

Marco Edward Gorelli

Data Scientist
pandas maintainer
University of Oxford MSc

Education

- 2016-2017** **MRes, Mathematics of Planet Earth**; 2016-2017; Imperial College & University of Reading; Distinction
- Courses: Computational Stochastic Processes (65), Partial Differential Equations (80), Dynamical Systems (84.2), Data and Uncertainty (63.1), Numerical Methods (67.13)
- Dissertation: Modelling the cloud and snow surface via KPZ equation (78.6)
- 2015-2016** **MSc, Mathematics and Foundations of Computer Science**; 2015-2016; University of Oxford
- Courses: Machine Learning (64), Categories Proofs and Processes (66), Quantum Computer Science (76), Networks (70), Computational Game Theory (70).
- Dissertation: Deductive verification of the s2n HMAC code (65)
- 2011-2015** **BSc, Mathematics with Professional Practice**; Brunel University London; First Class Honours
- Dissertation: The Tutte Polynomial and the Merino-Welsh Conjecture (91)

Open source

- pandas** Contributor since August 2019, maintainer since August 2020.
- Hosted pandas sprint at **PyData Festival Amsterdam**, where I mentored 15 attendees
- Added new method `DataFrame.to_markdown` (featured in the **v1.0.0 release notes**)
- Fixed issue **#27453** (featured in the **notable v1.1.0 bug fixes**)
- nbQA** I'm the authour of this code quality tool for Jupyter Notebooks.
- 100% test and documentation coverage + continuous integration.
- Source code: <https://github.com/nbQA-dev/nbQA>
- other** Minor code contributions to matplotlib, pytest, featuretools, and interrogate.

Experience

Data Scientist at Samsung R&D Institute UK; Nov18 - now

- Mentored fellow employees by co-organising workshops and reading groups on SHAP values, markov chains, bayesian methods, and `pre-commit`
- Improved accuracy/interpretability of hierarchical forecasting models by 5 percentage points using simpler models and more careful feature engineering
- Deployed end-to-end web-scraping dashboard and anomaly detection module for sensor data to web apps - included work on AWS, Google Cloud, CircleCI, MongoDB, PySpark
- Passed internal *Advanced Software Engineering Test (C++)* (£500 reward)

Data Scientist at Sedex; Jun18 - Nov18

Data Scientist at Sensium; Jan18 - May18

Maths tutor at Oxford Exclusif Tutorial Agency; Oct18 - Jan18

Risk Analyst Intern at General Electric Capital International; Jun13 - Jun14

Achievements

- Kaggle:
 - Competitions expert:
 - * top 8% in *Tweet Sentiment Extraction* : PyTorch RoBERTa NLP model
 - * top 8% in *M5 Forecasting - Uncertainty* : Keras quantile regression deep learning model
 - Discussions expert
- Coursera: “Deep Learning Specialization”, “Applied Data Science with Python Specialization”
- During BSc: Foster Award for “exceptional mathematical ability”, Level2 Award for highest grades

m.e.gorelli@gmail.com • <https://github.com/MarcoGorelli>