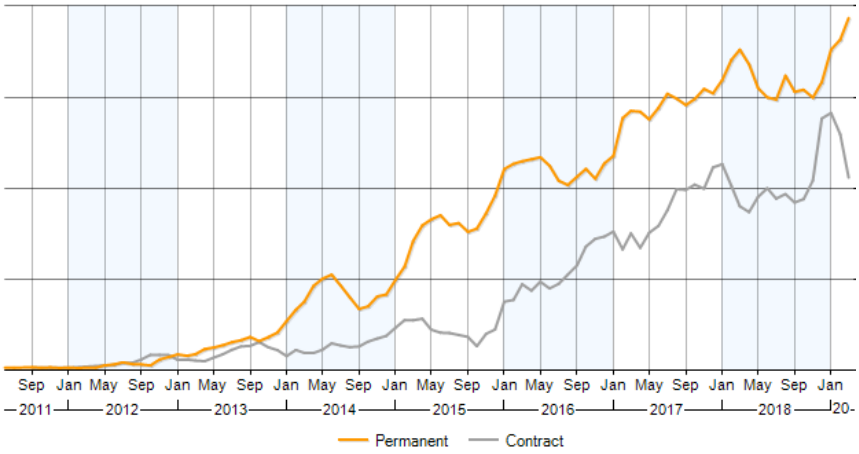
Solent University Unit Descriptor

Unit Code: COM618 **Unit title: Data Science**

Why is this unit important?
Data Science is a recent field which aims at making sense of data using data mining algorithms, machine learning tools and visualisation. "The amount of data we produce every day is truly mind-boggling. There are 2.5 quintillion bytes of data created each day" (Forbes, 2018). The study of this data area is growing rapidly as new technologies have enabled the creation, capture and analysis of this data.



Job postings for data scientist on itjobswatch.co.uk as of 2019.

What you will learn
Amongst other things you will learn about:

- Visualisations (Scatterplot, Boxplot, Distributions, Timeseries, maps)
- Data cleaning and preprocessing
- Using appropriate data mining tools with visual programming and interactive development environments
- Using Data Science programming libraries (such as Pandas, Scikit Learn and Keras)
- Unsupervised learning (kmeans clustering)
- Supervised learning (classification with SVM and Neural Networks)
- Reinforcement Learning (such as Deep Q learning and Markov decisional processes)
- Methods for recommender systems
- Prediction (Linear and polynomial regression, Neural Networks)

How you will learn
You will attend formal lectures to learn about the theory explaining the different algorithms you will be using. After the lecture you will experiment in hands-on sessions to apply these algorithms and techniques in order to answer questions using visual programming to create data flows or with programming. Practical sessions are also the place to have group discussions on topics related to data science.

How much time the unit requires
For a 20 credit unit, you are expected to study for 200 hours (which equates to 10 hours per credit). This total learning time is made up of contact time, directed learning tasks, independent study and assessment activity. Your tutor will offer you guidance on how you should best manage your study time on this unit

How you will be assessed
Tasks which help you learn and prepare you for summative tasks (formative):
Weekly exercises will prepare you for the unit assessment by using the algorithms seen in class on various datasets. Individual feedback in the class (formative) will help you to review and improve your final submission.

Tasks which count towards your degree (summative):
The assessment consists of a project in which you will be using the technologies seen in class on chosen datasets to solve real world problems. You will research the topic to find relevant literature and convey your understanding in your report. You will apply algorithms and create visualisations to support your findings. You will be given step by step tasks to develop your arguments.

When assessment does not go to plan:
If you are referred in this unit you will, subject to feedback, rework and resubmit their original submission, to a higher standard.

What you will be able to do after the unit:
1. Critically review data models, methods, techniques, statistics and tools applied in the modelling and analysis of data to meet enterprise operational and decision support requirements.
2. Autonomously analyse and evaluate enterprise data requirements, theoretical and applied data models and consequent data output.
3. Design data models to meet enterprise application requirements.
4. Choose and justify appropriate tools to model, implement, retrieve and analyse operational and decision-support data.
5. Articulate how legal and ethical issues concerning data impact data science applications.

How this relates to the dimensions of Solent's Real-world curriculum framework

Dimensions	How students learn	How students are assessed
Students are challenged to think in critical, creative and applied ways	Students are confronted with real-world problems and challenged to find solutions with data.	The assignment requires using technologies and methods creatively to solve problems.
Students are inspired to do research through inquiry, curiosity and problem-solving	Students search for research papers and discuss them in class	Students provide an annotated bibliography in their assignment
Students learn from authentic, engaging and programmatic assessment	The assignment enables students to use what they have seen in class on a motivating topic	Students are assessed on their capacity to create relevant visualisations and use machine learning and data mining methods

Summative assessment details

AE1	Weighting:	100%
	Assessment type:	Portfolio
	Aggregation:	N/A
	Length/duration:	2000 words plus artefacts
	Online submission:	Yes
	Grade marking:	Yes
	Anonymous marking:	No

Unit Author: Dr. Cédric Mesnage

Unit Title: Data Science			
Credit Points:	20	Unit Code:	COM618
FHEQ Level:		School/Service	SMAT
Unit Delivery Model:	CD	Max/Min student numbers	Not Applicable
Unit Leader:	Dr. Cédric Mesnage		
HECOS code	100751		

Unit change history:

Unit Approved/Year Implemented/Code	July 2019	2020/21	COM618
Unit modified/Year Implemented/Code			
Add extra rows as required			