Technology to communicate data science output.

DATA612 - Recommender Systems

William Outcault

18 June 2020

> "Time after time it happens that some ignorant or presumptuous member of a committee or a board of directors will upset the carefully-thought-out plan of a man who knows the facts, simply because the man with the facts cannot present his facts readily enough to overcome the opposition. . . .As the cathedral is to its foundation so is an effective presentation of facts to the data." - Willard Brinton author of *Graphic Methods for Presenting Facts*

Overview

In 2017 Kaggle surveyed 7,000 data scientists asking them for "barriers they face at work", four out of seven top answeres involved miscommunication between their team and non-technical personel. These barriers are derived from a lack of tranparency between numerous sectors including; the data science teams, financial managers and decisionmakers. This struggle to convey your saught after results to decisionmakers (known as the last mile) is surprisingly difficult to overcome. Not only does it take a multitude of soft-skills to persuade your audience, but team's also lack in the technology needed to present the findings. This technology has been overlooked by the science's technical part and is clearly lagging behind. The following material will cover the soft-skills and technologies needed to successfully display your data science output to decisionmakers.

Tailor Your Information

Who is your audience? Your team has completed a tremendous amount of work but don't give up yet. Do your homework and put yourself in your audience's shoes. Learn the background of your audience prior to pouring out different metrics and complex graphs that you used for your output, this could come across as a different language depending on your audience. In other words your team managers, engineering teams, and steakholders will need tailored reports so the information is not overcomplicated or oversimplified for that specific audience. Data science enables cross-functional work so be aware of what insights you are presenting to whom and how it will influence those personel. Always take into account your audience.

To ensure audience engagement create a compelling narative. Ask yourself questions before communicating outputs:

- What is the purpose of your data analysis? Why you have interpreted this data in the first place?
- What is your message?
- What problem are you proposing to solve with your insights?
- What value does your findings add to your organisation?

Lastly, summarize your insights clear and consisely. Only communicate information that directly addresses the objective of your work.

Pick Your Software

The technology will vary depending on who you are communicating your outputs to. I broke these variations into two parts presentations and reports. You should also consider who will be receiving these outputs, for instance a markdown file might be ideal for your team manager meanwhile decisionmakers won't have the software to read them.

Presentations

Three popular tools to use for presentations are PowerPoint, Apple's Keynote, Google Slides, and LaTeX Beamer. Each serve benefits, for instance PowerPoint integrates conveniently with the other Microsoft Office suite and practically every computer will be able to view the files. If your company gives you the option then strongly consider Keynote. They are known for the ease-of-use and intuitive interface. Like Google Slides they also allow for collaboration online if it is setup to use iCloud Drive. Google Slides might be able to limited in details such as slide effects, however the collaboration feature is personally my favorite. LaTeX Beamer is popular for its programmability. You can write code to generate slides or plots which is definitely a significant factor when choosing software.

Reports:

Several popular reporting methods are LaTeX, Markdown, Jupyter notebooks and Google Documents. Reports will be more team dependent than presentations, specifically because of the reliance on reproducability. You will want to make sure that the report's file type will be consistent with the team requirements for obvious reasons. All of the methods listed above are powerful enough to create clear and concise reports, rather consider your familiarity with the software and of course the language you are programming in.

Conclusion

Communication requires both soft-skills and the correct software. Unfortunately we are lagging behind in communication software compared to the technical side of data science, however so long as you present your outputs in a file that is readable for your audience you can focus on the bigger scope of things. Research your audience, create a compelling narative, and choose any software consistent with that of your teams'.

Question

Personally I use Microsofts Office suite with my current employer. Let me know what you use, cheers:)

References

https://hbr.org/2019/01/data-science-and-the-art-of-persuasion

https://medium.com/comet-ml/a-data-scientists-guide-to-communicating-results-c79a5ef3e9f1

https://towardsdatascience.com/tell-stories-with-data-communication-in-data-science-5266f7671d7

https://www.quora.com/What-tools-do-data-scientists-use-for-professional-presentations-e-g- diagrams-graphs-flow-charts-etc