

111 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：111 年 05 月 28 日

第 1 頁，共 11 頁

單選題 50 題 (佔 100%)

C	1. Python 語言中，執行「 <code>a = 1000; b = 1000</code> 」，下列敘述何者正確？ (A) <code>a == b</code> 結果為 False (B) <code>a is b</code> 結果為 True (C) <code>id(a) == id(b)</code> 結果為 False (D) <code>type(a)</code> 結果為 float
B	2. Python 的資料型態中，有可變 (Mutable) 與不可變 (Immutable)，以容器 (Container) 為例，下列何者是不可變物件？ (A) list (B) tuple (C) dict (D) set
D	3. 關於 Python 儲存數據的方式，下列何者相對較「不」適合？ (A) 轉成 JSON 檔案 (B) 轉成 txt 檔案 (C) 載入 MySQL 資料庫 (D) 轉為 HTML
D	4. 關於 R 語言的資料型態，下列敘述何者「不」正確？ (A) R 語言的基本資料單位稱作向量 (Vector) (B) 兩個維度的向量，稱為矩陣 (Matrix) (C) 多個維度的向量，稱為陣列 (Array) (D) 向量 (Vector) 可混合不同資料型態
C	5. 資料框 (Dataframe) 能包含多種向量類型，但「不」包含下列何者？ (A) 文字向量 (B) 邏輯值向量 (C) 影像向量 (D) 數值向量
D	6. 在 Python 3 執行環境中，執行附圖程式碼，下列敘述何者「不」正確？ <code>test_list = [(1,2,3), (4,5,6)]</code> (A) <code>test_list[1][2]</code> 的值為 6 (B) <code>test_list[0][2] + test_list[1][-2]</code> 的值為 8 (C) <code>test_list[1]</code> 之型態 (type) 為 tuple (D) 若再執行 <code>test_list[0][0] = 0</code> 後， <code>test_list</code> 將變為 <code>[(0,2,3), (4,5,6)]</code>
B	7. 在 Python 3 執行環境中，執行附圖程式碼後，得到的結果為何？ <code>x = 6</code> <code>y = '3'</code>

111 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：111 年 05 月 28 日

第 2 頁，共 11 頁

	<p><code>f'{str(x) * 2}{int(y) - 1}{x + 5}{y * 2}'</code></p> <p>(A) '122116'</p> <p>(B) '6621133'</p> <p>(C) 47916</p> <p>(D) TypeError</p>
A	<p>8. 在 Python 3 執行環境中，執行附圖程式碼後，得到的結果為何？</p> <pre>animal_list = [['hippo', 'cat', ['dog', 'alpaca'], 'bird']] animal_list[0][2][-1]</pre> <p>(A) 'alpaca'</p> <p>(B) 'bird'</p> <p>(C) 'cat'</p> <p>(D) 'dog'</p>
A	<p>9. 想把某資料表從 A 資料庫轉到 B 資料庫時，下列何者步驟是「不」必要的？</p> <p>(A) 確認 A 資料庫的剩餘空間</p> <p>(B) 在 B 資料庫上建立對應的表格及欄位</p> <p>(C) 將 A 資料庫的資料表轉出 (Export)</p> <p>(D) B 資料庫 Import 轉出的資料</p>
C	<p>10. 關於 MySQL 資料庫的檢視 (View)，下列敘述何者為正確？</p> <p>(A) SELECT 敘述不可以使用 FROM 檢視名稱</p> <p>(B) 使用檢視進行查詢時可以使用自行定義的變數</p> <p>(C) 新增檢視的語法為 CREATE VIEW 檢視名稱</p> <p>(D) 檢視名稱可以與資料表名稱相同</p>
B	<p>11. 「使用 MySQL 資料庫的「訂單」資料表，篩選「出貨日期」為 NULL 的所有欄位」，應使用下列選項何者語法？</p> <p>(A) <code>SELECT * FROM 訂單 WHERE 出貨日期 == NULL</code></p> <p>(B) <code>SELECT * FROM 訂單 WHERE 出貨日期 IS NULL</code></p> <p>(C) <code>SELECT * FROM 訂單 WHERE 出貨日期 = NULL;</code></p> <p>(D) <code>SELECT * FROM 訂單 WHERE 出貨日期 <> NULL</code></p>
B	<p>12. 關於 MongoDB 資料庫，下列敘述何者正確？</p> <p>(A) MongoDB 是 NoSQL 的鍵值資料庫 (Key-Value)</p> <p>(B) 使用 mongo shell 命令列介面的 show dbs 指令可以檢視所有資料庫</p> <p>(C) 使用 mongo shell 命令列介面的 revise 指令可以更新記錄</p> <p>(D) 使用 mongo shell 命令列介面的 drop 指令可以刪除記錄</p>
A	<p>13. 附圖為模擬 MySQL 資料庫之資料表，名稱為 countryTable。請問對其</p>

111 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：111 年 05 月 28 日

第 3 頁，共 11 頁

	<p>執行附圖程式碼後，得到的結果為何？</p> <table border="1"><thead><tr><th>id</th><th>country</th><th>continent</th></tr></thead><tbody><tr><td>1</td><td>Japan</td><td>Asia</td></tr><tr><td>2</td><td>Taiwan</td><td>Asia</td></tr><tr><td>3</td><td>United States</td><td>North America</td></tr><tr><td>4</td><td>Australia</td><td>Oceania</td></tr><tr><td>5</td><td>Russia</td><td>Europe</td></tr><tr><td>6</td><td>Ukraine</td><td>Europe</td></tr></tbody></table> <p>SELECT country FROM countryTable WHERE continent = 'Europe' OR id = 4;</p> <p>(A) Russia, Ukraine, Australia (B) Russia, Ukraine (C) Europe, Europe, Oceania (D) Australia, Ukraine</p>	id	country	continent	1	Japan	Asia	2	Taiwan	Asia	3	United States	North America	4	Australia	Oceania	5	Russia	Europe	6	Ukraine	Europe
id	country	continent																				
1	Japan	Asia																				
2	Taiwan	Asia																				
3	United States	North America																				
4	Australia	Oceania																				
5	Russia	Europe																				
6	Ukraine	Europe																				
D	<p>14. 關於 NoSQL，下列敘述何者「不」正確？</p> <p>(A) 也稱 Not Only SQL，為「不限定為關聯式資料庫」的資料庫管理系統的統稱</p> <p>(B) MongoDB、Neo4j 皆為常見的 NoSQL 資料庫</p> <p>(C) 具備擴充能力強、成本低特性</p> <p>(D) 在 Schema 欄位架構規劃上必須要有固定的結構</p>																					
D	<p>15. 關於使用 Python 進行資料讀取/存儲相關工作時，下列敘述何者「不」正確？</p> <p>(A) 要從 MySQL 資料表獲取資料，需要具有對該表讀取權限、才能進行操作</p> <p>(B) 可透過編寫 Python 程式碼，配合深度學習模型從影片中擷取特定圖像資訊</p> <p>(C) Python 可與 Google BigQuery 進行連接，以使用 Google 雲端的資料服務</p> <p>(D) Python 目前僅能與 SQL 類資料庫連接，尚無法與 NoSQL 類資料庫連接</p>																					
D	<p>16. 關於 SQL 相關語法指令，下列敘述何者「不」正確？</p> <p>(A) BETWEEN 常使用於「對某數值欄位進行區間選取」時</p> <p>(B) ORDER BY 常使用於「對篩選後資料進行排序」時</p> <p>(C) DISTINCT 常使用於「對欄位選出不重複值」時</p>																					

111 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：111 年 05 月 28 日

第 4 頁，共 11 頁

	(D) UPDATE 常使用於「挑選出某數值欄位之最新資料」時																				
D	<p>17. 關於非關聯式資料庫的特性，下列敘述何者「不」正確？</p> <p>(A) 可擴展性</p> <p>(B) 具高效能</p> <p>(C) 高功能性</p> <p>(D) 絕對安全性</p>																				
C	<p>18. 關於關聯式資料庫的 ACID 特性，下列敘述何者「不」正確？</p> <p>(A) A (Atomicity)：確保資料庫中的資料能夠從一個正確狀態接續下一個正確狀態</p> <p>(B) C (Consistency)：要求進行交易時資料就必須符合資料庫結構描述</p> <p>(C) I (Isolation)：確保同一筆資料可容許被兩個 Transaction 同時進行更改</p> <p>(D) D (Durability)：若在資料寫入的當下發生當機情況，引發寫入時的資料流失，RDBMS 也有機制在之後復原資料</p>																				
B	<p>19. 關於 Python 讀取檔案，下列敘述何者正確？</p> <p>(A) 使用 open('file', 'r+') 若 file 不存在，會創建 file</p> <p>(B) 使用 open('file', 'w+') 若 file 不存在，會創建 file</p> <p>(C) 使用 open('file', 'r') 為可讀寫的檔案模式</p> <p>(D) 使用 open('file', 'a+') 為附加讀寫檔案模式，寫入時採行覆寫方式，亦可以讀檔</p>																				
D	<p>20. 關於 Python 的識別符號，下列敘述何者「不」正確？</p> <p>(A) 可以包含字母（大寫或小寫）</p> <p>(B) 不能以數字來開頭</p> <p>(C) 沒有長度限制</p> <p>(D) 允許出現空格</p>																				
A	<p>21. 參考附圖，在 Python 中，x 為 Pandas 的 DataFrame 物件，請問如何取出 10？</p> <table border="1" style="margin-left: auto; margin-right: auto; border-collapse: collapse; text-align: center;"> <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</td> <td>12</td> </tr> <tr> <td>1</td> <td>orange</td> <td>8</td> <td>5</td> </tr> <tr> <td>2</td> <td>banana</td> <td>NaN</td> <td>5</td> </tr> <tr> <td>3</td> <td>lemon</td> <td>20</td> <td>10</td> </tr> </tbody> </table> <p>(A) x.iloc[3,2]</p> <p>(B) x.loc[3,2]</p> <p>(C) x.loc[3,3]</p>		name	price	cost	0	apple	30	12	1	orange	8	5	2	banana	NaN	5	3	lemon	20	10
	name	price	cost																		
0	apple	30	12																		
1	orange	8	5																		
2	banana	NaN	5																		
3	lemon	20	10																		

111 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：111 年 05 月 28 日

第 5 頁，共 11 頁

	(D) x.iloc[3,3]
D	<p>22. 關於 Python 的資料結構，下列敘述何者「不」正確？</p> <p>(A) pandas 有三種資料結構形式，分別是 Series、DataFrame 和索引物件</p> <p>(B) DataFrame 是最常使用也是 pandas 的核心資料形式</p> <p>(C) 在讀入 excel 檔案時，一般使用 read_excel 函式</p> <p>(D) Pandas 也能匯入 csv、dta 等檔案，使用 read_csv、read_stata 函數</p>
D	<p>23. 關於 R 語言，下列敘述何者「不」正確？</p> <p>(A) jsonlite 是 R 的一個 JSON 格式資料處理套件</p> <p>(B) 使用 jsonlite 套件，可以保留 R 物件完整的資料與屬性</p> <p>(C) jsonlite 套件使用 fromJSON() 函數讀取 JSON 文件的內容</p> <p>(D) 串列 (List) 內的所有資料形態必須要一致</p>
C	<p>24. 下列何者「不」是結構化資料？</p> <p>(A) JSON</p> <p>(B) CSV</p> <p>(C) TXT</p> <p>(D) XML</p>
C	<p>25. 在 Python 3 下進行資料處理時，下列何者「不」需要連線網際網路也可完成作業？</p> <p>(A) 以政府資料開放平臺之人口資料 JSON API，進行資料下載</p> <p>(B) 以 pip install -U pandas 安裝最新之 pandas 套件</p> <p>(C) 將處理完畢之資料，存出至本機之 mysql 資料庫</p> <p>(D) 將處理完畢之資料，以 Google Sheet API 將資料上傳</p>
B	<p>26. 附圖 Python 程式碼中，執行自定義函數 Show 時，會根據物件的不同而表現出不同的行為，請問此為物件導向編程中的何種特性？</p> <pre># 父類 class Human: name="人類" def sleep(self): print("睡覺") # 子類 class Taiwanese(Human): name="台灣人" def speak(self): print("中文")</pre>

111 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：111 年 05 月 28 日

第 6 頁，共 11 頁

	<pre># 子類 class American(Human): name="美國人" def speak(self): print("英文") # 定義父類作為引數，所有的子類都可以傳參進去 def Show(Human): Human.speak() Show(Taiwanese()) #輸出中文 Show(American()) #輸出英文 (A) 繼承 (Inheritance) (B) 多型 (Polymorphism) (C) 封裝 (Encapsulation) (D) 抽象 (Abstract)</pre>
D	<p>27. 下列何者「不」是 Python 的關鍵字？</p> <p>(A) class (B) assert (C) continue (D) dbf</p>
C	<p>28. 參考附圖，Python 語言中，執行「x[-1]」之結果何者正確？</p> <pre>1 city = ("台北市", "新北市", "台中市") 2 pop = (255, 402, 281) 3 rank = (3, 1, 2) 4 x = list(zip(city, pop, rank)) 5 x[-1]</pre> <p>(A) ('新北市', 402, 1) (B) [('台北市', 255, 3), ('新北市', 402, 1)] (C) ('台中市', 281, 2) (D) (3, 1, 2)</p>
C	<p>29. 參考附圖，Python 語言中，下列敘述何者「不」正確？</p> <pre>1 a = ["banana", "apple", "orange"] 2 b = [1, 2, "orange"]</pre> <p>(A) 執行 a + b 輸出結果為['banana', 'apple', 'orange', 1, 2, 'orange']</p>

111 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：111 年 05 月 28 日

第 7 頁，共 11 頁

	<p>(B) 執行 <code>a - b</code> 輸出結果為 <code>TypeError</code></p> <p>(C) 執行 <code>a.sort()</code> 輸出結果為 <code>['apple', 'banana', 'orange']</code></p> <p>(D) <code>len([a, b])</code> 輸出結果為 2</p>
D	<p>30. Python 語言中，執行附圖程式碼後，得到結果為何？</p> <pre>items = {k: v} for k in range(5) if k % 2 == 0 for v in range(5) if v % 2 == 0] len(items)</pre> <p>(A) 4</p> <p>(B) 6</p> <p>(C) 8</p> <p>(D) 9</p>
C	<p>31. 在 Python 3 下，執行附圖程式碼後，請問下列選項執行結果，何者「無法」顯示出 <code>Horror</code>？</p> <pre>class Movie: def __init__(self, movieType): self.movieType = movieType self._movieType = movieType self.__movieType = movieType self._movieType_ = movieType MovieA = Movie("Horror")</pre> <p>(A) <code>print(MovieA.movieType)</code></p> <p>(B) <code>print(MovieA._movieType)</code></p> <p>(C) <code>print(MovieA.__movieType)</code></p> <p>(D) <code>print(MovieA._movieType_)</code></p>
C	<p>32. 下列何者「不」屬於 R 語言的物件導向類別 (Class)？</p> <p>(A) S3</p> <p>(B) S4</p> <p>(C) S6</p> <p>(D) RC</p>
C	<p>33. 在 Python 3 下，定義 <code>mylambda()</code> 如附圖，執行附圖程式碼後，輸出結果為何？</p> <pre>def mylambda(): return [lambda x : i for i in range(3)] print([my(3) for my in mylambda()])</pre> <p>(A) <code>[0,1,2]</code></p>

111 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：111 年 05 月 28 日

第 8 頁，共 11 頁

	(B) 3 (C) [2,2,2] (D) 2
A	34. 一條流程控制語句中，最多可以有幾個 else 語句？ (A) 1 個 (B) 2 個 (C) 3 個 (D) 無限制
C	35. Python 語言中，執行附圖程式碼後，得到結果為何？ <pre>def Funct1(A=5, B=8): total=0 for i in range(A, B): total +=i return total</pre> Funct1() (A) 20 (B) 26 (C) 18 (D) 52
C	36. 在 Python 3 下，執行附圖程式碼後，得到結果為何？ <pre>def plus_two_integers(a: int, b: int): ans = a + b print(ans)</pre> v1 = 10.5 v2 = 20 plus_two_integers(v1, v2) (A) NameError: name 'a' is not defined (B) NameError: name 'a' & 'b' are not defined (C) 30.5 (D) ValueError('a is not int')
B	37. 在 Python 3 下，執行附圖程式碼後，輸出結果為何？ <pre>letters_collect = [] for letter in 'ALPACA IS SO CUTE': if letter == 'A' or letter == 'S' or letter == '':</pre>

111 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：111 年 05 月 28 日

第 9 頁，共 11 頁

	<pre style="margin: 0;">letters_collect.append(letter) elif letter == 'O': break else: continue print(letters_collect)</pre> <p>(A) ['A', ',', 'S'] (B) ['A', 'A', 'A', ',', 'S', ',', 'S'] (C) ['A', 'A', 'A', ',', 'S', ',', 'S', ','] (D) ['A', 'L', 'P', 'A', 'C', 'A', ',', 'T', 'S', ',', 'S', 'O', ',', 'C', 'U', 'T', 'E']</p>
C	<p>38. Python 語言中，下列何者為強制終止迴圈狀態的方法？</p> <p>(A) for loop (B) while loop (C) break (D) continue</p>
D	<p>39. Python 語言中，關於流程控制的 if 語句，下列敘述何者正確？</p> <p>(A) 其後必須至少一個 then 語句 (B) 其後必須至少一個 else 語句 (C) 其後可有多個 else 及 else if 語句 (D) 其後可有一個 else 及多個 else if 語句</p>
B	<p>40. 關於進行資料導向程式設計，下列敘述何者「不」正確？</p> <p>(A) 封裝 (Encapsulation) 為物件導向程式設計的方法之一 (B) 應減少自定義函數，避免模組化工作流程 (C) 留意元素型別可能被強制轉換的情況 (D) 注意引用函數其預設的引數值</p>
A	<p>41. Python 語言中，下列何者「不」可作為變數名稱？</p> <p>(A) 1Account (B) balance (C) Class (D) door5</p>
B	<p>42. 例外狀況處理函數允許程式設計師在例外狀況發生時，採取某些行動；例如結合 stop() 函數，以讓程式設計師在必要的時候傳回訊息，瞭解程式碼執行的狀況。下列何者「不」為 R 語言的例外狀況處理函數？</p> <p>(A) withCallingHandlers() (B) system.time() (C) tryCatch() (D) try()</p>

111 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：111 年 05 月 28 日

第 10 頁，共 11 頁

B	<p>43. 使用分散式資料庫或 Hadoop, Spark 在處理巨量資料 (>10G) 時，下列何種狀況最為消耗效能？</p> <p>(A) 計算整體資料筆數</p> <p>(B) 節點之間的資料洗牌 (Shuffle)</p> <p>(C) 對欄位的操作 (例如 sub_str, cast)</p> <p>(D) 根據單一條件篩選</p>
A	<p>44. 在 Python 3 下，執行附圖程式碼後，輸出結果為何？</p> <pre>def mydivide(a,b): assert b!=0, "除數不能為 0" return a/b r=0 try: r= mydivide(4,0) except: r=-1 print("印出 r 的值", r)</pre> <p>(A) 印出 r 的值-1</p> <p>(B) 印出 r 的值 0</p> <p>(C) 印出 r 的值"除數不能為 0"</p> <p>(D) 程式中斷，印出訊息："除數不能為 0"</p>
B	<p>45. R 語言中，關於 browser() 函數，下列敘述何者「不」正確？</p> <p>(A) 將一步一步執行程式碼，進入偵測模式</p> <p>(B) 偵測模式中，Q 指令可執行下一行</p> <p>(C) 偵測模式中，r 指令可重新啟動 resume</p> <p>(D) 偵測模式中，c 指令可繼續執行函數直到結束</p>
B	<p>46. Python 語言中，執行附圖程式碼後會產生錯誤，請問下列何者為主要原因？</p> <pre>x=[3] y=x+3 z=x*3 print(z)</pre> <p>(A) 第一行錯誤，變數 x 設定方式錯誤</p> <p>(B) 第二行錯誤，不可以使用「+」運算子</p> <p>(C) 第三行錯誤，不可以使用「*」運算子</p> <p>(D) 第四行錯誤，輸出指令錯誤</p>
D	<p>47. 在 Python 語言，所有程式錯誤 (Error) 都是透過例外 (Exception) 來處理，但並非所有例外都是程式錯誤，關於例外與其意義對應，下列</p>

111 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：111 年 05 月 28 日

第 11 頁，共 11 頁

	<p>何者正確？</p> <p>(A) ImportError：資料型態錯誤</p> <p>(B) SyntaxError：縮排錯誤</p> <p>(C) IndentationError：索引錯誤</p> <p>(D) NameError：識別字錯誤</p>
A	<p>48. 關於 Python 中的 traceback() 函數，下列敘述何者正確？</p> <p>(A) 可以協助獲得錯誤訊息的額外資訊</p> <p>(B) 若有錯誤發生，無法顯示錯誤發生於何處</p> <p>(C) 若有錯誤發生，不會顯示錯誤訊息</p> <p>(D) 未顯示錯誤訊息的程式碼，代表邏輯與語法均無誤</p>
<p>A B C 皆 給 分</p>	<p>49. 在 R 語言中，程式碼執行時會產生訊息，下列敘述何者正確？</p> <p>(A) 產生錯誤 (Errors) 訊息時，將繼續執行無錯誤的程式</p> <p>(B) 警告 (Warnings) 訊息會說明潛在的問題</p> <p>(C) 一般傳回的訊息在於說明代碼輸出的結果</p> <p>(D) 產生警告 (Warnings) 訊息時，必須把警告問題解決才可繼續執行程式</p>
B	<p>50. Python 語言中，關於提升程式效能，下列敘述何者「不」正確？</p> <p>(A) 採用向量化資料處理方式，避免額外迴圈執行</p> <p>(B) 使用多執行緒必定可以提升程式效能</p> <p>(C) 合適的平行處理技術可以提升程式效能</p> <p>(D) 在讀取資料庫時，僅取得必要資料，避免佔用過多記憶體</p>