



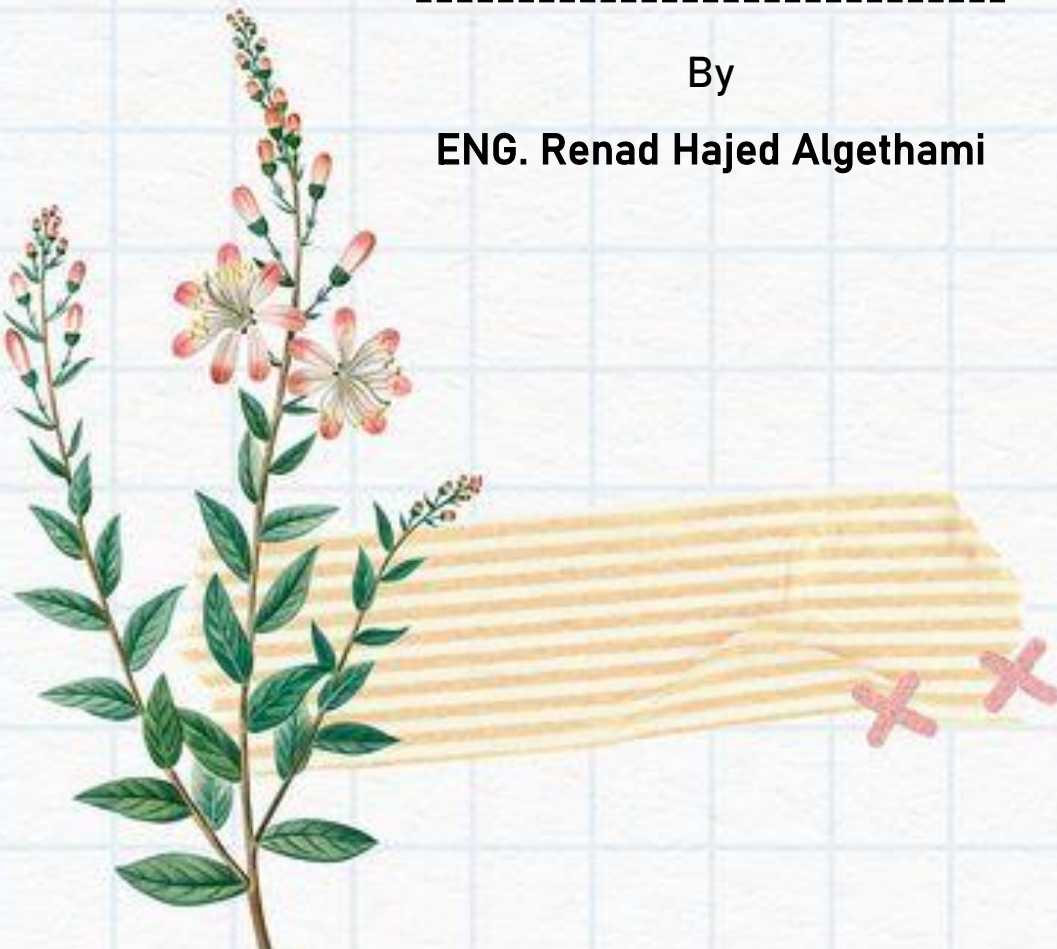
Report

**{Descriptive statistics on the number of deaths
from the Coronavirus within two months}**

Date: 10-2-2021

By

ENG. Renad Hajed Algethami



I conducted this report to illustrate the descriptive statistics and the methods I used to describe the number of deaths from Coronavirus during November and December 2020 in the Kingdom of Saudi Arabia.

At first, I collected data on the number of corona virus deaths from the Ministry of Health website, then I started with the first type of descriptive statistics, which is organizing the data in a table and a graph to describe it as follows:

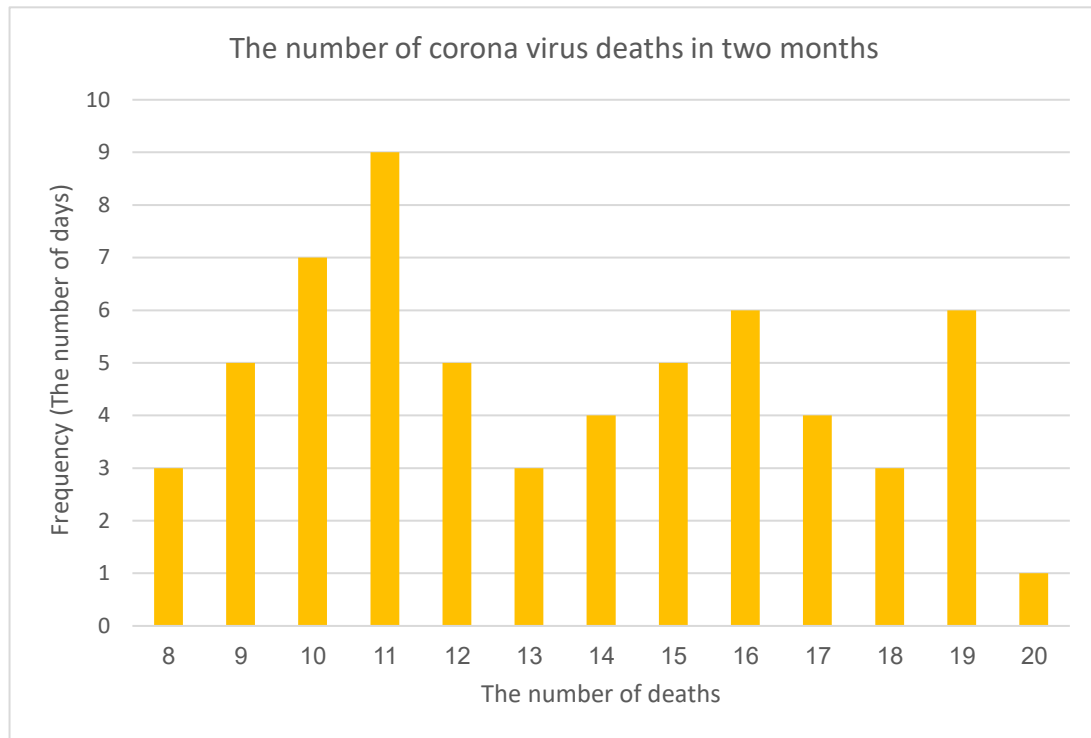
Data (Xi) The number of deaths	Frequency (Fi) The number of days	Relative Frequency (RFi) $= Fi / n$	Cumulative Frequency (CFi) $\sum_{i=1}^m Fi = n$	Cumulative Relative Frequency (CRFi) $\sum_{i=1}^m RFi = 1$
8	3	0.04918	3	0.04918033
9	5	0.081967	8	0.13114754
10	7	0.114754	15	0.24590164
11	9	0.147541	24	0.39344262
12	5	0.081967	29	0.47540984
13	3	0.04918	32	0.52459016
14	4	0.065574	36	0.59016393
15	5	0.081967	41	0.67213115
16	6	0.098361	47	0.7704918
17	4	0.065574	51	0.83606557
18	3	0.04918	54	0.8852459
19	6	0.098361	60	0.98360656
20	1	0.016393	61	1
Sum	n= 61	1	-	-

I organized the data in this table and set the formula for every column I extracted from the data. As for applying the formula, I applied them in Excel.

We conclude from this table that 8 deaths due to Coronavirus were recorded in 3 days of November and December, and that 20 deaths due to Corona virus were recorded in only one day of month November or December and did not recur after that.

The Relative Frequency column indicates the ratio of the number of days during which a specified number of deaths occur in relation to the remaining days of month November and December. For example, we note that 11 deaths due to Coronavirus were recorded during November and December, at a rate of 14.75%, which is the largest percentage. And there were 20 deaths due to the Coronavirus during November and December, at a rate of 1.64%, which is the smallest percentage.

We conclude from the cumulative relative frequency column that, at a rate of 24.59%, a number of deaths not exceeding 10 deaths were recorded per day in November and December. And the daily death toll from Coronavirus did not exceed 18 deaths during November and December, by 88.52%.



I organized the data in this Bar Diagram using Excel and now I will describe it as follows:

20 deaths due to the Corona virus were recorded with a frequency of only one day during the months of November and December, meaning that this number of deaths was not repeated. There were 8, 13 and 18 deaths due to the Coronavirus, at a frequency of 3 days during November and December. And there were 14 and 17 deaths due to the Corona virus, with a frequency of 4 days during the months of November and December. It recorded 9, 12 and 15 deaths due to the Corona virus, with a frequency of 5 days during the months of November and December. It also recorded 16 and 19 deaths due to the Corona virus, with a frequency of 6 days during November and December. It recorded 10 deaths due to the Corona virus, with a frequency of 7 days, during the months of November and December. 11 deaths due to the Corona virus were recorded with a frequency of 9 days during the months of November and December, which is the most frequent occurrence during these two months.

Now that I have finished with the first type of descriptive statistics, I will start with the second type, which is to summarize the data through the central tendency and the variation (dispersion) as follows:

The central tendency

it is the value around which the rest of the data is centered, so the more we move away from this value, the number of data begins to decrease and the closer we get to this value, the number of data begins to increase. To find out this value for the data of the number of corona virus deaths during November and December, I will do the calculations using the following formulas:

- **Mean:** $\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{820}{61} = 13.44$

Through this equation, we note that the average number of deaths due to Corona virus during November and December is 13.44 deaths.

- **Median:** $median = \frac{61}{2} = 30.5^{th} \rightarrow 13$

We notice that the total number of days in November and December is individual, so the median is the middle value of 30.5th, which means that 13 deaths are the median because it makes half the number of deaths greater than 13 deaths and the other half less than 13 deaths.

- **Mode:** $mode = \text{The data with the largest frequency} = 11$

Through this equation, we conclude that the number of deaths due to the Corona virus, which was more frequent during the months of November and December, is 11 deaths, as it was repeated 9 days out of 61 days.

<< The median number of deaths is more accurate than the mean number of deaths because it is not affected by the outliers of the number of deaths, whether they are very small or very large, while the mean is affected by the outliers because it is the sum of all values, whether extreme or not >>

I calculated the mean, median and mode using Excel using the formulas, and using the Data Analysis.

The variation (dispersion)

It is the extent of the spacing of the data from each other or from the central value. Therefore, the higher the value of the scale of dispersion, this indicates a large degree of divergence between the data values, and the smaller they are, this indicates that the difference between the data values is small. To find out this value for the data of the number of corona virus deaths during November and December, I will do the calculations using the following formulas:

- **Range:** Max=20, Min=8, Range= 20 – 8 = 12.

If we conclude that the extent of the dispersion of the Corona virus death numbers from each other is 12 deaths.

- **Variance:** $Var = \frac{\sum_{i=1}^m (x_i - \bar{x})^2 * F_i}{(n-1)} = \frac{729.049}{60} = 12.1508$

This means that the number of Corona virus deaths during the months of November and December is scattered around each other and around the mean number of Corona virus deaths, which equates to 13.44 deaths, and this dispersion reaches approximately 12 deaths meaning that it is not dispersed around one number only and the number of deaths is not constant.

- **Standard Deviation:** $s = \sqrt{var} = \sqrt{12.1508} = 3.49$

This means that the number of Corona virus deaths during the months of November and December is dispersed around the mean number of Corona virus deaths, which equals 13.44 deaths, and this dispersion reaches approximately 3 deaths meaning that the number of deaths is not fixed at a specific number.

- **Coefficient of Variation:** $CV = \frac{s}{\bar{x}} * 100 = \frac{3.49}{13.44} * 100 = 25.9\%$

The coefficient of variation for the number of corona virus deaths during November and December 2020 is equivalent to 25.9%. This ratio is compared with another ratio of the coefficient of variation for the number of corona virus deaths during another two months. Then, the number of deaths with the least coefficient of variation is the least dispersed and the number of deaths with the greatest coefficient of variation is the most dispersed.

I calculated the range, Variance and Standard Deviation using Excel using the formulas, and using the Data Analysis.

• **Five Number Summary:** I calculated the quadrants and the inter-quartile range using the Excel functions.

Q1: First quartile= 11

That is, 25% of Corona virus deaths during November and December were less than 11 deaths and 75% were more than 11 deaths.

Q2: Second quartile= 13

That is, 50% of Corona virus deaths during November and December were less than 13 deaths and 50% were more than 13 deaths.

Q3: Third quartile= 16

That is, 75% of Corona virus deaths during November and December were less than 16 deaths and 25% were more than 16 deaths.

IQR: Inter-quartile range= $Q3 - Q1 = 16 - 11 = 5$

The Inter-quartile Range of the number of deaths due to the Corona virus during the months of November and December is 5 deaths. That is, the middle half, after arranging the number of deaths in ascending order, will range between 11 and 16 deaths.

<< The inter-quartile range of the number of corona virus deaths during November and December is more precise than the range because it depends on the middle values and neglects the outliers in the number of deaths, that is the opposite of the range >>

In conclusion, I used descriptive statistics to describe the number of Coronavirus deaths during November and December 2020 in the Kingdom of Saudi Arabia. The description was not limited to formulas and numbers, but rather I described these numbers in a language that anyone could understand even if he was not a specialist in statistics. I also used Excel to make it easy for me to display results and solve formulas. I hope that you, dear reader, have benefited from the information contained in this report, which focuses on Coronavirus deaths, within only two months.