

巨量資料分析師職能基準

版本	職能基準代碼	職能基準名稱	狀態	更新說明	發展更新日期
V3	SMS2521-001v3	巨量資料分析師	最新版本	略	2022/12/30
V2	SMS2521-001v2	巨量資料分析師	歷史版本	已被《SMS2521-001v3》取代	2020/02/07
V1	SMS2521-001v1	巨量資料分析師	歷史版本	已被《SMS2521-001v2》取代	2019/12/31

職能基準代碼		SMS2521-001v3			
職能基準名稱 (擇一填寫)		職類			
		職業	巨量資料分析師		
所屬 類別	職類別	科學、技術、工程、數學 / 數學及科學		職類別代碼	SMS
	職業別	資訊及通訊專業人員 / 系統分析及設計師		職業別代碼	2511
	行業別	出版影音及資通訊業 / 資訊服務業		行業別代碼	J63
工作描述		依客戶或公司自訂目標，能依據商業需求進行分析並提出洞見，執行巨量資料加值的各項工程作業，協助產品建構與決策最佳化。			
基準級別		4			

主要職責	工作任務	工作產出	行為指標	職能 級別	職能內涵 (K=knowledge 知識)	職能內涵 (S=skills 技能)
T1 理解與 準備資料	T1.1 描述 與瞭解資 料收集流 程	O1.1.1 定義 商業問題 O1.1.2 流程 圖、報告或 相關資料表	P1.1.1 根據業務痛點擬定業務目標。 P1.1.2 能從業務流程中敏銳地理解各式資料 (結構與非結構) 生成的方式與商業邏輯。	4	K01 資料盤點 K02 企業資訊系統 (如企業資源 規劃、顧客關係管理、供應鏈管 理、產品/服務研發管理、人力資 源管理等資訊系統等) K03 商業與營運模式 K04 應用機率與統計	S01 業務流程分析 S02 商業模式分析
	T1.2 整合 與資料前	O1.2.1 降低 雜訊後的整	P1.2.1 能運用各式資料結構與圖表工具，整合出 合適的資料物件。	4	K01 資料盤點 K04 應用機率與統計 K05 資料庫	S03 資料分析工具 (如 SAS, SPSS 等)

主要職責	工作任務	工作產出	行為指標	職能級別	職能內涵 (K=knowledge 知識)	職能內涵 (S=skills 技能)
	處理 (含清理)	合資料表或資料物件			K06 資訊網路 (含網際網路與物聯網) K07 網頁資料擷取概念 K08 自然語言處理 K09 視覺音訊處理 K10 分散式運算與儲存概念	S04 程式語言 (如 R, Python 等) S05 資料庫技術 (如 SQL, NoSQL 等) S06 非結構化資料處理 (如文字、影音、影像與機器資料處理等) S07 資料探索技術 S08 版本控制 (如 Git, SVN 等)
	T1.3 轉換資料、挑選與萃取資料屬性	O1.3.1 屬性合宜且校準後的資料物件	P1.3.1 能從偽、雜、速、龐的巨量資料中，根據業務目標，運用資料解析思維，抓出關鍵的數據。	5	K04 應用機率與統計 K08 自然語言處理 K09 視覺音訊處理 K11 資料探勘基礎知識 (如多變量統計、機器學習、資料探勘等)	S03 資料分析工具 (如 SAS, SPSS 等) S04 程式語言 (如 R, Python 等) S08 版本控制 (如 Git, SVN 等) S09 資料視覺化技術 S10 資料縮減技術 S11 時空資料分析 (如時間序列分析、空間資料分析等)
T2 進行資料分析與建模	T2.1 選擇資料建模技術與快速雛形化	O2.1.1 巨量資料分析專案各模型開發邏輯說明書 O2.1.2 能就解決方案中	P2.1.1 能針對不同業務目標，運用各種資料建模技術，進行精簡合宜的資料解析與建模工作。	5	K04 應用機率與統計 K10 分散式運算與儲存概念 K11 資料探勘基礎知識 (如多變量統計、機器學習、資料探勘等)	S03 資料分析工具 (如 SAS, SPSS 等) S04 程式語言 (如 R, Python 等) S09 資料視覺化技術 S11 時空資料分析 (如時間序列分析、空間資料分析等)

主要職責	工作任務	工作產出	行為指標	職能級別	職能內涵 (K=knowledge 知識)	職能內涵 (S=skills 技能)
		的各個模組，混搭各式 IT 工具，測試所選建模技術之可行性報告書				S12 AutoML 操作能力 S13 機器學習框架 (如 Tensorflow, Keras, PyTorch, Scikit-Learn 等) S14 資料探勘技術 (如集群分析、頻繁型態分析、迴歸與分類、離群值分析等) S15 資料串流分析
	T2.2 調校與測試資料模型	O2.2.1 各個模組的程式說明書 O2.2.2 資料模型 O2.2.3 模型評估報告	P2.2.1 能避免模型過度配適，以有效與效率的機制調整巨量資料分析演算法的參數，挑選較佳的模型，並對模型進行驗證。 P2.2.2 能避免模型配適不佳，當發生時進行資料再蒐集。 P2.2.3 設計或組合資料分析演算法，以建置資料模型。	5	K01 資料盤點 K10 分散式運算與儲存概念 K11 資料探勘基礎知識 (如多變量統計、機器學習、資料探勘等) K12 資料結構與演算法 K13 特徵工程	S03 資料分析工具 (如 SAS, SPSS 等) S04 程式語言 (如 R, Python 等) S09 資料視覺化技術 S12 AutoML 操作能力 S13 機器學習框架 (如 Tensorflow, Keras, PyTorch, Scikit-Learn 等) S16 模型績效評估 S17 模型訓練與測試機制 S18 重抽樣方法 S19 集成學習 (如拔靴集成、效能提升等)
T3 參與系統部署與商業應用	T3.1 協助部署巨量	O3.1.1 依資料分析應用情境，修正	P3.1.1 協助資料工程師運用各式 IT 技術，測試與實現資料加值與決策支援的生產系統。 P3.1.2 巨量資料串流處理與模型生命週期管理。	5	K10 分散式運算與儲存概念 K14 系統分析與設計 K15 雲端運算概念	S04 程式語言 (如 R, Python 等)

主要職責	工作任務	工作產出	行為指標	職能級別	職能內涵 (K=knowledge 知識)	職能內涵 (S=skills 技能)
	資料分析 應用系統	或創新組織 的商業流程 與模式 O3.1.2 使用 資料導向程 式語言 (如 R, Python) · 協助開發系 統雛型			K16 UI/UX 設計概念 K17 決策支援系統	S05 資料庫技術 (如 SQL, NoSQL 等) S08 版本控制 (如 Git, SVN 等) S09 資料視覺化技術 S13 機器學習框架 (如 Tensorflow, Keras, PyTorch, Scikit-Learn 等) S20 單元與整合測試能力 S21 MLOps 工具操作
	T3.2 進行 商業溝通	O3.2.1 領域 別營運模式 彙整報告與 業務流程圖 O3.2.2 領域 別市場發展 趨勢分析 O3.2.3 解決 方案與業務 目標達成說 明書或視覺 化圖表	P3.2.1 能夠協助營運團隊分析比較不同領域的業務營運模式。 P3.2.2 能夠協助營運團隊了解巨量資料分析於不同領域的運用狀況、機會與未來發展趨勢。	5	K01 資料盤點 K02 企業資訊系統 (如企業資源 規劃、顧客關係管理、供應鏈管 理、產品/服務研發管理、人力資 源管理等資訊系統)	S01 業務流程分析 S02 商業模式分析 S09 資料視覺化技術 S22 簡報技巧 S23 成本效益分析

職能內涵 (A=attitude 態度)
A01 溝通協調、A02 團隊合作、A03 通盤思考、A04 成果導向、A05 跨領域學習、A06 細心謹慎、A07 彈性應變、A08 追求卓越、A09 主動積極、A10 自我管理、A11 自我學習、A12 抗壓性、A13 職業道德、A14 法遵與風險

說明與補充事項
● 建議擔任此職類/職業之學歷/經歷/或能力條件： • 大學或專科以上學歷，資訊或統計相關科系尤佳。 • 具備英文閱讀能力與跨領域學習興趣。