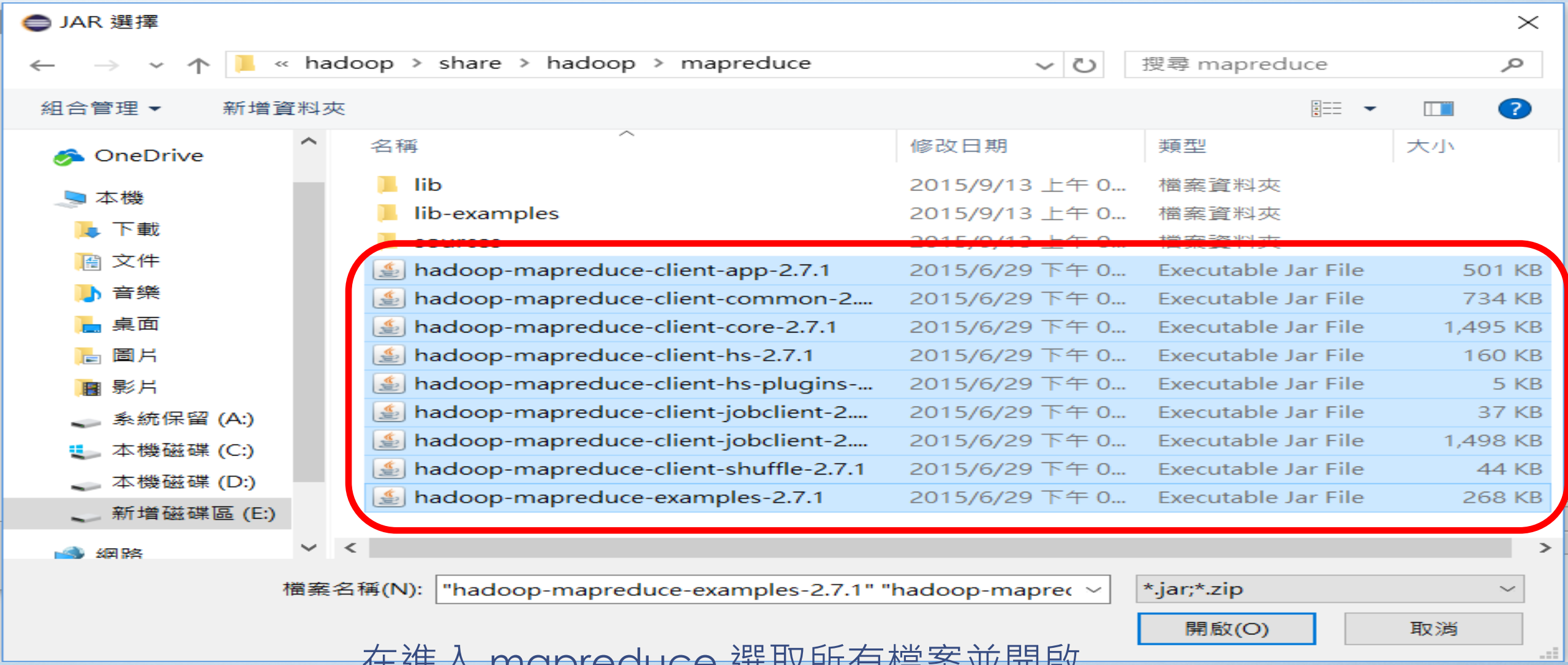

JAR 選擇
« hadoop > share > hadoop > common > lib
搜尋 lib
組合管理
新增資料夾
本機
下載
文件
音樂
桌面
圖片
影片
系統保留 (A:)
本機磁碟 (C:)
本機磁碟 (D:)
新增磁碟區 (E:)
網路
家用群組
名稱
修改日期
類型
大小
activation-1.1 2015/6/29 下午 0... Executable Jar File 62 K
apacheds-i18n-2.0.0-M15 2015/6/29 下午 0... Executable Jar File 44 K
apacheds-kerberos-codec-2.0.0-M15 2015/6/29 下午 0... Executable Jar File 676 K
api-asn1-api-1.0.0-M20 2015/6/29 下午 0... Executable Jar File 17 K
api-util-1.0.0-M20 2015/6/29 下午 0... Executable Jar File 79 K
asm-3.2 2015/6/29 下午 0... Executable Jar File 43 K
avro-1.7.4 2015/6/29 下午 0... Executable Jar File 297 K
commons-beanutils-1.7.0 2015/6/29 下午 0... Executable Jar File 185 K
commons-beanutils-core-1.8.0 2015/6/29 下午 0... Executable Jar File 202 K
commons-cli-1.2 2015/6/29 下午 0... Executable Jar File 41 K
commons-codec-1.4 2015/6/29 下午 0... Executable Jar File 57 K
commons-collections-3.2.1 2015/6/29 下午 0... Executable Jar File 562 K
commons-compress-1.4.1 2015/6/29 下午 0... Executable Jar File 236 K
"zookeeper-3.4.6" "activation-1.1" "apacheds-i18n-2.0.(
.jar;.zip
開啟(O)
取消
選取所有檔案
按下開啟



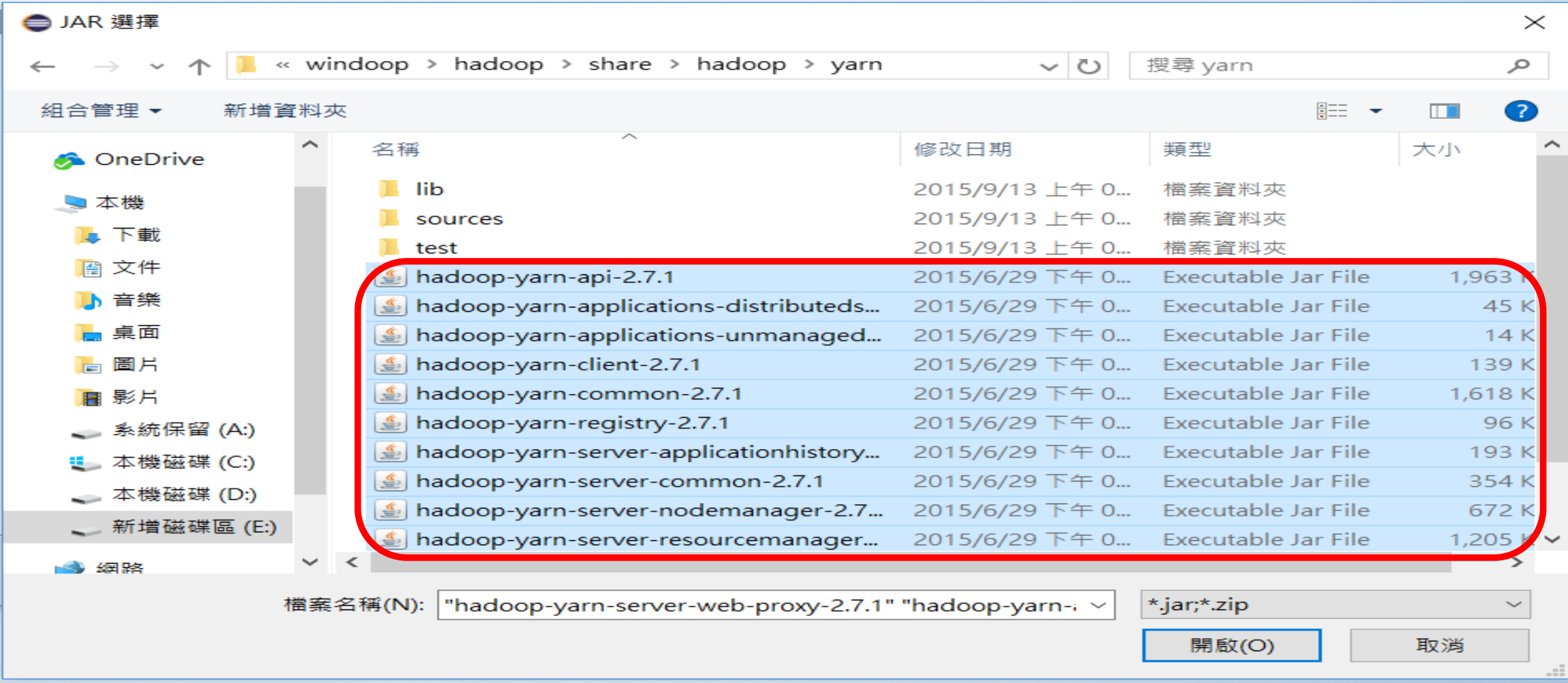
接著到Windoop -> hadoop -> share -> hadoop -> common，把這三個檔案也開啟



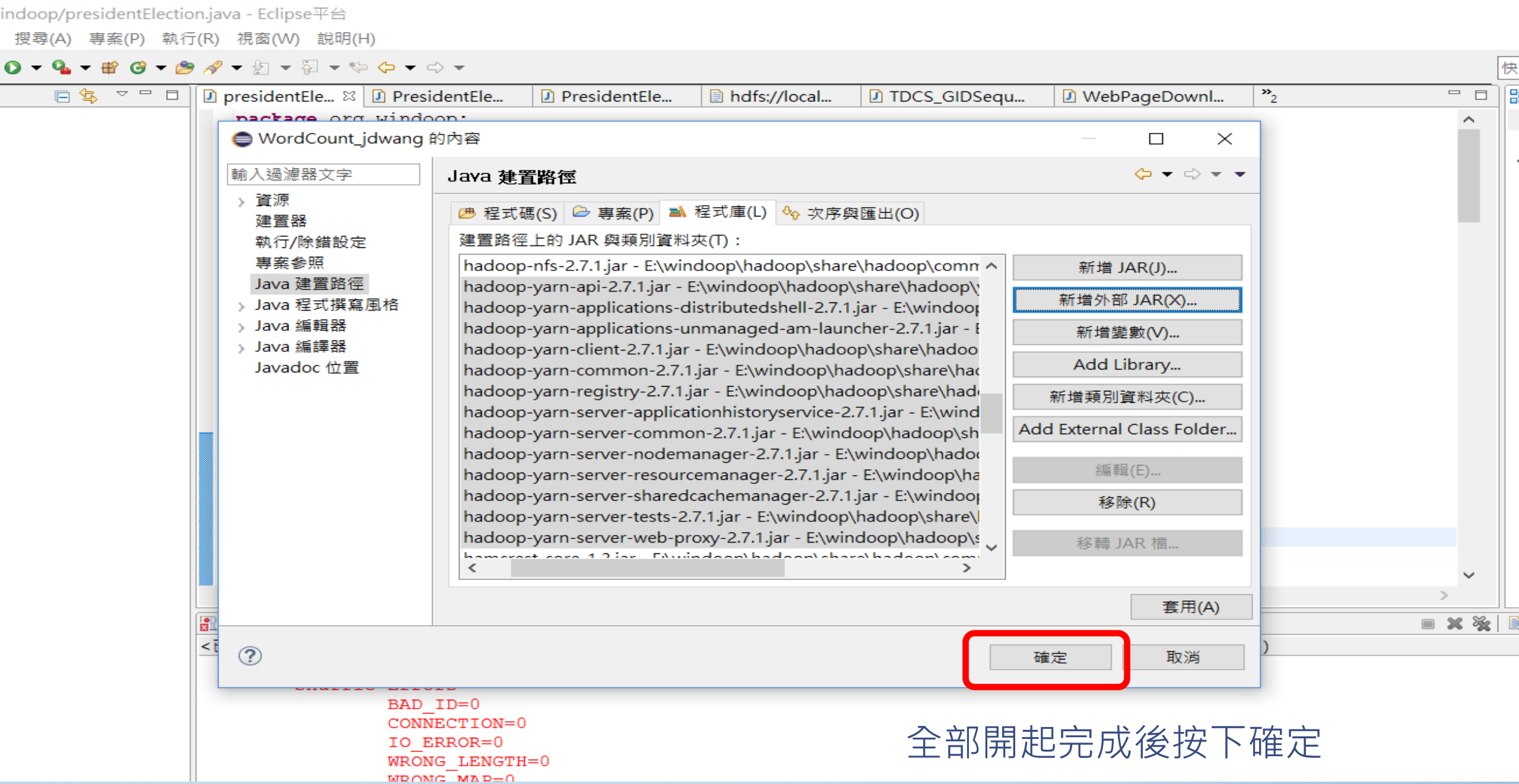
在到 hdfs 資料夾開啟這三檔案



在進入 mapreduce 選取所有檔案並開啟

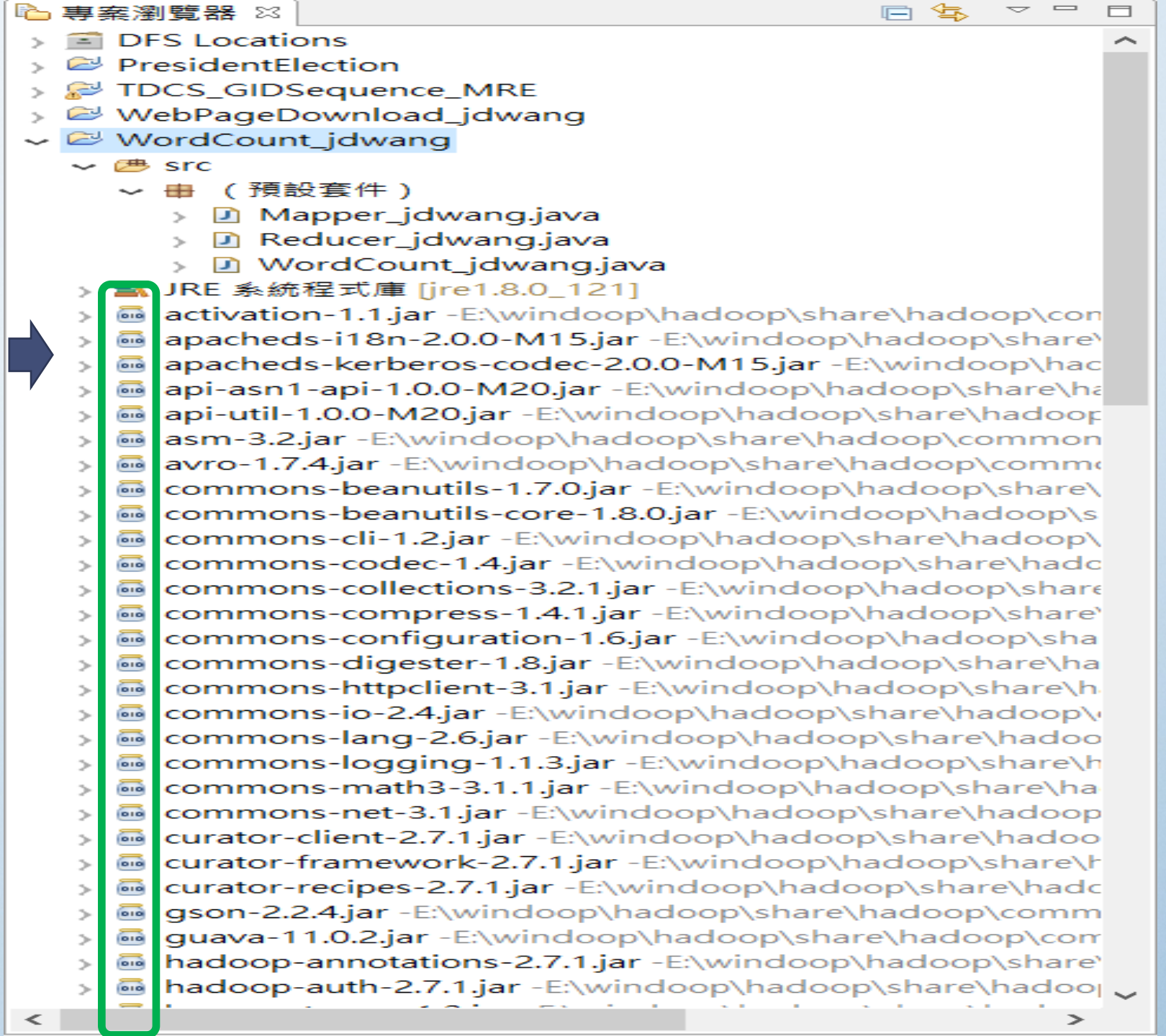


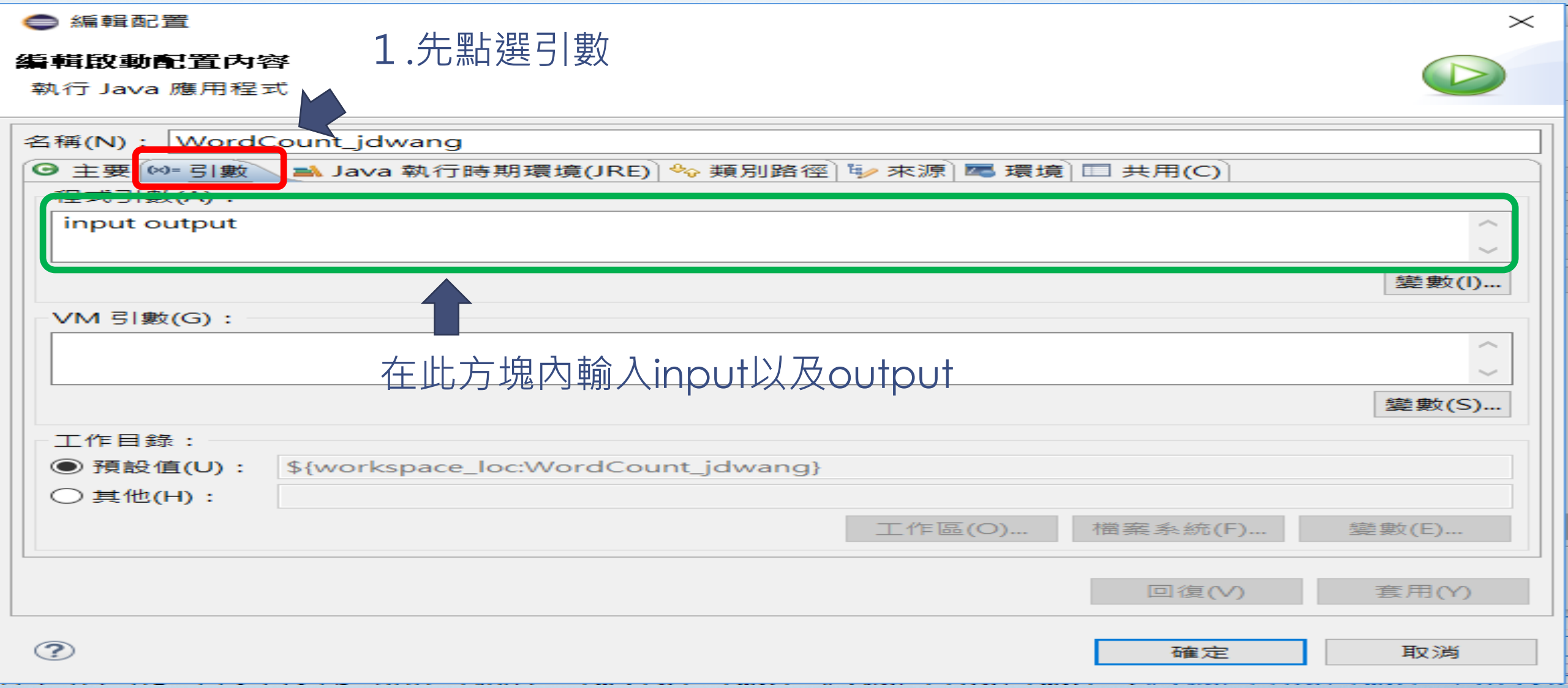
進入 yarn 選取所有檔案並開啟

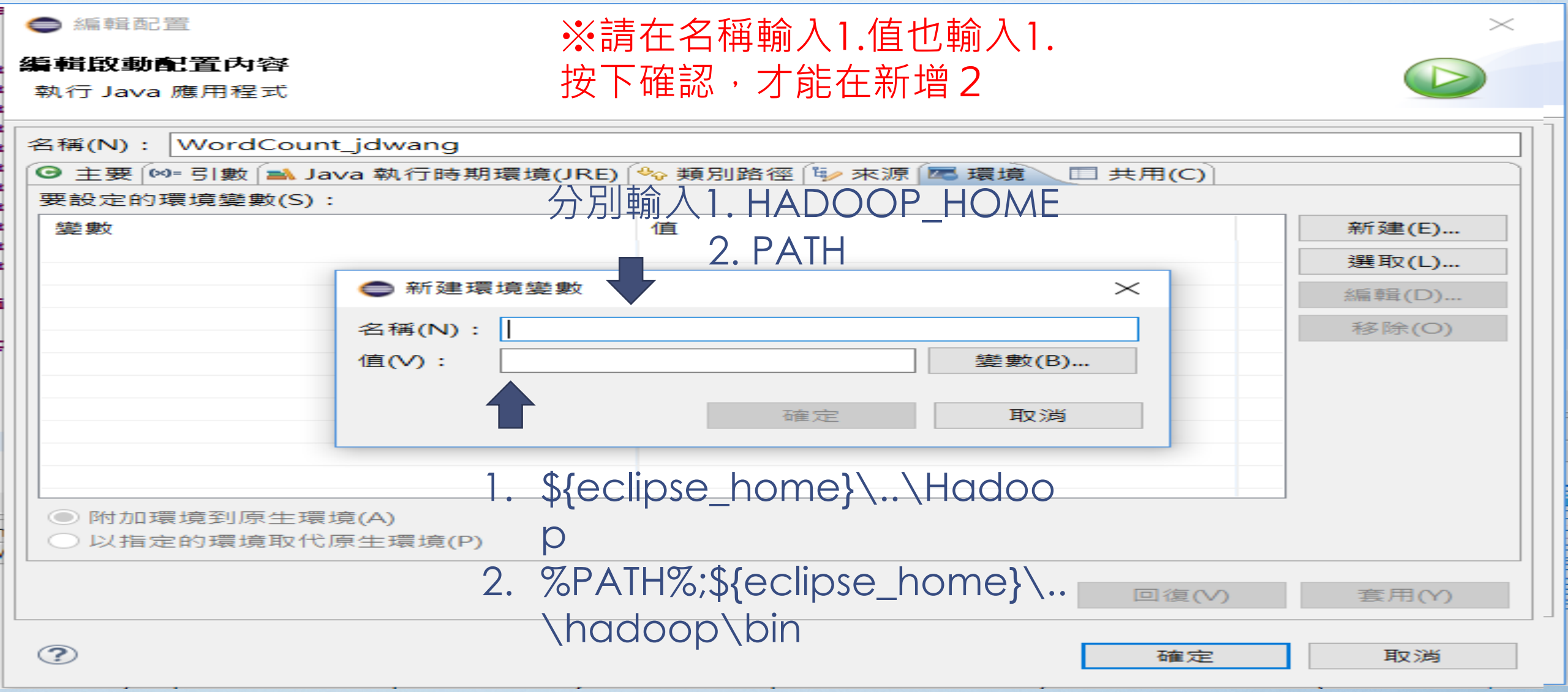


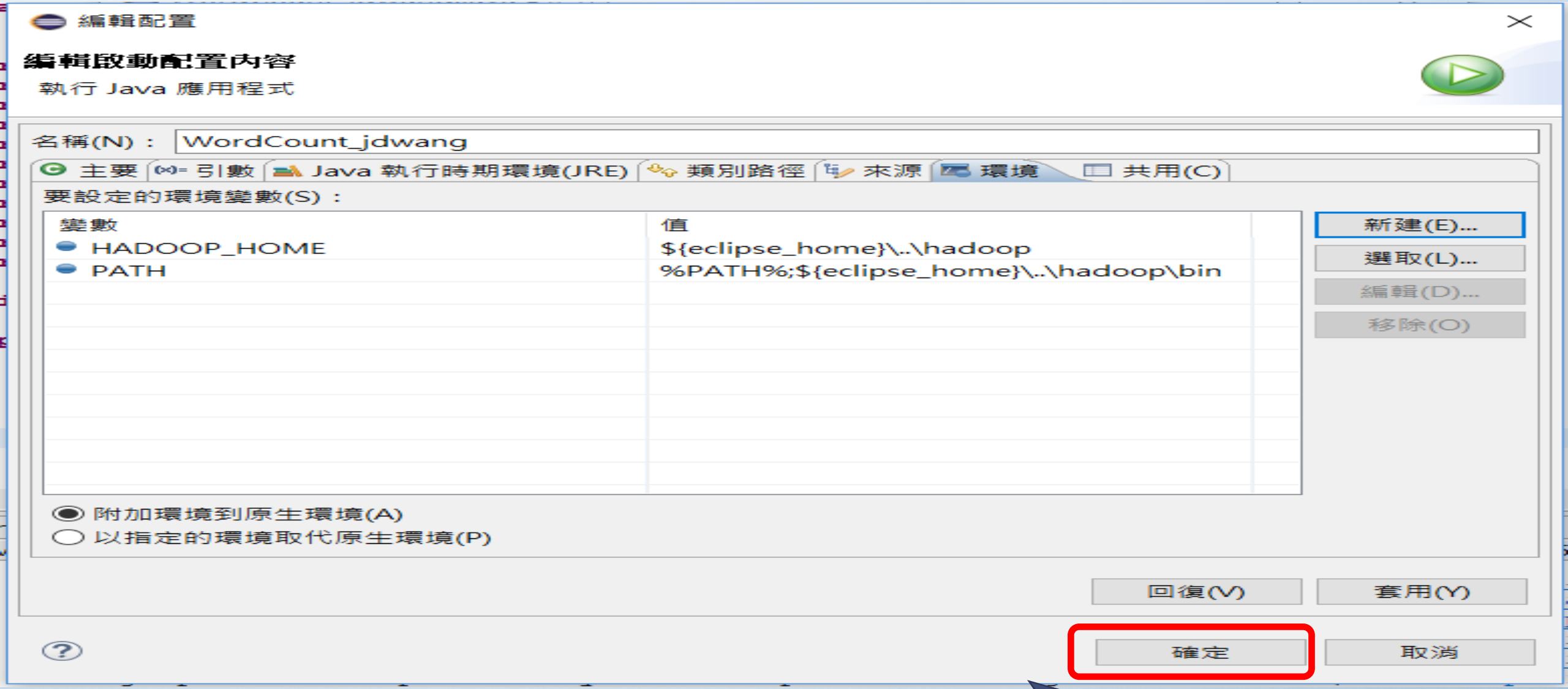
全部開起完成後按下確定

選擇完會發現驚嘆號都消失了

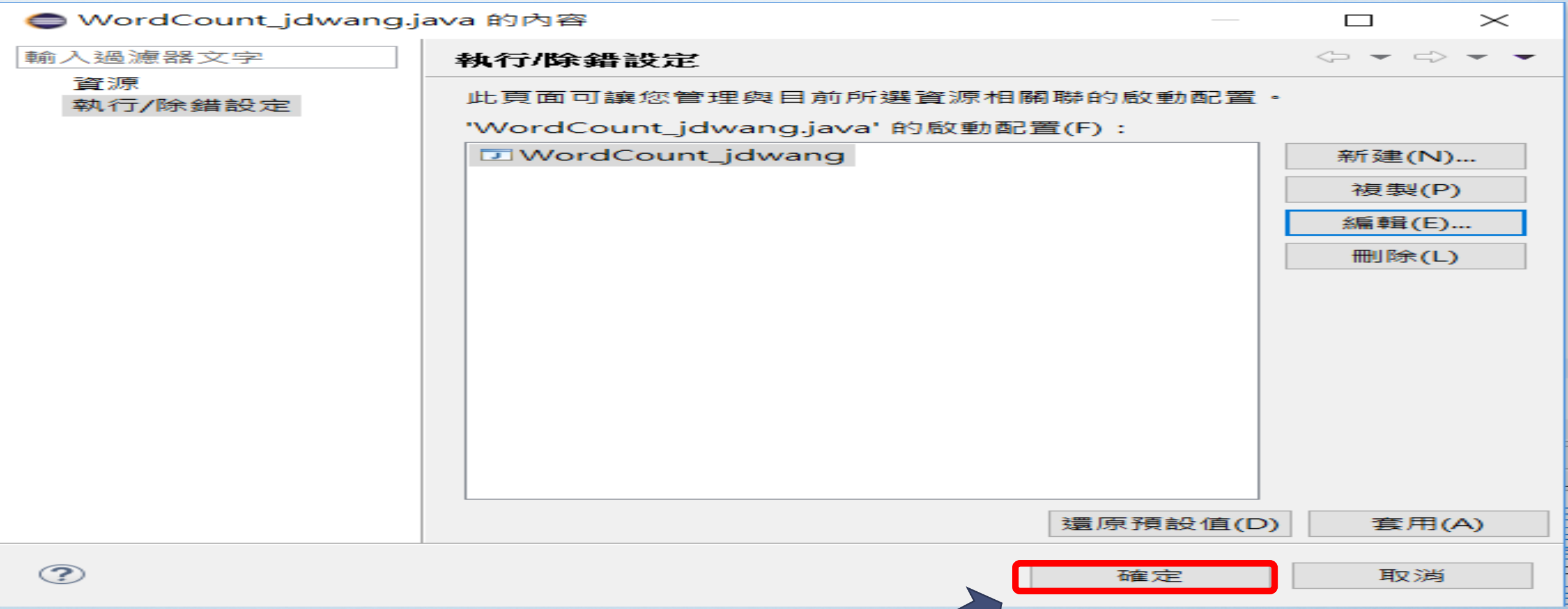



編輯配置
編輯啟動配置內容
執行 Java 應用程式
1.先點選引數
名稱(N)：
WordCount_jdwang
主要
引數
Java 執行時期環境(JRE)
類別路徑
來源
環境
共用(C)
input output
變數(I)...
VM 引數(G)：
變數(S)...
工作目錄：
預設值(U)：
${workspace_loc:WordCount_jdwang}
其他(H)：
工作區(O)...
檔案系統(F)...
變數(E)...
回復(V)
套用(Y)
確定
取消
在此方塊內輸入input以及output

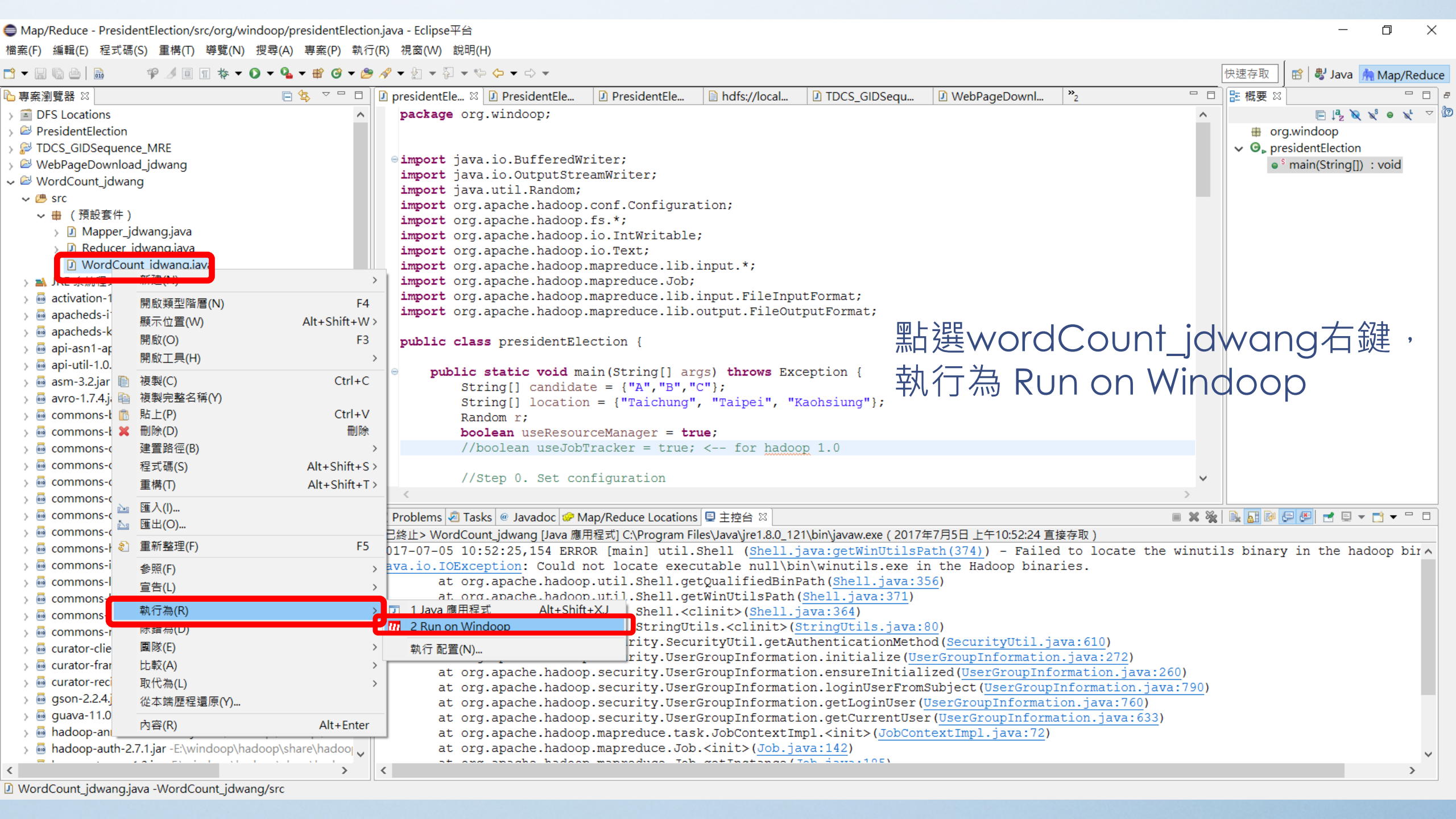

編輯配置
編輯啟動配置內容
執行 Java 應用程式
※請在名稱輸入1.值也輸入1.
按下確認，才能在新增 2
名稱(N)： WordCount_jdwang
主要
引數
Java 執行時期環境(JRE)
類別路徑
來源
環境
共用(C)
要設定的環境變數(S)：
變數
值
新建(E)...
選取(L)...
編輯(D)...
移除(O)
分別輸入1. HADOOP_HOME
2. PATH
新建環境變數
名稱(N)：
值(V)：
變數(B)...
確定
取消
1. ${eclipse_home}\..\Hadoop
2. %PATH%;${eclipse_home}\..\hadoop\bin
附加環境到原生環境(A)
以指定的環境取代原生環境(P)
回復(V)
套用(Y)
確定
取消


編輯配置
編輯啟動配置內容
執行 Java 應用程式
名稱(N)： WordCount_jdwang
主要
引數
Java 執行時期環境(JRE)
類別路徑
來源
環境
共用(C)
要設定的環境變數(S)：
變數
值
HADOOP_HOME
${eclipse_home}\..\hadoop
PATH
%PATH%;${eclipse_home}\..\hadoop\bin
新建(E)...
選取(L)...
編輯(D)...
移除(O)
附加環境到原生環境(A)
以指定的環境取代原生環境(P)
回復(V)
套用(Y)
確定
取消

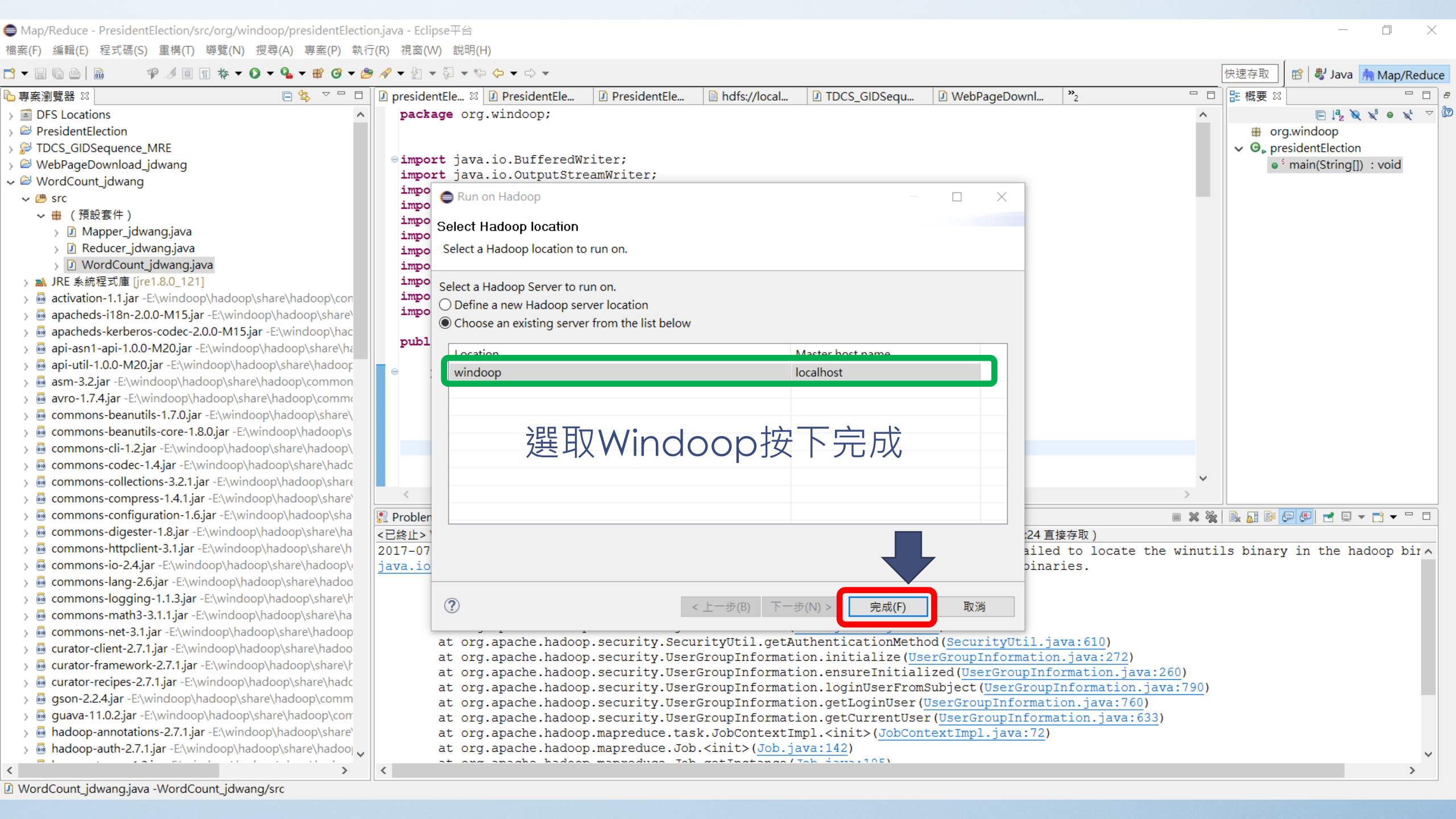
新增完後按下確認


WordCount_jdwang.java 的內容
輸入過濾器文字
資源
執行/除錯設定
執行/除錯設定
此頁面可讓您管理與目前所選資源相關聯的啟動配置。
'WordCount_jdwang.java' 的啟動配置(F)：
WordCount_jdwang
新建(N)...
複製(P)
編輯(E)...
刪除(L)
還原預設值(D)
套用(A)
確定
取消

在按確認



點選wordCount_jdwang右鍵，
執行為 Run on Windoop


Map/Reduce - PresidentElection/src/org/windoop/presidentElection.java - 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
PresidentElection
TDCS_GIDSequence_MRE
WebPageDownload_jdwang
WordCount_jdwang
src
（預設套件）
Mapper_jdwang.java
Reducer_jdwang.java
WordCount_jdwang.java
JRE 系統程式庫 [jre1.8.0_121]
package org.windoop;
import java.io.BufferedWriter;
import java.io.OutputStreamWriter;
Run on Hadoop
Select Hadoop location
Select a Hadoop location to run on.
Select a Hadoop Server to run on.
Define a new Hadoop server location
Choose an existing server from the list below
Location
Master host name
windoop
localhost
選取Windoop按下完成
< 上一步(B)
下一步(N) >
完成(F)
取消
概要
org.windoop
presidentElection
main(String[]) : void
at org.apache.hadoop.security.SecurityUtil.getAuthenticationMethod(SecurityUtil.java:610)
at org.apache.hadoop.security.UserGroupInformation.initialize(UserGroupInformation.java:272)
at org.apache.hadoop.security.UserGroupInformation.ensureInitialized(UserGroupInformation.java:260)
at org.apache.hadoop.security.UserGroupInformation.loginUserFromSubject(UserGroupInformation.java:790)
at org.apache.hadoop.security.UserGroupInformation.getLoginUser(UserGroupInformation.java:760)
at org.apache.hadoop.security.UserGroupInformation.getCurrentUser(UserGroupInformation.java:633)
at org.apache.hadoop.mapreduce.task.JobContextImpl.<init>(JobContextImpl.java:72)
at org.apache.hadoop.mapreduce.Job.<init>(Job.java:142)
WordCount_jdwang.java -WordCount_jdwang/src

Map/Reduce - PresidentElection/src/org/windoop/presidentElection.java - 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)

快速存取 | Java | Map/Reduce

專案瀏覽器

- DFS Locations
- PresidentElection
- TDCS_GIDSequence_MRE
- WebPageDownload_jdwang
- WordCount_jdwang
  - src
    - (預設套件)
      - Mapper_jdwang.java
      - Reducer_jdwang.java
      - WordCount_jdwang.java
  - JRE 系統程式庫 [jre1.8.0_121]
  - activation-1.1.jar -E:\windoop\hadoop\share\hadoop\con
  - apacheds-i18n-2.0.0-M15.jar -E:\windoop\hadoop\share\
  - apacheds-kerberos-codec-2.0.0-M15.jar -E:\windoop\hac
  - api-asn1-api-1.0.0-M20.jar -E:\windoop\hadoop\share\ha
  - api-util-1.0.0-M20.jar -E:\windoop\hadoop\share\hadoop
  - asm-3.2.jar -E:\windoop\hadoop\share\hadoop\common
  - avro-1.7.4.jar -E:\windoop\hadoop\share\hadoop\comm
  - commons-beanutils-1.7.0.jar -E:\windoop\hadoop\share\
  - commons-beanutils-core-1.8.0.jar -E:\windoop\hadoop\s
  - commons-cli-1.2.jar -E:\windoop\hadoop\share\hadoop\
  - commons-codec-1.4.jar -E:\windoop\hadoop\share\hadc
  - commons-collections-3.2.1.jar -E:\windoop\hadoop\share
  - commons-compress-1.4.1.jar -E:\windoop\hadoop\share\
  - commons-configuration-1.6.jar -E:\windoop\hadoop\sha
  - commons-digester-1.8.jar -E:\windoop\hadoop\share\ha
  - commons-httpclient-3.1.jar -E:\windoop\hadoop\share\h
  - commons-io-2.4.jar -E:\windoop\hadoop\share\hadoop\
  - commons-lang-2.6.jar -E:\windoop\hadoop\share\hadoo
  - commons-logging-1.1.3.jar -E:\windoop\hadoop\share\h
  - commons-math3-3.1.1.jar -E:\windoop\hadoop\share\ha
  - commons-net-3.1.jar -E:\windoop\hadoop\share\hadoop
  - curator-client-2.7.1.jar -E:\windoop\hadoop\share\hadoo
  - curator-framework-2.7.1.jar -E:\windoop\hadoop\share\h
  - curator-recipes-2.7.1.jar -E:\windoop\hadoop\share\hadc
  - gson-2.2.4.jar -E:\windoop\hadoop\share\hadoop\comm
  - guava-11.0.2.jar -E:\windoop\hadoop\share\hadoop\con
  - hadoop-annotations-2.7.1.jar -E:\windoop\hadoop\share\
  - hadoop-auth-2.7.1.jar -E:\windoop\hadoop\share\hadoo

presidentEle... | PresidentEle... | PresidentEle... | hdfs://local... | TDCS_GIDSequ... | WebPageDownl...

```java
package org.windoop;


import java.io.BufferedWriter;
import java.io.OutputStreamWriter;
import java.util.Random;
import org.apache.hadoop.conf.Configuration;
import org.apache.hadoop.fs.*;
import org.apache.hadoop.io.IntWritable;
import org.apache.hadoop.io.Text;
import org.apache.hadoop.mapreduce.lib.input.*;
import org.apache.hadoop.mapreduce.Job;
import org.apache.hadoop.mapreduce.lib.input.FileInputFormat;
import org.apache.hadoop.mapreduce.lib.output.FileOutputFormat;

public class presidentElection {

    public static void main(String[] args) throws Exception {
        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
```

概要

- org.windoop
- presidentElection
  - main(String[]) : void

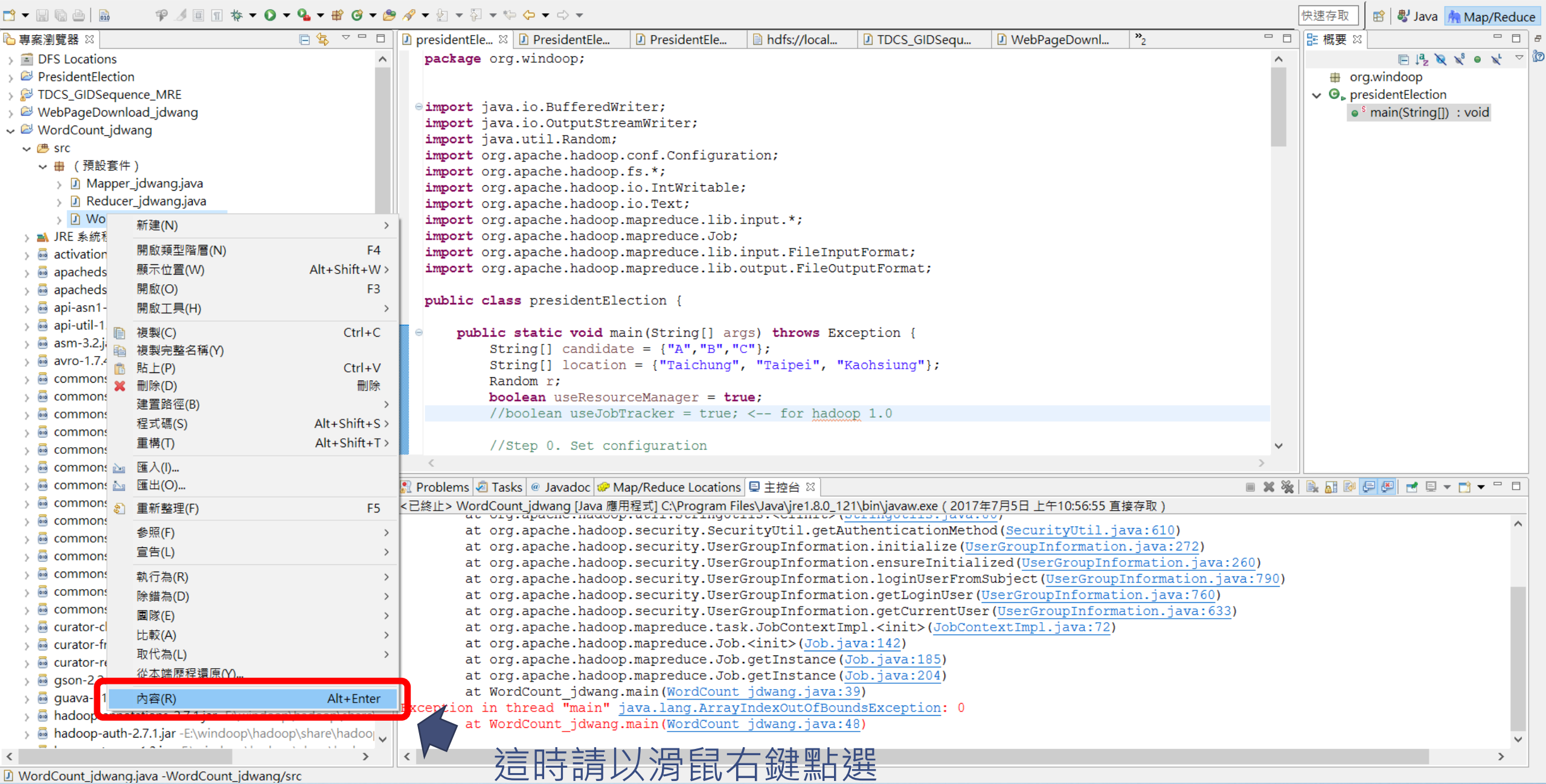
Problems | Tasks | Javadoc | Map/Reduce Locations | 主控台

<已終止> WordCount_jdwang [Java 應用程式] C:\Program Files\Java\jre1.8.0_121\bin\javaw.exe (2017年7月5日 上午10:56:55 直接存取)

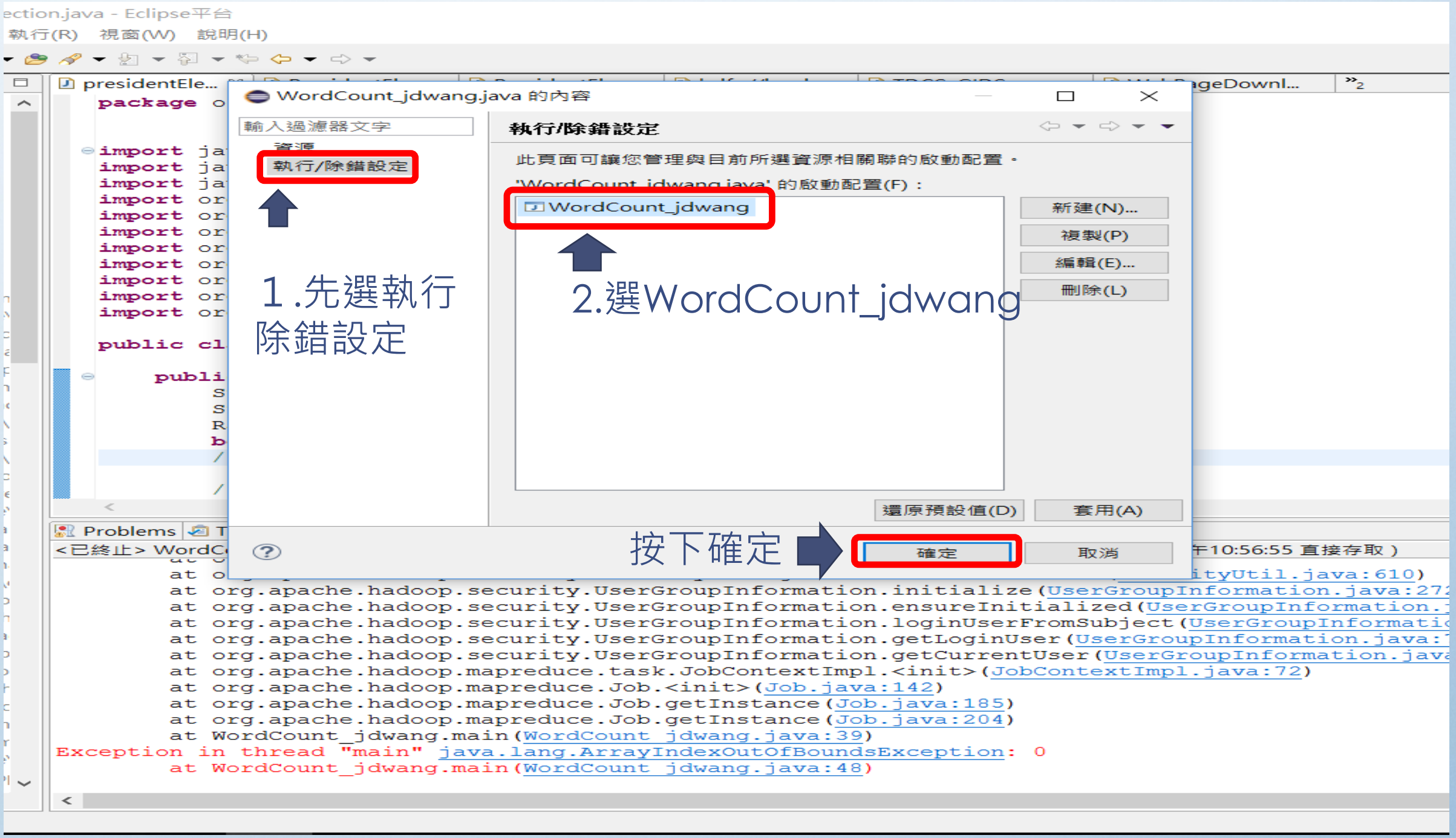
```
        at org.apache.hadoop.security.SecurityUtil.getAuthenticationMethod(SecurityUtil.java:610)
        at org.apache.hadoop.security.UserGroupInformation.initialize(UserGroupInformation.java:272)
        at org.apache.hadoop.security.UserGroupInformation.ensureInitialized(UserGroupInformation.java:260)
        at org.apache.hadoop.security.UserGroupInformation.loginUserFromSubject(UserGroupInformation.java:790)
        at org.apache.hadoop.security.UserGroupInformation.getLoginUser(UserGroupInformation.java:760)
        at org.apache.hadoop.security.UserGroupInformation.getCurrentUser(UserGroupInformation.java:633)
        at org.apache.hadoop.mapreduce.task.JobContextImpl.<init>(JobContextImpl.java:72)
        at org.apache.hadoop.mapreduce.Job.<init>(Job.java:142)
        at org.apache.hadoop.mapreduce.Job.getInstance(Job.java:185)
        at org.apache.hadoop.mapreduce.Job.getInstance(Job.java:204)
        at WordCount_jdwang.main(WordCount_jdwang.java:39)
Exception in thread "main" java.lang.ArrayIndexOutOfBoundsException: 0
        at WordCount_jdwang.main(WordCount_jdwang.java:48)
```

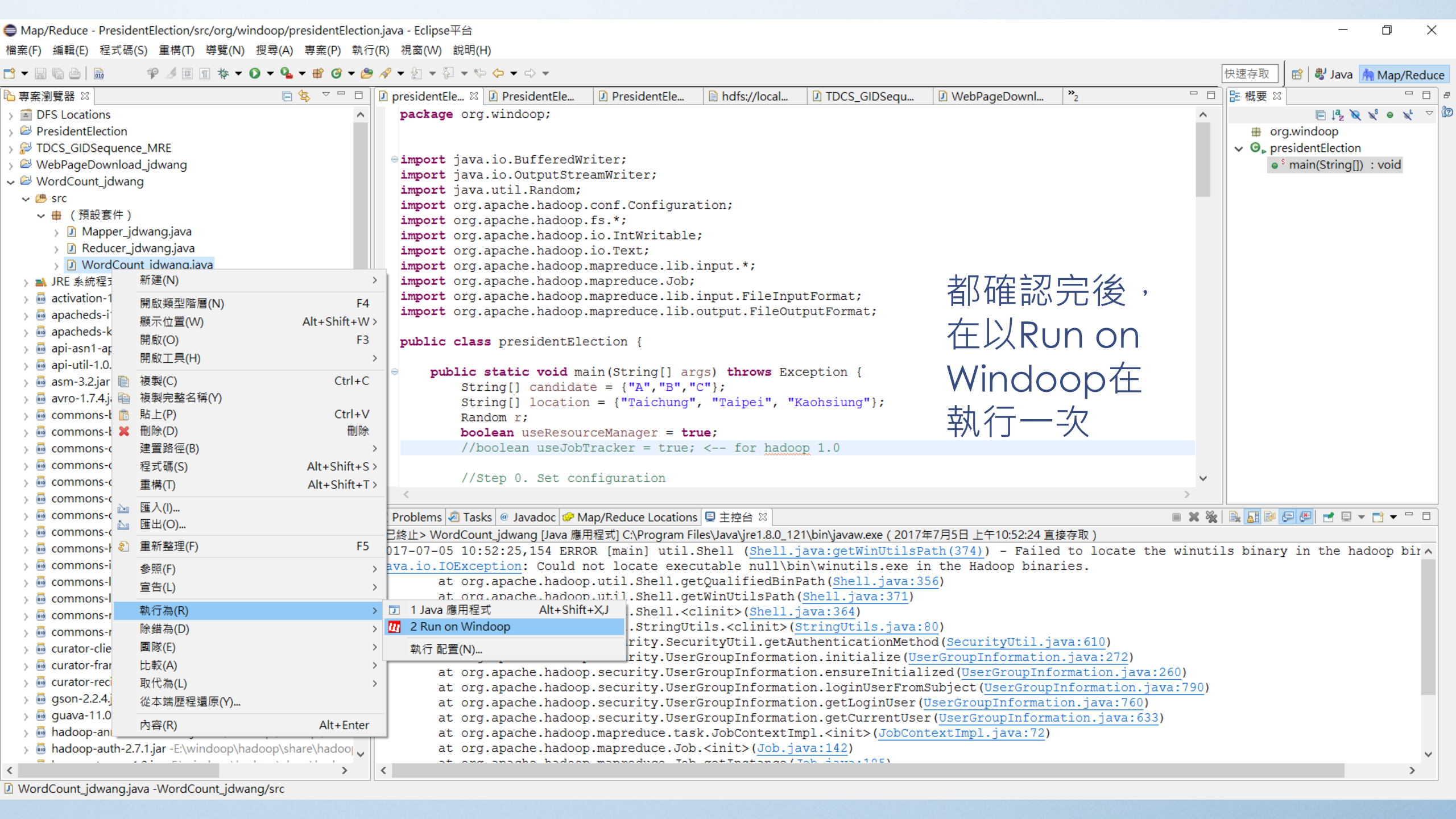
會發現主控台顯示錯誤

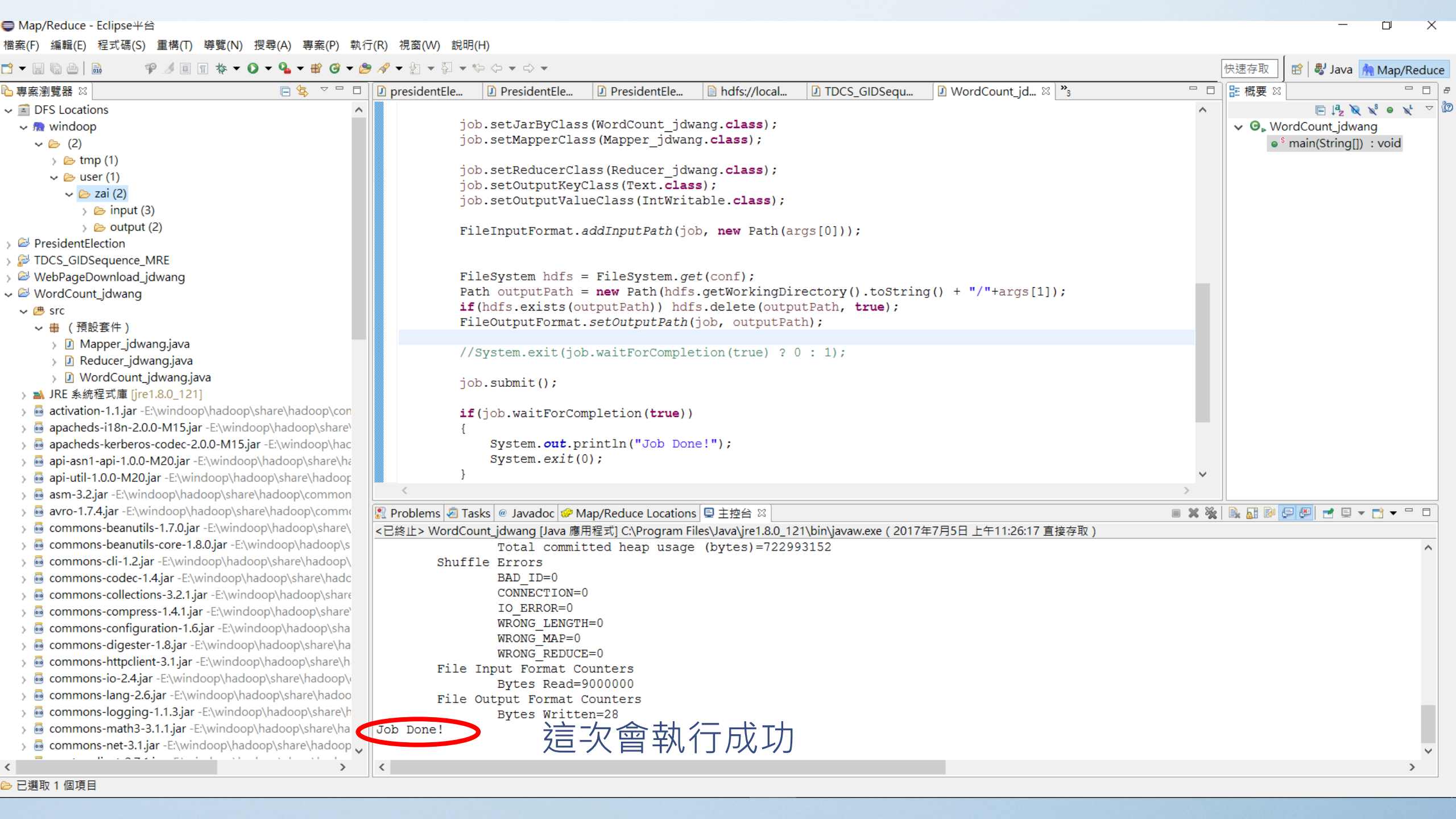
WordCount_jdwang.java -WordCount_jdwang/src

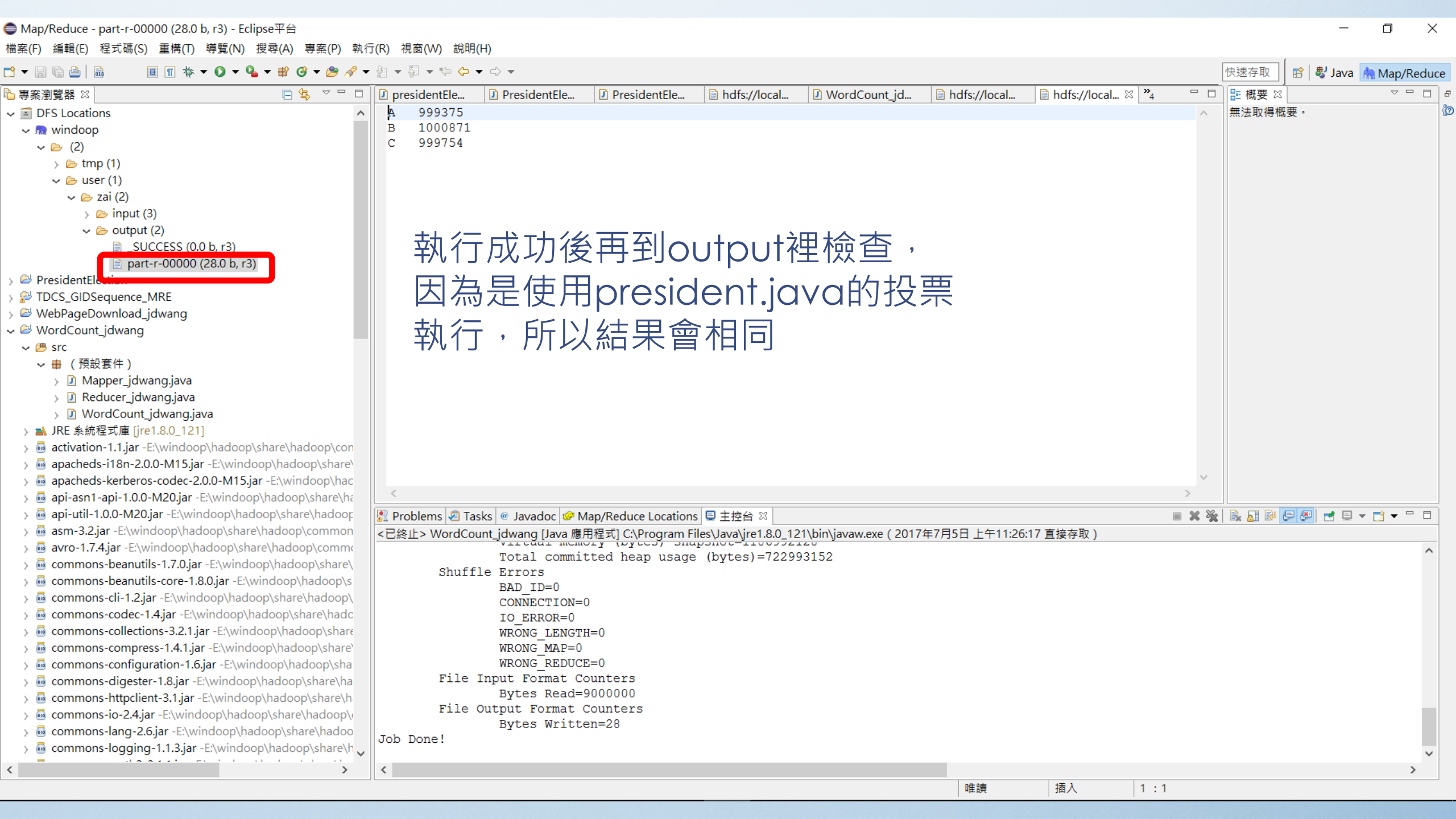


這時請以滑鼠右鍵點選Wordcount_jdwang，點選內容


WordCount_jdwang.java 的內容
輸入過濾器文字
資源
執行/除錯設定
執行/除錯設定
此頁面可讓您管理與目前所選資源相關聯的啟動配置。
WordCount_jdwang
新建(N)...
複製(P)
編輯(E)...
刪除(L)
還原預設值(D)
套用(A)
確定
取消
1.先選執行除錯設定
2.選WordCount_jdwang
按下確定
Problems
Exception in thread "main" java.lang.ArrayIndexOutOfBoundsException: 0
at WordCount_jdwang.main(WordCount_jdwang.java:48)


Map/Reduce - PresidentElection/src/org/windoop/presidentElection.java - 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
PresidentElection
TDCS_GIDSequence_MRE
WebPageDownload_jdwang
WordCount_jdwang
src
（預設套件）
Mapper_jdwang.java
Reducer_jdwang.java
WordCount_jdwang.java
新建(N)
開啟類型階層(N) F4
顯示位置(W) Alt+Shift+W
開啟(O) F3
開啟工具(H)
複製(C) Ctrl+C
複製完整名稱(Y)
貼上(P) Ctrl+V
刪除(D) 刪除
建置路徑(B)
程式碼(S) Alt+Shift+S
重構(T) Alt+Shift+T
匯入(I)...
匯出(O)...
重新整理(F) F5
參照(F)
宣告(L)
執行為(R)
除錯為(D)
團隊(E)
比較(A)
取代為(L)
從本端歷程還原(Y)...
內容(R) Alt+Enter
1 Java 應用程式 Alt+Shift+X,J
2 Run on Windoop
執行 配置(N)...
presidentEle...
PresidentEle...
PresidentEle...
hdfs://local...
TDCS_GIDSequ...
WebPageDownl...
package org.windoop;
import java.io.BufferedWriter;
import java.io.OutputStreamWriter;
import java.util.Random;
import org.apache.hadoop.conf.Configuration;
import org.apache.hadoop.fs.*;
import org.apache.hadoop.io.IntWritable;
import org.apache.hadoop.io.Text;
import org.apache.hadoop.mapreduce.lib.input.*;
import org.apache.hadoop.mapreduce.Job;
import org.apache.hadoop.mapreduce.lib.input.FileInputFormat;
import org.apache.hadoop.mapreduce.lib.output.FileOutputFormat;
public class presidentElection {
public static void main(String[] args) throws Exception {
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
都確認完後，
在以Run on
Windoop在
執行一次
概要
org.windoop
presidentElection
main(String[]) : void
Problems Tasks Javadoc Map/Reduce Locations 主控台
已終止> WordCount_jdwang [Java 應用程式] C:\Program Files\Java\jre1.8.0_121\bin\javaw.exe (2017年7月5日 上午10:52:24 直接存取)
017-07-05 10:52:25,154 ERROR [main] util.Shell (Shell.java:getWinUtilsPath(374)) - Failed to locate the winutils binary in the hadoop bir
ava.io.IOException: Could not locate executable null\bin\winutils.exe in the Hadoop binaries.
at org.apache.hadoop.util.Shell.getQualifiedBinPath(Shell.java:356)
at org.apache.hadoop.util.Shell.getWinUtilsPath(Shell.java:371)
at org.apache.hadoop.security.UserGroupInformation.initialize(UserGroupInformation.java:272)
at org.apache.hadoop.security.UserGroupInformation.ensureInitialized(UserGroupInformation.java:260)
at org.apache.hadoop.security.UserGroupInformation.loginUserFromSubject(UserGroupInformation.java:790)
at org.apache.hadoop.security.UserGroupInformation.getLoginUser(UserGroupInformation.java:760)
at org.apache.hadoop.security.UserGroupInformation.getCurrentUser(UserGroupInformation.java:633)
at org.apache.hadoop.mapreduce.task.JobContextImpl.<init>(JobContextImpl.java:72)
at org.apache.hadoop.mapreduce.Job.<init>(Job.java:142)
WordCount_jdwang.java -WordCount_jdwang/src


Map/Reduce - 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)
zai (2)
input (3)
output (2)
PresidentElection
TDCS_GIDSequence_MRE
WebPageDownload_jdwang
WordCount_jdwang
src
(預設套件)
Mapper_jdwang.java
Reducer_jdwang.java
WordCount_jdwang.java
JRE 系統程式庫 [jre1.8.0_121]
activation-1.1.jar -E:\windoop\hadoop\share\hadoop\con
apacheds-i18n-2.0.0-M15.jar -E:\windoop\hadoop\share\
apacheds-kerberos-codec-2.0.0-M15.jar -E:\windoop\hac
api-asn1-api-1.0.0-M20.jar -E:\windoop\hadoop\share\ha
api-util-1.0.0-M20.jar -E:\windoop\hadoop\share\hadoop
asm-3.2.jar -E:\windoop\hadoop\share\hadoop\common
avro-1.7.4.jar -E:\windoop\hadoop\share\hadoop\comm
commons-beanutils-1.7.0.jar -E:\windoop\hadoop\share\
commons-beanutils-core-1.8.0.jar -E:\windoop\hadoop\s
commons-cli-1.2.jar -E:\windoop\hadoop\share\hadoop\
commons-codec-1.4.jar -E:\windoop\hadoop\share\hadc
commons-collections-3.2.1.jar -E:\windoop\hadoop\share
commons-compress-1.4.1.jar -E:\windoop\hadoop\share\
commons-configuration-1.6.jar -E:\windoop\hadoop\sha
commons-digester-1.8.jar -E:\windoop\hadoop\share\ha
commons-httpclient-3.1.jar -E:\windoop\hadoop\share\h
commons-io-2.4.jar -E:\windoop\hadoop\share\hadoop\
commons-lang-2.6.jar -E:\windoop\hadoop\share\hadoo
commons-logging-1.1.3.jar -E:\windoop\hadoop\share\h
commons-math3-3.1.1.jar -E:\windoop\hadoop\share\ha
commons-net-3.1.jar -E:\windoop\hadoop\share\hadoop
presidentEle...
PresidentEle...
PresidentEle...
hdfs://local...
TDCS_GIDSequ...
WordCount_jd...
job.setJarByClass(WordCount_jdwang.class);
job.setMapperClass(Mapper_jdwang.class);
job.setReducerClass(Reducer_jdwang.class);
job.setOutputKeyClass(Text.class);
job.setOutputValueClass(IntWritable.class);
FileInputFormat.addInputPath(job, new Path(args[0]));
FileSystem hdfs = FileSystem.get(conf);
Path outputPath = new Path(hdfs.getWorkingDirectory().toString() + "/"+args[1]);
if(hdfs.exists(outputPath)) hdfs.delete(outputPath, true);
FileOutputFormat.setOutputPath(job, outputPath);
//System.exit(job.waitForCompletion(true) ? 0 : 1);
job.submit();
if(job.waitForCompletion(true))
{
System.out.println("Job Done!");
System.exit(0);
}
概要
WordCount_jdwang
main(String[]) : void
Problems
Tasks
Javadoc
Map/Reduce Locations
主控台
<已終止> WordCount_jdwang [Java 應用程式] C:\Program Files\Java\jre1.8.0_121\bin\javaw.exe (2017年7月5日 上午11:26:17 直接存取)
Total committed heap usage (bytes)=722993152
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!
這次會執行成功
已選取 1 個項目


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)
zai (2)
input (3)
output (2)
SUCCESS (0.0 b, r3)
part-r-00000 (28.0 b, r3)
TDCS_GIDSequence_MRE
WebPageDownload_jdwang
WordCount_jdwang
src
(預設套件)
Mapper_jdwang.java
Reducer_jdwang.java
WordCount_jdwang.java
JRE 系統程式庫 [jre1.8.0_121]
A 999375
B 1000871
C 999754
執行成功後再到output裡檢查，
因為是使用president.java的投票
執行，所以結果會相同
概要
無法取得概要。
Problems Tasks Javadoc Map/Reduce Locations 主控台
<已終止> WordCount_jdwang [Java 應用程式] C:\Program Files\Java\jre1.8.0_121\bin\javaw.exe (2017年7月5日 上午11:26:17 直接存取)
Total committed heap usage (bytes)=722993152
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!