



# R and LSAF – Challenges and Solutions

John Hendrickx  
Senior Statistical Programmer  
Danone Nutricia Research  
Utrecht, The Netherlands

# R and LSAF

- LSAF
  - Statistical Programming Environment by SAS
  - Audit history
  - Version control
  - Access control
  - Cloud based

# LSAF Basics

- LSAF
  - Repository for final results
  - Workspace for program development
  - Job file to run program in repository
    - Required files only

# R and LSAF

- R programs on LSAF 5.3 and higher
- Much like running SAS
  - F3 (or Ctrl+Enter) to run
  - Console output to the log window
  - Job to run in repository

# R and LSAF

- Rough edges
  - No data viewer (yet)
  - No run and populate
  - Installing packages with no internet

# Data Viewer

- DT::dataviewer
  - Creates an html page
  - Options for hiding columns, sorting, filtering

# Data Viewer

- Requirements
  - DT package
  - htmlwidgets package
  - Pandoc software (OS level)

# Pandoc

- Software to convert between output formats
- Also used by rmarkdown, arsenal for creating Word output
- Requires an RFC submission to SAS

# Data Viewer

```
library(DT)
library(htmlwidgets)

mytbl <- datatable(mtcars)

saveWidget(mytbl, paste0(LSAF_WS, LSAF_FILELOCATION, "/mtcars.html"))
```

# Data Viewer

|                          |                                                                                   |                        |
|--------------------------|-----------------------------------------------------------------------------------|------------------------|
| <input type="checkbox"/> |  | <u>mtcars_files</u>    |
| <input type="checkbox"/> |  | <u>DT-datatables.R</u> |
| <input type="checkbox"/> |  | <u>mtcars.html</u>     |

Column visibility

Show 10 entries

Search:

|                   | mpg  | cyl | displacement | hp  | drat | wt    | qsec  | vs  | am  | gear | carb |
|-------------------|------|-----|--------------|-----|------|-------|-------|-----|-----|------|------|
|                   | All  | All | All          | All | All  | All   | All   | All | All | All  | All  |
| Mazda RX4         | 21   | 6   | 160          | 110 | 3.9  | 2.62  | 16.46 | 0   | 1   | 4    | 4    |
| Mazda RX4 Wag     | 21   | 6   | 160          | 110 | 3.9  | 2.875 | 17.02 | 0   | 1   | 4    | 4    |
| Datsun 710        | 22.8 | 4   | 108          | 93  | 3.85 | 2.32  | 18.61 | 1   | 1   | 4    | 1    |
| Hornet 4 Drive    | 21.4 | 6   | 258          | 110 | 3.08 | 3.215 | 19.44 | 1   | 0   | 3    | 1    |
| Hornet Sportabout | 18.7 | 8   | 360          | 175 | 3.15 | 3.44  | 17.02 | 0   | 0   | 3    | 2    |
| Valiant           | 18.1 | 6   | 225          | 105 | 2.76 | 3.46  | 20.22 | 1   | 0   | 3    | 1    |
| Duster 360        | 14.3 | 8   | 360          | 245 | 3.21 | 3.57  | 15.84 | 0   | 0   | 3    | 4    |
| Merc 240D         | 24.4 | 4   | 146.7        | 62  | 3.69 | 3.19  | 20    | 1   | 0   | 4    | 2    |
| Merc 230          | 22.8 | 4   | 140.8        | 95  | 3.92 | 3.15  | 22.9  | 1   | 0   | 4    | 2    |
| Merc 280          | 19.2 | 6   | 167.6        | 123 | 3.92 | 3.44  | 18.3  | 1   | 0   | 4    | 4    |

Showing 1 to 10 of 32 entries

Previous

1

2

3

4

Next

# Data Viewer

- `showData`: R function to view dataframe
- But:
  - Can't handle very large dataframes
  - Need to delete temporary files manually

## R job files

- LSAF can run R programs
- LSAF can create a job file
- But:
  - No “run and populate”

# R job files

- Solution: generic R job files
  - Tied to Nutricia standard folder structure
- Standard input locations
- Standard output locations
- Use “Include subfolders” option judiciously
- “Check in” the results

# Sample Job

Details

Programs

Parameters

Publish Setup

Check-in

Files

Folders

# Sample Job

Details

Programs

Parameters

Publish Setup

Check-in

 Version limits may result in versions being permanently deleted. Unversioned files will be overwritten. 

☐ Enable versioning on all new output files

Version to be applied to all versioned output files:

☒ Major ☐ Minor ☐ Custom

1.0

|   |                   |          |                                     |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------|----------|-------------------------------------|
| <input type="checkbox"/>                                                                                                                                                | Name              | Location | Include Subfolders                  |
| <input type="checkbox"/>                                                                                                                                                | 55 Output Results | ../..    | <input checked="" type="checkbox"/> |
|                                                                                                                                                                         |                   |          |                                     |
|                                                                                                                                                                         |                   |          |                                     |
|                                                                                                                                                                         |                   |          |                                     |

# Generic Job File

- Job files are XML (text)
- SAS (!) macro to write the job file
  - SAS program to call the macro
  - Job file to run the SAS program
- User runs job from job folder
  - Creates generic R jobs

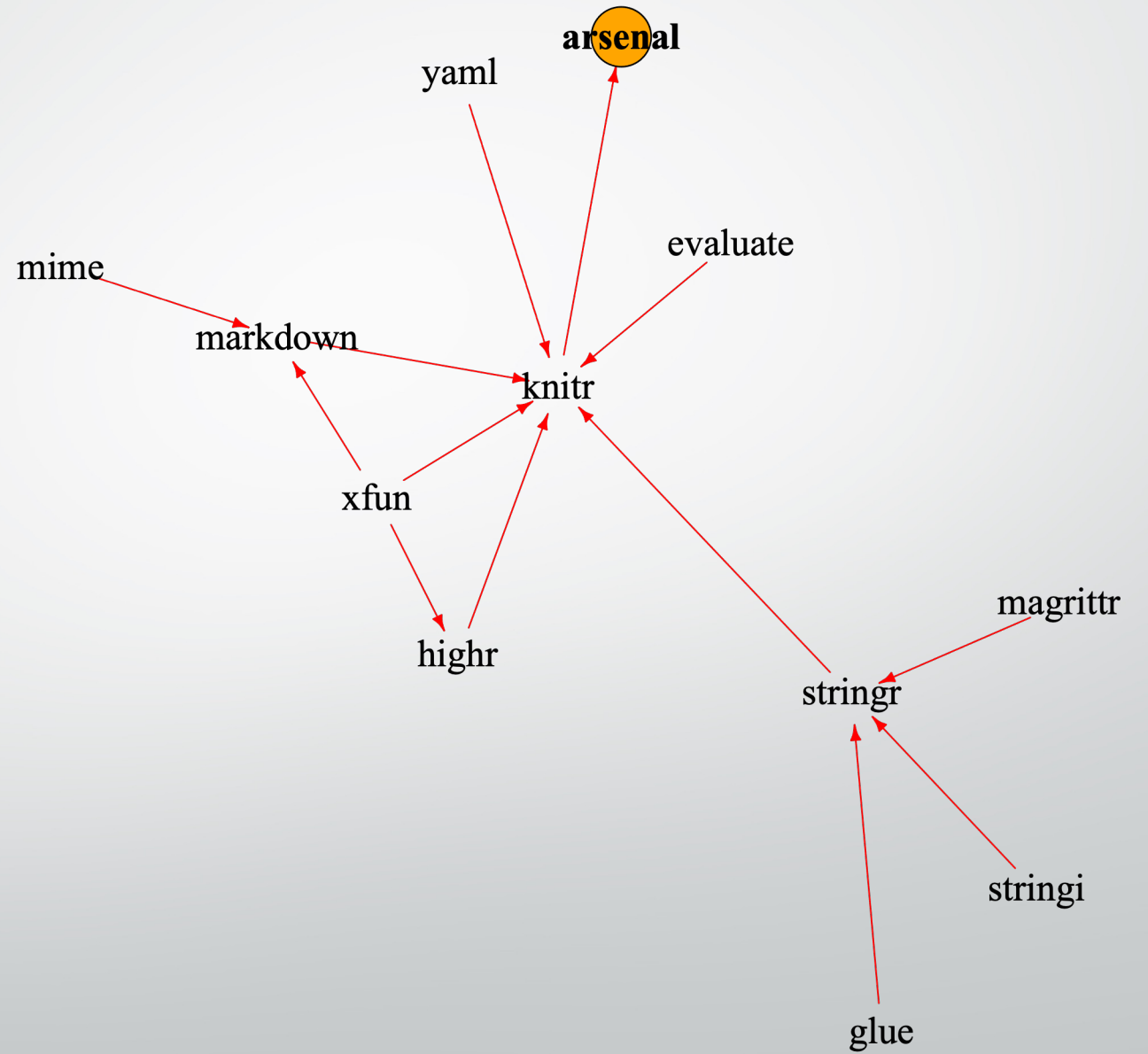
# Generic Job File

- Works as intended
  - In Repository
  - Workspace requires copying program
- Unnecessary files copied to transient area
  - (Small) speed penalty

# R packages

- Complicated
  - No direct download
  - Download from CRAN mirror, upload to LSAF
  - Packages require packages
  - Install order trial error

## Dependencies for the arsenal package



# miniCRAN

- Creates an offline CRAN repository for selected packages
- Upload miniCRAN repository to LSAF
- `install.packages()`, `uninstall.packages` work

```
install.packages("arsenal", repos=CRAN_local, lib=RLibrary, type="source")
```

- Package dependencies handled automatically

# Local Repository

LSAF R develop

03 miniCRAN test

.Rproj.user

CRAN\_offline

src

contrib

| Name                        | Status | Date modified   | Type       | Size     |
|-----------------------------|--------|-----------------|------------|----------|
| PACKAGES                    | ✓      | 16-7-2021 16:50 | File       | 56 KB    |
| PACKAGES.gz                 | ✓      | 16-7-2021 16:50 | gz Archive | 15 KB    |
| PACKAGES.rds                | ✓      | 16-7-2021 16:50 | RDS File   | 14 KB    |
| nutriciaconfig_0.1.4.tar.gz | ✓      | 16-7-2021 16:50 | gz Archive | 16 KB    |
| rmarkdown_2.9.tar.gz        | ✓      | 15-7-2021 16:41 | gz Archive | 3,171 KB |
| tinytex_0.32.tar.gz         | ✓      | 15-7-2021 16:41 | gz Archive | 30 KB    |

# Local Repository

- Create using `miniCRAN::makeRepo()`
  - Packages are downloaded
  - `PACKAGES*` files are updated
  - Contain information on dependencies

# LSAF miniCRAN

- Upload local repository
- Reference with "file:///"
  - URL required

```
CRAN_local <- paste0("file:///",LSAF_WS,"/SAS/Files/00_miniCRAN_test/CRAN_offline")  
RLibrary <- paste0(LSAF_WS,"/SAS/Files/00_miniCRAN_test/R Library")  
install.packages("DT", repos=CRAN_local, lib=RLibrary, type="source")
```

- Dependencies are detected and installed in the correct order

# R Packages in LSAF

- Dependencies aren't a problem
- Still no walk in the park
  - Newer version of pre-installed
  - Latex libraries
  - Linux programming libraries

# Summary

- R in LSAF 5.3
  - Dataframes can be viewed
  - Pre-populated job files can be created
  - Package dependencies aren't a problem

# Conclusion

