

巨量資料分析師職能基準

職能基準代碼		SMS2521-001			
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	資料庫設計師及管理師	職業別代碼	2521	
	行業別	出版、影音製作、傳播及資通訊服務業 / 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，能具體通過資料加值的各項作業，完成產品建構與服務過程改善的解決方案系統開發專案。			
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 流程圖、報告 或資料表	P1.1 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 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) K02 商業與營運模式	S01 資料蒐集 S02 流程分析 S03 商業模式分析
	T1.2 整合與清 理資料	O1.2 移除雜訊後的 整合資料表或資料 物件	P1.2 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	K03 資料結構 K04 程式語言 K05 資料庫 K06 應用統計學 K07 資訊網路(含網際網路 與物聯網) K08 網頁資料擷取概念	S04 資料分析工具(如 SAS, SPSS, WEKA, R, Python...等) S05 資料庫技術(如 SQL, NoSQL...等) S06 分散運算與儲存概念(如 Hadoop, Hive 等) S07 資料視覺化技術

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)
					K09 自然語言處理 K10 影像處理	S08 低結構化資料組織技術(如HTML/XML/JSON等) S09 低結構化資料處理(如文字、影音、影像與機器資料處理等)
	T1.3 轉換資料、挑選與萃取資料屬性	O1.3 屬性合宜且校準後的資料物件	P1.3 能從偽/雜/速/龐的巨量資料中，運用資料解析思維抓出關鍵的數據	5	K09 自然語言處理 K10 影像處理 K11 矩陣代數 K12 資料探勘基礎知識(如多變量統計、機器學習、資料探勘等) K13 應用機率模型	S04 資料分析工具(如 SAS, SPSS, WEKA, R, Python...等) S07 資料視覺化技術 S10 資料縮減技術 S11 問題解決方法
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	P2.1 能針對不同應用情境，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	K11 矩陣代數 K12 資料探勘基礎知識(如多變量統計、機器學習、資料探勘等) K13 應用機率模型	S04 資料分析工具(如 SAS, SPSS, WEKA, R, Python...等) S07 資料視覺化技術 S12 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) S13 時空資料分析(如時間序列分析、空間資料分析、序列資料分析等) S14 資料串流分析 S15 圖形資料分析

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)
						S16 深度學習技術
	T2.2 調校與測試資料模型	O2.2 解決方案各個模組的程式說明書	P2.2 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	5	K12 資料探勘基礎知識(如多變量統計、機器學習、資料探勘等) K14 作業研究與演算法 K15 實驗設計	S04 資料分析工具(如 SAS, SPSS, WEKA, R, Python...等) S06 分散運算與儲存概念(如 Hadoop, Hive 等) S07 資料視覺化技術 S17 重抽樣方法 S18 模型訓練與測試機制 S19 模型績效評量 S20 拔靴集成 S21 效能提升
T3 參與系統部署與商業應用	T3.1 協助部署巨量資料分析應用系統	O3.1 依巨量資料分析應用情境，修正或創新組織的商業流程與模式	P3.1 協助巨量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	4	K04 程式語言 K05 資料庫 K16 系統分析與設計 K17 雲端運算概念 K18 平行運算概念 K19 決策支援系統 K20 UI/UX 設計概念	S22 版本控制軟體(如 Github, Subversion...等) S23 編譯程式語言(如 Java, C/C++...等) S04 資料分析工具(如 SAS, SPSS, WEKA, R, Python...等) S06 分散運算與儲存概念(如 Hadoop, Hive 等) S24 整合測試能力
	T3.2 瞭解領域知識	O3.2.1 領域別營運模式彙整報告與作	P3.2.1 能夠分析比較不同領域的業務營運模式。	4	K02 商業與營運模式 K21 線上廣告	S01 資料蒐集 S02 流程分析

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)
		業流程圖 O3.2.2 領域別市場發展趨勢分析	P3.2.2 能夠深入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。		K22 推薦系統 K23 社群網絡分析	S03 商業模式分析 S25 企業流程再造
	T3.3 進行商業溝通	O3.3 解決方案與分析目標達成說明書及/或視覺化圖表	P3.3 能運用財會、行銷與管理等商業語言，向各方利害關係人溝通	4	K01 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) K24 商業管理(如財務管理、會計學、行銷管理、生產與作業管理等)	S07 資料視覺化技術 S26 簡報技巧 S27 商業計畫書撰寫 S28 成本效益分析

職能內涵 (A=attitude 態度)

- A01 溝通協調能力
- A02 團隊意識
- A03 策略性思考
- A04 成果導向
- A05 價值判斷
- A06 跨領域企圖心
- A07 細心謹慎
- A08 彈性
- A09 追求卓越

職能內涵 (A=attitude 態度)

A10 執行力

A11 主動積極

A12 自我管理

A13 自我學習

A14 抗壓性

A15 問題分析與解決

A16 職業道德標準

A17 法規與風險意識

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：
 1. 大學或專科以上學歷，具有文法商理工農醫相關應用領域背景，資訊統計相關科系尤佳。
 2. 具備英文閱讀能力與跨領域學習興趣。
- 基準更新紀錄
 - 因應 2017/05/25 公告職能基準品質認證作業規範修訂版，將原「入門水準」內容移至「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】。