# Lab2

## R

### 1. How do you ask if a file exists?

```
file.exists('datafile.1.a.txt') && !dir.exists('datafile.1.a.txt')
```

```
## [1] FALSE
```

### 2. How to you ask if a file is a directory?

```
dir.exists('datafile.1.a.txt')
```

```
## [1] FALSE
```

### 3. How do you remove (delete) a file?

```
file.remove('datafile.1.a.txt')
```

```
## Warning in file.remove("datafile.1.a.txt"): cannot remove file 'datafile.
## 1.a.txt', reason 'No such file or directory'
```

```
## [1] FALSE
```

### 4. How do you get the size of a file?

```
file.size('datafile.1.a.txt')
```

```
## [1] NA
```

### 5. How do you get all the file names matching a pattern?

```
list.files(pattern = "\\.txt")
```

```
## character(0)
```

### 6. How do you get all the file names matching a pattern recursively?

```
list.files(pattern="*.[0-9*].[ab]*", recursive = TRUE)
```

```
##   [1] "AdvInformatics_Lab2.html"
##   [2] "AdvInformatics_Lab2.pdf"
##   [3] "AdvInformatics_Lab2.Rmd"
##   [4] "Assignment1tbatarse.md"
##   [5] "Assignment2_AdvInf.html"
##   [6] "Assignment2_AdvInf.log"
##   [7] "Assignment2_AdvInf.Rmd"
##   [8] "Assignment2_AdvInf.tex"
##   [9] "assignment2.html"
##  [10] "assignment2.pdf"
##  [11] "assignment2.Rmd"
##  [12] "lab2/lab2data.tar"
##  [13] "lab2/lab2data/data/datafile.1.a.txt"
##  [14] "lab2/lab2data/data/datafile.1.b.txt"
##  [15] "lab2/lab2data/data/datafile.10.a.txt"
##  [16] "lab2/lab2data/data/datafile.10.b.txt"
##  [17] "lab2/lab2data/data/datafile.11.a.txt"
##  [18] "lab2/lab2data/data/datafile.11.b.txt"
##  [19] "lab2/lab2data/data/datafile.12.a.txt"
```

```
##  [20] "lab2/lab2data/data/datafile.12.b.txt"
##  [21] "lab2/lab2data/data/datafile.13.a.txt"
##  [22] "lab2/lab2data/data/datafile.13.b.txt"
##  [23] "lab2/lab2data/data/datafile.14.a.txt"
##  [24] "lab2/lab2data/data/datafile.14.b.txt"
##  [25] "lab2/lab2data/data/datafile.15.a.txt"
##  [26] "lab2/lab2data/data/datafile.15.b.txt"
##  [27] "lab2/lab2data/data/datafile.16.a.txt"
##  [28] "lab2/lab2data/data/datafile.16.b.txt"
##  [29] "lab2/lab2data/data/datafile.17.a.txt"
##  [30] "lab2/lab2data/data/datafile.17.b.txt"
##  [31] "lab2/lab2data/data/datafile.18.a.txt"
##  [32] "lab2/lab2data/data/datafile.18.b.txt"
##  [33] "lab2/lab2data/data/datafile.19.a.txt"
##  [34] "lab2/lab2data/data/datafile.19.b.txt"
##  [35] "lab2/lab2data/data/datafile.2.a.txt"
##  [36] "lab2/lab2data/data/datafile.2.b.txt"
##  [37] "lab2/lab2data/data/datafile.20.a.txt"
##  [38] "lab2/lab2data/data/datafile.20.b.txt"
##  [39] "lab2/lab2data/data/datafile.21.a.txt"
##  [40] "lab2/lab2data/data/datafile.21.b.txt"
##  [41] "lab2/lab2data/data/datafile.22.a.txt"
##  [42] "lab2/lab2data/data/datafile.22.b.txt"
##  [43] "lab2/lab2data/data/datafile.23.a.txt"
##  [44] "lab2/lab2data/data/datafile.23.b.txt"
##  [45] "lab2/lab2data/data/datafile.24.a.txt"
##  [46] "lab2/lab2data/data/datafile.24.b.txt"
##  [47] "lab2/lab2data/data/datafile.25.a.txt"
##  [48] "lab2/lab2data/data/datafile.25.b.txt"
##  [49] "lab2/lab2data/data/datafile.26.a.txt"
##  [50] "lab2/lab2data/data/datafile.26.b.txt"
##  [51] "lab2/lab2data/data/datafile.27.a.txt"
##  [52] "lab2/lab2data/data/datafile.27.b.txt"
##  [53] "lab2/lab2data/data/datafile.28.a.txt"
##  [54] "lab2/lab2data/data/datafile.28.b.txt"
##  [55] "lab2/lab2data/data/datafile.29.a.txt"
##  [56] "lab2/lab2data/data/datafile.29.b.txt"
##  [57] "lab2/lab2data/data/datafile.3.a.txt"
##  [58] "lab2/lab2data/data/datafile.3.b.txt"
##  [59] "lab2/lab2data/data/datafile.30.a.txt"
##  [60] "lab2/lab2data/data/datafile.30.b.txt"
##  [61] "lab2/lab2data/data/datafile.31.a.txt"
##  [62] "lab2/lab2data/data/datafile.31.b.txt"
##  [63] "lab2/lab2data/data/datafile.32.a.txt"
##  [64] "lab2/lab2data/data/datafile.32.b.txt"
##  [65] "lab2/lab2data/data/datafile.33.a.txt"
##  [66] "lab2/lab2data/data/datafile.33.b.txt"
##  [67] "lab2/lab2data/data/datafile.34.a.txt"
##  [68] "lab2/lab2data/data/datafile.34.b.txt"
##  [69] "lab2/lab2data/data/datafile.35.a.txt"
##  [70] "lab2/lab2data/data/datafile.35.b.txt"
##  [71] "lab2/lab2data/data/datafile.36.a.txt"
##  [72] "lab2/lab2data/data/datafile.36.b.txt"
##  [73] "lab2/lab2data/data/datafile.37.a.txt"
```

```
##  [74] "lab2/lab2data/data/datafile.37.b.txt"
##  [75] "lab2/lab2data/data/datafile.38.a.txt"
##  [76] "lab2/lab2data/data/datafile.38.b.txt"
##  [77] "lab2/lab2data/data/datafile.39.a.txt"
##  [78] "lab2/lab2data/data/datafile.39.b.txt"
##  [79] "lab2/lab2data/data/datafile.4.a.txt"
##  [80] "lab2/lab2data/data/datafile.4.b.txt"
##  [81] "lab2/lab2data/data/datafile.40.a.txt"
##  [82] "lab2/lab2data/data/datafile.40.b.txt"
##  [83] "lab2/lab2data/data/datafile.41.a.txt"
##  [84] "lab2/lab2data/data/datafile.41.b.txt"
##  [85] "lab2/lab2data/data/datafile.42.a.txt"
##  [86] "lab2/lab2data/data/datafile.42.b.txt"
##  [87] "lab2/lab2data/data/datafile.43.a.txt"
##  [88] "lab2/lab2data/data/datafile.43.b.txt"
##  [89] "lab2/lab2data/data/datafile.44.a.txt"
##  [90] "lab2/lab2data/data/datafile.44.b.txt"
##  [91] "lab2/lab2data/data/datafile.45.a.txt"
##  [92] "lab2/lab2data/data/datafile.45.b.txt"
##  [93] "lab2/lab2data/data/datafile.46.a.txt"
##  [94] "lab2/lab2data/data/datafile.46.b.txt"
##  [95] "lab2/lab2data/data/datafile.47.a.txt"
##  [96] "lab2/lab2data/data/datafile.47.b.txt"
##  [97] "lab2/lab2data/data/datafile.48.a.txt"
##  [98] "lab2/lab2data/data/datafile.48.b.txt"
##  [99] "lab2/lab2data/data/datafile.49.a.txt"
## [100] "lab2/lab2data/data/datafile.49.b.txt"
## [101] "lab2/lab2data/data/datafile.5.a.txt"
## [102] "lab2/lab2data/data/datafile.5.b.txt"
## [103] "lab2/lab2data/data/datafile.50.a.txt"
## [104] "lab2/lab2data/data/datafile.50.b.txt"
## [105] "lab2/lab2data/data/datafile.6.a.txt"
## [106] "lab2/lab2data/data/datafile.6.b.txt"
## [107] "lab2/lab2data/data/datafile.7.a.txt"
## [108] "lab2/lab2data/data/datafile.7.b.txt"
## [109] "lab2/lab2data/data/datafile.8.a.txt"
## [110] "lab2/lab2data/data/datafile.8.b.txt"
## [111] "lab2/lab2data/data/datafile.9.a.txt"
## [112] "lab2/lab2data/data/datafile.9.b.txt"
## [113] "lab2/lab2data/output/outfile.1.a.out"
## [114] "lab2/lab2data/output/outfile.1.b.out"
## [115] "lab2/lab2data/output/outfile.10.a.out"
## [116] "lab2/lab2data/output/outfile.13.a.out"
## [117] "lab2/lab2data/output/outfile.15.b.out"
## [118] "lab2/lab2data/output/outfile.16.a.out"
## [119] "lab2/lab2data/output/outfile.19.a.out"
## [120] "lab2/lab2data/output/outfile.22.a.out"
## [121] "lab2/lab2data/output/outfile.22.b.out"
## [122] "lab2/lab2data/output/outfile.25.a.out"
## [123] "lab2/lab2data/output/outfile.28.a.out"
## [124] "lab2/lab2data/output/outfile.29.b.out"
## [125] "lab2/lab2data/output/outfile.31.a.out"
## [126] "lab2/lab2data/output/outfile.34.a.out"
## [127] "lab2/lab2data/output/outfile.36.b.out"
```

```
## [128] "lab2/lab2data/output/outfile.37.a.out"
## [129] "lab2/lab2data/output/outfile.4.a.out"
## [130] "lab2/lab2data/output/outfile.40.a.out"
## [131] "lab2/lab2data/output/outfile.43.a.out"
## [132] "lab2/lab2data/output/outfile.43.b.out"
## [133] "lab2/lab2data/output/outfile.46.a.out"
## [134] "lab2/lab2data/output/outfile.49.a.out"
## [135] "lab2/lab2data/output/outfile.50.b.out"
## [136] "lab2/lab2data/output/outfile.7.a.out"
## [137] "lab2/lab2data/output/outfile.8.b.out"
## [138] "labweek1_tiffanybatarseh.pdf"
## [139] "labweek1_tiffanybatarseh.Rmd"
## [140] "Prelabwork_Week1.html"
## [141] "Prelabwork_Week1.Rmd"
```

**8. How do you open gzip-compressed files for reading and for writing?**

```
gunzip("file.gz")
```

Where the default behavior is to remove the original compressed file after the unzipped file has been created and completed.

**9. Generate a list of those output files that were not created**

Stringr package makes it better, data files are .txt and ouput are .out, replace input patterns so maybe string replacement; use stringr if doing string replacement

Rstudio.com/resources/cheatsheet

## Python

Get current working directory

```
os.getcwd()
```

List contents of a directory

```
os.listdir("/Users/tiffanybatarseh/AdvInformatics/lab2/lab2data")
```

Change into a different directory

```
os.chdir("/Users/tiffanybatarseh/AdvInformatics/lab2/lab2data")
```

**1. How do you ask if a file exists?**

```
import os
os.path.exists("datafile.1.a.txt")
```

**2. How to you ask if a file is a directory?**

```
os.path.isdir("datafile.1.a.txt")
```

**3. How do you remove (delete) a file?**

```
os.remove("datafile.1.a.txt")
```

**4. How do you get the size of a file?**

```
os.path.getsize('datafile.1.a.txt')
```

**5. How do you get all the file names matching a pattern?**

```
import glob
glob.glob("data*.txt")
```

**6. How do you get all the file names matching a pattern recursively?**

```
for filename in glob.glob('/Users/tiffanybatarseh/AdvInformatics/lab2/lab2data/**/[do]*', recursive=Tru
    print(filename)
```

**7. How do you get an iterator to all files matching a pattern, as opposed to returning a potentially huge list?**

```
glob.iglob('/Users/tiffanybatarseh/AdvInformatics/lab2/lab2data/**/[do]*', recursive=False)
```

**8. How do you open gzip-compressed files for reading and for writing?**

```
gzip.open(filename)
```

**9. Generate a list of those output files that were not created**

```
listhere = glob.glob('/Users/tiffanybatarseh/AdvInformatics/lab2/lab2data/**/*.[0-9*][0-9*].[ab].*', re
comparelist = set(listhere)
```

```
listhere = glob.iglob('/Users/tiffanybatarseh/AdvInformatics/lab2/lab2data/**/datafile.[0-9*].[ab].txt'
dcompare = set(listhere)
list2 = glob.iglob('/Users/tiffanybatarseh/AdvInformatics/lab2/lab2data/**/outfile.[0-9*].[ab].out', re
ocompare = set(list2)

for ocompare in dcompare
```