

# Student projects provided by employers

Craig McNeile<sup>1</sup>  
`craig.mcneile@plymouth.ac.uk`

<sup>1</sup>University of Plymouth, Centre for Mathematical Sciences

June 29, 2022

- Short overview of the workshop. <https://sites.google.com/view/employersandmathscurriculum/home>
- Some experiences with assessments inspired by business problems.
- Background to the panel session.

The workshop is supported by the IMA/RSS and in collaboration with ESF Smart Specialisation.



Workshop co-organized with Malgorzata Wojtys (Gosia).

- In case of fire alarm we will move outside to Portland Villas road to the right of the entrance.
- There are toilets at the back of this building, next to the vending machine.
- There is information about people trained in first aid on the walls at the entrance.
- Lunch is not provided, but I suggest we go to the Reservoir Cafe.

# Data science competition

In 2017 I ran a data science competition for the 2nd year mathematics students.

*In association with MATH2605 there will be a workshop developed by a Big Data company, based in Plymouth, called Iotec (<https://www.iotecglobal.com/>). The competition is voluntary and is not part of the assessment. It will be an excellent experience (and good for your CV), if you are interested in data analysis related to business applications and how programming and statistics are used in the real world.*

- Iotech was a local company who used data science to analyze web advertising data.
- A couple of staff members had to try to work with the company before, but the company had terminated the collaboration.

# Structure of the competition

- The session was run in one afternoon.
- The dataset was simulated advertising data in a csv file with 14 columns and 10,000 rows
- Two people came from the company to give a presentation and supervise the students.
- The statistical analysis didn't involve any machine learning.
- The company gave me Jupyter python notebook, which I recoded in Matlab.

The company thought this would help recruitment.

- It was a lot of work for me to convert the python code to Matlab.
- Only about 8 students from a class size of 50 turned up.
- There was no prize and the session was not part of assessment.
- The students that attended the session enjoyed it, but I am not sure there was any lasting benefit.

About a year after this event the company either left Plymouth or went bust.

# Project using data from BT research

- In 2019 I ran a group project using data from BT research.
- A graduate from Mathematical Sciences at Plymouth works at BT research <https://atadastral.co.uk/bt/>.
- There were two students in the group project.
- Midway through this project the covid pandemic lockdowns started.

We have continued the collaboration with BT research using data for MSc and PhD projects.

# Project using data from BT research

- The goal of the project was to build a machine learning model to find the factors that cause a customer to leave a broadband provider. (In the jargon this is known as “churn”.)
- When a company has a lot of customers it is better for profits to keep existing customers than to recruit new customers.



# Project using data from BT research

The original data set was open data from the regulator Ofcom  
[https://www.ofcom.org.uk/research-and-data/  
telecoms-research/broadband-research/  
broadband-speeds/home-broadband-performance-2018](https://www.ofcom.org.uk/research-and-data/telecoms-research/broadband-research/broadband-speeds/home-broadband-performance-2018)

- The data from the project were 5 csv files with over 50 columns.
- There are simpler data sets on kaggle for the same problem.

Although this was not a big data project, the csv files were typically much bigger than seen by students before.

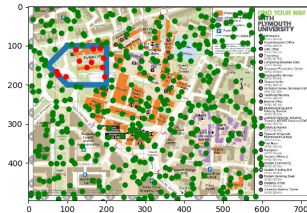
- The students did well with the visualizations of the data using Excel.
- It would have been good if the students could have talked with our contacts in BT research, but they were super busy keeping the broadband running when Covid struck.
- There was probably too much data cleaning required.

For example there are many different broadband packages offered by BT: "BT8", "BT20", "BT40".

For an initial analysis you might want to convert BT8, BT20, BT40 to BT. This is best done using "regular expressions", but although I use regular expressions as a tool in my research, I don't teach it.

## Geofencing project

- A local company wanted an investigation of geofencing. A trigger should happen if the GPS goes outside a specified area.
- Part of ESF Smart Specialisation project.
- The company wanted the code written in embedded c.



# Reflection on the geofencing project

- The students delivered the python code and a review of the mathematics, and some scaling of the different algorithms.
- The company was happy with the student's work.
- The students enjoyed the project. Particular a visit to the factory.
- The students did have a lot of communication with the staff at the company.
- It might have been better to include a UG computer science student in the team to do the embedded programming.

A little more work is required to convert the python into c-code.

# What did I learn from the 3 case studies?

- It is a lot of work to set up employer based projects, so I can only realistically supervise a couple of such projects year. It doesn't scale well with a large number of students.
- An assessment has to be at the correct level both academically and the amount of time required to complete the project.
- It can be hard to get data from employers, because the data is valuable to the employers ("value added.")
- There is always the risk of the employer not providing something.
- The bulk of our employer engagement is with statistics and data science projects.
- It could be that I need to build joint teams with engineers, maths students and computer science students.

Prof. Kurt Langfeld used to run projects such as, solving the heat equation on a model of human tissue.

# Should we do more industry based projects?

Over Easter I read: The Power of Explicit Teaching and Direct Instruction by Greg Ashman.

- Although it sounds very exciting to do “problem based learning,” (or the Moore Method) as soon as possible, it may be better to build up a student’s knowledge first.
- There is evidence that at first the students should work on more structured smaller problems, to reduce the cognitive load.

There has to be careful cooperation between industrial partners and the University. Some larger companies run “graduate programs” <https://www.lloydsbankinggrouptalent.com/graduates/why-lbg/training-and-development/>.

# Teaching programming is hard

Below is a lab exercise I set in a 2nd year module in 2015

*An arithmetic series has the  $n$ th term of the form where  $a_j = a_1 + (j-1)d$ ,  $a_1$  is the first term and  $d$  is the common difference. Write a function to sum up the first 1000 numbers in the series for any  $a_1$  and  $d$ .*

- Many students in the class had problems solving the above. (It didn't help that, at that time, we taught a mixture of R/Maple and Matlab.)
- I am not convinced that it is better to ask them to solve a more complicated industrial problem, which may be more interesting.

Now at the start of the module I give them a code with the lines in the wrong order, and they have to rearrange them (Parson problem).

# Designing a Curriculum is hard

The California Department of Education (CDE), Instructional Quality Commission (IQC), and State Board of Education (SBE) have commenced the revision process for the Mathematics Framework for California Public Schools: Kindergarten Through Grade Twelve (Mathematics Framework).

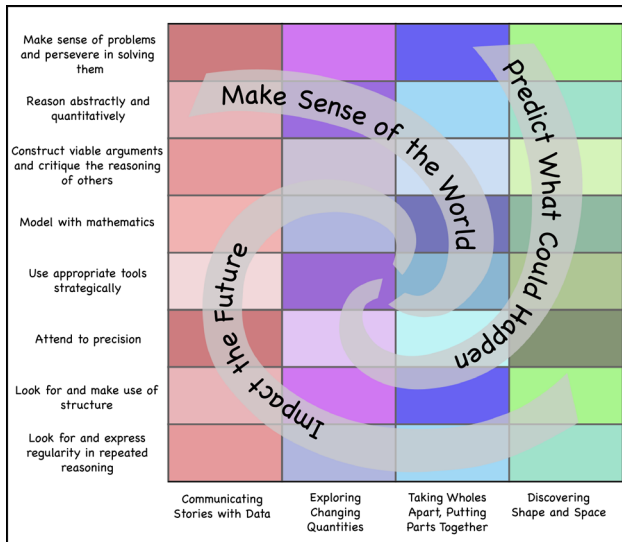
<https://www.cde.ca.gov/ci/ma/cf/>

*The ability to work with and understand data is an essential life skill in a world continuously inundated with data. Data drives students lives, whether they see them or not; making sense of data, being able to identify data that is misleading, and using data to make decisions are all important for their role as global citizens*



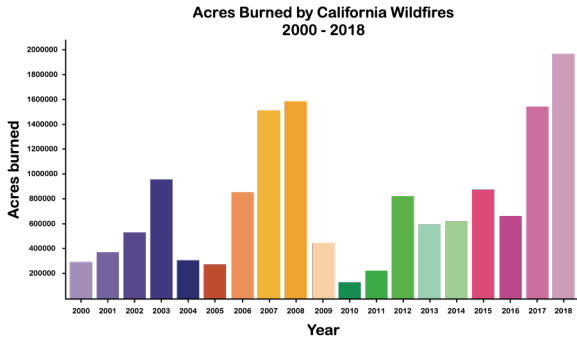
# Diagram explaining data science in California

<https://www.cde.ca.gov/ci/ma/cf/>



# Data set for California schools

I am not sure what the colours are for in the plot.



# Some suggestions for improved mathematics content

*Once children were birched at school. Now they are taught maths.*

The columnist Simon Jenkins in the Guardian newspaper is a good source for books on what mathematics should be taught, given his “interesting views” on mathematics.

**The Math(s) Fix: An Education Blueprint for the AI Age** by Conrad Wolfram. Conrad is the CEO of the company that makes Mathematica and Wolfram Alpha, so he is against hand calculation in favour of more computer based calculations of harder problems.

**The Math Myth: And Other STEM Delusions** by Andrew Hacker. One of his suggestions was “I asked my class to find 27% of a day.”

# Final year projects

Some of the titles of the final year projects this year:

- Cellular Automata
- Physics of rotating black holes.
- Solving differential equation with Neural Networks
- Algebraic solutions of hypergeometric equations
- Group theory and knots
- An analysis of increasing global supply chain dependency and methods of mitigation
- Phase Transitions in the Ising Model: from Monte Carlo Simulations to Automised Classification using Machine Learning

Most of the projects are academic.

Panel members: Tony Mann, Sarah Littler and Jo Byrne

- Q1 Do we need to move beyond asking employers for “skills” of graduate applicants to employers and University staff co-designing modules?
- Q2 How do we make sure students with academic interests in for example: pure maths or theoretical physics, have the skills that employers want.
- Q3 What is the role of the library/careers service in teaching employability skills and how do we get the students to engage?