

General Notes

Cloudera's training courses use a Virtual Machine (VM) running the CentOS Linux distribution. Already installed in the VM are tools, data, and examples that you will use to complete the exercises.

Points to Note while Working in the VM

1. The VM is set to automatically log in as the user `training`. If you log out, you can log back in as the user `training` with the password `training`.
2. If you need it, the root password is `training`. You may be prompted for this if, for example, you want to change the keyboard layout. In general, you should not need this password since the `training` user has unlimited `sudo` privileges.
3. In some command line steps in the exercises, you will see lines like this:

```
$ mkdir \  
/home/training/training_materials/jes/my_work
```

The dollar sign (\$) at the beginning of each line indicates the Linux shell prompt. The actual prompt will include additional information (for example, `[training@localhost ~]$`) but this is omitted from these instructions for brevity.

The backslash (\) at the end of the first line signifies that the command is not completed, and continues on the next line. You can enter the code exactly as shown (on two lines), or you can enter it on a single line. If you do the latter, you should *not* type in the backslash.

4. Although most students are comfortable using UNIX text editors like `vi` or `emacs`, some might prefer a graphical text editor. To invoke the graphical editor from the command line, type `gedit` followed by the path of the file you wish to edit. Appending `&` to the command allows you to type additional commands while the editor is still open. Here is an example of how to edit a file named `myfile.txt`:

```
$ gedit myfile.txt &
```

If you prefer to work in an IDE, *Eclipse* is installed in the VM and has been configured with Scala projects for the course. The Eclipse workspace can be found in the `/home/training/workspace` directory, once you have run the setup script described in the first exercise.

Points to Note During the Exercises

5. For some exercises, two directories are provided. Each contains files that you will need to edit in order to complete the exercise. The `solution` directory contains the complete solution, while the `hints` directory provides hints to help you build the solution. For each exercise, you may choose the one that best matches how you would like to work:
 - `hints`: includes hints about the changes you will need to make
 - `solution`: contains a complete implementation of the solution. You may use these “as-is,” or compare them with your own solution
6. The VM defines a few environment variables that are used in place of longer paths in the instructions. Since each variable is automatically replaced with its corresponding values when you run commands in the terminal, this makes it easier and faster for you to enter a command. You can always use the `echo` command if you would like to see the value of an environment variable:

```
$ echo $SCA
$ echo $SCADATA
```

7. All files and data used in the course are under the main course directory which can be referenced using `$SCA`:

```
/home/training/training_materials/jes
```

8. Under the course directory you will find subdirectories which you will use when completing the exercises:

| Directory | Description |
|------------------------------------|--|
| <code>data/</code> | <ul style="list-style-type: none">Contains data files used in the exercisesCan be referenced using <code>\$SCADATA</code> |
| <code>exercises/*/solution/</code> | <ul style="list-style-type: none">Contains programs (<code>.scala</code> files) and Scala shell code (<code>.scala.txt</code> files) solutions for the exercises in this manualThe <code>*</code> is a placeholder for the specific exercise directory name |
| <code>exercises/*/hints/</code> | <ul style="list-style-type: none">Contains programs (<code>.scala</code> files) and Scala shell code (<code>.scala.txt</code> files) hints for developing exercise solutionsThe <code>*</code> is a placeholder for the specific exercise directory name |
| <code>other_examples/</code> | <ul style="list-style-type: none">Contains programs and Scala shell code snippets for general reference |
| <code>scripts/</code> | <ul style="list-style-type: none">Contains scripts that are used for initializing the environment |