Your Name here

CSCIE63 Big Data Analytics

Assignment 01

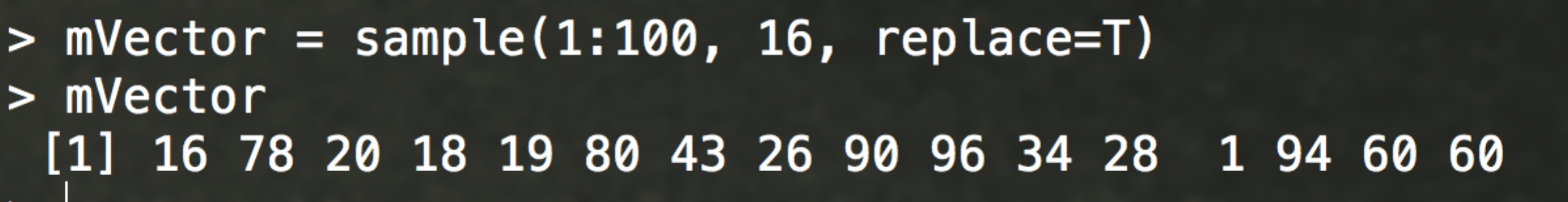
**Assignment 01 Solution**

Problem 1

“**Not a real problem: do something with matrices ….”**

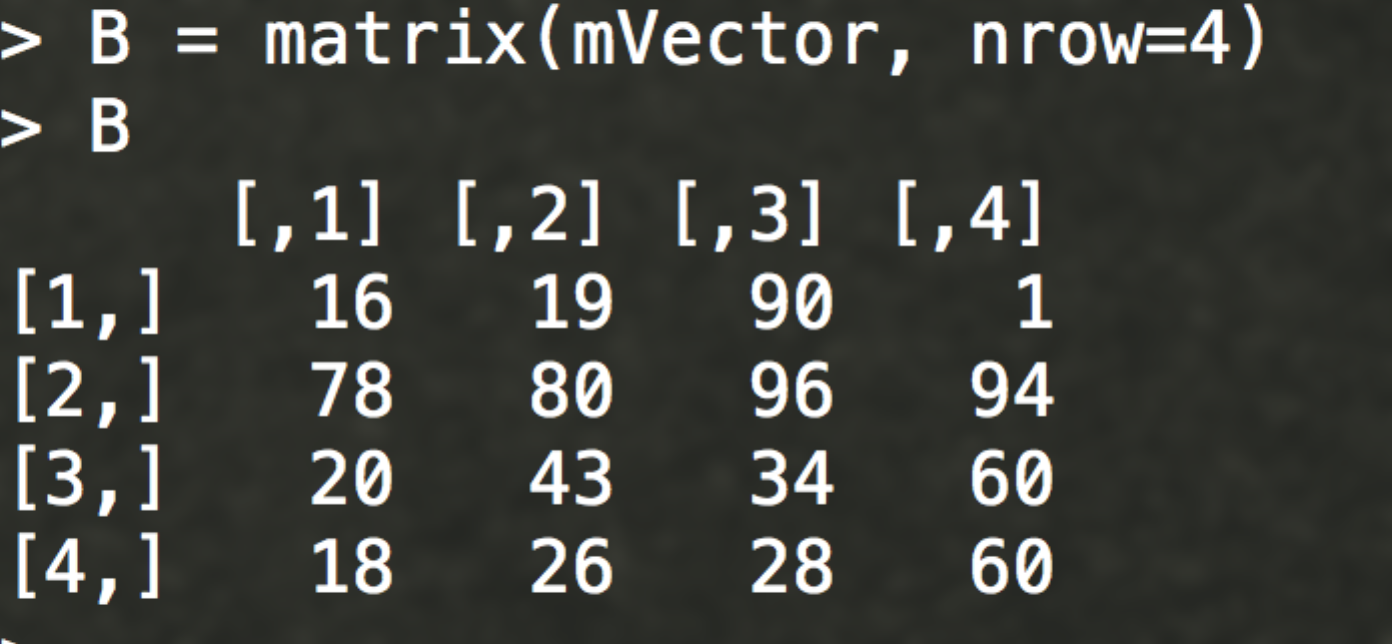
Solution:

1. First we create a vector containing 16 randomly generated numbers



We do this by using the sample function and specifying a range between 1 and 100 and allowing replacement. We can also alternatively use runif to generate numbers including decimals.

1. Next, we use this vector to generate a 4x4 matrix

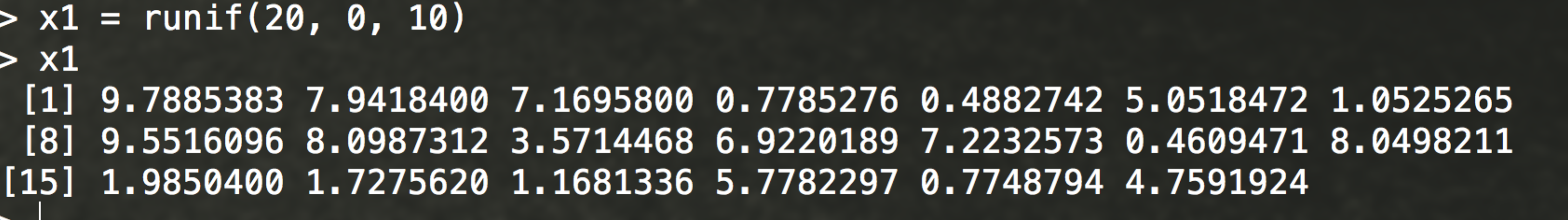


Problem 2

**“Not a real problem: do some work with matrices and plots …”**

Solution:

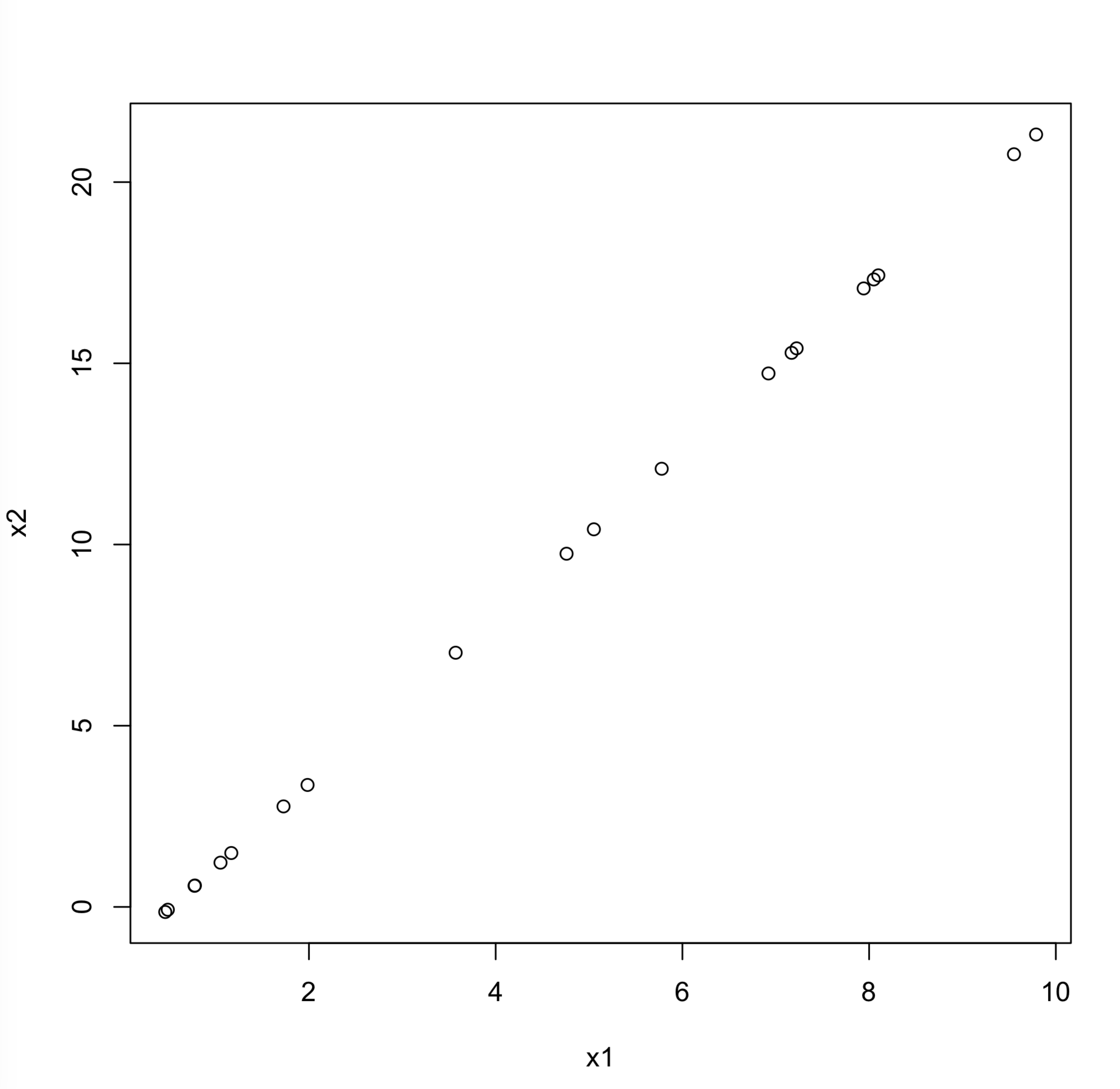
1. First we want to create a vector of 20 values between 0 and 10. We will use runif for this.



1. More steps ….
2. Next we plot the scatter plot

plot(x1,x2 )

|  |
| --- |
| > plot(x1,x2) |

1. 

Problem 3

**“Not a real problem – an example of a problem that requires Java programming and some AWS interactions”**

Solution:

1. Using SimpleDBSample as a template, I modified createSampleData to add the movie stars and Nobel laureates. Here is the createSampleData code modified with stars and nobels:

/\* Creates an array of SimpleDB ReplaceableItems populated with sample data. \*/

**private** **static** List<ReplaceableItem> createSampleData() {

List<ReplaceableItem> sampleData = **new** ArrayList<ReplaceableItem>();

sampleData.add(**new** ReplaceableItem("Chris-Hemsworth").withAttributes(

**new** ReplaceableAttribute("Full Name", "Chris Hemsworth", **true**),

**new** ReplaceableAttribute("Most Popular Movie", "Thor The Dark World", **true**);

…

**return** sampleData;

}

2. Then I modified the try statement as shown below. I changed myDomain from Store to People. I modified the selectExpression to select all people. I deleted everything after querying the domain. So the new try statement looks like this:

**try** {

// Create a domain

String myDomain = "People";

System.***out***.println("Creating domain called " + myDomain + ".\n");

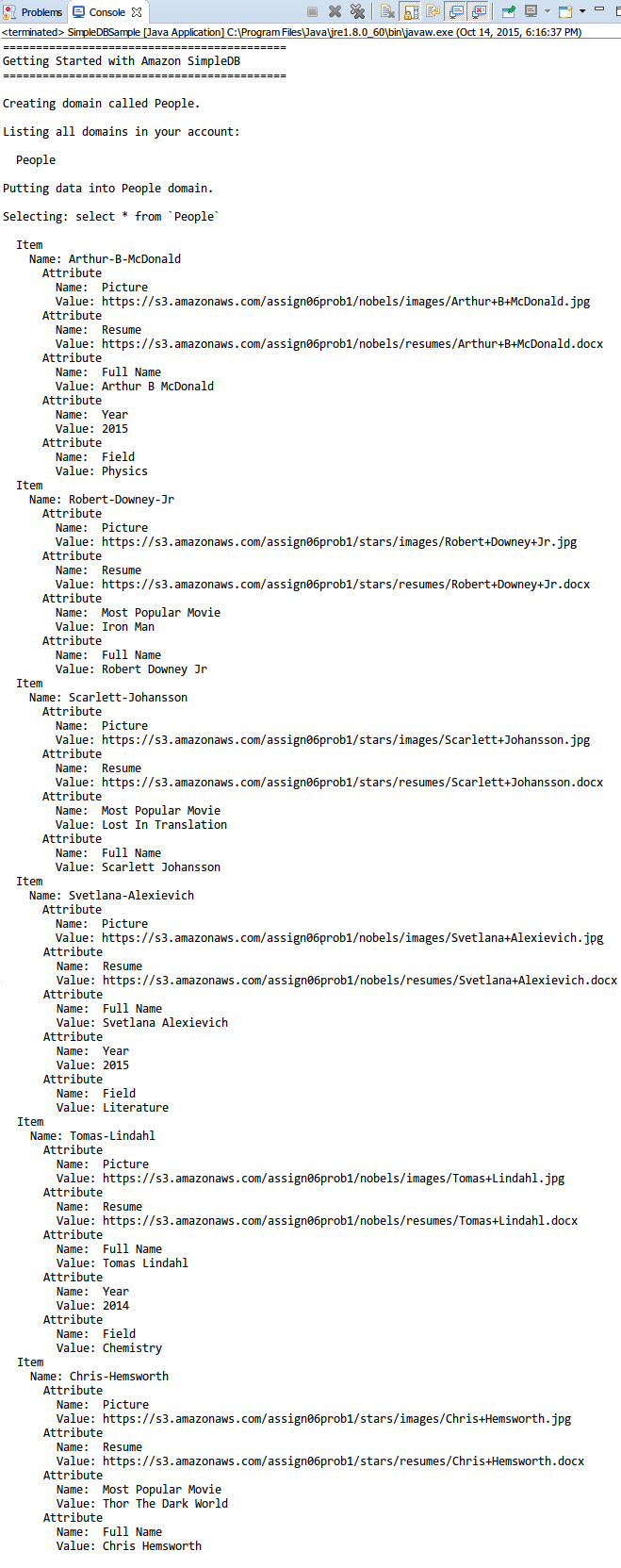
sdb.createDomain(**new** CreateDomainRequest(myDomain));

// List domains

// more code goes here

}

3. After changing up the AWS credentials, run the program. Here’s the console output:



Here you can see the 6 people inserted successfully into the domain.

3.a) Demonstrate that you can change (correct) the year of one Nobel prize award, programmatically.

Steps:

1. When I inserted the data into the domain, I purposely inserted the wrong year (2014) in one of the Nobel prize winners, Tomas Lindahl. The correct year is actually 2015, so I will change it from 2014 to 2015. In the try statement, I modified the section where the code replaces an attribute as shown:

// Replace an attribute

System.***out***.println("Replacing Year of Thomas-Lindahl with 2015.\n");

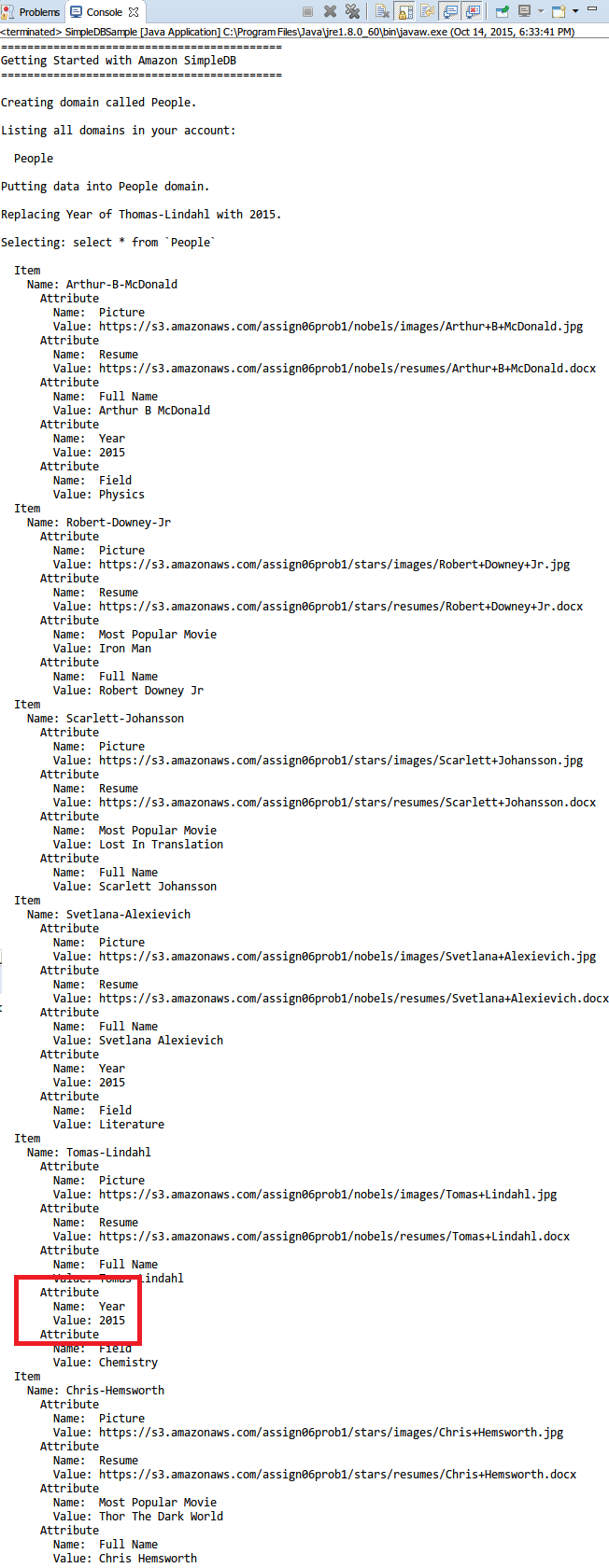
List<ReplaceableAttribute> replaceableAttributes = **new** ArrayList<ReplaceableAttribute>();

replaceableAttributes.add(**new** ReplaceableAttribute("Year", "2015", **true**));

sdb.putAttributes(**new** PutAttributesRequest(myDomain, "Tomas-Lindahl", replaceableAttributes));

And I place this code before the query statement.

2. Run the program. Here’s the console output:



Here you can see that the year is changed from 2014 to 2015 for Tomas Lindahl.

3.b) Demonstrate that you can delete one movie star from SimpleDB domain programmatically

Steps:

1. Let’s delete movie star “Chris Hemsworth” from the domain. The following code deletes an item and all of its attributes:

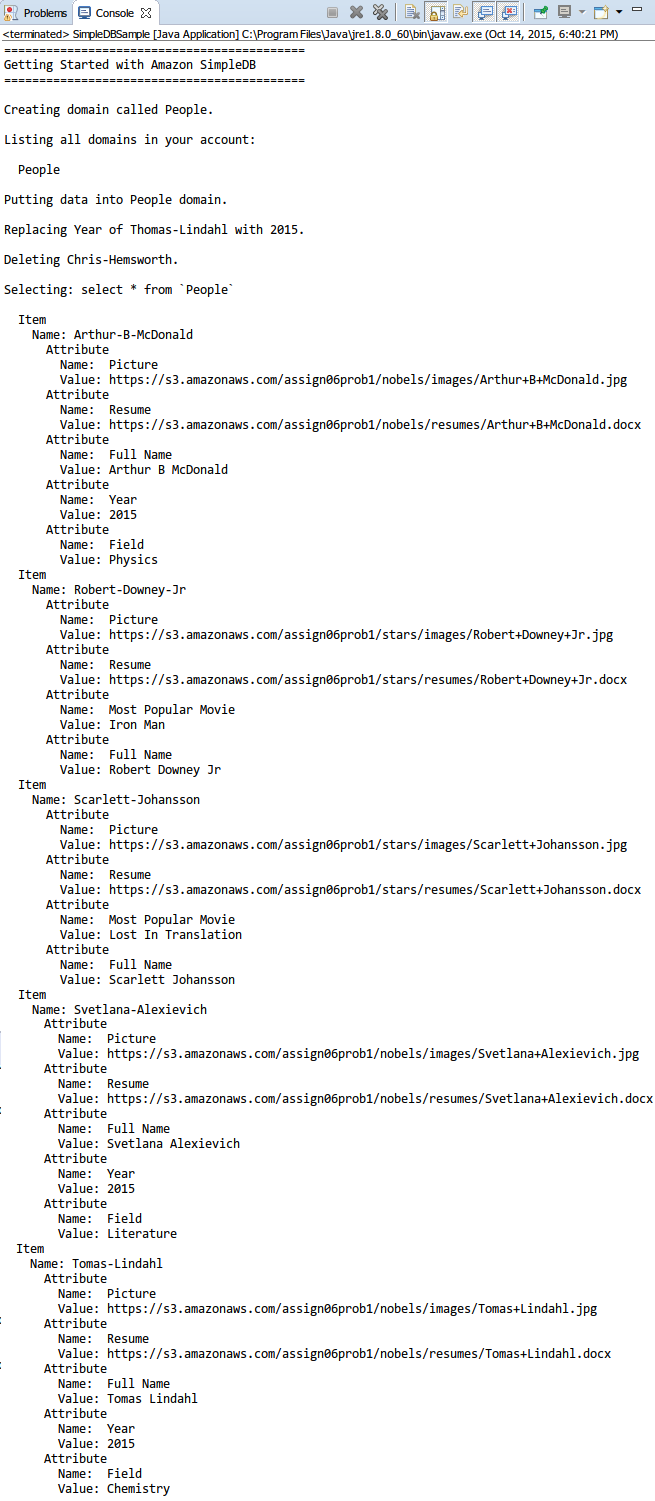
// Delete an item and all of its attributes

System.***out***.println("Deleting Chris-Hemsworth.\n");

sdb.deleteAttributes(**new** DeleteAttributesRequest(myDomain, "Chris-Hemsworth"));

And I place the code before the query statement.

2. Run the program. Here’s the console output:

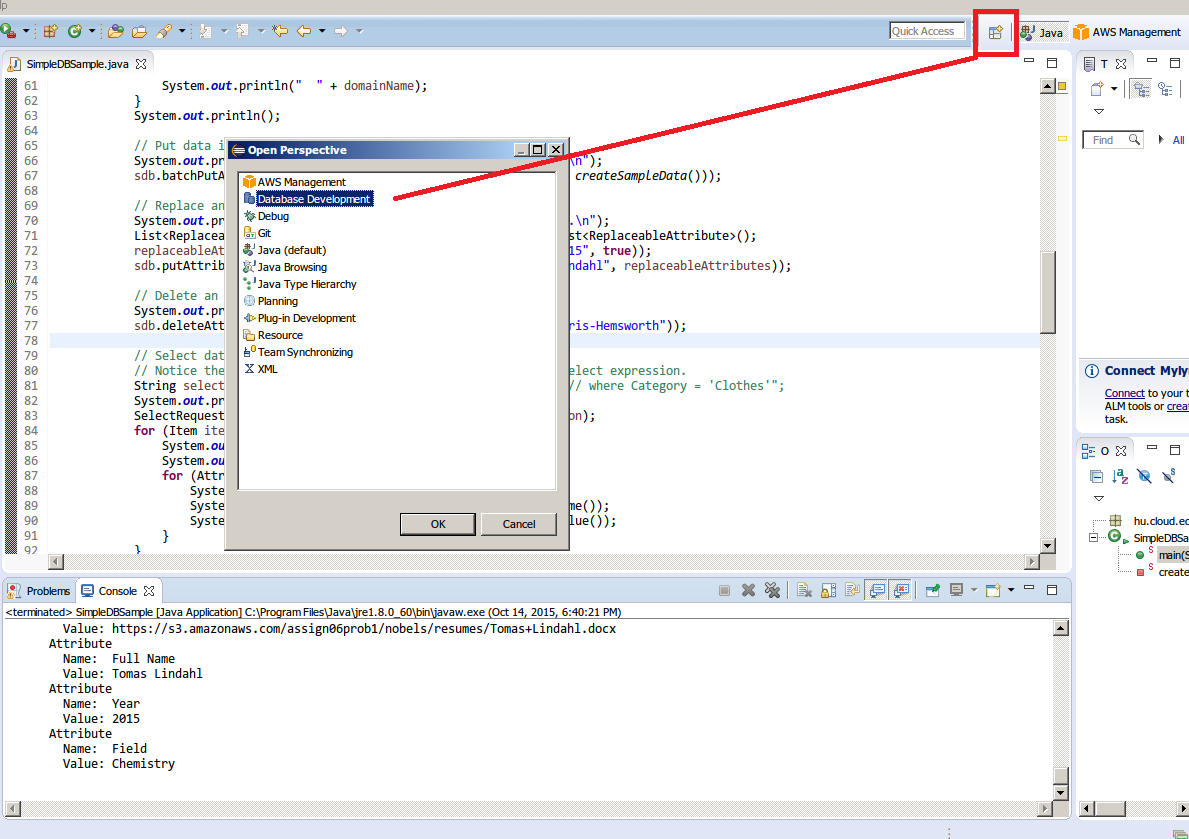


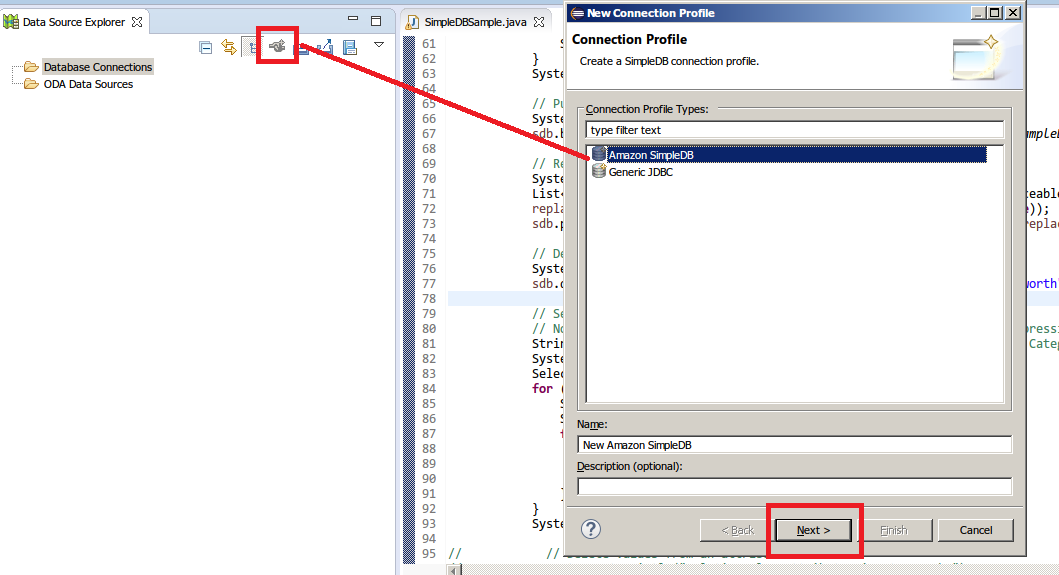
Here you can see that there are only five people now and Chris Hemsworth is removed from the domain.

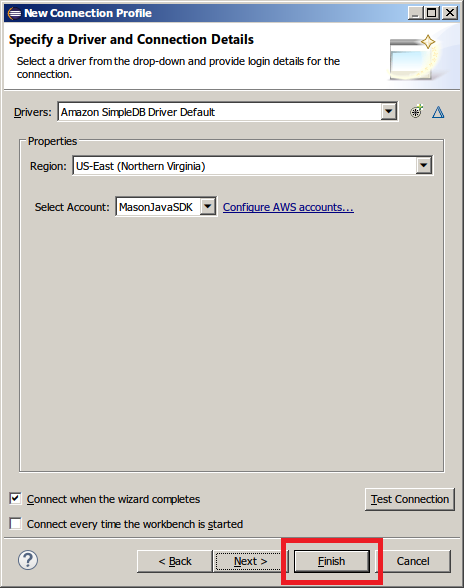
3.c) Capture the content of your database as displayed in the Database Development perspective.

Steps:

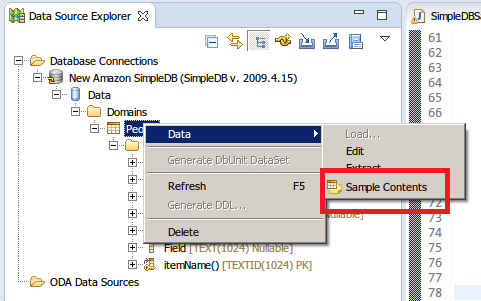
1. Change perspective to Database Development and make a new connection profile (Amazon SimpleDB).



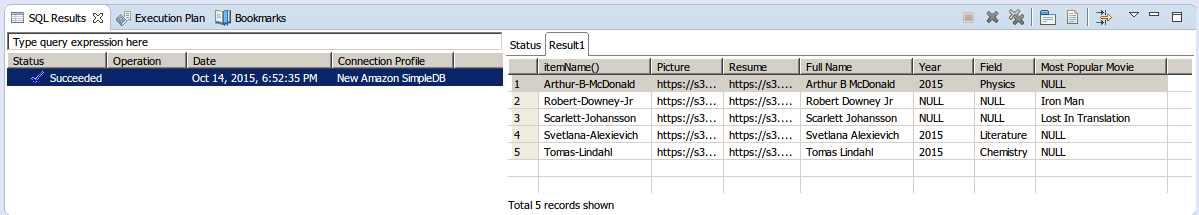




2. Sample the data contents.



Here’s the SQL Results output:



Here you can see the correct year for Tomas (2015) and no Chris Hemsworth in the domain.

Final code is uploaded as SimpleDBSample.java