HW_6_Q#1

Tulsi Ram Gompo

11/18/2016

Question 1

Read the article *Good Enough Practices in Scientific Computing* **by Wilson et al. (available here). In a half page, describe which of these "pretty good practices" your last homework incorporated. Also list one or two practices that you did not follow in your last homework but that would have made sense and how you could have followed them**

Description of pretty good practices that we incorporated in the last home work

- Data Management

a) *Saving the raw data* is necessary and we saved the raw FARS data in the seperate directory to prevent overwriting on it.So, we are able to get back it and see it when needed.

b) *Creating the data we wish to see* includes converting to file formats for R like csv, dbf etc and rename the files for each year to avoid any easy acess fora particular year. The *variables names* of were in the upper case and we converted to lower case and we recoded some of them.

c) *Create analysis friendly data*: We made the tidy data by making each column a variable and each row an observation.

d) *Record all the step to process data*: We wrote the script for what we did for each of data and recored in them in seperate files to access easily for future.

e) *Submit data to reputable DOI -issuing reppository to acess by others and cite it.* We pushed all our work in github such that others could easily acess and reproduced it.

- Software: It is writing creating and sharing the scripts in readble,reusable and testable form with explicitly. Firstly,we should write a brief comment at start of every program with examples.Secondly, the programs should be decomposed in functions .Also, meaningful naming of the functions and variables should be done to avoid duplication.Next we should provide a toy data and small examples.Later the codes should be submitted to open acess DOI-repository. We did these steps in our last homework.

- Collaboration: The project should be worked with collaboartions and every member could contibute something on the projet.It is necessary to create a README.txt file to give a brief over view of the the project containing title, uptodate contact informations etc.

- Project organizations: *The sepertate project directory for a project should be ceated* We made a separate project directory for our projects and organized the documents in sub directories.We have Raw directory for raw data and meta data,doc for text document,results directory. We also maminatin src directory to keep all the project codes.Also we named all the files representing the functions.For insatance,clean.data.R,fars.analysis,fars_function.R etc.

- Keeping all tarck changes: There are use of version control system in software like Git to track the changes. We are able to see the changes in the data from the colloborators from the commit window of the github.

- Manuscripts We should make the mauscript easy accessible to us and co-authors making a master document.Good manuscrpits has advantages over loss the chances of lost or over writing,avoid the duplications,make easy to final pdf,and make easy to share the final version with colloboatos and submit to journal.We made the manuscript of the project in pdf form and submitted for grading.

Practices that we did not followed in last home work

i) Use of non-commercial liscence: If we are doing the big projects like the government or the company data we should avoid the non-commerical version.As these type of liscence may impede in reuse.

ii) Collaboration: We are working sepertaely on own project directory and there were no collaborations for the HW 4 and 5.

We could tried to apply as good practices as possible to make the tidy project management but we are just learning the projects so that seems ok.