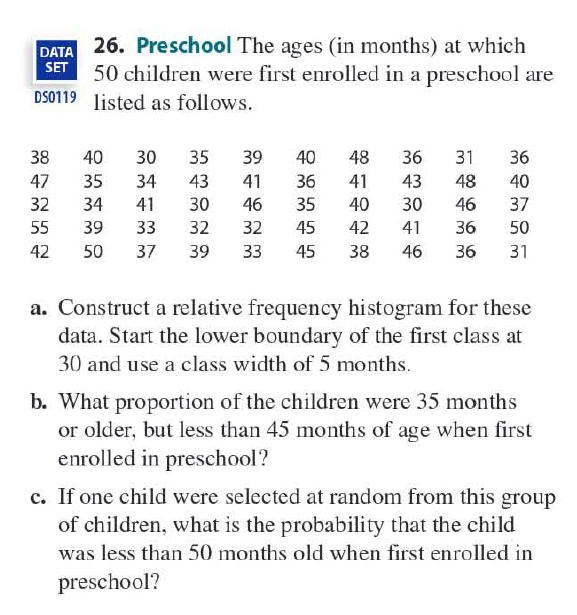
HW0913

CH01.04\_Q26 編碼的意思  Chapter 1, Section 4, Question 26.

這次的作業有6題 如下

|  |  |
| --- | --- |
| 1 | CH01.04\_Q26 |
| 2 | CH02.01\_Q02 |
| 3 | CH02.02\_Q14 |
| 4 | CH02.04\_Q02 |
| 5 | CH02.04\_Q20 |
| 6 | CH02.04\_Q24 |

1.



50 筆數據，範圍從 **30 到 55**

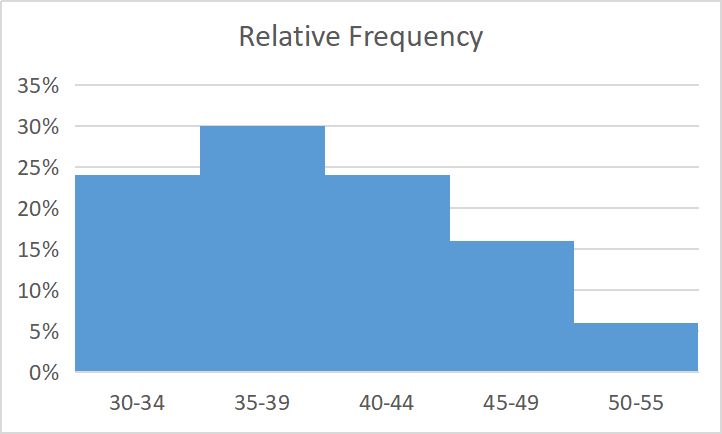
|  |  |
| --- | --- |
| AVG | 39.08 |
| MAX | 55 |
| MIN | 30 |

第一類的下限設為30，類寬設為5個月。可以將數據分成 5 個區間（bins），

相對頻率 = 該區間筆數 ÷ 總筆數（50）

|  |  |  |
| --- | --- | --- |
| Range | f | Relative Frequency |
| 30-34 | 12 | 24% |
| 35-39 | 15 | 30% |
| 40-44 | 12 | 24% |
| 45-49 | 8 | 16% |
| 50-55 | 3 | 6% |
|  | 50 |  |

a.



b. 年齡在 35 個月及以上，但未滿 45 個月的兒童佔比是多少？54%

|  |  |  |
| --- | --- | --- |
| Range | f | Relative Frequency |
| 35-39 | 15 | 30% |
| 40-44 | 12 | 24% |
| Total |  | 54% |

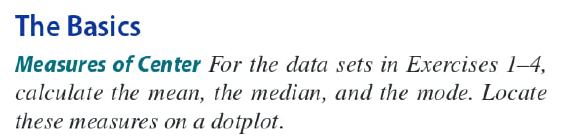
c. 如果從這組孩子中隨機選擇一個孩子，那麼這個孩子第一次入讀幼稚園時，

年齡不到 50 個月的機率是多少？

從資料中找出小於 50 的數值：總人數 = 50 小於 50 的人數 = 47

所以機率為：47/50=94%

2.



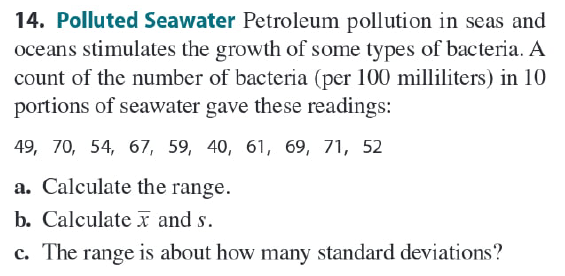


"mean" 指的是平均數（或稱平均值、算術平均數）: 4

“median”（中位數）是一組按順序排列的數據中居於中間位置的數值計算方法，是先將所有數值由小到大排列，若為偶數，則為中間兩個數值的平均數: 4

“mode” 眾數指一組數據中出現次數最多的數據值: 3, 4, 5都各出現2次

3.



1. 全距（Range, R） 指的是一組數值資料中，最大值與最小值之間的差

最大值: 是71。

最小值: 是 40。

全距: 為71 - 40 = 31

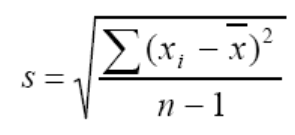
b. 樣本平均數(Sample Mean): x̄ (x-bar)

加總：將樣本中的所有數據值加總。

除以數量：將總和除以樣本的數量（即數據的個數）。

592/10=59.2

s =樣本標準差(standard deviation)用於表示資料之離散程度，若由母體中抽取n個  
樣本，其值分別為n x , x ,..., x 1 2 ，其樣本標準差計算如下：



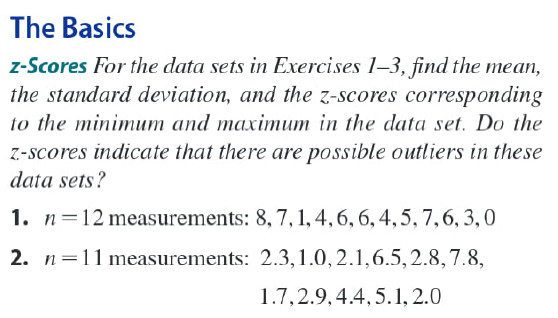
計算每個數據值與平均值之間的偏差，並將這些偏差平方故平方和: 967.6

變異數: 967.6/(總樣本數-1)=107.511111

平方根結果為標準差10.36875649

c. Range 31/ standard deviation 10.36875649 = 2.989750992

4.



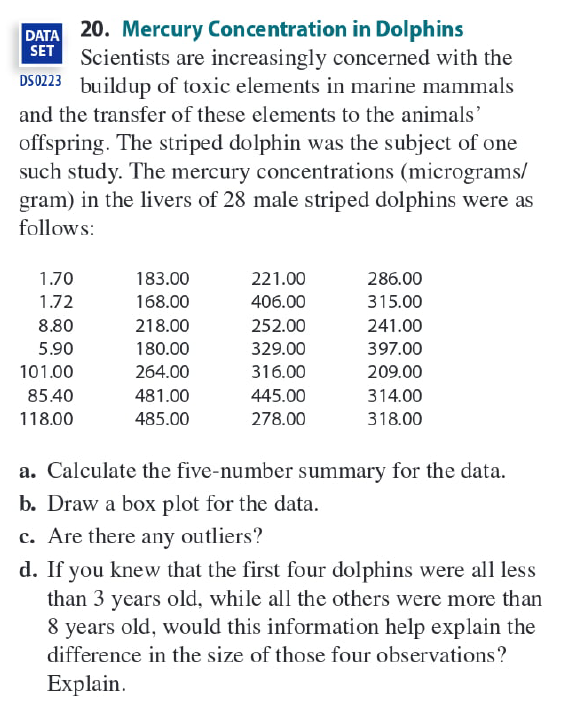
|  |  |
| --- | --- |
| mean | 3.509090909 |
| s | 4.684909091 |

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  | z-scores |
| min | 1 |  | -0.535568751 |
| max | 7.8 |  | 0.915900182 |

|  |  |
| --- | --- |
| Position | Value |
| Q1=0.25(11+1)=3 | 2.1 |
| Q3=0.75(11+1)=9 | 4.4 |
| IQR=Q3-Q1=6 | 7.8 |
| Q3+1.5\*7.8=16.1 |  |
| Q1-1.5\*7.8=-9.6 |  |

沒有outliers, 因數據中沒有大於16.1或小於-9.6的數值.

5.

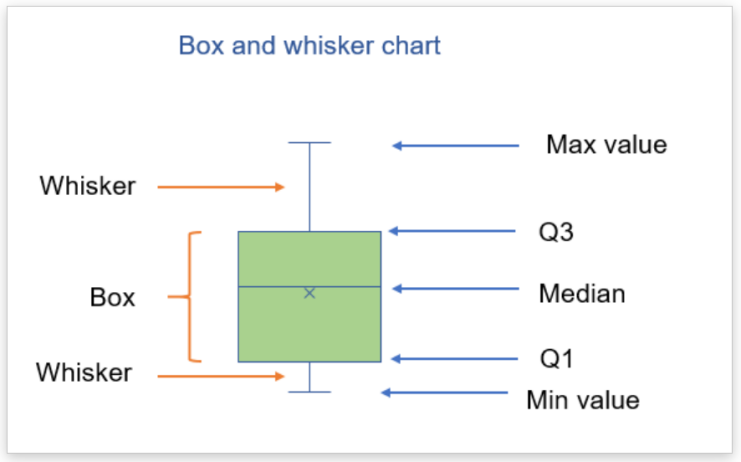
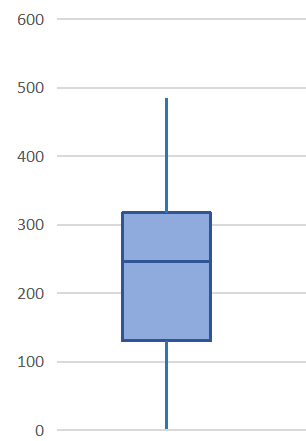


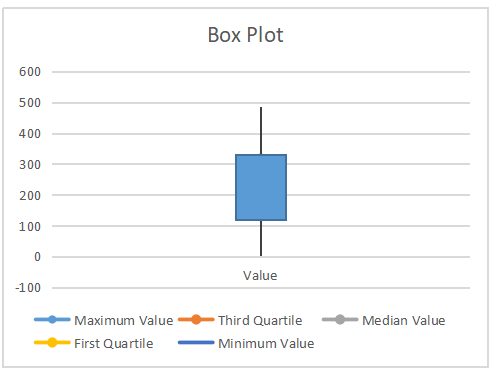
a.

|  |  |
| --- | --- |
| Min | 1.7 |
| Position | Value |
| Q1=0.25(28+1)=7.25 | 130.5 |
| Q2=0.5(28+1)=14.5 | 246.5 |
| Q3=0.75(28+1)=21.75 | 317.5 |
| Max | 485 |

b.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Value** | |  | **Differences** | |
| **Minimum Value** | 1.7 |  | **Minimum Value** | 1.7 |
| **First Quartile** | 130.5 |  | **First Quartile-Minimum** | 128.8 |
| **Median Value** | 246.5 |  | **Median-First Quartile** | 116 |
| **Third Quartile** | 317.5 |  | **Third Quartile-Median** | 71 |
| **Maximum Value** | 485 |  | **Maximum-Third Quartile** | 167.5 |



c.

|  |  |
| --- | --- |
| lower\_fence = Q1-1.5\*IQR | -150 |
| upper\_fence =  Q3 +1.5\*IQR | 598 |

沒有outliers, 因數據中沒有大於598或小於-150的數值.

d. 年齡資訊是否有助於解釋前四筆異常值？

其中 **1.70 μg/g** 顯著低於其他值。

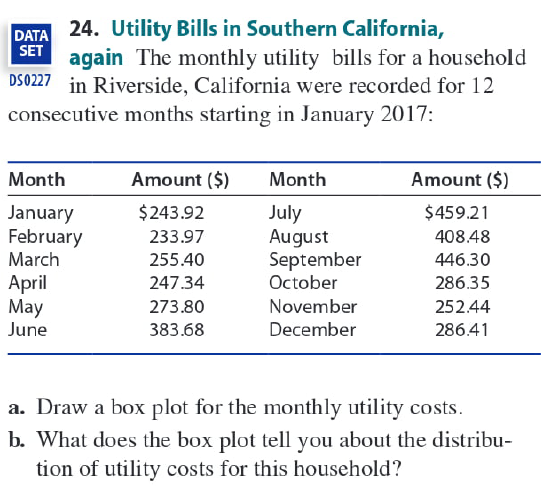
若知道前四隻海豚年齡 < 3 歲，而其他皆 > 8 歲，這可能反映：

年幼海豚尚未累積汞 → 生理代謝與暴露時間短

年長海豚因食物鏈累積汞 → 濃度較高

✅ **結論**：年齡資訊提供了生物學上的合理解釋，支持這些低值並非測量錯誤，而是年齡差異造成的自然變異。

6.



a.

|  |  |
| --- | --- |
| **Value** | |
| **Minimum Value** | 233.97 |
| **First Quartile** | 248.615 |
| **Median Value** | 280.075 |
| **Third Quartile** | 402.28 |
| **Maximum Value** | 459.21 |

b. 盒鬚圖告訴我們什麼？

**偏態分布**：中位數偏向下半部，表示資料右偏（右尾長），夏季水電費較高。

**變異大**：Q1 到 Q3 的範圍（IQR）約為 $134.31，顯示水電費波動明顯。

**高值集中在夏季**：最大值 $459.21 出現在 7 月，反映冷氣使用量大幅提升。

**無明顯離群值**：所有值都在合理範圍內，無極端異常。

✅ 結論：這個家庭的水電費呈現季節性波動，夏季顯著偏高，整體分布右偏，反映出氣候對能源消耗的影響。