

Candidate Information

Position:	Data Engineer
School/Department:	Mathematics and Physics
Reference:	20/108153
Closing Date:	Wednesday 25 March 2020
Salary:	£33,797 to £40,322 per annum
Anticipated Interview Date:	Tuesday 7 April 2020

JOB PURPOSE:

To play a key role in providing technical expertise and support in developing the Statistical Data Analytics core platforms. Working within the Mathematical Sciences Research Centre the role holder will be a point of reference for all data related technical issues and as such will be required to network more broadly across the University and with key external stakeholders.

MAJOR DUTIES:

1. Design, develop, maintain & support the data science processing infrastructure across a diverse range of technologies and platforms using Python, R, Java, C++ and/or other appropriate software languages.
2. Develop query tools to analyse large data-sets by using relevant concepts, sourcing, manipulating and storing data from a wide variety of data sources using SQL and Cloud based 'big data' technologies.
3. Contribute to the development of new and emerging concepts by maintaining own expertise through continuous personal development, championing new approaches to big data and striving for greater functionality in data systems.
4. To conduct data audits and cleansing maintaining calculating and recording test results; devises and run tests to ensure data quality.
5. Create data tools in support of pioneering Data Analytics Research, to facilitate data storage, data extraction; data verification, data manipulation and data analysis.
6. To provide guidance and expertise in support of student projects and academic research.
7. To input to the Data Analytics curriculum through the identification of emerging technologies and developing supporting teaching guides for both students and staff, imputing to assessment of relative modules and student projects.
8. Perform failure verification and troubleshoot across a range of data management tools, carrying out all routine equipment maintenance and managing the contract with service providers.
9. To contribute to the development of a Quality Documentation Manual procedures to run data analytics tools and handle all lab equipment.
10. Carry out routine administrative duties as requested, e.g. arranging GDPR compliance, maintaining research group website etc.
11. Contribute as required to the development of research concepts and projects, including managing project work streams under the direction of PI or CI's providing specialist knowledge in the preparation of grant funding applications and throughout the project life cycle.
12. Assist with project placement supervision for the MSc Data Analytics students and tutorials and deliver lab demonstrations for the MSc Data Analytics course.

Planning and Organising:

1. Plan own work over the short to medium term with an awareness of longer term issues, in response to manager's general instructions.
2. Organise own day to day activities, and meetings.
3. Ensure regular reporting and meeting with Professor Adele Marshall.
4. Be able to work using agile practises.
5. Be able to understand business and technical requirements, as well as to design, implement and document appropriate solutions.

Resource Management Responsibilities:

1. To maintain all equipment associated with the data analytics labs.

Internal and External Relationships:

1. Liaise with QUB colleagues and support staff on routine matters.
2. Work closely with internal and external data providers both locally and internationally.
3. Attend and contribute to relevant meetings.

ESSENTIAL CRITERIA:

1. 2:1 Honours Degree, or equivalent, in Computer Science, Electrical Engineering, or related discipline.
2. At least 3 years relevant experience (e.g. as a Data Technician or in a closely related role).
3. Experience of using and setting up cloud based infrastructure in platforms such as AWS, Azure or GCP.
4. Advanced working SQL knowledge and experience designing and working with relational databases, query authoring (SQL) as well as working familiarity with a variety of database engines.
5. Experience building and optimizing 'big data' pipelines, architectures and data sets.
6. Demonstrable experience of successfully manipulating, processing and extracting value from large disconnected datasets.
7. Experience developing RESTful APIs using frameworks such as Flask, Django Rest Framework, Spring Boot, PlumbR or similar.
8. Demonstrate experience working with Python and/or R and/or any of JVM (Java, Kotlin, Groovy, Scala) programming languages.
9. Ability to interact with research colleagues and support staff.
10. Strong project management and organizational skills.
11. Ability to communicate complex information clearly in both written and spoken English.
12. Ability to communicate with public sector organisations such as healthcare providers, and to industry.
13. A team player who is open-minded and is prepared to work closely with researchers and industry partners as needed.
14. Be able to work off site with collaborators when extracting / managing data.

DESIRABLE CRITERIA:

1. MEng, MSc or a Ph.D. in a relevant area.
2. Experience working with big data technologies: Hadoop, Spark, Kafka, etc.
3. Experience building processes supporting data transformation, data structures, metadata, dependency and workload management.
4. Experience building microservice architectures.
5. Experience developing dashboards using the frameworks such as Shiny, Dash or Google Data Studio.
6. Working knowledge of message queuing, real-time data stream processing and highly scalable 'big data' data stores.
7. Experience with containerisation / orchestration technologies such as Docker, or kubernetes.
8. Working knowledge of designing and implementing automated DevOps infrastructure that utilises best CI/CD practises.
9. Evidence of strong presentation skills and ability to prepare clear and concise presentation materials.
10. Experience of collaborative research or working in a team is desirable.