

# EMILY M. TAYLOR PHD

Mobile: +49 15218253686 ▪ [emily.michie.taylor@gmail.com](mailto:emily.michie.taylor@gmail.com) ▪ [linkedin.com/in/emilytaylor203](https://www.linkedin.com/in/emilytaylor203) ▪ [github.com/ETay203](https://github.com/ETay203)

---

## DATA SCIENTIST | DATA ANALYST | PROJECT MANAGER

*Works in R and Python, with skills in statistical analysis, predictive modelling and data wrangling*

Data analyst professional who enjoys exploring and visualising data, finding patterns and insights. Highly numerate with great math and statistical skills. Strong analytical ability.

**Proactive problem solver** who is methodical and objective with a commitment to accuracy. Relentless at finding the solution to problems.

**Insatiable learner** unintimidated by new data types or packages. Enjoys the challenge of finding answers and gaining new knowledge, rapidly assimilating new and diverse topics. Continuously loves developing and broadening skills and experience.

**Data visualizer** with innate talent for finding the simplest description. Creates graphics from raw data that are both clear and informative. Enables fast and easy communication of technical information.

**Drives enhancements** by developing tools and reports that help users access and analyze data resulting in informed decision making and a better customer experience.

## PROGRAMMING AND DEVELOPMENT SKILLS

---

**Languages:** R, Python, Visual Basic for Applications (VBA), SQL, XPath XQuery and XSLT, Perl and LaTeX

**Tools:** R-Markdown, Knitr, git, Bash and Unix

## CONTINUED PROFESSIONAL DEVELOPMENT AND CERTIFICATIONS

---

Currently undertaking the John Hopkins University Data Science Certification composed of ten courses. Courses were assessed through peer-reviewed assignments (my project submissions can be viewed on [github.com/ETay203](https://github.com/ETay203)) case studies and online tests. Courses completed to date include:

The Data Scientist's Toolbox: Completed a conceptual introduction to the ideas behind turning data into actionable knowledge. Learned to use version control tools like GitHub with, markdown, git, R, and RStudio.

R Programming: Completed with 98% by writing my own R programs. Demonstrated my understanding of loops, data structures, a range of functions and lexical scoping as well as learning to identify bottlenecks and optimizing my code for faster running and more efficient programs.

Getting and Cleaning Data: Completed with 97%, demonstrating my ability to obtain and read data from a range of sources, formats and file types (web, APIs, databases using MySQL), transforming raw and messy data by using R to handle missing values, organize, merge and manage it in a reproducible way. Tidy or clean datasets are easy to manipulate, model and visualize saving time on data analysis tasks.

Exploratory Data Analysis: Completed with 95.3%. Mastered essential exploratory techniques for summarizing data, including clustering (Hierarchical and K means) and dimension reduction techniques as well as hands on experience of advanced graphing systems such as Lattice and ggplot2. These skills save time by quickly focusing in on what data is showing.

Reproducible Research: Completed with 97.2%. Interpretable results are the final step in a series of often complex processes. Can the results be reproduced by others? How did we get there? I used tools such as knitr to document two reports that included both my code and the raw data in a way that others can clearly follow my work flow, verify my findings and build upon them.

Statistical Inference: Completed the course with 98.3%. Showed my ability to draw conclusions about data populations in a two-part project. First demonstrated the central limit theorem through a simulation of the exponential distribution in R and calculating sample means and variance. Second analysed the effect of Vitamin

C supplements on tooth growth by applying my understanding of fundamental inferential data analysis techniques such as confidence intervals, hypothesis testing, and p-values, as well as permutation tests.

Regression Modelling: Completed with 92.6%. Applied my knowledge of regression to answer “Are Manual or Automatic transmission cars more fuel efficient?”. My report explored the relationship between a set of variables and miles per gallon, with least squares and inference using regression of linear and multivariable models.

Machine Learning: Completed with 97%. Using a dataset that recorded participants performing weight lifting exercises I built and tested a model to predict how well they did the exercise. The peer-reviewed report, including the R code used, describes how the prediction model was built, how cross validation was used, the expected out of sample error, and why various choices were made.

Accomplished the following additional courses, all with distinction;

- Data Analysis and Statistical Inference, Duke University, February 2014,
- An Introduction to Interactive Programming in Python, Rice University, June 2013
- Introduction to Databases, Stanford University, March 2013.

---

## EXPERIENCE

### **Consultant and Associate Consultant, Galson Sciences Ltd. (GSL), Oakham, UK** **2009 –2016**

GSL provides a range of nuclear decommissioning and radioactive waste management services. Independently managed and supported a range of projects for both domestic and overseas clients. Developed models, and coded programmes to support environmental impact assessments, for example:

- **Enabled client to quantify and specify** radioactive doses by writing code, testing, debugging, implementing and documenting a bespoke computational tool.
- **Demonstrated compliance** with nuclear safety regulations, presented complex data to non-technical audience as easy-to-interpret graphics and reports.
- **Established data integrity** ensuring government compliance and saving the client time by designing and developing a relational database to track a list of features, events and processes.
- **Optimized strategies** for monitoring and inspection by waste store operators by developing and coding a statistical analysis tool using VBA. The tool enabled different sites to plan and tailor their most efficient schedule for monitoring and inspecting waste canisters.

### **Graduate Student, Imperial College, London, UK** **2000 –2009**

- Utilized numerous statistical techniques, including Markov Chain Monte Carlo for predicting material structures.
- Optimized high volumes of simulations by running batch calculations across different processors in parallel, maximising speed and efficiency.
- Published work in peer reviewed papers
- Presented work publically at several international conferences.

---

## EDUCATION

|                    |  |
|--------------------|--|
| <b>2005 – 2009</b> | PhD, Department of Materials, Imperial College London. Focused on using high performance computing techniques and predictive modelling to better understand the structure and behaviour of materials. Supervisor Prof. R. W. Grimes. |
| <b>2000 – 2005</b> | MEng with Upper Second Class Honours, in Materials Science and Engineering, Department of Materials, Imperial College London, with one year placement at CENIM, Madrid, 2002-2003.   |
| <b>1992 – 1999</b> | A levels in Maths, Physics, and Chemistry, GCSEs in 9 subjects, South Hampstead High School for Girls, London NW3 5SS.   |

---

## LANGUAGES

English (Native), Spanish (fluent), French (basic), German (basic – recently completed Level A2.2).