IGNITE
YOUR
CAREER

IGNITION
contact centre

IGNITION GROUP

DATA SCIENCE
PROSPECTUS

POWERED BY:
EXPLORE

Table of Contents

IGNITION GROUP

POWERED BY:
EXPLORE

Ignition Overview

Ignition Group was founded in 2002, by brothers Sean and Donovan Bergsma. We are a globally recognised Customer Acquisition, Sales, Performance Marketing and Distribution company, focused on our People and our Partners.

WE ARE FUTURE FORWARD

As a business we are always looking for opportunities to improve and develop. By embracing the digital age and opportunities within the Artificial Intelligence, Machine Learning and Big Data space, we utilise next generation skills to improve our Business Intelligence and processes, making our business more adaptable and cutting edge in terms of performance and delivery.

WE ARE PEOPLE FOCUSED

We know our most valued asset is our People. That's why we recognise them, reward them and give them the opportunity to work in an environment that helps maintain a great work-life balance. We strive to develop our People's potential, offering key benefits that assist in helping them achieve, as well as continually providing them with opportunities to develop: personally and professionally.

WE BUILD PARTNERSHIPS

Not only do we employ the best in the business, we also work with the best in the business. We create Partnerships with both local and international Telecommunications, Home Services, and Financial Services brands and provide them with our solution-based services.

With the help of our People, we provide our Partners with highly skilled in-house teams in the following:

- Complete Customer Life-cycle Management
- Business Applications and Process Enablement
- Data Science and Machine Learning
- Brand and Digital marketing
- Lead Generation
- Sales

WE ARE #IGNITIONPROUD

IGNITION GROUP

Powered by possibility.

EXPLORE Overview

EXPLORE is a next generation Learning Institution that teaches students the skills of the future. From Data Science to Data Engineering to Machine Learning to Deep Learning we deliver cutting edge courses to satisfy your hunger to learn. Our Programmes are built by an amazing Faculty - we employ some of the world's most talented Scientists who have experience solving difficult problems on a global stage.

Our philosophy is to teach our students how to solve problems in the real world. We emphasise teamwork, collaboration and working within constraints, under pressure, with deadlines while understanding context, audience and implementation challenges. We are not a theoretical institution (although we cover the theory) - we are a 'practical, hands-on, roll-up-your-sleeves and get stuff done' kind of institution. As real-world Scientists who have delivered impact in the world of work we're well positioned to deliver these skills.

EXPLORE launched during 2013 and since then has taught 1,000's of students and solved many problems for businesses across multiple Industries across the world. We're reinventing education and invite you to join us to change things for the better.
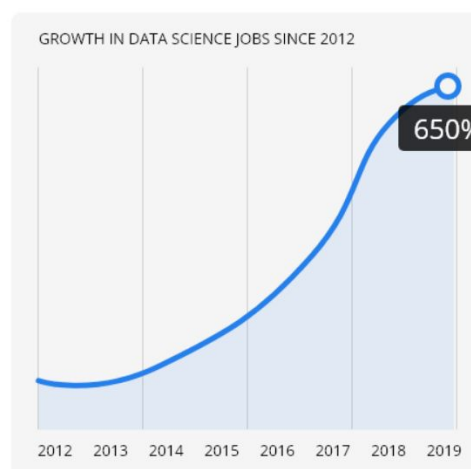
IGNITION GROUP

Why Data Science?

Four megatrends are fundamentally changing the shape of our world:

1. Vast amounts of data are being generated every minute.
2. The processing speed of our machines is increasing exponentially.
3. We now have cloud providers who can store insane amounts of data for a few dollars.
4. Powerful open source algorithms that can read, write, translate and see are now available to everyone.

Data Science is the skillset used to harness the power of these tectonic shifts in our world. With the Data Scientist's tool-kit almost any problem can now be solved.

It is no surprise that Data Scientists are one of the most highly paid and in demand skills of the 21st Century. In fact, the job 'Data Scientist' was rated the 'sexiest profession of the 21st Century' by the Harvard Business Review.

GROWTH IN DATA SCIENCE JOBS SINCE 2012

650%

2012 2013 2014 2015 2016 2017 2018 2019

IGNITION GROUP

Curriculum Overview

This course will provide students with the knowledge, skills and experience to get a job as a Data Scientist - which requires a mix of programming and statistical understanding. The course will teach students to gather data, visualise data, apply statistical analysis to answer questions and make their insights and information as actionable as possible. We cover the fundamentals of the Data Scientist toolkit as well as a broad set of machine learning algorithms.

Duration: 24 weeks   **Time Per Day:** 6 hours

Course Difficulty: ⬤⬤⬤ Advanced

Pre-Requisite Skills: Basic analytical background

Tools Learnt: python   jupyter   Microsoft SQL Server   Power BI   aws

Phase	Module	Weeks
Fundamentals	Data Science 101	2 weeks
	Analysing Data	4 weeks
	Gathering Data	3 weeks
	Visualising Data	2 weeks
	Communicating Insights	1 week
Machine Learning	Regression Techniques	4 weeks
	Classification Techniques	4 weeks
	Unsupervised Learning	4 weeks

Data Science 101

Learn about the process of solving problems with data

What is covered in Module 1:

Analyse
What is analysing?

➜ Where does analyse fit in the data science process?

➜ Tools we use to analyse data

Gather
What is gathering?

➜ Where does gather fit in the data science process?

➜ Tools we use to gather data

Visualise
What is visualising?

➜ Where does visualise fit in the data science process?

➜ Tools we use to visualise data

Communicate
What is communicating?

➜ Where does communicate fit in the data science process?

➜ Tools we use to communicate data

IGNITION GROUP

POWERED BY:
EXPLORE

Analysing Data

①②③④⑤⑥⑦⑧⑨⑩⑪⑫⑬⑭⑮⑯⑰⑱⑲⑳㉑㉒㉓㉔

Project: Create a report of the insights of the recent Cape Town water crisis

What is covered in Module 2:

Python

Programming Fundamentals
➔ Set-up environments - local notebooks & cloud
➔ Syntax, variables and expressions
➔ Conditional statements, and loops
➔ Building functions

Data Structures
➔ Python data types
➔ Lists, tuples, sets and dictionaries
➔ Using Numpy and Pandas libraries
➔ Store, retrieve and transform data

Statistics

Probability Theory
➔ Basics of probability theory
➔ Experiments, outcomes and sample space
➔ Calculating probability

Conditional Probability
➔ Independence in probability
➔ Events that are conditional on one another
➔ Calculating conditional probability

IGNITION GROUP

POWERED BY:
EXPLORE

Module 3
Gathering Data

Project: Set-up and extract valuable information from the TMDB
movie database

What is covered in Module 3:

SQL Server

Querying from the data
➔ Retrieving records from a database
➔ Specifying conditions for your queries
➔ Joining data across multiple sources
➔ Performing calculations on the data

Changing the data
➔ Add additional data to the database
➔ Making changes to existing data
➔ Making changes to the structure(schema) of the database

Statistics

Probability Theory
➔ What is a probability distribution?
➔ Discrete vs continuous distributions
➔ Normal distribution

Set Theory
➔ What is set theory?
➔ Venn diagrams
➔ Unions and intersections

IGNITION
GROUP

POWERED BY:
EXPLORE

Project: Build a dashboard to tell a story of team and player performance of an IPL cricket team

What is covered in Module 4:

Power BI

Setting up the data	➔ Loading data into Power BI
	➔ Linking datasets together with unique key
	➔ Create columns using DAX formulas
	➔ Create measures using DAX formulas

Visualising the data	➔ Create summary statistic cards
	➔ Create appropriate graphs
	➔ Using filters
	➔ Formatting visuals

Statistics

Sampling	➔ What is sampling?
	➔ t-distributions
	➔ Confidence intervals

Hypothesis Testing	➔ Null hypothesis vs alternative hypothesis
	➔ Significance levels
	➔ One-tailed & two-tailed tests

IGNITION GROUP

POWERED BY:
EXPLORE

Module 5

Communicating Insights

Project: Build and deliver a presentation

What is covered in Module 5:

Problem Identification

Defining a problem

➔ Problem identification - find out what is really needed
➔ Setting up the correct problem statement

Articulate solution

➔ Producing relevant solutions - answer the question asked
➔ Actionable insights - recommendations should be actionable

Effective Communication

Visual aids

➔ Structure of a slide deck
➔ Creating ghost decks to plan presentation
➔ What to include and not include on a slide

Presenting

➔ Body language when presenting
➔ Presenting principles - what to say and how to say it

IGNITION GROUP

POWERED BY:
EXPLORE

Module 6
Regression Techniques

Project: Using features of a house alone, predict the sale price

What is covered in Module 6:

Pre-processing and Model Building

Preparing data for modelling
- ➔ Basics of Scikit Learn library in Python
- ➔ Define features and labels
- ➔ Train/test split
- ➔ Encoding categorical features
- ➔ Standardizing features

Steps to build a model
- ➔ Overview of regression
- ➔ Linear regression models
- ➔ Train models and make predictions
- ➔ Regularization techniques - Ridge, Lasso
- ➔ Test model output - MSE, MAE, RMSE

Regression Algorithms

Algorithms for regression models
- ➔ K nearest neighbors
- ➔ Decision trees
- ➔ Random forests
- ➔ Support vector machines

Optimizing model performance
- ➔ Optimizing model performance by tuning hyperparameters
- ➔ Gridsearch techniques
- ➔ Which models to use when

IGNITION GROUP

POWERED BY:
EXPLORE

Module 7
Classification Techniques

Project: Predict if someone will take out a personal loan or not

What is covered in Module 7:

Classification Algorithms

Steps to build a model
- ➜ Overview of classification
- ➜ Logistic regression models
- ➜ Build and train models in python
- ➜ Make predictions using trained model
- ➜ Test model output - confusion matrix, classification report

Algorithms for classification models
- ➜ K nearest neighbors
- ➜ Decision trees
- ➜ Random forests
- ➜ Support vector machines

Dimensionality Reduction

Feature selection
- ➜ Determining significant features - linear and non-linear methods
- ➜ Variance Thresholds to drop features
- ➜ Correlations Thresholds

Features extraction
- ➜ Principal component analysis
- ➜ Linear discriminant analysis
- ➜ Interpreting extracted features and transforming them back

IGNITION GROUP

POWERED BY:
EXPLORE

Unsupervised Learning

Project:　Cluster customer reviews into themes

What is covered in Module 8:

Natural Language Processing

Overview of NLP
- ➜ Overview of NLTK library in python - for NLP
- ➜ Removing punctuation and symbols
- ➜ Tokenizing text
- ➜ Stopwords and removing them

Analysing text
- ➜ Lemmatization of words
- ➜ Bag of words
- ➜ Sentiment analysis

Clustering

Hard and Hierarchical Clustering
- ➜ What is clustering?
- ➜ Various forms of clustering - hard, hierarchical, soft
- ➜ K-Means Clustering
- ➜ Hierarchical Clustering

Soft Clustering
- ➜ Gaussian Mixture Models
- ➜ LDE clustering of text
- ➜ Labelling data using cluster output

IGNITION GROUP

EXPLORE Philosophy:
Solving problems in the real world

At EXPLORE we focus on building our student's ability to solve problems in the real world. Building things that work and make a difference is hard - that's what we teach.

We're not a traditional learning institution that spends weeks teaching matrix multiplication on a whiteboard (although understanding that is useful) - we're a practical, solution-oriented institution that teaches our students to work in teams, under pressure, with deadlines while understanding context, constraints and the audience.

Our courses are typically broken into Sprints where we teach a core set of concepts within the framework of solving a problem in a team with a tight deadline.

Students cycle from Sprint to Sprint solving different problems in different teams as they build this core muscle over the course.

Contact Information

For any **admissions related enquiries** mail us on:

admission@explore-ai.net

For any **general enquiries** mail us on:

general@explore-ai.net

Visit: www.explore-datascience.net

IGNITION GROUP

POWERED BY:

EXPLORE