Introduction to R
Spring 2021

Wim Cardoen & Brett Milash
Center for High-Performance Computing (CHPC)
University of Utah

Overview

- What is R?
- Why R?
- R language basics, in casu:
  - Basic Data Types
  - Vector, Matrix & Array
  - List, Data Frame
  - Functions
  - Loops, Conditionals,..
  - IO
- R @ CHPC
- Interesting sites/Links
- **NO Stats will be covered here. => MATH Dept.**

What is R?

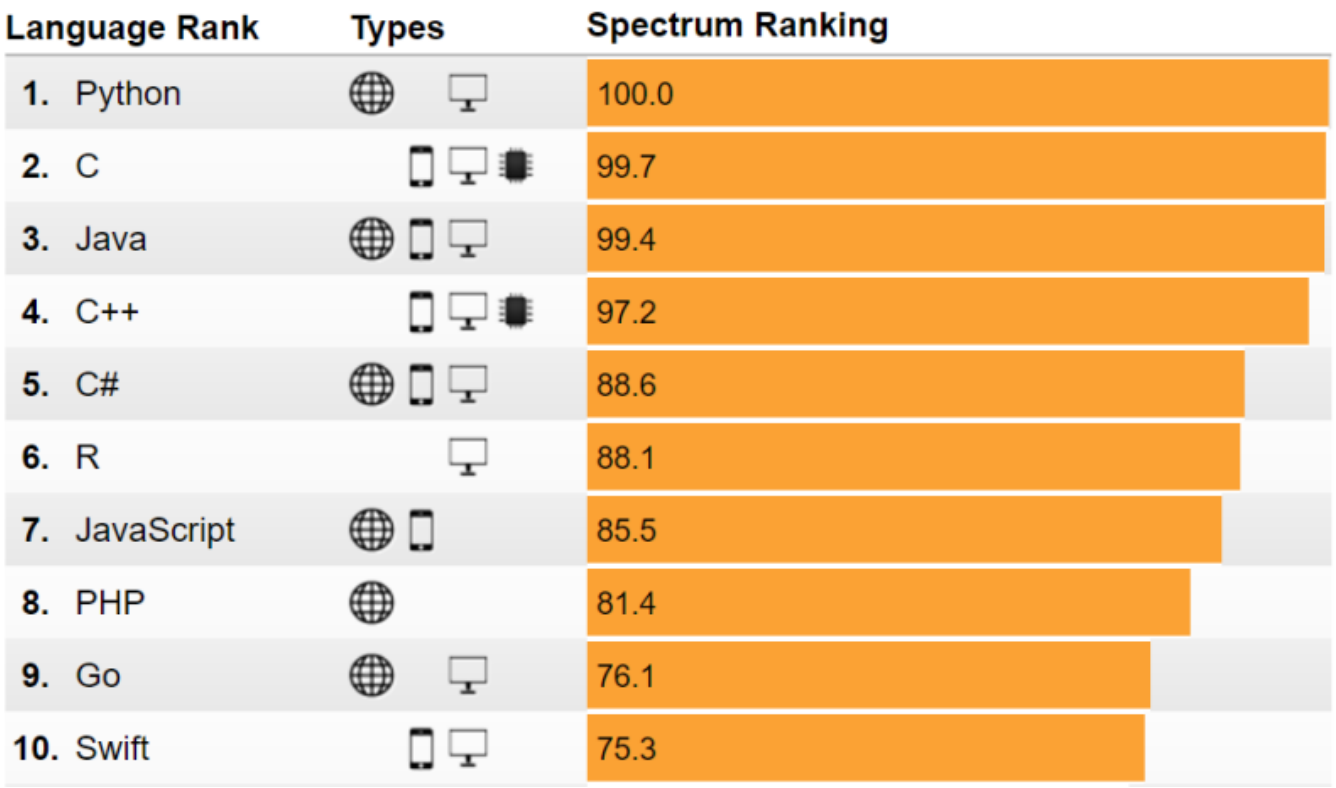
- Implementation of S (stat. prog. lang. developed @ Bell Labs by John Chambers)
- Original authors: Ross Ihaka & Robert Gentleman (Auckland, NZ) around 1992.
- Two facets:
  - Scripting language (vs. compiled language)
  - Data Analysis environment
- R-code can run on different OS's (Linux, Windows, MacOs)
- Under the hood: relies on C/C++, Fortran for comp. expensive tasks

  (e.g. linear algebra,…)
- Free & Open-Source

Why R?

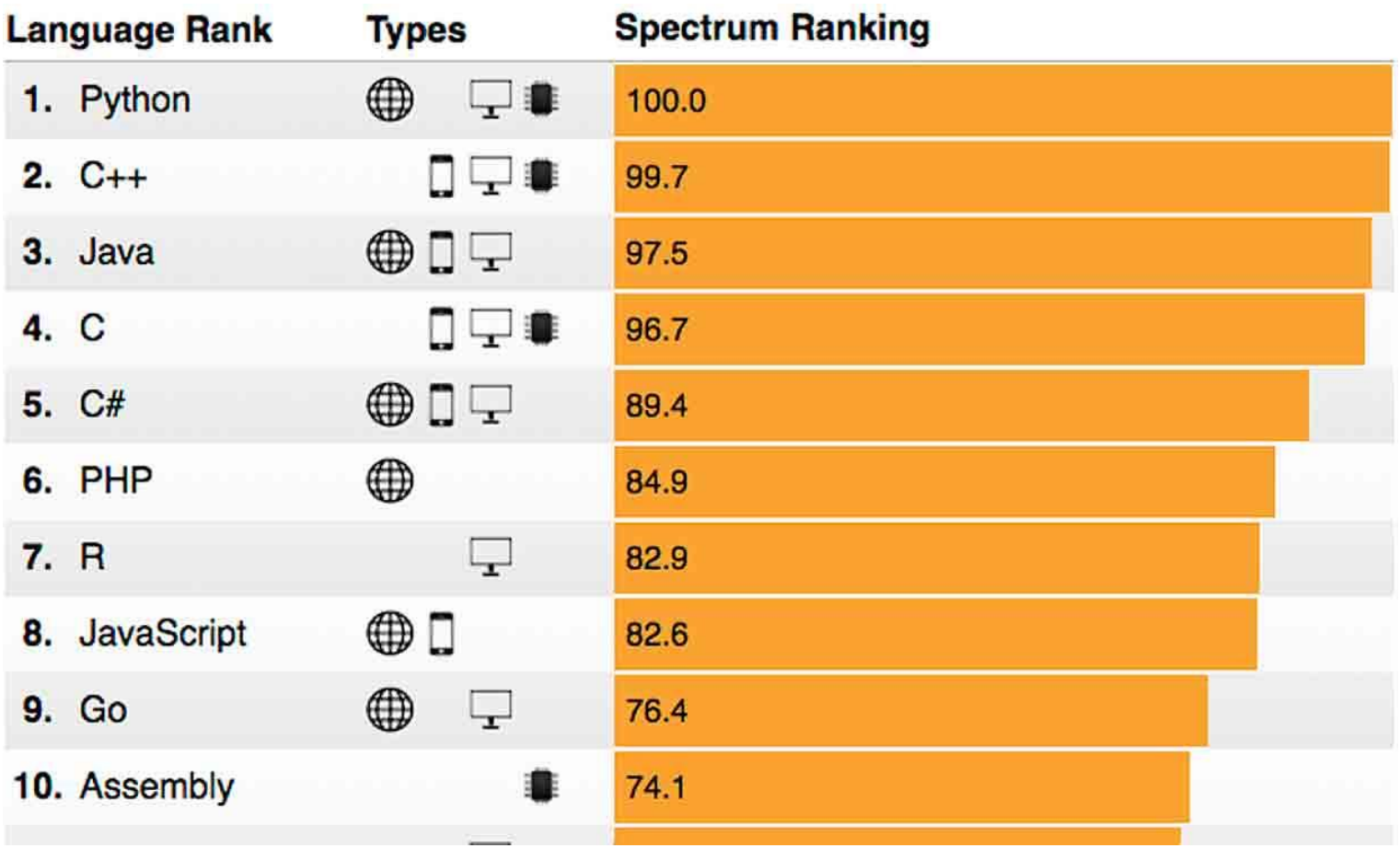
- Scripting language -> rapid proto-typing
- Most diverse set of statistical tools
- A lot of pre-canned packages (libraries)
- Relatively easy to add new packages
- Large Community
- Free
- Job “security”?

IEEE: The Top 10 Prog. Languages (2017)



Language Rank	Types	Spectrum Ranking
1. Python		100.0
2. C		99.7
3. Java		99.4
4. C++		97.2
5. C#		88.6
6. R		88.1
7. JavaScript		85.5
8. PHP		81.4
9. Go		76.1
10. Swift		75.3

IEEE: The Top 10 Prog. Languages (2018)



Using R @ CHPC

```
# Simple Approach (to start)
module avail
module load R/4.0.2      # Centos 7
# Invoke Command Line Interpreter (CLI)
R
R>

# If you want to use an IDE (e.g. Rstudio)
module load RStudio
rstudio                  # Load IDE
```

R @ CHPC cont'd

- Using an R Batch script (CMD Line -> Testing)
  - Options:
    - Rscript yourfile.R # Redirects output to stdout
    - R CMD BATCH yourfile.R # Redirects output to yourfile.Rout
    - R –no-save < yourfile.R
    - ./yourfile2.R # Add '#!/usr/bin/env Rscript' as top line

R @ CHPC cont'd

- Submit an R Batch script on the cluster (using SLURM):

  Example SLURM scripts to be found in:

https://www.chpc.utah.edu/documentation/software/r-language.php

R @ CHPC cont'd

- R library:= A location where R packages are installed
- R package := Fundamental Unit of Reproducible R code
- => R packages are installed in R libraries (vs. C/C++/Fortran,..)
- Different libraries:
  - R-Core installation
    - *R>.Library / echo $R_HOME*
  - Libraries installed by the CHPC
    - *echo $R_LIBS_SITE*
  - User Libraries (if the user decides to do this – default: absent)
    - *echo $R_LIBS_USER*
- Check existing libraries: *R>.libPaths()*

R @ CHPC cont'd

- Installation of packages:
  - install.packages() (**high level**)

    Example: (maRketSim package)

    install.packages(c("maRketSim"),
    lib=c("$YOUR_LIBRARY"),
    repos=c(http://cran.us.r-project.org),verbose=TRUE)
  - R CMD INSTALL (**low level**)

    Example: (RNetCDF package)

    export PATH=/uufs/chpc.utah.edu/sys/installdir/netcdf-c/4.3.2i/bin:$PATH
    export PATH=/uufs/chpc.utah.edu/sys/installdir/udunits/2.2.20/bin:$PATH

R @ CHPC cont'd

```
R CMD INSTALL \
--library=/uufs/chpc.utah.edu/common/home/$USER/RLibs/3.5.2i \
--configure-args="CPPFLAGS='-I/uufs/chpc.utah.edu/sys/installdir/udunits/2.2.20/include'\
      LDFLAGS='-Wl,-rpath=/uufs/chpc.utah.edu/sys/installdir/netcdf-c/4.3.2i/lib \
      -L/uufs/chpc.utah.edu/sys/installdir/netcdf-c/4.3.2i/lib -lnetcdf \
      -Wl,-rpath=/uufs/chpc.utah.edu/sys/installdir/udunits/2.2.20/lib\
      -L/uufs/chpc.utah.edu/sys/installdir/udunits/2.2.20/lib -ludunits2 ' \
      --with-nc-config=/uufs/chpc.utah.edu/sys/installdir/netcdf-c/4.3.2i/bin/nc-config " \
RNetCDF_1.8-2.tar.gz
```

Note:

R CMD install calls ./configure under the hood

Questions?

Links

- https://www.r-project.org/
- https://cran.r-project.org/ (**C**omprehensive **R** **A**rchive **N**etwork)
- https://www.r-bloggers.com/
- The Art of R Programming (Norman Matloff)
- Hadley Wickham
- R mailing-list
- Stack Overflow (R Channel)