



MASTERCLASS

DATA ANALYSTS: WHAT SKILLSETS DO THESE JOBS REQUIRE?

Dr Patricia Menéndez

Department of Econometrics and Business Statistics

Monash University

My story in data science

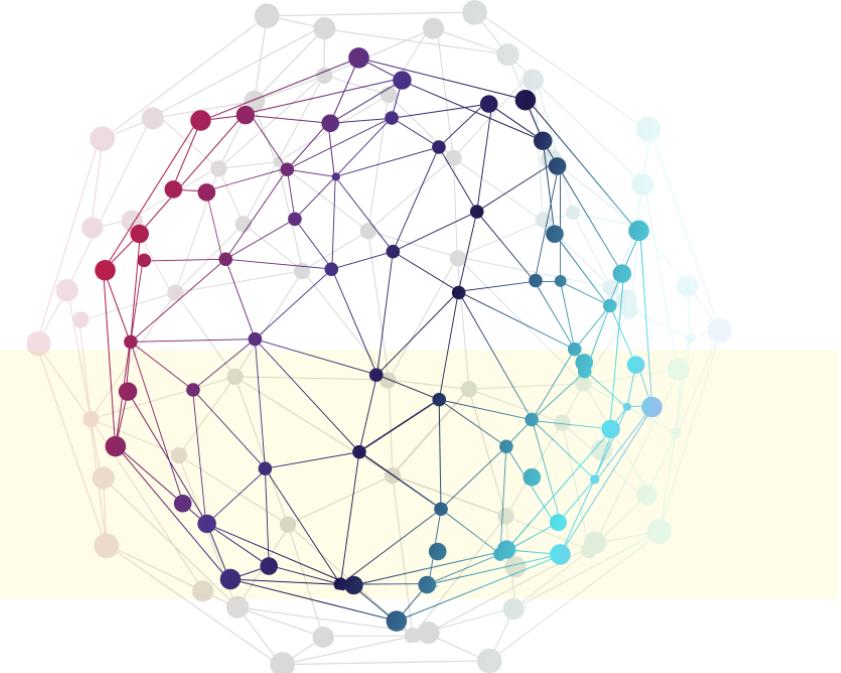


"4th Industrial revolution"

Data is driving the 4th Industrial Revolution

Top jobs

- Data Scientists 19% expected growth
- Statistician 33% expected growth



Source: <https://www.careercast.com/jobs-rated/best-jobs-of-2019>

Cartoon by Hiclipart.com

Data Analysts/Scientists?



My definition

*"Data analyst/scientist is a professional who analyzes, (sometimes collects) and interprets **data**, using statistical, mathematical and computational methods"*

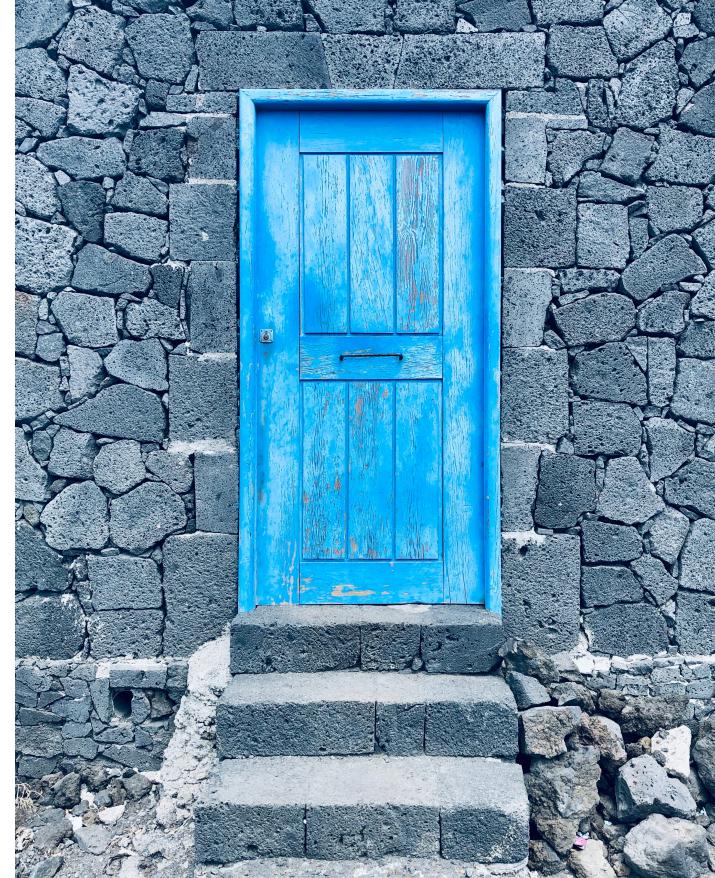
Patricia Menéndez



The multiple doors to data science



The multiple doors to data science Photo by [Dil](#) on [Unsplash](#)



Mathematics and Statistics
Photo by [Sies Kranen](#) on [Unsplash](#)

Strong quantitative/computational training



Photo by [Lachlan Donald](#) on [Unsplash](#)

The academic skills that I have found very useful

- Good mathematical and statistical training
- Algorithmics knowledge --> "Algorithmics is the systematic study of the design and analysis of algorithms" [Wikipedia definition](#)
- Strong computational skills --> different programming languages
- Problem solving skills
- The capacity to --> translate a real problem into a quantitative /data science one

Other skills that I have found very useful

- Set clear goals and expectations for a project
- Adapt to changes
- Being able to listen and to learn from different viewpoints
- Capacity to communicate your findings to specialized and broad audiences
- Embrace criticism and use it to get better!
- Professional integrity and high standards
- Ethics

Cartoons by Hiclipart.com



The learning never ends!



Photo by [Erwan Hesry](#) on [Unsplash](#)

Communicating to broad audiences: my first experience

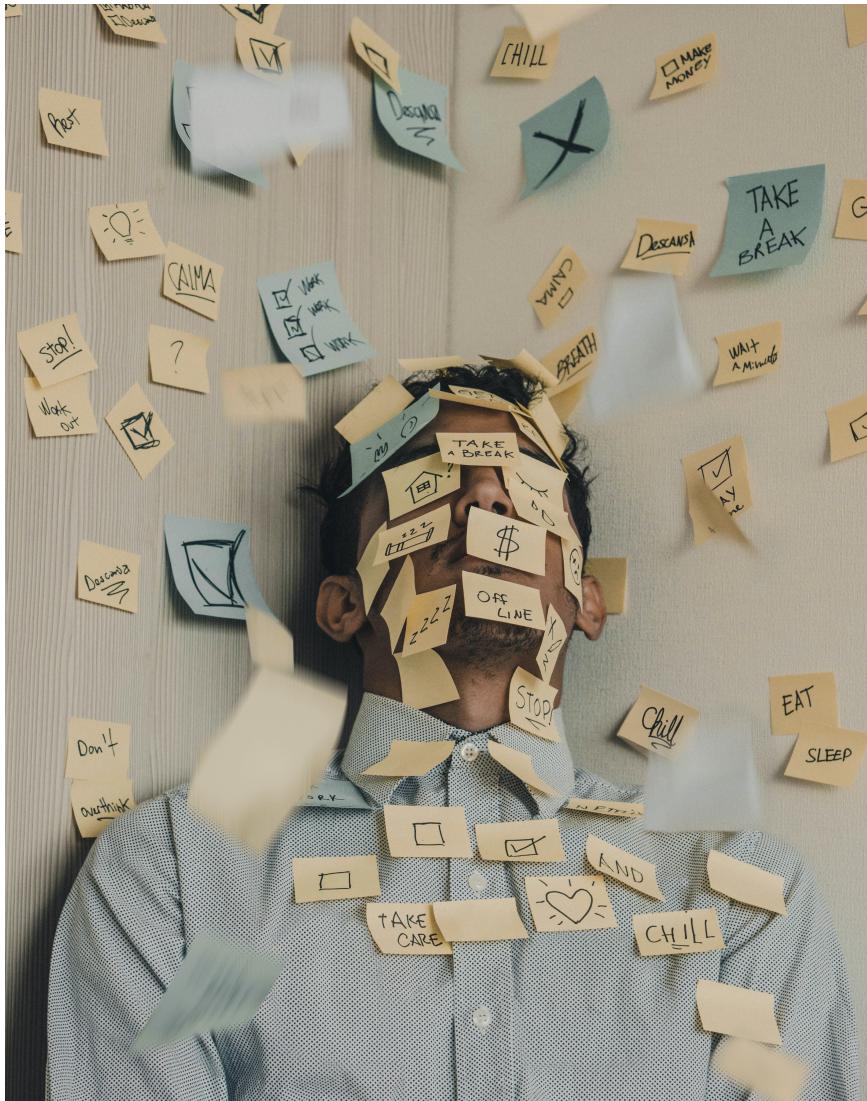


Photo by Luis Villasmil on [Unsplash](#)

The skills that I had to develop



Photo by [Dil](#) on [Unsplash](#)

- Theory and practice are two different things!
- Communicating with people in other disciplines is **not** easy and requires **lots** of practice --> **master it!**
- Ask questions!
- Set real expectations and deadlines
- Fresh air is always good :-)!

Companies and governments use data and want ...

Answers!

- Use your data analyst toolbox --> stats/math methods + computational implementation
- Create solutions and know what you are doing!

They want to be able to understand the answers:

- Specialized audiences --> know your methods well
- Broad audiences --> focus on the answers

Each data set has its own story

- Talk to the people who collect the data
- Their knowledge is invaluable
- Learn about the context
- Context is fundamental to the analysis & essential for the story telling of your project
- Know how to interact with databases
- Data wrangling
- Plot, plot, plot, plot your data!



Reproducibility and open source

- Reproducibility means obtaining consistent computational results using the same input data, computational step, methods, code and conditions of analysis.



Definition by the USA National Academies of Science, Engineering and Medicine

- Open source code: It is code where its authors make it available to others (under some open license) so that the code can be viewed, copy, inspect and alter by others.

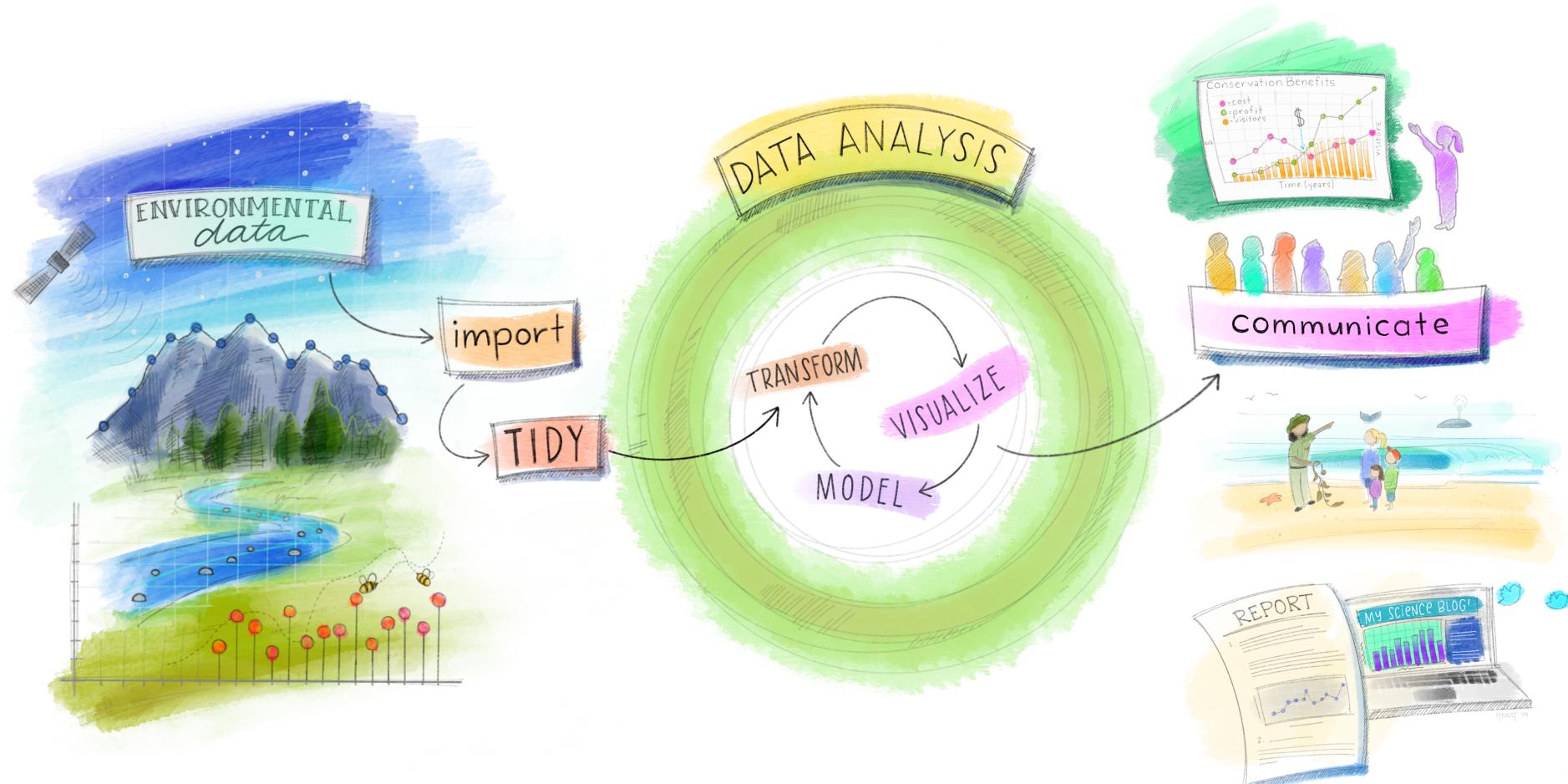


Cartoons by Hiclipart.com

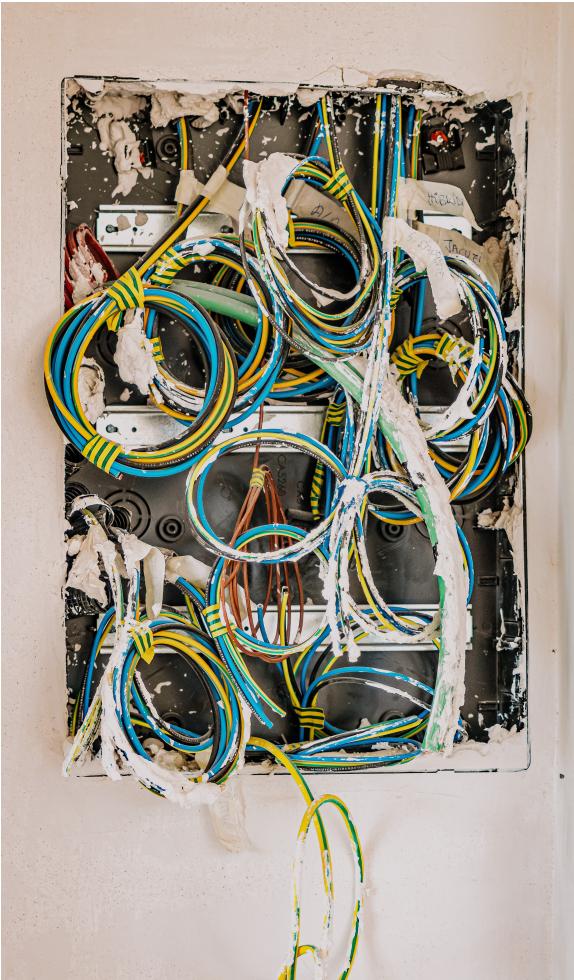
Version control and open source software examples



Reproducible workflows



Documentation is essential to reproducibility



- Data science projects can be very complex
- Large multidisciplinary team
- Each piece is part of the puzzle
- The instructions to connect the pieces are essential
- Documentation to ensure reproducibility and longevity of the project is critical!

The infinity of angles in data science



Photo by [Nicola Fioravanti](#) on [Unsplash](#)

Data science depth

- Data scientists come from different backgrounds
- Huge diversity in expertise, specialization and also in available jobs
- Know your strengths and also your limitations
- **Don't autopilot** --> Know what you are doing
- Keep learning!



Photo by David Holifield on Unsplash

Know how to --> but most importantly why to!



Thank you very much!



Photo by Alex Azabache on Unsplash

patricia.menendez@monash.edu
@PM_maths

<https://github.com/okayama1/Master-Class>

- Interested in collaborations with industry including contract research and training
- Multidisciplinary research

Feel free to reach out!