Interacting with an Open Source Community

Jasmine Dumas

September 21, 2016

## Tips & Methodology

* As Insurance Data Scientist's we have the opportunity to use [open-source](https://opensource.org/) software to model and communicate ideas across the business. Open-source software is surpassing commercial software due to cost, restriction, feature access and is being adopted by hobbyist for weekend data projects to being used for FDA submissions by researchers.
* The open source community and its developers are fearlessly passionate about the tools they create and that leads emerging new features, such as new functions for loss functions in xgboost - so it is advantageous to know how to properly interact with the developers of the tools we use (and love) to do our work, request new features, report bugs and submit technical questions.

### Google

* Google your error!
* ... or use <http://rseek.org/> which provides some **R** specific resources when searching
* as always ?FunctionName on Google will not work well as it does in **R**!

### StackOverflow (SO)

* Interacting on stackoverflow does not require signing up for a account, but you can log in from another Google account to be invovled with the voting process.
* Before asking a new question make sure that it has not already been answered, in part.
* The key to writing a great question and recieveing a answer back (without sassy responses) is in creating a [minimal reproducible example](http://stackoverflow.com/questions/5963269/how-to-make-a-great-r-reproducible-example). As Data Scientist's with access to sensitive material it can be hard to generate a minimal exmaple suitable enough for world view, so look at the "native" datasets in the base **R** distribution, and the ever useful iris dataset to craft an example to submit to SO. Screenshots of the error can be helpful too for visual defects but the code should support your question fully.

### GitHub

* same deal as SO, except you will need an account to ask questions, but you will be able to search the repository of the **R** package of interest.
* questions on GitHub are in the form of submitting a "issue", such as on this link for [knitr](https://github.com/yihui/knitr/issues) which can be categorized and assigned by the developer to priotizie feature request and bug fixes which are linked to specific commits or version of the **R** package or software.
* Be kind on GitHub as you will have direct access to the developer and they usually are volunteers and have limited time.

### Email & Listserves

* Also be kind when emailing **R** package developers using personal or academic emails that are provided in the CRAN description for each package.
* [mailing lists](https://www.r-project.org/mail.html) are another avenue for seeking out help and interacticing with the R community.

## Further reading

* [Mastering Issues on GitHub](https://guides.github.com/features/issues/)
* <https://opensource.com/>
* [Outreachy](https://www.gnome.org/outreachy/)
* [The Journal of Open Source Software](http://joss.theoj.org/)

fin.