**Task #07 作業 01【簡答題】統計學基本觀念**

1. 請問「質化資料 (qualitative data) 」與「量化資料 (quantitative data) 」差別是什麼？

質化資料通常不能用數字直接衡量，表示類別、屬性或特徵描述(ex:性別、顏色、品牌…)。

量化資料由數字組成，可運用數學運算進行分析(ex:身高、收入、年齡…)。

2. 描述統計中有三組指標「最大值、最小值和總和」、「眾數、均值和中位數」和「變異數、標準差、斜方差和分位數」 ，請問這三組分別可以看出資料的哪些趨勢？

「最大值、最小值和總和」可協助了解資料的範圍大小，以及數據的總體規模或總量。

「眾數、均值和中位數」可反映數據的集中趨勢。

「變異數、標準差、斜方差和分位數」可反映數據的分散趨勢。

3. 相關係數分析能夠看出兩組資料間的關係，請問什麼叫做「正相關」、「無相關」與「負相關」？

「正相關」指兩組數據的關係呈現同方向變動的趨勢，當一組數據增加時，另一組也會增加。

「負相關」指兩組數據的關係呈現反方向變動的趨勢，當一組數據增加時，另一組會減少。

「無相關」指兩組數據並無明顯的線性關係，一組數據的變動不會受另外一組數據的影響。

4. 在統計中有一句話叫做「Correlation is not causation.」，這句話的想表達的含義是什麼呢？

「相關性不代表因果關係」代表兩組數據即使存在相關性，也不一定代表一組數據的改變會影響另一組數據的變化。

例如直播課程中提到的冰淇淋銷售及防曬乳銷售的關係，兩者存在著正相關(天氣炎熱大家都會去買冰淇淋或防曬乳，但不代表冰淇淋銷售增加會導致防曬乳的銷售也增加)。

5. 模型的誤差可以分為「Bias」和「Variance」，請解釋這兩者分別代表什麼誤差？

「Bias」表示模型預測值與實際真實值之間的差異，通常是代表模型過於簡單，無法捕捉數據的複雜性而引起。

「Variance」指模型在數據上的預測值之間的變動程度，反映模型對於數據的敏感性。