

110 年度初級巨量資料分析師能力鑑定試題

科目 1：資料導向程式設計

考試日期：110 年 09 月 26 日

第 1 頁，共 12 頁

單選題 50 題 (佔 100%)

B	<p>1. Python 語言中，附圖程式碼輸入「6」並執行後，輸出結果為下列何者？</p> <pre>x=input("請輸入數字") y=x*2 z=x+y print(z)</pre> <p>(A) 可正常輸出，結果為 18 (B) 可正常輸出，結果為 666 (C) 不可輸出，因為資料型態不對 (D) 不可輸出，因為運算式有誤</p>																				
A	<p>2. Python 語言中，x 為 Pandas 的 DataFrame 物件（如附圖），請問下列何者為 type(x['name'])的回傳值？</p> <table><thead><tr><th></th><th>name</th><th>price</th><th>cost</th></tr></thead><tbody><tr><td>0</td><td>Apple</td><td>30.0</td><td>12</td></tr><tr><td>1</td><td>Banana</td><td>8.0</td><td>5</td></tr><tr><td>2</td><td>Orange</td><td>NaN</td><td>5</td></tr><tr><td>3</td><td>Lemon</td><td>20.0</td><td>10</td></tr></tbody></table> <p>(A) pandas.core.series.Series (B) numpy.ndarray (C) str (D) pandas.core.frame.DataFrame</p>		name	price	cost	0	Apple	30.0	12	1	Banana	8.0	5	2	Orange	NaN	5	3	Lemon	20.0	10
	name	price	cost																		
0	Apple	30.0	12																		
1	Banana	8.0	5																		
2	Orange	NaN	5																		
3	Lemon	20.0	10																		
C	<p>3. 請問執行附圖程式碼後，下列何者為輸出結果？</p> <pre>math_grade = 80 history_grade = 75 str(math_grade) + str(history_grade)</pre> <p>(A) 為字串（string）型態"155" (B) 為整數（int）型態 155 (C) 為字串（string）型態"8075" (D) TypeError</p>																				
B	<p>4. 請問執行附圖程式碼後，下列敘述何者「不」正確？</p> <pre>t = (1, 2, 3) t[0] = 0</pre> <p>(A) t 之現值為(1, 2, 3) (B) t 之現值為(0, 2, 3)</p>																				

110 年度初級巨量資料分析師能力鑑定試題

科目 1：資料導向程式設計

考試日期：110 年 09 月 26 日

第 2 頁，共 12 頁

	<p>(C) t 之型態 (type) 為 tuple</p> <p>(D) 會產生 TypeError</p>
A	<p>5. Python 語言中，關於 tuple 的觀念，下列敘述何者「不」正確？</p> <p>(A) $(1, 2) * 2 = (2, 4)$</p> <p>(B) tuple 與 list 相似，最大的差別是 tuple 每個索引的參考物件不可變動</p> <p>(C) (1, 2) 的資料型態為 tuple</p> <p>(D) Data = (0, 1, 2)，則 Data[0:2] = (0, 1)</p>
A	<p>6. 資料大多是由樣本及其屬性所形成的集合，下列哪一對用語「並非」樣本與屬性的別名？</p> <p>(A) 資料點與樣例</p> <p>(B) 物件與其函數</p> <p>(C) 記錄與欄位</p> <p>(D) 實體與特徵</p>
C	<p>7. 關於不同變數尺度運算，下列敘述何者「不」正確？</p> <p>(A) 4 公尺是 2 公尺的兩倍，所以長度是比例尺度</p> <p>(B) 35 攝氏度比 25 攝氏度高 10 度，所以溫度是區間尺度</p> <p>(C) 代表男性的 1 大於代表女性的 0，所以性別是順序尺度</p> <p>(D) 白色與黑色截然不同，所以顏色是名目尺度</p>
C	<p>8. 參考附圖之網路通用資料格式 (Network Common Data Form, NetCDF) 與 R 語言的 ncdf4 套件，下列敘述何者「不」正確？</p> <pre> > ncnew File TaiwanCity_population.nc (NC_FORMAT_CLASSIC): 1 variables (excluding dimension variables): integer Pop[City] units: count _FillValue: -1 long_name: Population 1 dimensions: City Size:20 units: count long_name: City > class(ncnew) [1] "ncdf4" </pre> <p>(A) NetCDF 是一種自我描述、跨平台與基於陣列的科學資料格式</p> <p>(B) NetCDF 已經有支援 HDF5 資料格式 (Hierarchical Data Format, HDF)</p> <p>(C) ncdf4 套件定義 NetCDF 變數的方法為 nc_open</p> <p>(D) ncnew 資料物件的檔案名稱為「TaiwanCity_population.nc」</p>
B	<p>9. 在 SQL 中，可使用何者去除重複資料？</p>

110 年度初級巨量資料分析師能力鑑定試題

科目 1：資料導向程式設計

考試日期：110 年 09 月 26 日

第 3 頁，共 12 頁

	(A) having (B) distinct (C) count (D) order by																								
C	10. 附圖資料表 Data，請問下列何者無法同時取得 Celine、Shawn 的資料？ <table><tr><th>Name</th><th>Sales</th><th>Age</th><th>CityName</th></tr><tr><td>Aries</td><td>1500</td><td>29</td><td>Taipei</td></tr><tr><td>Celine</td><td>250</td><td>17</td><td>Tainan</td></tr><tr><td>Shawn</td><td>300</td><td>39</td><td>NewTaipei</td></tr><tr><td>Jiyu</td><td>700</td><td>35</td><td>Pingtung</td></tr></table> (A) SELECT * FROM Data WHERE Sales <= 300 (B) SELECT * FROM Data WHERE Age < 20 OR Age >38 (C) SELECT * FROM Data WHERE Name == 'Celine' OR Name == 'Shawn' (D) SELECT * FROM Data WHERE CityName IN ('Tainan', 'NewTaipei')	Name	Sales	Age	CityName	Aries	1500	29	Taipei	Celine	250	17	Tainan	Shawn	300	39	NewTaipei	Jiyu	700	35	Pingtung				
Name	Sales	Age	CityName																						
Aries	1500	29	Taipei																						
Celine	250	17	Tainan																						
Shawn	300	39	NewTaipei																						
Jiyu	700	35	Pingtung																						
D	11. 關於 Spark 叢集計算框架，下列敘述何者正確？ (A) 是一個封閉的叢集計算框架 (B) 搭配低階程式語言編寫叢集運算程式 (C) 執行效率低 (D) 支援 GraphX 圖形處理演算法																								
A	12. 請問附圖資料表已達到哪一個正規化階段？ <table><tr><th>學號</th><th>導師</th><th>導師辦公室</th><th>課程 1</th><th>課程 2</th><th>課程 3</th></tr><tr><td>101</td><td>Amy</td><td>123</td><td>108</td><td>153</td><td>169</td></tr><tr><td>205</td><td>Bob</td><td>246</td><td>213</td><td>123</td><td>312</td></tr><tr><td>307</td><td>Cindy</td><td>369</td><td>213</td><td>155</td><td>312</td></tr></table> (A) 未正規化 (B) 第一正規化 (C) 第二正規化 (D) 第三正規化	學號	導師	導師辦公室	課程 1	課程 2	課程 3	101	Amy	123	108	153	169	205	Bob	246	213	123	312	307	Cindy	369	213	155	312
學號	導師	導師辦公室	課程 1	課程 2	課程 3																				
101	Amy	123	108	153	169																				
205	Bob	246	213	123	312																				
307	Cindy	369	213	155	312																				
C	13. 下列指令何者「不」屬於資料操作語言（data manipulation language）？ (A) INSERT (B) UPDATE (C) CREATE (D) DELETE																								
B	14. 附圖資料表為 productTable。請問對其執行																								

110 年度初級巨量資料分析師能力鑑定試題

科目 1：資料導向程式設計

考試日期：110 年 09 月 26 日

第 4 頁，共 12 頁

	<p>SELECT COUNT(productName) FROM productTable WHERE price>18; 語句後，得到之 COUNT(productName)結果為何？</p> <table><tr><th>id</th><th>productName</th><th>price</th></tr><tr><td>1</td><td>lemon</td><td>10</td></tr><tr><td>2</td><td>apple</td><td>15</td></tr><tr><td>3</td><td>banana</td><td>20</td></tr><tr><td>4</td><td>guava</td><td>30</td></tr></table> <p>(A) 1 (B) 2 (C) 3 (D) 4</p>	id	productName	price	1	lemon	10	2	apple	15	3	banana	20	4	guava	30
id	productName	price														
1	lemon	10														
2	apple	15														
3	banana	20														
4	guava	30														
C	<p>15. 關於使用 Python 進行資料讀取/存儲相關工作，下列敘述何者「不」正確？</p> <p>(A) 可透過編寫 Python 程式碼，從文件/圖片/影片中擷取文字資訊 (B) 可透過編寫 Python 程式碼，從.csv 檔案讀入資訊後，存入 MySQL 資料庫 (C) Python 無法與 NoSQL 資料庫（如 MongoDB）進行連接 (D) 依據使用情境，Python 可做到將資料存儲於靜態檔案（如.csv、.json）或資料庫</p>															
D	<p>16. 請問下列何者屬於半結構化資料格式？</p> <p>(A) CSV (B) EXCEL (C) TXT (D) JSON</p>															
C	<p>17. 關於 ETL（Extract-Transform-Load），下列敘述何者「不」正確？</p> <p>(A) 是建置或更新資料倉儲中的內容時所需的過程 (B) Extract：資料擷取，從資料來源處擷取所需之數據資料 (C) Transform：資料轉換，針對結構資料轉換，非結構資料則無法處理 (D) Load：資料載入，最後將已作適當轉換過的數據資料載入到目的地</p>															
C	<p>18. 關於 R 與 Python 語言，下列敘述何者「不」正確？</p> <p>(A) 在 R 語言中，可使用 rm()函數刪除不需要的物件 (B) 在 Python 中，Jupyter Notebook 可輸入%reset 刪除所有變數 (C) 在 R 語言中，無法使用 rm()函數刪除自訂函數 (D) 在 Python 中，可輸入 clear，以清除 IPython console 中所有的先前</p>															

110 年度初級巨量資料分析師能力鑑定試題

科目 1：資料導向程式設計

考試日期：110 年 09 月 26 日

第 5 頁，共 12 頁

	訊息
C	19. 關於讀取.csv 逗號分隔檔案 (comma separated values, csv)，下列敘述何者「不」正確？ (A) 在 R 語言中，可使用 read.csv() 函數 (B) 在 Python 語言中，可使用 pandas 套件的 read_csv() 方法 (C) 在 R 語言中，可使用 open() 函數 (D) 在 Python 語言中，可使用 open() 函數
C	20. 關於 R 語言與 Python 語言，下列敘述何者正確？ (A) 在 R 語言中，無法使用 write.table() 函數匯出.csv 逗號分隔檔案 (comma separated values, csv) (B) 在 Python 語言中，可使用 numpy 的 to_csv() 方法匯出.csv 逗號分隔檔案 (C) 在 R 語言中，可使用 write.csv() 函數匯出.csv 逗號分隔檔案 (D) 在 Python 語言中，可使用 open("file.csv", "r") 讀寫.csv 逗號分隔檔案
B	21. 關於讀取.xlsx 檔，下列敘述何者正確？ (A) 在 R 語言中，只有一個套件{readxl}可以使用 (B) 在 Python 語言中，可使用 pandas 套件中的 read_excel() 方法 (C) 在 R 語言中，可使用{readxl}套件中的 readxl() 函數 (D) 在 Python 語言中，可使用 pandas 套件中的 open_xlsx() 方法
D	22. Python 中常使用 open(filename, mode, ...) 函數來開啟文件，關於其函數內參數 - 開啟模式之作用，下列敘述何者「不」正確？ (A) mode=r，以唯讀模式 (read-only) 開啟文件 (B) mode=a，以追加模式 (appending) 開啟文件 (C) mode=w+，以讀寫模式 (writing and reading) 開啟文件 (D) mode=r+，以唯讀模式 (read-only) 開啟文件
C	23. JSON 是目前在資料匯入與匯出時很常見到的格式，請問若是以 Python 語言讀入 JSON 的資料時，其含有的資料型態可能為何？ (A) list (B) dictionary (C) 可能是 list，也可能是 dictionary (D) 不可能是 list 或是 dictionary
B	24. 關於 JSON，下列敘述何者正確？ (A) JSON 的格式較 XML 來得複雜，但所需的儲存容量相對較小 (B) JSON 的資料格式是以「鍵-值」(key:value) 方式配對儲存 (C) 在 JSON 格式中，鍵可以是一個數值或字串，字串需用雙引號表

110 年度初級巨量資料分析師能力鑑定試題

科目 1：資料導向程式設計

考試日期：110 年 09 月 26 日

第 6 頁，共 12 頁

	<p>示</p> <p>(D) 在 JSON 格式中，值可以是數值、字串、布林值、陣列，但不能是 null 值</p>
D	<p>25. 請問附圖資料為何種格式？</p> <pre>{ "name": "Julie", "age": 30, "cars": ["HONDA", "BMW", "LEXUS"] }</pre> <p>(A) CSV (B) XLS (C) XML (D) JSON</p>
B	<p>26. Python 語言中，下列何者「不」是合法的命名規則？</p> <p>(A) int32 (B) 8XL (C) G_U_I (D) __apple__</p>
A	<p>27. 參考附圖程式碼，請問註解處應填入選項中何者程式敘述方能使最後 print(s) 印出 'john'？</p> <pre>a = 'nice' b = 'john' d = {} for x,y in zip(a,b): #這裡應填寫哪個程式敘述 s = '' for c in a: s += d[c] print(s)</pre> <p>(A) d[x] = y (B) d[y] = x (C) d[a] = b (D) d[b] = a</p>
C	<p>28. Python 語言中，執行附圖程式碼後，下列何者為輸出結果？</p> <pre>s = 'ABC' n = 1 for c in s:</pre>

110 年度初級巨量資料分析師能力鑑定試題

科目 1：資料導向程式設計

考試日期：110 年 09 月 26 日

第 7 頁，共 12 頁

	<pre>print(c * n) n += 1</pre> <p>(A) A B C (B) A B C (C) A BB CCC (D) A BB CCC</p>
C	<p>29. Python 語言中，執行附圖程式碼後，x 為下列何者？</p> <pre>import numpy as np x = np.array([4,2] * 2)</pre> <p>(A) $\begin{bmatrix} 4 & 4 \\ 2 & 2 \end{bmatrix}$ (B) $\begin{bmatrix} 4 \\ 2 \\ 4 \\ 2 \end{bmatrix}$ (C) [4 2 4 2] (D) [8 4]</p>
D	<p>30. Python 語言中，執行附圖程式碼後，下列選項中之程式碼何者無法返回值'mary'？</p> <pre>class Student: def __init__(self, student_name): self._student_name = student_name def get_name(self): return self._student_name def _get_name(self): return self._student_name def __get_name(self): return self._student_name Mary = Student('mary')</pre> <p>(A) Mary._student_name (B) Mary.get_name() (C) Mary._get_name() (D) Mary.__get_name()</p>
C	<p>31. 在 Python 3 環境中，根據附圖程式碼，slicelist 的內容為何？</p>

110 年度初級巨量資料分析師能力鑑定試題

科目 1：資料導向程式設計

考試日期：110 年 09 月 26 日

第 8 頁，共 12 頁

	<pre>mylist=[1,2,3,4,5] slicelist=mylist[5:]</pre> <p>(A) 5 (B) [1,2,3,4,5] (C) 空串列 (D) 程式異常，slicelist 不存在</p>
B	<p>32. 在 Python 3 環境中，根據附圖程式碼，a、b、c、d 何者與 t 是同一個物件？</p> <pre>import copy t=[1,2,[3,4]] a=[1,2,[3,4]] b=t c=copy.copy(t) d=copy.deepcopy(t)</pre> <p>(A) a (B) b (C) c (D) d</p>
D	<p>33. 下列何者「不」屬於程式語言中的流程控制結構 (Control Structures)？</p> <p>(A) while 迴圈 (B) if else (C) for 迴圈 (D) 變數宣告</p>
D	<p>34. 附圖 Python 語言的自訂函數結果為何？</p> <pre>1 def travelverify(param1, **param2): 2 print(type(param2)) 3 travelverify('travel', Taipei='101大樓', 嘉義='阿里山', 南投='日月潭')</pre> <p>(A) None (B) <class 'tuple'> (C) <class 'list'> (D) <class 'dict'></p>
C	<p>35. Python 語言中，執行附圖程式碼後，下列何者為輸出結果？</p> <pre>def test_func(): a = 10 print('目前 a 值為: {}'.format(a)) a = 5 test_func()</pre>

110 年度初級巨量資料分析師能力鑑定試題

科目 1：資料導向程式設計

考試日期：110 年 09 月 26 日

第 9 頁，共 12 頁

	<pre>print('目前 a 值為: {}'.format(a))</pre> <p>(A) 目前 a 值為: 10 目前 a 值為: 10</p> <p>(B) 目前 a 值為: 5 目前 a 值為: 5</p> <p>(C) 目前 a 值為: 10 目前 a 值為: 5</p> <p>(D) 目前 a 值為: 5 目前 a 值為: 10</p>
A	<p>36. 如附圖所示，假設 R 語言有一個自訂函數，執行 A(5)後回傳的值為何？</p> <pre>A = function(x) { if (x<4) return(x) else return(A(x-1) - A(x-2) + A(x-3)) }</pre> <p>(A) 1 (B) 3 (C) 5 (D) 11</p>
B	<p>37. Python 語言中，執行附圖程式碼後，下列何者為輸出結果？</p> <pre>for n in range(1,5,2): pass print(n)</pre> <p>(A) 1 (B) 3 (C) 5 (D) 7</p>
D	<p>38. Python 語言中，執行附圖程式碼後，下列何者為輸出結果？</p> <pre>def xyz(): a = 10 print(a) a = 'abc' print(a) a = 20 print(a)</pre> <p>(A) 10 (B) abc (C) 20</p>

110 年度初級巨量資料分析師能力鑑定試題

科目 1：資料導向程式設計

考試日期：110 年 09 月 26 日

第 10 頁，共 12 頁



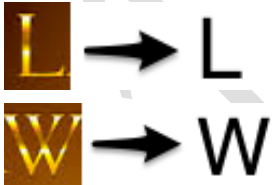
	(D) 執行錯誤
C	<p>39. Python 語言中，執行附圖程式碼後，下列何者「不」會出現於執行結果？</p> <pre>def test_func(user_name, *args, **kwargs): print(f"user_name: {user_name}") for arg in args: print(f'arg: {arg}') for k, v in kwargs.items(): print(f'key arg : {k} value {v}')</pre> <p>test_func("Ken", 1, 2, height=44, weight=55)</p> <p>(A) user_name: Ken (B) key arg: height value 44 (C) key arg: arg value 44 (D) key arg: weight value 55</p>
B	<p>40. 關於自訂函數，下列敘述何者正確？</p> <p>(A) 必須有一個回傳值 (B) 可以有任意個參數 (C) 可以存在相同名稱的參數 (D) 沒有使用指令可直接結束函數</p>
C	<p>41. 自訂函數中可以自定義幾個參數？</p> <p>(A) 1 個 (B) 2 個 (C) 無限制 (D) 實際用到幾個就能定義幾個</p>
C	<p>42. 當需求為「循環重複地執行一個語句，直到條件不為真為止」時，可單獨使用下列何種流程控制結構實作？</p> <p>(A) for 迴圈 (B) if else (C) while 迴圈 (D) do 迴圈</p>
D	<p>43. 下列敘述何者「不」正確？</p> <p>(A) 撰寫程式時經常發生不可預期的錯誤程式 (B) 有狀況發生時，系統會回報錯誤、警告或訊息 (C) 透過例外處理機制，可避免錯誤發生時造成程式停止執行 (D) 警告訊息（Warning）說明程式輸出的結果；訊息（Message）說</p>

110 年度初級巨量資料分析師能力鑑定試題

科目 1：資料導向程式設計

考試日期：110 年 09 月 26 日

第 11 頁，共 12 頁

	明潛在的問題
D	<p>44. 「將 data 複製到每個運算單元 (worker) 內，使每個 process 都有自己的 memory space」，請問此敘述為下列何者？</p> <p>(A) 共享記憶平行處理 (Shared Memory Parallelism)</p> <p>(B) 資料平行處理 (Data Parallelism)</p> <p>(C) 工作平行處理 (Task Parallelism)</p> <p>(D) 分散記憶平行處理 (Distributed Memory Parallelism)</p>
A	<p>45. 下列敘述何者正確？</p> <p>(A) 良好的程式可讀性有助於除錯的進行</p> <p>(B) 程式的結果正確就代表程式沒有錯誤</p> <p>(C) 邏輯錯誤比起語法錯誤更容易被發現</p> <p>(D) 偶發的程式錯誤沒有除錯的必要</p>
A	<p>46. 如附圖所示，小明想實作一個文字辨識系統，目標如下：(1)文字偵測→(2)字元切割→(3)字元辨識。完成後，小明發現系統總是無法正確辨識，請問應從何處開始除錯？</p> <p>(1)文字偵測</p>  <p>(2)字元切割</p>  <p>(3)字元辨識</p>  <p>(A) 文字偵測</p> <p>(B) 字元切割</p> <p>(C) 字元辨識</p> <p>(D) 都可以</p>
A	<p>47. 關於資料導向程式設計，下列敘述何者「不」正確？</p> <p>(A) 需多撰寫迴圈，避免使用向量化函數，以增加運算效率</p> <p>(B) 多採用邏輯值索引，避免運用條件式語法</p> <p>(C) 善用自定義函數模組化工作流程</p>

110 年度初級巨量資料分析師能力鑑定試題

科目 1：資料導向程式設計

考試日期：110 年 09 月 26 日

第 12 頁，共 12 頁

	(D) 注意引用函數其預設的引數值
C	48. 使用 Python 語言 try...except 處理例外時，下列何者子句 (Clause)「不」是一定要使用到的？ (A) try (B) except (C) finally (D) 以上敘述皆需要
B	49. 一般而言，程式設計師透過下列程序偵查與解決代碼的錯誤，請問流程順序應為何者？1.找出錯誤的原因、2.尋找相似的錯誤並修正之、3.讓錯誤可以重製、4.修正錯誤並測試、5.意識到錯誤的存在 (A) 1>2>3>4>5 (B) 5>3>1>4>2 (C) 5>4>3>2>1 (D) 3>5>1>2>4
D	50. 在 Python 3 環境中進行變數命名，下列何者「不」會產生錯誤 (SyntaxError)？ (A) get-the-data (B) totalCaculated\$ (C) 10yearsSum (D) print_all_relatedMemberInfo_cn