Driving R adoption in an NHS information service: barriers and solutions

Chris Mainey

Intelligence Analyst
University Hospitals Birmingham NHS Foundation Trust

chris.mainey@uhb.nhs.uk
@chrismainey

EARL
ENTERPRISE APPLICATIONS OF THE R LANGUAGE

Health Informatics
Information. Analysis. Statistics. Reporting.

University Hospitals Birmingham NHS FT

- Second largest acute hospital trust in England
- Four hospital sites seeing approx. 440,000 in-patients per year

- Large informatics team(s):
  - Core Information
  - Analytics
  - SQL engineers/ Dashboard developers
  - Research
  - Commercial

So we're all R-wielding data ninjas, right?

Healthcare Evaluation Data (HED)

www.hed.nhs.uk

- Online hospital benchmarking system
- Statistical models and analysis tools
- Activity, Mortality, Readmissions, Length-of-Stay, Marketshare etc.
- Used by ~60 NHS and other organisations
- Training and support

- **Using national NHS data**

Data analysis tools in the NHS

- Excel and Access were baseline 10-years ago
- Variable reliance on IT departments and external suppliers
- Huge differences in skills across, and within, organisations
- Increasing use of Qlik, PowerBI and Tableau

UHB:

- Recruited SQL devs and slowly moved whole department
- Excel, SSRS and TIBCO Spotfire used for outputs

Keenies...

- R was used only for rare 'statistical' calculations
- A few keen researchers or medical students
- Simlar at other Trusts
- Also the case in our team
- **But...**

One keen user...

Allowed space to develop:

- PhD project, funded by UHB
- Developed statistical, visualisation and data wrangling skills
- Chose to use `R` from the outset, rather than SAS or Stata.
- Python training harder to find when I started

Took a risk

- Moved some HED models to `R`, from SAS
- Supported other team members to use it
- This generated a demand for basic training

Followed established patterns:

- Set up a local user group
- Trained senior team
- Encouraged new starters to use `R`
- External relationships:
  - Twitter has been key
  - NHS-R Community: https://nhsrcommunity.com/
  - External `R` user groups
  - Regional relationships: Other Trusts, PHE, City Council
  - Advice from Mango, RStudio and others
- ***Demonstrated ROI !!!***

Barriers

- **Open Source fears:** IT teams sometimes reluctant to install
- **Wrestling with default settings:**
  - Mapped user folders that aren't writable
  - Can't write to package library, or download temp files
- **Poor hardware**
  - Cheap desktops, low on disk space with minimal RAM
- **Lack of understanding of fundamentals:**
  - Networking/file locations
  - Suspicions around security
  - `R` coding, or its application to problems

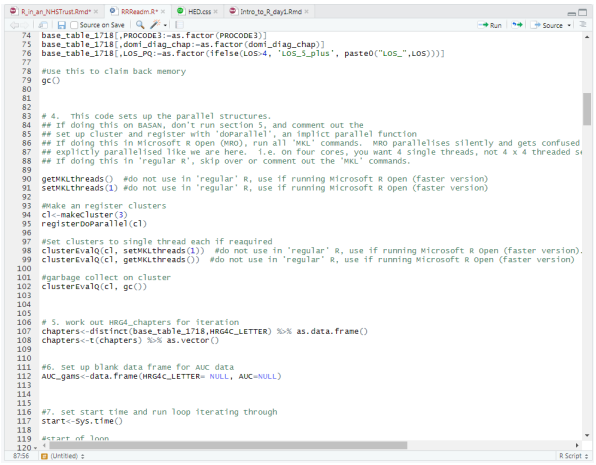
Solutions

- **Permissions:** *Explicitly identify, and change, default settings*
  - Set a folders: `R_libs_user` or `.Renviron` file
- **Hardware:**
  - Make the case for more RAM! It doesn't cost much
  - Consider server approaches: VMs or 'Desktop' on a 'Server'
  - RStudio Server / RStudio Connect
- **Skill/Awareness:**
  - Protect time
  - Windows users often don't know file paths, SSH etc.
  - Central guidance on R/Python implementation

How do we use it?

1. Regression, and related models (`mgcv`, `lme4`, `glmmTMB`, `glmnet`)
2. Machine Learning technique (`caret`, `randomForest`, `gbm`)
3. Optimisation problems (solving equations using `nloptr`)
4. Summary stats and Exploration
5. Data Visualisation (`ggplot2`, `FunnelPlotR`)
6. `RMarkdown` Reports and Presentations
7. API interaction (NHS ODS service, CQC registered organisations)

Initial adoption:

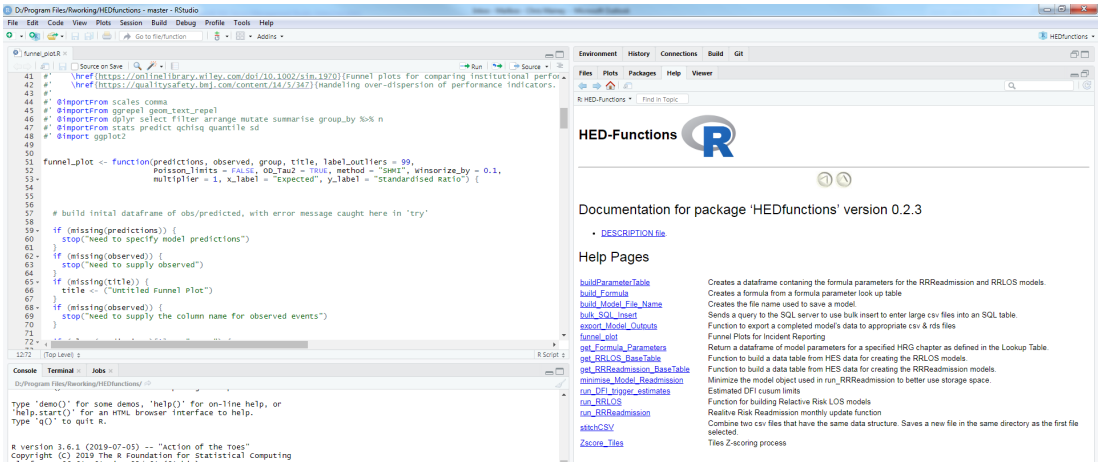


- Collections of `R` scripts saved to directory
  - Required lead analyst to write
  - Required team skills to run and/or debug
- Constant debugging:
  - Data changing
  - Littered with comments, hard to decipher

Creating a local package

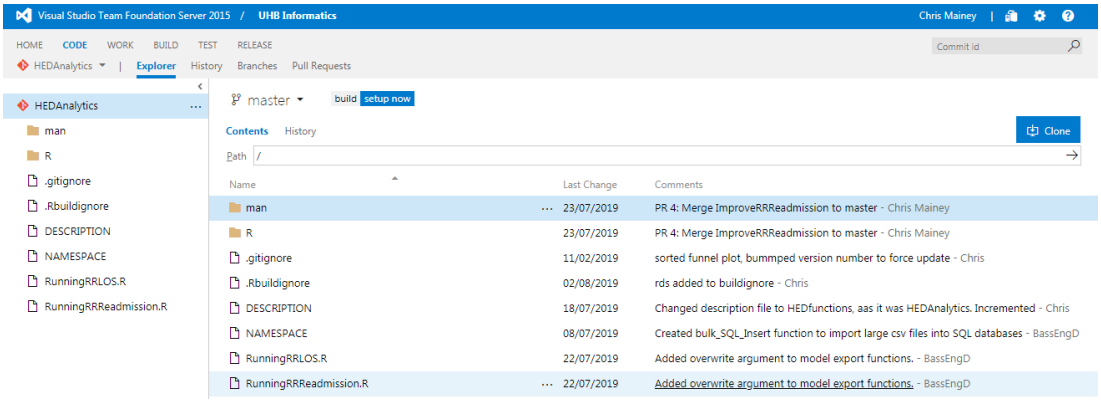
Wrapped functions in to an `R` package structure

- Built binary package and saved to shared folders
- Used `roxygen2` to write support material



Source control

- Moved to source control
  - Install git for windows
  - Git learning curve
  - Happy Git with R! https://happygitwithr.com/
- Git repo on TFS
  - Figure out https connections
  - Steph Locke's `tfsr` https://github.com/lockedata/tfsR



NHS-R community

- Promote R use in NHS
- Led by Prof. MA Mohammed
- Funded by grant from the Health Foundation
- Training provision
- Blog posts
- Share best practice
- Development of tools
- National Conference: 4th & 5th November

NHS-R COMMUNITY

What's next?

- Continuing to develop analyst skills in `R`
- Spreading to different teams: control charts, Rmarkdown
- Contributing to, or building, Open Source tools with NHS-R community:
  - `FunnelPlotR`
  - `NHSRdatasets`
- Pushing further into ML techniques for prediction
  - Improving current models: Ensemble techniques etc.
  - Unsupervised techniques for cluster analysis.

Thanks for your time!

chris.mainey@uhb.nhs.uk

http://www.hed.nhs.uk

@chrismainey

chrismainey

http://www.mainard.co.uk

`FunnelPlotR` now available on CRAN!