DATA SCIENCE

Your Career Pathway

```
                    Specialist              Management
                    (S6,S7)                 (S6,S7)


        Principal Data      Principal Data      Principal Data      Data Architect
        Scientist           Analyst             Engineer            (S4)
        (S5,S6)             (S5,S6)             (S5,S6)

        Sr. Data Scientist  Sr. Data Analyst    Sr. Data Engineer
        (S4,S5)             (S4,S5)             (S4,S5)

        Jr. Data Scientist  Data Analyst        Data Engineer
        (S3)                (S3)                (S3)
```

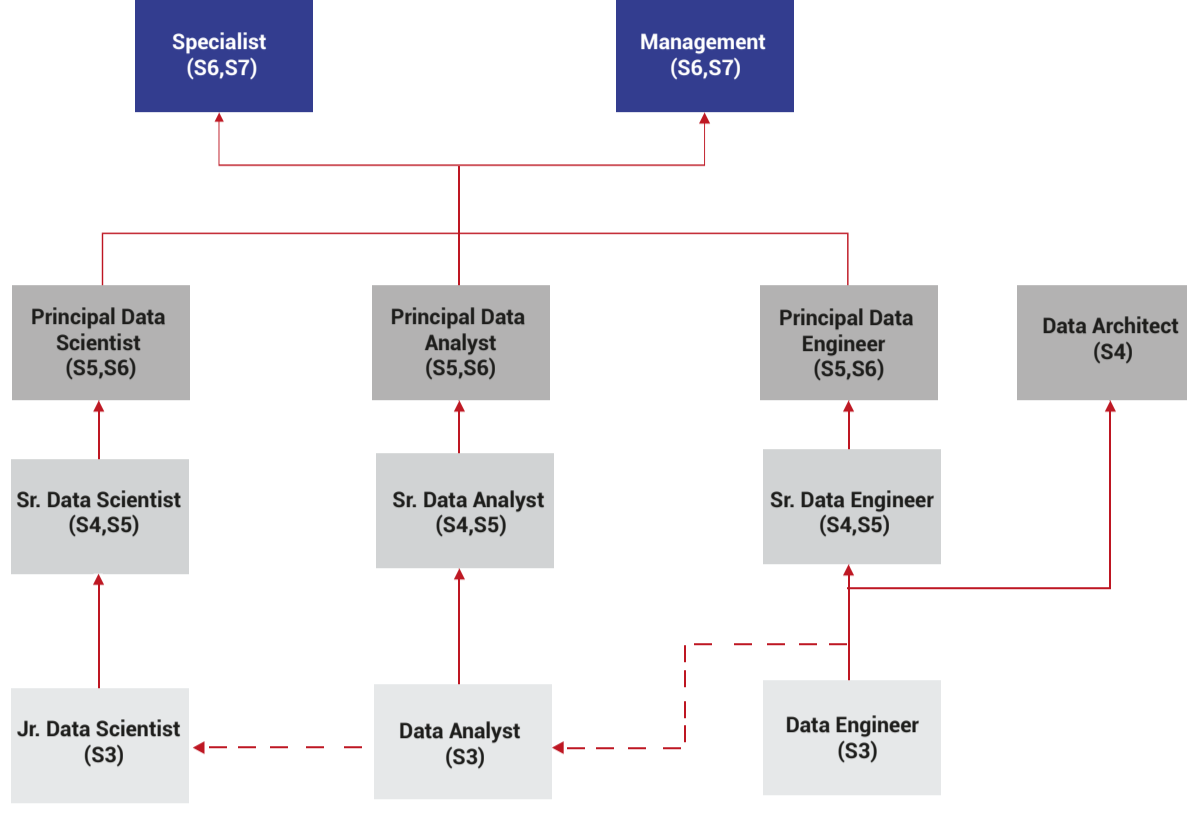**LEGEND (LEVELS OF RESPONSIBILITY IN SFIA)‡ :**
S1 - FOLLOW
S2 - ASSIST
S3 - APPLY
S4 - ENABLE
S5 - ENSURE, ADVISE
S6 - INITIATE, INFLUENCE
S7 - SET STRATEGY, INSPIRE, MOBILISE

SFIA stands for Skills Framework for the Information Age. It is a model for describing and managing skills and competencies for professionals working in the field of Information and Communication Technologies (ICT), software engineering and digital transformation. Published in 2000 by the British Computer Society (BCS).

Click here to find out more

DATA SCIENTIST

Job Description

- Use analytical techniques combined with data skills to develop scalable and robust analytical models.
- Analyses data through application of scientific methods and data-discovery tools which integrates and prepares large and varied datasets, and models complex business problems, business insights and identifies opportunities through the use of statistical, algorithmic, mining and visualisation techniques.
- It is a combination of analytic, machine learning, data mining and statistical skills as well as experience with algorithms and coding to solve business problem.

Competencies

- Knowledge of hypothesis forming and testing
- Curation of relevant data using visualisations that assists stakeholders' understanding
- Explore and analyse data for insights using visualisation
- Explain insights with clear and compelling written, verbal and visual communication
- Extraction of knowledge and insights generation
- Map data sources to data visualisation libraries
- Ability to write code to read data, access packages, apply logic
- Debugging, profiling and optimization
- Ability to clean data through statistical approaches, such as identifying outliers
- Ability to transform data into machine-readable formats
- Create, read, update and delete on databases and apply data normalisation
- Deep knowledge of statistical and mathematical concepts Identify trends and behaviours with descriptive statistics
- Knowledge of common models for prediction such as linear and logistic regression
- Ability to research latest methods to improve the accuracy and results
- Knowledge of big data tools and platforms to access data and run models
- Ability to manage the deployment of model lifecycles
- Identify, monitor, and measure quality of models over time
- Data Governance

Common Certifications*

- Microsoft Certified Azure Data Scientist Associate
- Microsoft Professional Program in Data Science
- SAS Certified Big Data Professional
- SAS Certified Data Scientist
- Huawei Certified Information Communication Technology Associate (HCIA-Big Data)
- Cloudera Certified Associate (CCA) Spark and Hadoop Developer
- Amazon Web Services (AWS) Certified Big Data – Specialty
- AWS Certified Machine Learning - Specialty
- IBM Watson IoT Data Science Certificate
- Dell Technologies Data Scientist Associate (DCA-DS)
- Dell Technologies Data Scientist Advance Analytics Specialist (DCS-DS)
- EMC Data Science and Big Data Analytics Certification
- eCornell Machine Learning Certificate
- Certified Analytics Professional (CAP)
- Coursera Data Science Specialisation by Johns Hopkins University
- DataCamp Data Science Career Track
- Senior Data Scientist (SDS™)
- Principal Data Scientist (PDS™)

Essential / Core Skills

- Data Design and Strategy
- Business Needs Analysis
- Data Governance
- Analytics and Computational Modelling
- Emerging Technology Synthesis
- Solution Architecture
- Data Visualisation
- Programming
- Stakeholder Management

Commonly Used Tools / New Technologies

- Python
- R
- Keras
- Pandas
- TensorFlow
- Scikit-Learn
- NumPy
- Machine Learning & AI
- Database Knowledge (eg. SQL, NoSQL)
- Cloud Technologies (eg. MS Azure, AWS, Google Cloud)
- Visualisation tools (eg. Tableau, Qlik, Power BI, ggplot, d3.js)
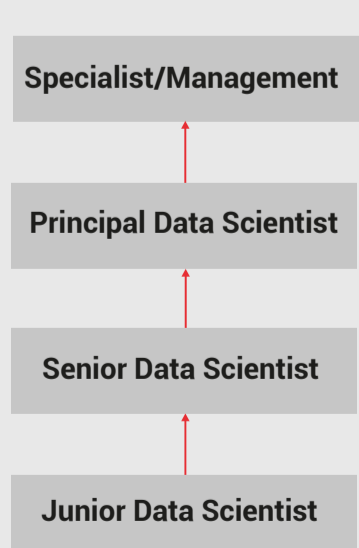
Soft Skills

- Problem solving
- Critical and analytical thinking
- Creativity
- Strong Communication
- Self-starting motivation
- Teamwork
- Pro-active
- Data collaboration
- Agile and fast learner
- Collaborative skills
- Intellectual curiosity
- Project Management
- Business Acumen

*This list is not exhaustive and serves only as a guide

Career Pathway

```
        Specialist/Management

        Principal Data Scientist

        Senior Data Scientist

        Junior Data Scientist
```

	SFIA Level[1]	NOSS[2]
Specialist/Management	6 & 7	
Principal Data Scientist	5 & 6	
Senior Data Scientist	4 & 5	
Junior Data Scientist	3	

[1] SFIA Level:
SFIA stands for Skills Framework for the Information Age. It is a model for describing and managing skills and competencies for professionals working in the field of Information and Communication Technologies (ICT), software engineering and digital transformation. Published in 2000 by the British Computer Society (BCS).

[2] NOSS:
The acronym NOSS stands for National Occupational Skills Standards. NOSS is a document that outlines the dexterity required of an employee working in Malaysia at a certain level of employment to achieve specific skills.

CONNECT WITH

DATA ENGINEER

Job Description

- Supports the design, implementation and maintenance of data flow channels and data processing systems that support the collection, storage, batch and real-time processing, and analysis of information in a scalable, repeatable and secure manner.
- Focuses on defining optimal solutions to data collection, processing, and warehousing.
- Designs, codes and tests data systems and works on implementing those into the internal infrastructure.
- Design and develop high-performance infrastructure and tools to enable users to consume and understand data more effectively. (typically by data/systems architect)

Competencies

- Ability to work with stakeholders to idenitify data sources
- Understanding of where data is stored, secured, access right, etc.
- Understanding of how to do sizing and backend design
- Debugging, profiling and optimisation for code related to Extract, Transform, Load (ETL) processes
- Ability to create various reporting/analytics data marts
- Knowledge of Parallel Programming Techniques
- Create, read, update, delete on databases and apply data normalisation and manage database schema
- Ability to collect data from variety of sources including APIs and web
- Ability to manage data and work with data sources like data streaming, real-time data, etc.
- Ability to make data easily accessible for analysis
- Ability to set up the entire ecosystem of big data and fast data
- Data Governance

Common Certifications*

- Microsoft Certified Azure Data Engineer Associate
- Google Professional Data Engineer
- IBM Certified Data Architect
- IBM Certified Data Engineer
- MapR Certified Hadoop Developer
- MapR Certified Spark Developer
- Data Science Council of America (DASCA)[1] Associate Big Data Engineer
- Cloudera Certified Professional (CCP): Data Engineer
- HDP Apache Spark Developer
- HDP Certified Developer Big Data Hadoop
- Hortonworks Certified Associate (HCA)
- Hortonworks HDP Certified Developer
- MongoDB Certified Developer Associate
- DataCamp Python Programmer Career Track
- Associate Big Data Engineer (ABDE™)
- Senior Big Data Engineer (SBDE™)

Career Pathway
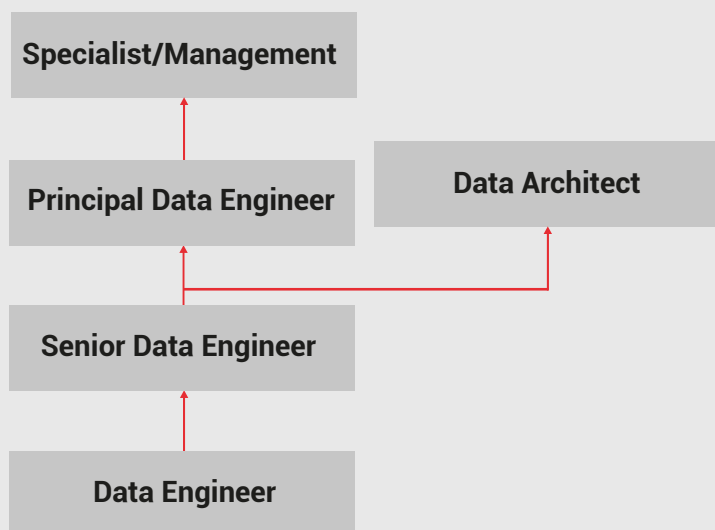
Essential / Core Skills

- Analytics and Computational Modelling
- Programming
- Data Design and Migration
- Data Analysis
- System Integration
- Database Management
- Emerging Technology Synthesis
- Business Intelligence (BI)

Commonly Used Tools / New Technologies

- Python
- Java
- Database Knowledge (eg. SQL, NoSQL)
- Pipelines
- Docker Products
- Cloud Technologies (eg. MS Azure, AWS, Google Cloud)
- Data Warehousing (eg. Hadoop, MapReduce, HIVE, PIG, Apache Spark, Kafka)
- Basic Machine Learning familiarity

Soft Skills

- Problem solving
- Critical and analytical thinking
- Creativity
- Strong Communication
- Self-starting motivation
- Teamwork
- Pro-active
- Data collaboration
- Agile and fast learner
- Collaborative skills
- Intellectual curiosity

*This list is not exhaustive and serves only as a guide

[1] DASCA is the World's Foremost Standards & Credentialing Body for the Data Science Profession.

	SFIA Level[1]	NOSS[2]
Specialist/Management	7	
Principal Data Engineer	5 & 6	
Data Architect	4	
Senior Data Engineer	4 & 5	
Data Engineer	3	

[1] SFIA Level:
SFIA stands for Skills Framework for the Information Age. It is a model for describing and managing skills and competencies for professionals working in the field of Information and Communication Technologies (ICT), software engineering and digital transformation. Published in 2000 by the British Computer Society (BCS).

[2] NOSS:
The acronym NOSS stands for National Occupational Skills Standards. NOSS is a document that outlines the dexterity required of an employee working in Malaysia at a certain level of employment to achieve specific skills.

DATA ANALYST

Job Description

- Communicate insights that deliver business value based on exploratory analysis. They blend historical data from available industry reports, public information, field reports or purchased sources and performs analysis to support business and product decisions.
- Develop projection on the outcome of implementing certain business strategies that result in actionable insights.

Competencies

- Ability to curate data relevant to business outcomes
- Ability to present a data science project to stakeholders
- Ability to identify areas of business impact
- Explain insights with clear and compelling written, verbal and visual communication
- Ability to map data sources to data visualisation libraries
- Knowledge of database concepts
- Identify trends and behaviours with descriptive statistics
- Apply inferential statistics on randomised samples to understand characteristics of a population
- Knowledge of hypothesis testing (A/B Test, etc.)
- Knowledge of outlier detection using statistical methods
- Ability to normalise a shared database
- Data Governance

Common Certifications*

- Microsoft Certified Data Management and Analytics
- Microsoft Certified Power BI Reporting
- SAS Certified Advanced Analytics Professional
- eCornell Data Analytics Certificate
- MapR Certified Data Analyst
- RapidMiner Analyst Certificate
- Cloudera Certified Associate Data Analyst
- Oracle Business Intelligence Certification
- Certified Analytics Professional (CAP)
- Coursera Business Analytics Specialisation by University of Pennsylvania
- DataCamp Data Analytics Career Track
- Associate Big Data Analyst (ABDA™)
- Senior Big Data Analyst (SBDA™)

Career Pathway

Specialist/Management

↑

Principal Data Analyst

↑

Senior Data Analyst

↑
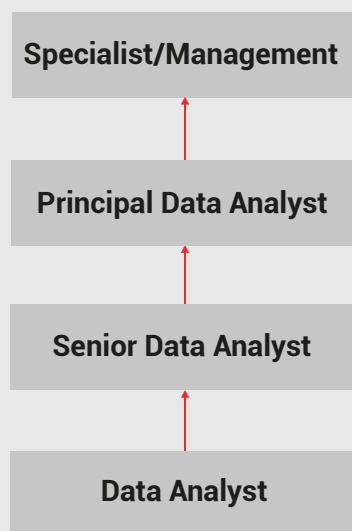
Data Analyst

Essential / Core Skills

- SQL
- Programming
- Data Analysis
- Data Visualisation
- Database Management
- Analytics
- Business Intelligence (BI)
- Business Analysis

Commonly Used Tools / New Technologies

- Python
- R
- Database Knowledge (eg. SQL, NoSQL)
- Visualisation tools (eg. Tableau, Qlik, Power BI, ggplot, d3.js)
- Google Analytics

Soft Skills

- Problem solving
- Critical and analytical thinking
- Creativity
- Strong Communication
- Self-starting motivation
- Teamwork
- Pro-active
- Data collaboration
- Agile and fast learner
- Collaborative skills
- Intellectual curiosity

*This list is not exhaustive and serves only as a guide

	SFIA Level [1]	NOSS [2]
Specialist/Management	7	Level 5, J631 Information Service Activities - Data Processing
Principal Data Analyst	5 & 6	
Senior Data Analyst	4 & 5	
Data Analyst	3	

[1] SFIA Level:
SFIA stands for Skills Framework for the Information Age. It is a model for describing and managing skills and competencies for professionals working in the field of Information and Communication Technologies (ICT), software engineering and digital transformation. Published in 2000 by the British Computer Society (BCS).

[2] NOSS:
The acronym NOSS stands for National Occupational Skills Standards. NOSS is a document that outlines the dexterity required of an employee working in Malaysia at a certain level of employment to achieve specific skills.