**Class Reflections for UP 494**

Written by - Arpita Banerjee, for Tuesday Session No. 3 (02-09-2021)

Session Name: *Introduction to R-Visualizing Data*

Reflection Prompts:

* *How is working with data via a command line interface different from using a graphical user interface?*
* *How do we balance reproducibility of our analysis with accessibility of the analytical tools?*
* *What’s challenging for you right now regarding learning and using R? What are the “a-ha” moments that you’re celebrating?*

Write-Up:

R enables the user to write a set of commands (in a sub-language/somewhere between high-level and machine level languages) to conduct data analysis. The difference between analysing in R and Excel is the fact that one cannot press buttons on the GUI to execute a function. R steers away from the intuitive representation of data possibilities and bestows the responsibility of analytical imagination on the user. This increased sense of autonomy and agency over one’s data can be both liberating and intimidating simultaneously. Operating on a command line interface takes away the burden of discovering accurate buttons/options/menus from the GUI , enabling us to find the types of analysis we want to perform as we keep writing – preventing interruptions in our streams of analytical thought. However, at the same time, analytical imagination itself (for beginners) becomes a constraint – we only know a few things and can do, just them. In case of graphical user interface, there is a plethora of options, for both actual quantitative analysis and data visualization – one that is very accessible and enables us to play with representations of simple analytics more intuitively. However, it is easy to get lost in the cornucopia of design options and colorful chart junk and lose that sense of individual agency over one’s representations.

I have often caught myself guilty of selecting easy and aesthetically pleasing data visualizations ( which are quick, easy and pretty) and not paid much attention to how they describe my analysis or guide the viewer towards its overarching rationales. Therefore, the accessibility of easy graphics can be counterintuitive. In case of R, it is deliberately designed to enable the user to think, ideate, imagine, and learn – I find myself searching constantly for new commands/bug fixes that could enable me to engage deeper with my analytical process. I have come to discover several visual biases that I endorsed, through my earlier representations (same colors for graphs depicting SWOC analyses, using only bar-charts and sometimes changing colors for visual variety) and a number of stories that I had therefore misrepresented. Although difficult to adjust to, I believe the command line interface will help me painstakingly steer away from some of the visual dilemmas that I have driven dangerously close to, unhinge my analytical freedom and propagate a new modality of analytical thought that I was unaware of. Overall, I think command line interface > GUI.

Since R makes us think in nitty-gritty details about the conventions of our analysis, it is also difficult to learn, and mostly inaccessible to planning analysts who are not a part of the developed world. Accessible analytical tools (which mostly have GUIs) can mislead both the analyst and their audience, whereas command line interface can be difficult to access, despite being highly reproducible. It is important to mention here that R analyses can be extremely reproducible owing to how they are generated – with intricate details of the analytical framework embedded within the data handling, manipulation, and visualization. It is not the results of the analysis that can be replicated, but the analytical thought, the analytical ideology – which often, if designed well, is independent of context/setting. The classic dilemma between accessibility and reproducibility is not something that can be solved easily – according to me it is the model of pedagogy that needs a transformation, a model of thinking that needs alteration. Although I have no clear answer on how to balance the two equally important values of analytical tools, from my personal experience I can say that intentionality is key. If I was thinking about each individual data item in such sharp detail, a command line interface would be no different from a GUI. But we have not, and I have not and instead I have chosen accessible modes of analysis, over analytical nuance or even fit – which is unfair to the analytical process altogether. So, I believe, thinking long and hard about our data and analytical norms and challenging the core values that they hold, could be an important step towards marrying accessibility with reproducibility.

The most challenging thing about R is the way it makes us introspect about who we are as data analysts. I have had insightful (and sad) realizations about all the ways in which my analytical processes (which have been very rigorous by my standards) are sorely inadequate. The number of choices that I must keep making now in R reminds me of the number of choices I never made before. It has also made me acutely aware of the geography of information inequity – one that has been a cause I have vehemently advocated for, since the time I learnt to think for myself. How truth distorts, lies embedded within how we experience truth and how truth unfolds before our analytical minds. And yes, my “aha” moment in all of this existential mayhem has been the possibility of inching closer to that truth, to that comprehensiveness and neutrality which has always been my goal as an planning analyst.