Module:	Data Analytics
Assignment No:	2
Hand-Out Date:	Monday 4th November 2019
Hand-In Date:	Friday 6th December 2019 **Note:** *Late assignments will **NOT** be accepted.*
Lecturer:	Andrew Shields (andrew.shields@staff.ittralee.ie)

Requirements

Working individually, you are to engage in the analysis of a dataset using data analytics technologies or a combination of. For example, Python, Juypter Notebooks, Scikit-Learn, Pandas or other data analysis tools. This should demonstrate your ability to apply the skills developed in the area of data analysis technologies. Accompany this with a written description that reflects on your work; and relates to your experiences and the available literature.

You should follow a standard data analytics project life cycle discussed in class. KDD, SEMMA and CRISP-DM are all examples. These process can vary in the number of steps.

While candidates may NOT copy a tutorial, notebook or book chapter which they have not developed themselves you may reference existing tutorials and this is expected as part of you research into the topic once fully acknowledged and referenced. Candidates are required to demonstrate their ability to work under their own initiative.

The research paper title must be agreed in advance with your lecturer

- This assignment is much more about quality rather than quantity
- Research (find) and review several research papers at least four.
- The formulation of your paper's "theme" (title, thesis)
- Your approach **MUST** be "action" oriented (i.e. applied) but may contain "theory" oriented (concept, investigation, State of the Art review)

Deliverables:

Notebook and Files: Any practical input and outputs files from your data analysis should form part of the assignment. Jupyter Notebook, data files, and scripts result files etc. must be available on Azure Notebooks, or GitHub. Your Notebook should demonstrate the analysis and technologies developed/used as part of the assignment. (A copy of the notebook must be submitted to blackboard) talk to me in class if you want to use another repository not listed above.
Your notebook should have a section for each stage of data analytics life cycle. Plase reference the attached notebook template on Blackboard.

Each of the following sections must be accompanied with markdown content in the notebook file describing the results of you analysis and the experiments that you ran. (see attached notebook template on Blackboard)

Grading Criteria

- Active and engaged participation in the development of the deliverables.
- Configuration and setup of a Data Analytics environment using Python.
- Clear evidence of work on each stage of the analytics life cycle stages.
- Application of technologies and insights into their use in a broader context.
- Evidence of logically structured and clear demonstration of data analysis with evidence of significant amount of outside reading which is integrated appropriately into the reflection and analysis content.
- Clear findings and conclusions from the data
- Clear analysis and interpretation of the process and experience with evidence of managing own learning through illustrated with examples where appropriate.
- Clear application of the scientific method to data analysis.
- Working data files to accompany analysis.
- Referencing as per standard college style (Harvard, Anglia Ruskin)