LaTeX

Info 98: Practical Data Science Skills for Internships

Data Science Society at Berkeley

datasciencedecal@gmail.com

September 10th, 2018

Berkeley
UNIVERSITY OF CALIFORNIA

First, some important logistical stuff

1.) Check to see that you are on bCourses. This is *critical.* The syllabus should be on there as well.
2.) Join our Piazza using this link:

https://piazza.com/berkeley/fall2018/info98

Attendance policy: You must submit a code to bCourses every lecture to ”prove” that you attended. The code will be written on the board at some point during class.

Why LaTeX?

Disclaimer:

LaTeX is actually not a critical skill for data science

So why the hell are we teaching this?

- It requires no prior coding experience.
- For the purposes of this course, we needed a coding language that everyone could use for the collaborative git project.
- It is very useful for writing papers that involve math or code.
- It can be used to display math elegantly in Jupyter notebooks.
- *This presentation was written entirely in LaTeX*

Introduction to Overleaf

LaTeXis actually quite annoying to learn on your own if you are used to just using Google Docs or Microsoft Word.

For the purposes of this demonstration, we will be easing you into the process by introducing you to the **Overleaf editor** and, later on, showing off some really cool templates.
You can follow along this presentation using this link:

https://v2.overleaf.com/read/zthjmxpnhpvp

Berkeley
UNIVERSITY OF CALIFORNIA

Readable Mathematics

Let $X_1, X_2, \ldots, X_n$ be a sequence of independent and identically distributed random variables with $\mathrm{E}[X_i] = \mu$ and $\mathrm{Var}[X_i] = \sigma^2 < \infty$, and let

$$S_n = \frac{X_1 + X_2 + \cdots + X_n}{n} = \frac{1}{n}\sum_{i}^{n} X_i$$

denote their mean. Then as n approaches infinity, the random variables $\sqrt{n}(S_n - \mu)$ converge in distribution to a normal $\mathcal{N}(0, \sigma^2)$.

Tables and Figures

- Use `tabular` for basic tables — see Table 1, for example.
- You can upload a figure (JPEG, PNG or PDF) using the files menu.
- To include it in your document, use the `includegraphics` command (see the comment below in the source code).

Item	Quantity
Widgets	42
Gadgets	13

Table 1: An example table.

Incorporating figures and pictures

This is the three way Venn diagram for ”Computer Science using Big Data,” ”Math and Statistics,”

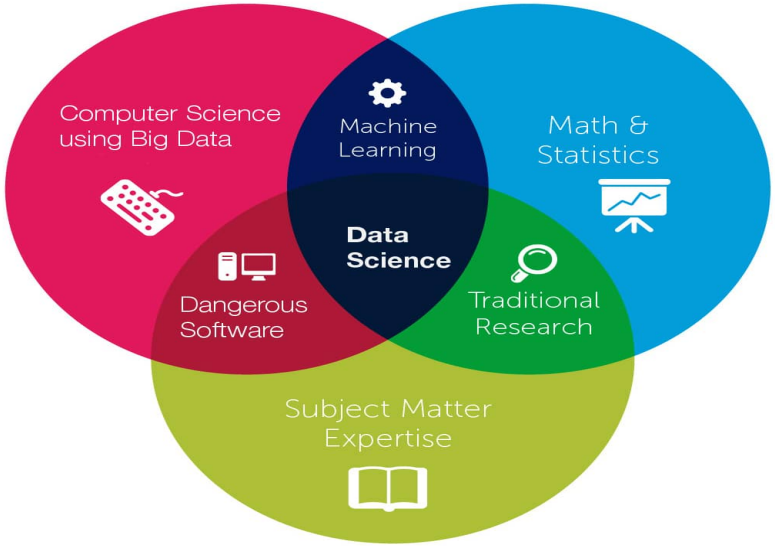

Figure 1: Data Science = CS + Math and Stats + Subject Matter Expertise.

Multiple columns

LaTeX can feel extremely clunky at times. But, if you know how to use it, it is extremely versatile.

- First bullet goes here
  - Secondary bullet goes here
    - Tertiary bullet goes here

Theorem

Theorem ("Best" Linear Predictor for 2 Variables)

$$L(Y|X) = E[Y] + \frac{cov(X,Y)}{var(X)}[X - E[X]]$$

Verbatim

Example (Code for Previous Slide)

```
\begin{frame}
\frametitle{Theorem}
\begin{theorem}["Best" Linear Predictor for 2 Variables]
\begin{center}
L(Y|X) = E[Y] + $\frac{cov(X,Y)}{var(X)}[X-E[X]]$
\end{center}
\end{theorem}
\end{frame}
```

INTERACTIVE DEMO TIME :D

Next Two Assignments

Update Your Résumé

This is a *solo assignment* due 9/16/18 11:59 PM. You **must** use LaTeX to create your résumé.

Design a DeCal Syllabus

This is a *group assignment* due 9/30/18 11:59 PM. Groups can only be in sizes of **3 or 4**.

Rubrics for both of these assignments can be found on the syllabus.

Additional Resources

- https://en.wikibooks.org/wiki/LaTeX/
- https://tex.stackexchange.com
- https://www.latex-tutorial.com/tutorials/
- https://texblog.org
- https://www.sharelatex.com/learn