Assignment 1 Solution (Week 1 - 2)

STAT 2601 - Business Statistics

School of Mathematics and Statistics, Carleton University

Total Marks: 27

Q1: [18] Revenue of Canadian Banks/Financial Institutions

(a) **Descriptive Statistics:**

Mean: $\bar{x} = \frac{\sum x}{n}$[.5] $= \frac{8+9+\cdots+48}{10}$[.5] $\approx$ 19.4 billion$[.5]

Median:

(a) **Step 1: Sort** 8, 9, 11, 14, 15, 17, 19, 24, 29, 48. [.5]

(b) **Step 2: Position** $(n+1)\frac{1}{2} = (10+1)\frac{1}{2}$[.5] $= 5.5$[.5]

(c) **Step 3: Median** $15+0.5(17-15)$[.5] $=$ 16 billion $[.5] (If anyone calculates the median based on the observation halfway between 5th and 6th in the sorted data without method of linear interpolation, it will also be accepted)

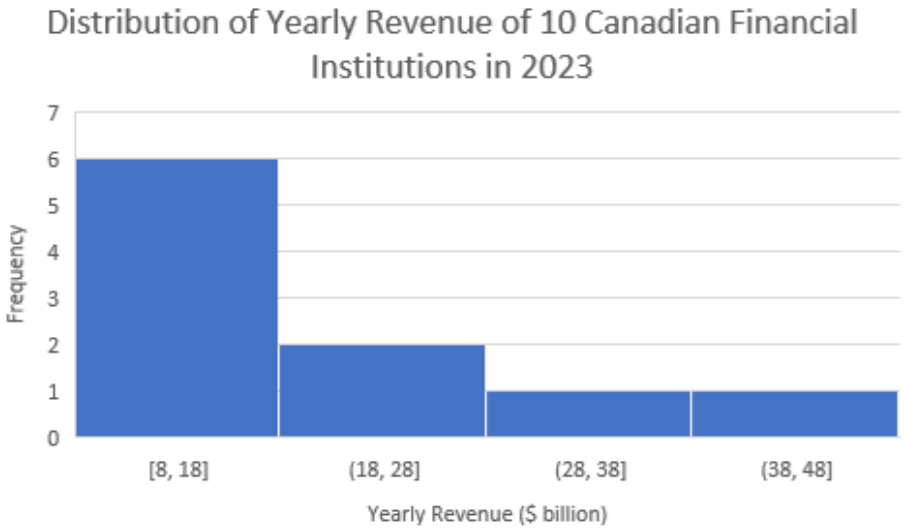
Standard Deviation:

$$s = \sqrt{\frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n-1}}[.5] = \sqrt{\frac{5058 - \frac{(194)^2}{10}}{10-1}}[.5] \approx 11.99 \text{ billion } \$[.5]$$

Shape: Since mean (19.4 billion \$) $>$ median (16 billion \$)[.5], the distribution is right-skewed (or positively skewed)[.5].

Recommended Measure of Central Tendency: For skewed distribution, median [.5] is the recommended measure of central tendency as it is not affected by extreme values.[.5]

(b) **Histogram:** [2; .5 for each - chart title, axis title, number of bins (4), histogram]



(c) Steps of Calculation of Q_3:

(i) Step 1 (Sort): 8, 9, 11, 14, 15, 17, 19, 24, 29, 48.

(ii) Step 2 (Position): $(n+1)\frac{3}{4} = (10+1)\frac{3}{4}$[0.5] $= 8.25$[0.5].

(iii) Step 3 (Method of Linear Interpolation): $Q_3 = 24 + 0.25(29-24)$[0.5] $= 25.25$ billion$[0.5].

Box-and-Whiskers Plot:

(i) Lower Limit: $Q_1 - 1.5(Q_3 - Q_1) = 10.5 - 1.5(25.25 - 10.5)$[0.5] $= -11.625$[0.5].

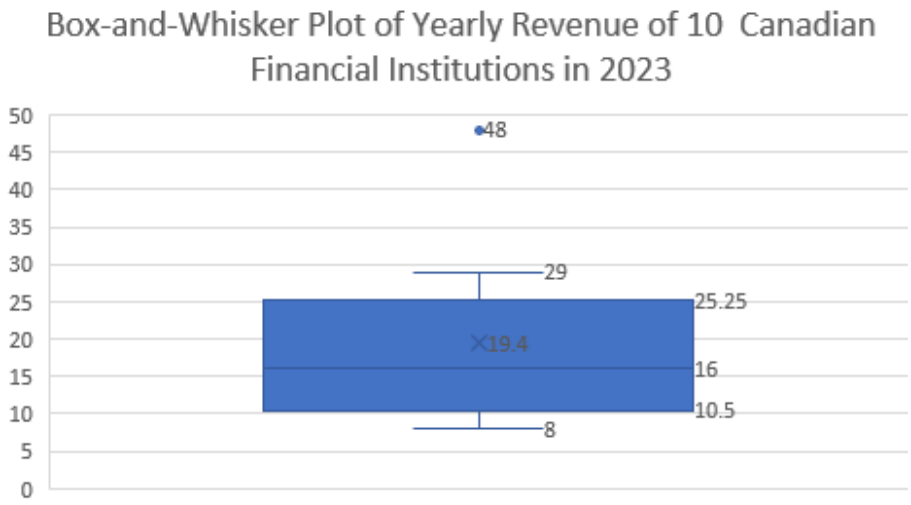
(ii) Upper Limit: $Q_3 + 1.5(Q_3 - Q_1) = 25.25 + 1.5(25.25 - 10.5)$[0.5] $= 47.375$[0.5].

(iii) Lower Whisker: Extends up to 8[0.5].

(iv) Upper Whisker: Extends up to 29[0.5].

(v) Outlier: Yearly revenue of RBC in the sampled Canadian financial institutions in 2023 is considered outlier as it is beyond (-11.625 billion$, 47.375 billion$)[0.5].

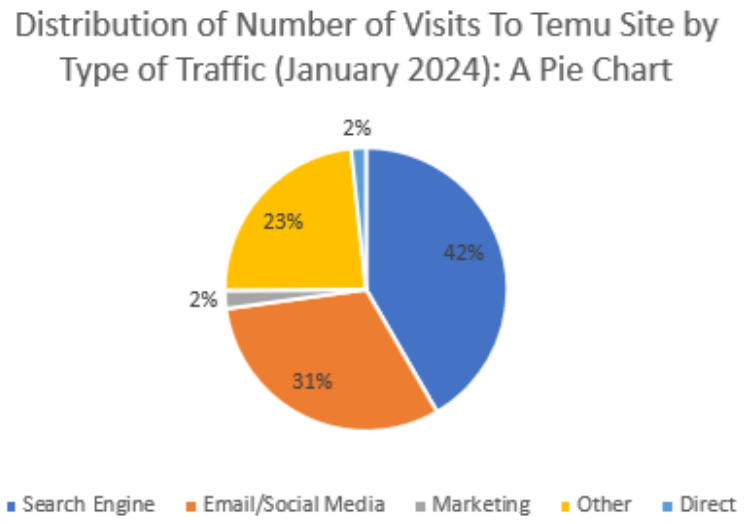
Box-and-Whisker Plot: [2; 0.5 for each - Q_1, Median, Q_3, Box]
(Only hand-drawn boxplot is accepted)



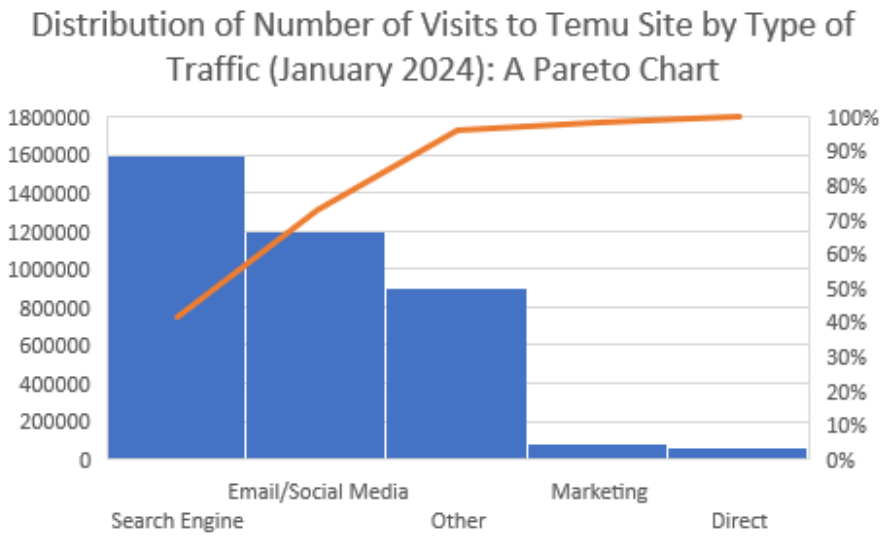
(d) Shape: The box plot in step (c) exhibits right-skewed (or positively skewed) distribution [0.5] because $Q_3 - Q_2(25.25 - 16) > Q_2 - Q_1(16 - 10.5)$ [0.5].

Q2: [4] Temu Visitors

(a) **Pie Chart** [1.5; .5 for each - chart, title, percentage of categories]



Pareto Chart[1.5; .5 for each - chart, title, axis title]



(b) The Pareto chart [.5] will be more useful as it is easy to see the most popular type of traffic (i.e., Search Engine) and the least least popular type of traffic (i.e., Direct)[.5].

Q3: [5] Voter Turnout

(a) Population: All [.5] provinces and territories [.5] (any answer along the line 10 provinces and 3 territories will also be accepted).
Variable: Percentage of voters. [.5].
Unit of Analysis: Each province/territory [.5].

(b) Variable (Percentage of voters): Quantitative ratio [.5] and continuous[.5].

(c) Stem-and-leaf Plot: Leaf Unit = 0.1[.5], [1.5: .5 for each - plot, stem, ordered leaves]