28 Madison Street Jeppestown PBO: 930 033 432

**Data Science (Junior) Job Specification**

**Abilities**

* Object-oriented coding using Python
* Data importing from a variety of sources (relational databases, JSON, REST APIs)
* Cleaning, manipulating and merging data using Python
* Visualising categorical and continuous data in Python and Tableau / Redash
* Summarising data
* Interpreting graphs and models
* Hypothesis testing
* Survey design
* Modelling data: basic regression, classification and clustering techniques
* Performing cross-validation and model diagnostics
* Creating and querying relational databases using Python and SQL
* Using version-control software, individually and as part of a team

**Python**

* Ability to create and debug functions
* Object-oriented programming
* Comfortable with scientific libraries such as numpy, pandas, statsmodels, scikit-learn, matplotlib, seaborn
* Familiar with Jupyter Notebooks

**SQL**

* Database, table and key creation
* Querying data
* Joining tables

**Linux (Ubuntu)**

* File and directory navigation and manipulation
* Basic bash scripting
* Setting user permissions and file ownership
* Standard input/out, stream redirection, piping and filters
* Installing, updating and upgrading packages

**HTML5**

* Ability to create a basic HTML page
* Ability to link pages in a website
* Create a responsive site that will suit all devices and keep content fixed
* Create an html page that is user friendly

**Methodologies and Professional Skills**

* Scrum and Agile
* Ability to break down tasks and use productivity tools such as Trello
* Presentation skills

**Tech Stack**

* Linux, Python, MySQL, Tableau / Redash, HTML5, CSS, Git

The successful Candidate will

* Have a National Certificate NQF level 5 in IT
* Be ready for three months’ work experience
* Have experience working in teams, producing and reviewing code for real-life products