



AppWorks School

Data Engineering Class

從 Python 程式語言出發，從 AppWorks School 精心規劃的實作訓練中，
深入學習關連式及非關聯式資料庫、資料整理與操作，成為 Data Engineer。

遠距預習

- 期間不須進駐 AppWorks School, 每週約 15 - 20 小時學習量。
- 透過 AppWorks School 規劃的線上學習資源, 循序漸進地學習基礎語法, 以及基本開發功能。

指定練習專案

- 開始進駐 AppWorks School, 每週約 70 - 75 小時學習量。
- 透過精心設計的練習專案, 大量實作開發, 並建立工程思維、解決問題的能力。

專題研討

- 透過分組專題研討、釐清實作開發中遇到的各項重要觀念。
- 學習各式數據的處理技術, 同時練習架設雲端伺服器、初探基礎機器學習與實作資料視覺化。

協作練習專案

- 解讀資深工程師的程式碼架構, 合作於既有程式碼架構中增加新功能。
- 與不同班級學員協作, 學習跨平台開發中溝通協調、程式碼衝突解決與版本控制。

個人自主專案

- 獨立架構與開發, 大量除錯、效能優化, 並完成產品展現, 實現個人創意與想法。
- 學習 Scrum 開發流程、專案管理、開發時程預估。

求職就業

- 以履歷撰寫工作坊、模擬面試工作坊等活動, 準備求職。
- 透過 AppWorks School 舉辦的 Hiring Day 活動, 優先面試 Hiring Partners 提供的工作機會。

Programming Language

1. Python
2. HTML
3. CSS
4. Javascript
5. Shell Script (Basic)

Environment and Framework

1. Linux
2. Flask
3. PySpark

Relational Database

1. CRUD Operations: MySQL
2. Indexing, Primary Key, Foreign Key and Joins
3. Transaction and ACID
4. Data Model: One-to-One, One-to-Many, Many-to-Many
5. Database Normalization
6. Backup and Migration

NoSQL Database

1. Introduction to NoSQL
2. MongoDB
3. CAP Theorem (Basic)
4. Column-based database
5. OLAP v.s. OLTP

Cloud Service

1. AWS EC2, S3, RDS
2. AWS Data Pipeline
3. AWS Lambda
4. AWS EMR
5. AWS command line tools

Data Related Topics

1. Data Cleaning
2. Data Wrangling
3. Data Pipeline Automation
4. Concept of Machine Learning (Basic)
5. Concept of Map-Reduce (Basic)
6. Data Visualization (Basic)
7. Dashboard
8. Recommendation System

Key Concepts

1. Version Control: Git, Github
2. SSH commands and remote access
3. Extract, Transform, Load (ETL)
4. Parallel Programming
5. OOP and Functional Programming
6. Comparison of databases and data format
7. Batch and Real Time data processing