R Hands-on Introduction

David Walling, Weijia Xu
Texas Advanced Computing Center
The University of Texas at Austin
Goal

What you will do

• Start interactive R session on Stampede compute node
• Follow along as we explore basic R syntax
  o Variables & Types
  o Data structures
  o Control Flow
  o Functions
  o Data I/O
• Objective: Know enough R to follow the more advanced R examples to follow
• Reference script available in course materials: RIntroduction.R
Start Interactive Session

- From stampede login node:
  - login1$> idev -A 20130927DataIntensive -p normal-mic -m 60

```bash
login4@ idev -A 20130927DataIntensive -p normal-mic
System     : Stampede
Using Project     : -A 20130927DataIntensive
Using queue     : -p normal-mic
We found an ACTIVE reservation request for you, named TACC-Training-2013-09-27. Do you want to use it for your interactive session? Enter y/n [default y]: Reservation     : --reservation=TACC-Training-2013-09-27 (ACTIVE)
Using res. queue : -p normal-mic
Welcome to the Stampede Supercomputer

--> Verifying valid submit host (login4)...OK
--> Enforcing max jobs per user...OK
--> Verifying availability of your home dir (/home/00791/xwj)...OK
--> Verifying availability of your work dir (/work/00791/xwj)...OK
--> Verifying availability of your scratch dir (/scratch/00791/xwj)...OK
--> Verifying access to desired queue (normal-mic)...OK
--> Verifying job request is within current queue limits...OK
--> Checking available allocation (20130927DataIntensive)...OK
Submitted batch job 1006326

After your idev job begins to run, a command prompt will appear, and you can begin your interactive development session. We will report the job status every 4 seconds: (P=pending, R=running).

Job status: R
--> Job is now running on master node- c412-502...OK
--> Sleeping for 7 seconds...OK
--> Checking to make sure your job has initialized an env for you....OK
--> Creating interactive terminal session (login) on master node c412-502.
TACC Stampede System
LosP 0.40.0 (Top Notch)
Provisioned on 07-Sep-2012 at 11:51
```

TXA - Texas Advanced Computing Center
Get the sample data and script

- `c412-502$ > cp –r /scratch/00791/xwj/R_lab_09272013`

- `c412-502$ > cd R_lab_09272013`

```
total 128
-rw-r--r-- 1 xwj G-80788 2526 Sep 27 12:20 cluster.csv
-rw-r--r-- 1 xwj G-80788 248 Sep 27 12:20 kmeans.R
-rw-r--r-- 1 xwj G-80788 224 Sep 27 12:20 mtcars2.R
-rw-r--r-- 1 xwj G-80788 193 Sep 27 12:20 mtcars.R
-rw-r--r-- 1 xwj G-80788 3038 Sep 27 12:20 RIntroduction.R
-rw-r--r-- 1 xwj G-80788 129 Sep 27 12:20 sampleData.csv
-rw-r--r-- 1 xwj G-80788 171 Sep 27 12:20 Sample.R
-rw-r--r-- 1 xwj G-80788 95 Sep 27 12:20 SampleSource.R
-rw-r--r-- 1 xwj G-80788 32268 Sep 27 12:20 sonar_test.csv
-rw-r--r-- 1 xwj G-80788 53850 Sep 27 12:20 sonar_train.csv
-rw-r--r-- 1 xwj G-80788 328 Sep 27 12:20 svm.R
-rw-r--r-- 1 xwj G-80788 748 Sep 27 12:20 worms.txt
```
Start R session

- `c412-502$ > module load R_mkl`
- `c412-502$ > R`
Rintroduction.R

• Basic practice on R syntax and types.
• Including seven part
  • Using existing libraries and source code
  • Data Type
  • Data Structures
  • Control Flow
  • Functions
  • Data Input/output
  • Create Plots
On Running R in batch mode and with parameters

• mtcars.R mtcars2.R sample.R
On kmeans clustering

- Kmeans.R, cluster.csv
On SVM classification

• svm.R, sonar_train.csv, sonar_test.csv