

初級巨量資料分析師能力鑑定樣題

科目 1：資料導向程式設計

第 1 頁，共 10 頁

單選題 50 題 (佔 100%)

C	<p>1. 矩陣中，線性獨立的行向量或線性獨立的列向量稱為？</p> <p>(A) 相關矩陣 (Correlation matrix)</p> <p>(B) 單位矩陣 (Identity matrix)</p> <p>(C) 秩 (Rank)</p> <p>(D) 反矩陣 (Inverse matrix)</p>																																				
B	<p>2. 在 Python 語言中，已知「<code>a=numpy.array([[1, 3],[2, 4]])</code>」，則 <code>3*a</code> 意義為何？</p> <p>(A) <code>a*a*a</code></p> <p>(B) 3 乘以每個元素</p> <p>(C) 3 乘以第 1 行元素</p> <p>(D) 3 乘以第 3 行元素</p>																																				
B	<p>3. 請問在觀察分類模型之分類結果，通常會使用下列何種矩陣？</p> <p>(A) 稀疏矩陣</p> <p>(B) 混淆矩陣</p> <p>(C) 相似度矩陣</p> <p>(D) 對角矩陣</p>																																				
A	<p>4. 下面之矩陣最可能為何種矩陣？</p> <table style="margin-left: auto; margin-right: auto; border-collapse: collapse;"> <thead> <tr> <th></th> <th>a</th> <th>b</th> <th>c</th> <th>d</th> <th>e</th> </tr> </thead> <tbody> <tr> <td>0</td> <td>0</td> <td>1</td> <td>0</td> <td>6</td> <td>0</td> </tr> <tr> <td>1</td> <td>0</td> <td>0</td> <td>0</td> <td>0</td> <td>0</td> </tr> <tr> <td>2</td> <td>1</td> <td>0</td> <td>0</td> <td>0</td> <td>0</td> </tr> <tr> <td>3</td> <td>0</td> <td>0</td> <td>0</td> <td>0</td> <td>0</td> </tr> <tr> <td>4</td> <td>9</td> <td>0</td> <td>0</td> <td>0</td> <td>1</td> </tr> </tbody> </table> <p>(A) 稀疏矩陣</p> <p>(B) 混淆矩陣</p> <p>(C) 相似度矩陣</p> <p>(D) 對角矩陣</p>		a	b	c	d	e	0	0	1	0	6	0	1	0	0	0	0	0	2	1	0	0	0	0	3	0	0	0	0	0	4	9	0	0	0	1
	a	b	c	d	e																																
0	0	1	0	6	0																																
1	0	0	0	0	0																																
2	1	0	0	0	0																																
3	0	0	0	0	0																																
4	9	0	0	0	1																																
B	<p>5. 關於稀疏矩陣 ($n \times m$)，下列敘述何者不正確？</p> <p>(A) 多數的元素為零</p> <p>(B) 二維陣列 ($n \times m$) 為常見且有效率之儲存方式</p> <p>(C) 常使用於文字探勘與推薦系統</p> <p>(D) 使用不同的儲存方式可能影響記憶體之使用空間</p>																																				
B	<p>6. 考慮 R 語言運算，使用 <code>read.table</code> 函數匯入文字檔的資料結構為何？</p> <p>(A) 串列 (List)</p>																																				

初級巨量資料分析師能力鑑定樣題

科目 1：資料導向程式設計

第 2 頁，共 10 頁

	<p>(B) 資料框 (Data Frame)</p> <p>(C) 矩陣 (Matrix)</p> <p>(D) 向量 (Vector)</p>
C	<p>7. 考慮 Python 語言運算「<code>a=list(range(3));b=list(range(4));c=[a,b]</code>」，則「<code>c[1][2]</code>」結果為何？</p> <p>(A) 0</p> <p>(B) 1</p> <p>(C) 2</p> <p>(D) 3</p>
C	<p>8. 請問 R 語言內建的資料結構中，何者與 JSON (JavaScript Object Notation) 最為相近？</p> <p>(A) 向量 (Vector)</p> <p>(B) 矩陣 (Matrix)</p> <p>(C) 串列 (List)</p> <p>(D) 資料框架 (Date Frame)</p>
B	<p>9. 關於 Python 中內建字典 (dict) 結構，下列敘述何者不正確？</p> <p>(A) 與 JSON 的結構相似</p> <p>(B) 是一有序集合</p> <p>(C) 可儲存任意類型的資料</p> <p>(D) 由鍵 (Key) 與值 (Value) 對應所組成的</p>
A	<p>10. 下列何者所描述的是結構化資料 (Structured Data) ？</p> <p>(A) 存在關聯式資料庫中的營養成分資料集</p> <p>(B) SHP 格式的河川地理圖資</p> <p>(C) JSON 型態儲存的鐵路時刻表</p> <p>(D) 以 XML 來紀錄的觀光資料</p>
C	<p>11. ACID，是指資料庫管理系統 (DBMS) 在寫入或更新資料的過程中，為保證事務 (transaction) 是正確可靠的，所必須具備的四個特性：原子性 (atomicity，或稱不可分割性)、一致性 (consistency)、隔離性 (isolation，又稱獨立性)、持久性 (durability)。關於此四個特性，下列何者不正確？</p> <p>(A) 原子性：一個交易 (transaction) 中的所有操作，只有全部完成或者全部不完成，不會結束在中間某個環節</p> <p>(B) 一致性：在交易開始之前和交易結束以後，資料庫的完整性沒有被破壞</p> <p>(C) 隔離性：資料庫與資料庫之間的隔離性可以防止互相衝突導致數據錯誤</p> <p>(D) 持久性：交易處理結束後，對數據的修改就是永久的，即便系統</p>

初級巨量資料分析師能力鑑定樣題

科目 1：資料導向程式設計

第 3 頁，共 10 頁

	故障也不會丟失
A	<p>12. 某大型電商需儲存的資料量一天大約為 1TB，選用 HDFS (Hadoop Distributed File System) 進行儲存，下列敘述何者不正確？</p> <p>(A) HDFS 無法儲存圖片</p> <p>(B) 若為二維陣列資料，CSV (Comma-Separated Values) 是可以儲存至 HDFS 的資料格式</p> <p>(C) Hadoop 系統的資料儲存都與 HDFS 有關</p> <p>(D) HBase (Hadoop database) 是架構在 HDFS 上的分散式資料庫</p>
D	<p>13. 在資料表正規化的過程 (1NF 到 BCNF) 中，每一個階段都是以欄位的「相依性」作為分割資料表的依據之一，關於正規化步驟之敘述，下列何者不正確？</p> <p>(A) 第一正規化型式：除去重覆群</p> <p>(B) 第二正規化型式：除去部分相依</p> <p>(C) 第三正規化型式：除去遞移相依</p> <p>(D) Boyce-Code 正規化型式：除去多值相依</p>
C	<p>14. 您是一位資料庫設計師，在 NoSQL 資料庫設計特性中，下列敘述何者不正確？</p> <p>(A) 資料存儲可以不需要固定的表格模式</p> <p>(B) 資料平台具有水平擴充能力</p> <p>(C) 採用標準 SQL 語法來查詢</p> <p>(D) 使用 Key-Value 資料模式</p>
D	<p>15. 您是一位巨量資料分析師，關於 NoSQL 資料庫的設計特性，下列敘述何者不正確？</p> <p>(A) 使用 Key-Value 資料模式</p> <p>(B) 資料查詢透過 API 查詢</p> <p>(C) Google Big Table 為 NoSQL 類型</p> <p>(D) 需預先設計固定的 Schema 欄位</p>
D	<p>16. 以下何者不是 NoSQL 資料庫？</p> <p>(A) MongoDB, CouchDB</p> <p>(B) Apache Cassandra, Apache HBase</p> <p>(C) Redis, Memcached</p> <p>(D) MariaDB, SQLite</p>
C	<p>17. 關於以下 SQL 語法的排列，請問順序應為何？</p> <ol style="list-style-type: none"> 1. FROM 選課資料表 2. ORDER BY 成績 Desc 3. SELECT 學號, 課號, 成績 4. WHERE 課程代號='A01'

初級巨量資料分析師能力鑑定樣題

科目 1：資料導向程式設計

第 4 頁，共 10 頁

	<p>(A) 1324</p> <p>(B) 2134</p> <p>(C) 3142</p> <p>(D) 4312</p>
B	<p>18. 關於 NoSQL 的敘述，下列敘述何者不正確？</p> <p>(A) NoSQL 全名為 Not Only SQL</p> <p>(B) NoSQL 的資料儲存有固定的表格模式，方便後續查詢</p> <p>(C) NoSQL 資料庫較能處理大量讀寫週期、眾多用戶的資料</p> <p>(D) NoSQL 資料庫是半結構和非結構式資料的理想資料庫</p>
C	<p>19. 請問要在學生資料表中查詢姓名結尾為"凱"的學生基本資料，其 SQL 語法為何？</p> <p>(A) <code>SELECT * FROM 學生資料表 WHERE 姓名 is '凱'</code></p> <p>(B) <code>SELECT * FROM 學生資料表 WHERE 姓名 is '凱%'</code></p> <p>(C) <code>SELECT * FROM 學生資料表 WHERE 姓名 Like '%凱'</code></p> <p>(D) <code>SELECT * FROM 學生資料表 WHERE 姓名 Like '凱%'</code></p>
A	<p>20. 關於關聯式資料庫與 NoSQL 資料庫的比較，下列敘述何者不正確？</p> <p>(A) 對於複雜的查詢而言，NoSQL 有較強的查詢能力</p> <p>(B) MariaDB 屬於關聯式資料庫，而 MongoDB 則是屬於 NoSQL 資料庫</p> <p>(C) 關聯式資料庫必須透過資料庫的 Schema 欄位架構來確立資料表之間的關聯，NoSQL 資料庫則是改用 Key-Value 資料模式來解決龐大資料的異動困難</p> <p>(D) 關聯式資料庫與 NoSQL 資料庫並不是互斥的概念，我們可以在系統裡面用關聯式資料庫系統儲存文章、評論，同時也用 NoSQL 資料庫來搜集使用者資訊</p>
B	<p>21. Java 中，關於建構元 (Constructor)，下列敘述何者不正確？</p> <p>(A) 是產生物件實例時，用來作為初始化工作的函式</p> <p>(B) 會回傳整數值 0 或 -1，以表示物件產生是否成功</p> <p>(C) 建構元函式名稱和類別名稱一樣</p> <p>(D) 不會回傳任何結果</p>
A	<p>22. Java 中，封裝 (Encapsulation) 為物件導向程式設計的方法之一，下列敘述何者不正確？</p> <p>(A) 透過 <code>private</code> 將函數包裝起來，外部其他物件無法存取</p> <p>(B) 被定義為 <code>public</code> 的函數，可以被其他物件呼叫使用</p> <p>(C) 基於安全考量，被定義為 <code>private</code> 的屬性的變數，只能被同類別下的 <code>private</code> 方法所讀取，無法被同類別下的 <code>public</code> 方法存取</p> <p>(D) 被定義為 <code>private</code> 的方法，可以透過繼承的方式覆寫</p>

初級巨量資料分析師能力鑑定樣題

科目 1：資料導向程式設計

第 5 頁，共 10 頁

C	<p>23. Java 中，可以呼叫相同函式，卻可以因參數屬性不同而出現不同行為，此種現象是物件導向特性中的哪一項特性？</p> <p>(A) 封裝 (Encapsulation)</p> <p>(B) 繼承 (Inheritance)</p> <p>(C) 多型 (Polymorphism)</p> <p>(D) 實例 (Instance)</p>
B	<p>24. 下列為 Python 程式的物件導向概念：</p> <pre>class c1: def __init__(self): self.i = 0 def add(self): self.i += 2</pre> <p>c = c1() c.add() print(c.i)</p> <p>請問顯示出的 c.i 值為下列何者？</p> <p>(A) 0</p> <p>(B) 2</p> <p>(C) 4</p> <p>(D) 8</p>
C	<p>25. 下列為 Python 程式的物件導向概念：</p> <pre>class C1: a = 10 class C2(C1): a = 20</pre> <p>c2 = C2()</p> <p>請問 c2.a 的值為下列何者？</p> <p>(A) 0</p> <p>(B) 10</p> <p>(C) 20</p> <p>(D) 40</p>
B	<p>26. 在物件導向程式設計中，父類別 (Parent class) 或稱基礎類別 (Base class) 與子類別 (Child class) 或稱衍生類別 (Derived class) 的關係比較接近於下列何選項？</p> <p>(A) 西瓜與蘋果</p> <p>(B) 哺乳類與人類</p> <p>(C) 鍵盤與滑鼠</p>

初級巨量資料分析師能力鑑定樣題

科目 1：資料導向程式設計

第 6 頁，共 10 頁

(D) Word 與 Excel	
C	<p>27. 以下為 Java 的物件建構流程，請問其排列順序應為何？</p> <ol style="list-style-type: none"> 1. 初始化 2. 儲存物件參考 3. 實體化 4. 宣告 5. 執行建構式 <p>(A) 41235 (B) 41253 (C) 43152 (D) 43125</p>
B	<p>28. 關於 Java 繼承，下列敘述何者不正確？</p> <p>(A) Java 使用 extends 來表示繼承的關係 (B) 在繼承關係的架構下，子類別變數名稱= new 父類別建構子() (C) Java 的類別只能單一繼承 (D) 當子類別繼承父類別，就擁有父類別中的屬性、方法與建構式</p>
B	<p>29. 在 Python 語言中可用 open() 函數開啟檔案，若僅要讀取檔案，mode 引數需使用下列何項設定值？</p> <p>(A) a (B) r (C) w (D) wb</p>
A	<p>30. 下列何者為換行符號？</p> <p>(A) \n (B) \r (C) \s (D) \t</p>
A	<p>31. 關於 Python 語言讀取檔案，下列敘述何者不正確？</p> <p>(A) 使用 open("file", "r ") 若 file 不存在，會創建 file (B) 使用 open("file", " w+") 若 file 不存在，會創建 file (C) 使用 open("file", " w+") 為可讀寫的檔案模式 (D) 使用 open("file", " a+") 為附加讀寫檔案模式，寫入時添加於後，亦可以讀檔</p>
C	<p>32. 正規表示式可運用於多處，可使用下列何者搜尋出空白字元？</p> <p>(A) \n (B) \r (C) \s</p>

初級巨量資料分析師能力鑑定樣題

科目 1：資料導向程式設計

第 7 頁，共 10 頁

	(D) \t
A	<p>33. 在使用分散式資料庫提升系統平行效率的技術，下列敘述何者正確？</p> <p>(A) 使用分段 (partition) 技術</p> <p>(B) 使用覆寫 (overwrite) 技術</p> <p>(C) 使用串流 (streaming) 技術</p> <p>(D) 使用主從 (master-slave) 技術</p>
C	<p>34. JSON 是目前在資料匯入與匯出時很常見到的格式，關於 JSON 的敘述，下列何者不正確？</p> <p>(A) JSON 的格式較 XML 來的精簡，所需的儲存容量也比較小</p> <p>(B) JSON 的資料格式是以「鍵 - 值」(key:value) 方式配對儲存</p> <p>(C) 在 JSON 格式中，鍵 (key) 可以是一個數值或字串，字串需用雙引號表示</p> <p>(D) 在 JSON 格式中，值可以是數值、字串、布林值、陣列或是 null 值</p>
D	<p>35. 請問下列何者不是 Python 語言中內建的資料型態？</p> <p>(A) list</p> <p>(B) tuple</p> <p>(C) dict</p> <p>(D) array</p>
C	<p>36. 關於資料導向程式設計，下列敘述何者不正確？</p> <p>(A) 精進記憶體管理以提升程式效能</p> <p>(B) 經常注意經函數處理後輸出的物件其類別型態</p> <p>(C) 多採用條件式語法，避免運用邏輯值索引</p> <p>(D) 加強撰寫平行化程式的技巧可有助提升效能</p>
B	<p>37. 關於程式碼錯誤處理，下列敘述何者不正確？</p> <p>(A) 撰寫程式時讓錯誤可以重製，可便於偵查</p> <p>(B) 可執行的程式即為無誤的程式</p> <p>(C) 可先要求程式碼的邏輯正確，再設法提升執行效率</p> <p>(D) 可藉由函數輔助處理例外狀況與瞭解錯誤訊息</p>
B	<p>38. 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，請問流程順序何者正確？</p> <ol style="list-style-type: none"> 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 <p>(A) 1>2>3>4>5</p>

初級巨量資料分析師能力鑑定樣題

科目 1：資料導向程式設計

第 8 頁，共 10 頁

	<p>(B) 5>3>1>4>2</p> <p>(C) 5>4>3>2>1</p> <p>(D) 3>5>1>2>4</p>
A	<p>39. 在 R 語言中，程式碼執行時會產生訊息，下列敘述何者不正確？</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，可繼續執行程式</p>
A	<p>40. 關於進行資料導向程式設計時，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</p> <p>(B) 留意元素型別可能被強制轉換的情況</p> <p>(C) 善用自定義函數模組化工作流程</p> <p>(D) 注意引用函數其預設的引數值</p>
C	<p>41. 關於 Python，下列敘述何者不正確？</p> <p>(A) range(10)可以產生 0, 1, ...9 的數字序列</p> <p>(B) continue 語法可以進入下一次循環</p> <p>(C) pass 語法可以產生判斷邏輯</p> <p>(D) 使用 if...else 進行條件判斷</p>
C	<p>42. 關於 Python 自訂函數，下列敘述何者不正確？</p> <p>(A) 關鍵字 def 可以定義一個自訂函數</p> <p>(B) 關鍵字 return 可以傳回值</p> <p>(C) 關鍵字 pass 可以傳回值</p> <p>(D) 敘述 def A(n=10)可以預設參數 n 等於 10</p>
B	<p>43. 「例外狀況處理函數」之目的為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()</p> <p>(B) suppressMessages()</p> <p>(C) tryCatch()</p> <p>(D) try()</p>
A	<p>44. 關於程式除錯與效率監測 R 函數，下列敘述何者不正確？</p> <p>(A) debug()可關閉函數的偵錯模式</p> <p>(B) suppressMessages()可用來壓制錯誤訊息</p> <p>(C) traceback()可以協助獲得錯誤訊息的額外資訊</p> <p>(D) system,time()可以衡量程式碼的執行時間</p>
D	<p>45. 處理 R 語言程式碼的例外狀況、錯誤與訊息時，通常不會使用下列何</p>

初級巨量資料分析師能力鑑定樣題

科目 1：資料導向程式設計

第 9 頁，共 10 頁

	<p>種函數？</p> <p>(A) stop()</p> <p>(B) warning()</p> <p>(C) message()</p> <p>(D) system.time()</p>
C	<p>46. 請問以下哪個程式錯誤類型是指程式可以正常執行並且正常輸出，但輸出的結果與預期不一致？</p> <p>(A) 語法錯誤 (Syntax Error)</p> <p>(B) 執行期間錯誤 (Runtime Error)</p> <p>(C) 邏輯錯誤 (Logic Error)</p> <p>(D) 編譯時期的錯誤 (Compile Error)</p>
C	<p>47. 請問以下敘述執行後輸出的結果為何？</p> <pre>try: x=0 y.split() y=0 except: z=0 print(z) print(y)</pre> <p>(A) y 與 z 皆可正常輸出</p> <p>(B) y 與 z 皆會發生錯誤</p> <p>(C) y 會發生錯誤，z 可正常輸出</p> <p>(D) y 可正常輸出，z 會發生錯誤</p>
D	<p>48. 關於 Python 程式效能提升方法，下列敘述何者不正確？</p> <p>(A) 可以優化迴圈執行</p> <p>(B) 使用較新版本 Python</p> <p>(C) 改進程式演算法</p> <p>(D) 減少函式的使用</p>
C	<p>49. 某專案採用 Python 語言來設計，請問 Python 的變數名稱中，「10account」是屬於哪一種程式錯誤類型？</p> <p>(A) 程式邏輯設計錯誤</p> <p>(B) 程式執行結果錯誤</p> <p>(C) 程式語法錯誤</p> <p>(D) 程式判斷流程錯誤</p>
B	<p>50. 請問以下 Python 程式碼執行後其結果為何？</p> <pre>try:</pre>

初級巨量資料分析師能力鑑定樣題

科目 1：資料導向程式設計

第 10 頁，共 10 頁

```
X = 10
Y = 0
Z = X / Y
except ZeroDivisionError:
    print("除數不得為 0")
except Exception as e1:
    print(e1.args)
else:
    print("沒有捕捉到例外！X 除以 Y 的結果等於 Z ")
finally:
    print("離開 try...except 區塊")
```

- (A) 除數不得為 0
- (B) 除數不得為 0；離開 try...except
- (C) 除數不得為 0；沒有捕捉到例外！X 除以 Y 的結果等於 Z；離開 try...except
- (D) 無法執行