Mtg Notes 2021-03-25

Christopher Desjardins, Lisa Lendway, Matthew Beckman

Author Order

I want to suggest alphabetical author order, and I feel like there's a running joke at USCOTS that Allan Rossman is always suggesting folks go alphabetical by **first** name.

Title:

- Reframing EDA to keep up with the times (Matt sort of likes this one)
- Closing the gap between real & perceived value of EDA
- Make EDA Great Again! (n.b. 90% chance of hats if we choose this, though it makes Matt uncomfortable that it would appear on his CV)

(not the title): Is Exploratory Data Analysis Dying When We Need it Most?

This 1-2-paragraph description should include a condensed version of what you provided in your proposal:

- Overview of topic
- Specific goals of session
- Indication of how session will be interactive

Session Description (250 words):

Exploratory Data Analysis (EDA) perhaps suffers a wider gap between real & perceived value than any contribution the data analyst offers. The popularity of Data Science has ushered in a rising tide of open-data initiatives and elegant software solutions readily accessible to non-specialists, thereby democratizing data analysis for all manner of subjects. Initially, some novices may feel overwhelmed and others may risk spurious results utilizing opaque methods with fancy names, yet all can be empowered by careful EDA based on lucid methods powered by curiosity, creativity, and critical thinking to build fruitful intuition and responsible insights. However, EDA too often seems a cursory obligation to be minimized.

The session will include several focus group style conversations in small groups reflecting on issues including the evolution of contemporary EDA, top priorities & best practices for a careful EDA to be reinforced with our students, and other topics.

As a result, the session aims to provoke a discussion that would (1) recast/affirm goals and considerations for EDA in a world laden with found data, big data, etc. (2) rethink where/how EDA should be addressed in data analysis courses at all levels.

Connection to Theme

In a very real sense, I would argue that EDA democratizes data analysis. The prerequisites creativity and critical thinking rather than fancy models with sophisticated underpinnings. Through careful attention to EDA, students can be equipped with powerful tools for learning from data from the first weeks of a first

course with very few technical obstacles. Open-data initiatives (e.g., data.gov) and remote hosting services (e.g., GitHub, Kaggle) have made all manner of data on an enormous range of subjects available to the public, and increasingly powerful software tools are designed to enable the non-specialist (as well as the specialist) to explore data to build intuition for sophisticated questions. In many cases, a thoughtful framework for responsible EDA may well differentiate whether the combination of rich open data and novice-friendly tools results in fruitful citizen (data) science or spurious conclusions that loosely appeal to the authority of "the data."