

## PERSONAL INFO

### Email

heonjaeshintech@gmail.com

### Phone Number

+61 403 015 104

### Address

6C Brockhoff Drive VIC 3125

### Linkedin

<https://www.linkedin.com/in/heonjae->

## KEY SKILLS

- Python
- Spark (Databricks)
- Cloud Computing
- AWS Solution Architecture
- Big Data handling (Numpy, Pandas, xarray, Dask, matplotlib)
- CloverDX (ETL)
- SQL (MS SQL, Oracle, MySQL)
- Unix/Linux Shell Scripting
- Power BI
- Machine Learning (sckit-learn, SciPy, AutoML, Pytorch)
- Git collaboration
- Stakeholder Management
- Agile Methodologies

## EDUCATION

### BSc of Science

#### Majoring Data Science

The University of Melbourne

WAM: 78.722

Mar 2019 – Dec 2021

#### Key Subjects:

- Machine Learning
- Applied Data Science
- Database Systems
- Probability
- Statistics
- Linear Statistical Model
- Modern Applied Statistics

**CSL Project:** Applied data quality metrics, machine learning, and built dashboards on medical research data

**NYC Taxi Project:** Visualised, evaluated and predicted performance of NYC satellite taxi data

#### John Monash Science School

ATAR: 98.15

Dec 2018

## Heonjae Shin

## Data Scientist

Technically driven data scientist with over two years of industry experience in the development of ETL processes for end to end data migration projects. Specialises in Python and SQL programming to handle big data analytics, visualisation and machine learning. I am a quick learner in adopting new technology and actively work to deepen my knowledge in programming.

## EXPERIENCE

### Powered Data Consultant

KPMG Australia (FDM contractor)

Dec 2021 – Present

Enabled the extraction, transformation and load process of data migration projects

**Main clients:** ASX, Department of Defence

- Orchestrated the build and debugging of transformation processes for big data migration projects for **SQL databases** and **SAP ERP systems**
- Managed face-to-face delivery of **stakeholder engagement** across progress of ETL milestones on daily basis
- Automated the generation of 100+ SQL metadata and queries, shortening duration of build activities and minimising errors by creatively implementing **Python** programming into ETL process
- Constructed the data reconciliation (verification) processes for pre-load and post-load phases between source, load and target data utilising **Java** based ETL tool **CloverDX**
- Designed **Power BI** dashboards for clients to seamlessly navigate results of data migration process
- Created an effective **SQL Database** structure to support scalability and effectiveness of ETL processes

### Machine Learning Modelling Intern

Bureau of Meteorology

July 2021 – Nov 2021

Tested and evaluated performance of numerous machine learning models to enhance prediction of lightening forecasting in Australia

- Remotely controlled NCI supercomputer with **UNIX** to write and execute **PBS scripts** for parallel computing of big data
- Generated models and predictions for lightening forecasting for large climate data using libraries such as **xarray**, **numpy**
- Tested performance of variations of training data with methods such as transformations, feature engineering and feature selection through libraries such as **SciPy/sckit-learn**
- Evaluated and visualised (**matplotlib**) performance of various machine learning models with **AutoML** to identify most accurate and applicable model for lightening forecasting
- Collaborated with senior scientists and developers using **Git**

## CERTIFICATIONS

### AWS

[AWS Certified Solutions Architect, Associate | Amazon Web Services | June 2023](#)

### Databricks

[Academy Accreditation – Databricks Lakehouse Fundamentals](#)

### FDM Academy Training:

#### Data Engineering Certificate

International Tower, Sydney NSW

Dec 2021 – Feb 2022

- Advanced Python
- Data Analytics with Hive and Spark
- Python foundations
- Distributed Data Storage
- SQL
- UNIX
- Java

## INTERESTS

- Mixed Martial Arts
- Brazilian Jiu Jitsu

### Workshop: Assistant Python Programming

#### Tutor/Lecturer

Korean Academy of Scientists and Engineers in Australasia

3 days - Dec 2021

Lectured segments of the Python big data handling workshop and assisted with the development of tutoring material

- Delivered short lecture centred on implementation of **JupyterLab** fundamentals for Python
- Aided workshop participants during workshop to follow along with fast-paced lectures and programming tasks
- Provided tuition to Korean scientists requiring a strong understanding on **Unix, Git and Python libraries Numpy, Pandas, xarray, Dask**

### Private VCE Mathematical Tutor

Private tutor

Dec 2018 – Dec 2021

Educated high school students on improving their skills on VCE subjects Mathematical Methods and Specialist Maths

- Experienced in both in-person and remote methods of communication and education
- Delivered hundreds of both group and one-on-one tutoring sessions to highschool students
- Learnt the necessity of building a deep level of understanding on any given topic/subject in order to excel