

Marco Edward Gorelli

Education

MRes, Mathematics of Planet Earth; 2016-2017; Imperial College & University of Reading

MSc, Mathematics and Foundations of Computer Science; 2015-2016; University of Oxford

BSc, Mathematics with Professional Practice; 2011-2015; Brunel University London

Open source (pandas)

- **Volunteer maintainer** since Jan 2020, have reviewed 60+ pull requests
- Added new method `DataFrame.to_markdown`, which was featured in the **v1.0.0 release notes**
- Fixed bugs in `groupby.apply`, `groupby.agg`, `groupby.nunique`, `explode`, `Categorical`, visualisation, and **more**

Experience

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

- Mentored fellow employees by co-organising internal machine learning reading group: presented on SHAP values, markov chains, and bayesian methods
- Improved accuracy/interpretability of hierarchical forecasting models and maintainability of source code (*Prophet*, *statsmodels*, *Jupyter Lab*, *CircleCI*, *pre-commit*, *Google Cloud Storage*, *wily*, *SHAP*)
- Deployed end-to-end web-scraping dashboard, including tests, automated documentation, static type-checking (*Google Cloud Platform*, *streamlit*, *scrapy*, *ScrapingHub*, *pymongo*)
- Deployed end-to-end anomaly detection module for sensor data (*pytest*, *sphinx*, *MongoDB*, *mongomock*, *PySpark*, *AWS EMR+S3*, *Docker*, *mypy*)
- Debugged Python and C++ code, allowing for Python ML model to be reliably deployed to wearable (*Tizen Studio*, *opencv*, *scikit-learn*, *pandas*)
- Passed internal “Advanced Software Engineering Test (C++)” (£500 reward)

Data Scientist at Sedex; Jun18 - Nov18

- Produced interactive dashboards and performed analytics on survey data (*Tableau*, *PostgreSQL*)

Data Scientist at Sensium; Jan18 - May18

- Rewrote internal visualisation/reporting tool from Matlab to Python (*PyQt*, *reportlab*, *matplotlib*)

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

- Reporting on data from credit risk agencies (*SAS*)

Other achievements

- During BSc: Foster Award for “exceptional mathematical ability”, Level2 Award for highest grades
- Kaggle: 2019 Data Science Bowl: top 11% (time series, tabular data, *LightGBM*); Google QUEST Q&A Labeling: top 13% (NLP, *tensorflow*, *keras*, *BERT*); University of Liverpool - Ion Switching: top 15% (signal processing, *WaveNet*)

- Coursera: “Deep Learning Specialization”, “Applied Data Science with Python Specialization”
- Holder of Tutoring Standards’ Professional Tutoring Certificate

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