PYU33C01 Numerical Methods Assignment 2

Completed by Seán Alexander 21364546

Introduction:

This assignment will examine the plotting of national Covid-19 daily reported cases, linear regression as a method of fitting exponential growth and the Gaussian curve as a proposed fit for impure exponential curve behaviour due to vaccination factors.

Plotting National Covid-19 Cases

The data set provided detailing daily cases of Covid-19 since February 29$^{\text{th}}$, 2020 was utilised to make a list. This array was used to create a scatter plot shown below:

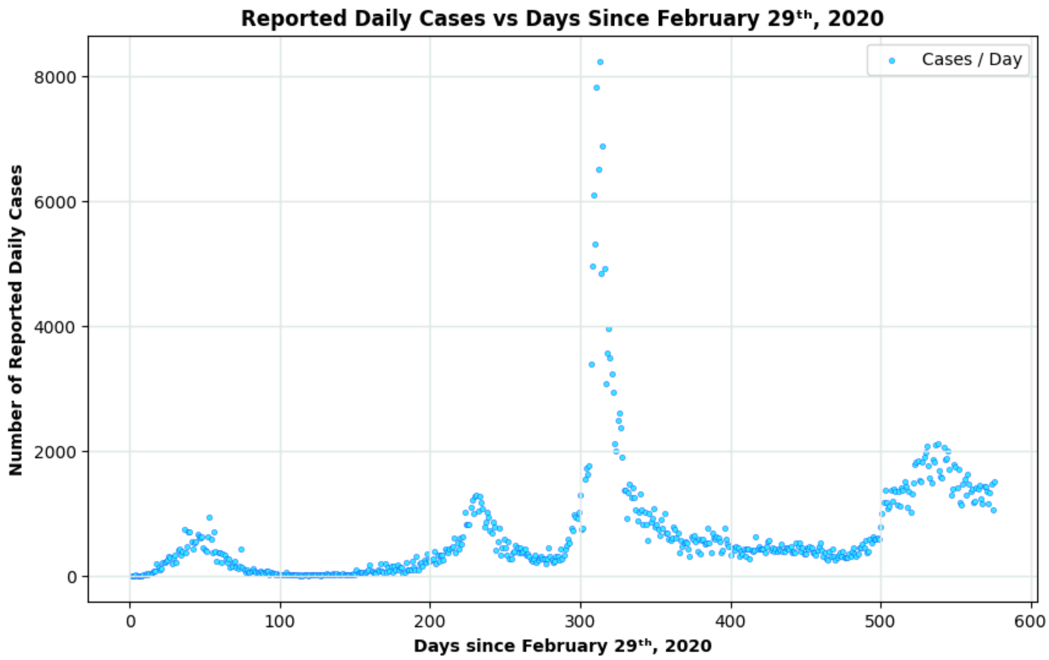

Figure 1: *Reported Daily Cases vs Days Since February 29th, 2020*

This scatter plot can be characterised by 4 "waves" of of cases which can be modelled as exponential growth and shrinkage as the waves of infection "ignite" and "die down". In order to demonstrate the exponential growth and shrinkage of the reported case numbers the natural logarithm of the reported daily cases were plotted:

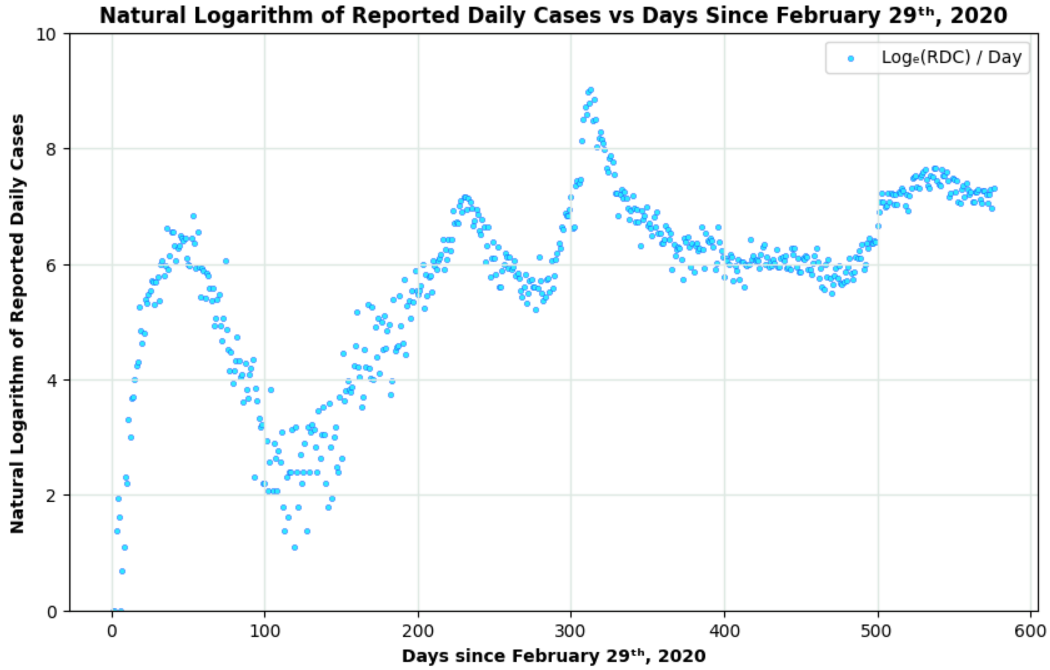

Figure 2: *Natural Logarithm of Reported Daily Cases vs Days Since February 29th, 2020*

The waves demonstrate a roughly linear rise and fall when plotted in this fashion which is characterised by an associated exponential rise and fall when plotted conventionally. This relationship can be used to model the data using linear regression.

Linear Regression

In order to apply linear regression to the data, it was expressed in a linear form; Where $\ln n(t) = a + bt$ such that $a = \ln n_0 - \lambda t_0$ and $b = \lambda$. In order to determine the two constants to plot these linear regressions they had to have been solved for.
The simultaneous equations:

$$a_N + b\sum_{i=1}^{N} t_i = \sum_{i=1}^{N} x_i \tag{0.1}$$

$$a\sum_{i=1}^{N} t_i + b\sum_{i=1}^{N} t_i^2 = \sum_{i=1}^{N} x_i t_i \tag{0.2}$$

- were solved and formatted for computational ease:

$$a = \frac{1}{N}\sum_{i=1}^{N}(y_i - b\cdot x_i) \tag{0.3}$$

$$b = \frac{\sum_{i=1}^{N}(N\cdot x_i\cdot y_i - x_i\cdot\sum_{i=1}^{N} y_i)}{\sum_{i=1}^{N}(N\cdot x_i^2 - (x_i)^2)} \tag{0.4}$$

This method was used to plot the linear regressions for $\ln n(t) = a + bt$ for the first, second and third waves as shown below:

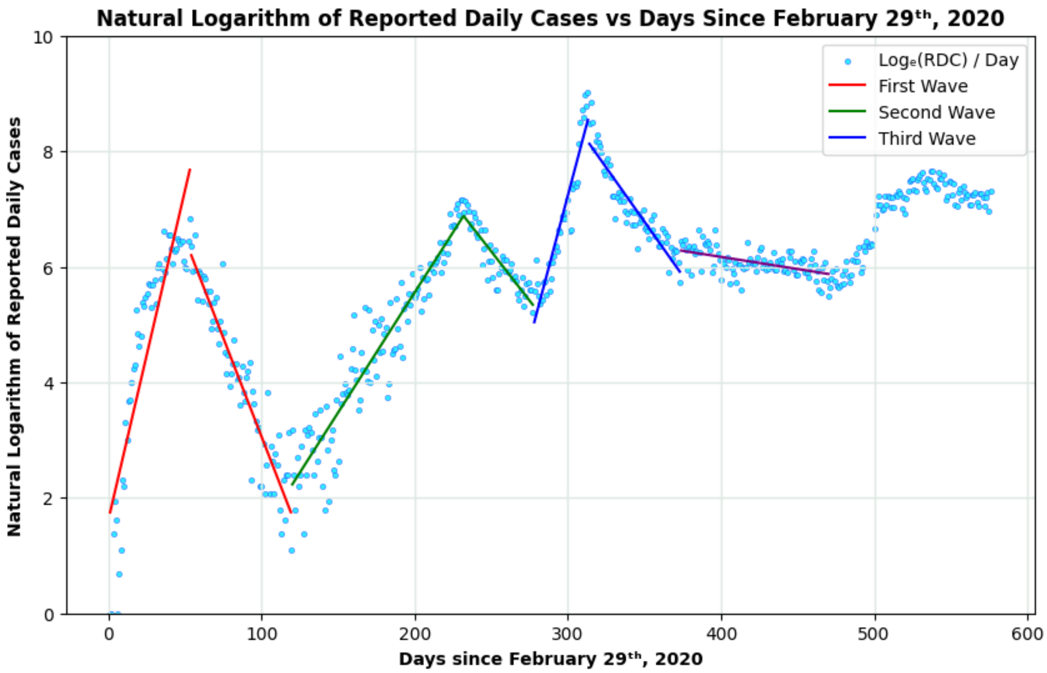

Figure 3: *Linear Regression Analysis of Natural Logarithm of Reported Daily Cases vs Days Since February 29*th*, 2020*

Wave	Incline	t_0	a	b	N_0	λ
1	Rise	1.635659259	1.635659259	0.11396754	1	0.11396754
1	Fall	9.892823594	9.892823594	-0.068412944	577	-0.068412944
2	Rise	-2.756388866	-2.756388866	0.041584091	24	0.041584091
2	Fall	14.77378576	14.77378576	-0.034047183	1031	-0.034047183
3	Rise	-22.7542868	-22.7542868	0.099984929	265	0.099984929
4	Fall	19.89681993	19.89681993	-0.037482255	4842	-0.037482255

Table 1: Associated Parameters

These linear fits resulted in the associated parameters shown in table 1.

These parameters were used to plot exponential fits for the conventionally plotted data as included below:

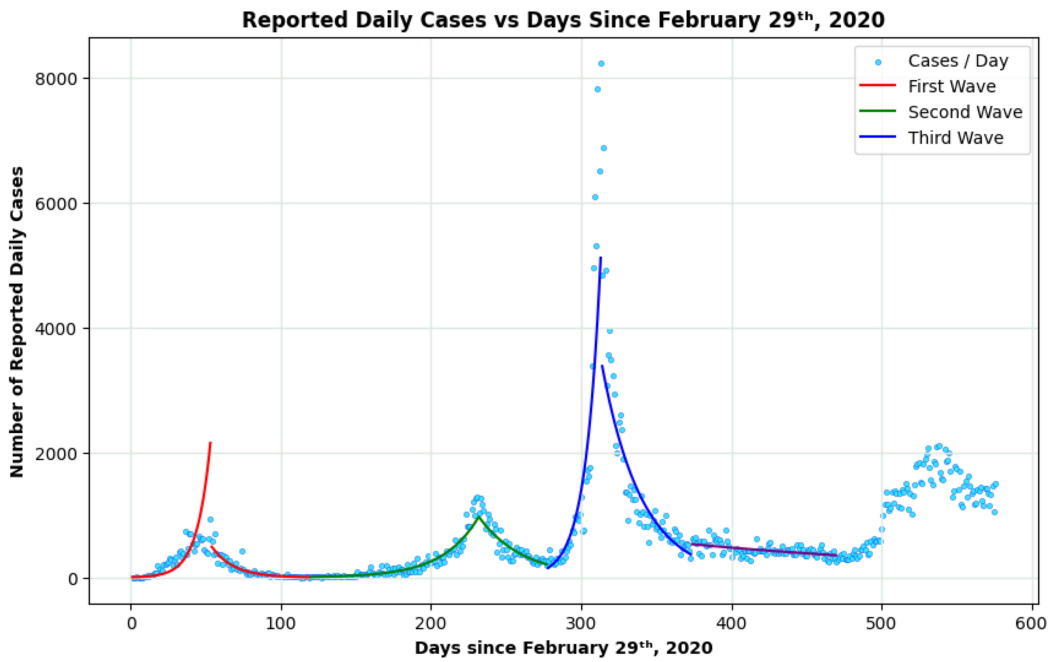

Figure 4: *Exponential Fit Analysis of Natural Logarithm of Reported Daily Cases vs Days Since February 29th, 2020*

Fitting the Fourth Wave

The exponential relationships used to plot fittings derived from the linear relationship $\ln n(t) = a + bt$ do not best describe the data for the advent of the fourth wave.

The relationship for this curve must be more complicated as the vaccine was a factor in reducing the transmission rate, which provided a vaccination rate variable to the transmission modeling expression. As this data is not available, a good fit for the fourth wave of infection is a Gaussian due to the following equation:

$$\text{gaussian}(x, k, \sigma) = k \cdot \exp\left(-\frac{(x - \overline{x})^2}{2\sigma^2}\right) \tag{0.5}$$

The Gaussian expression represents a probability distribution. It defines the value of a Gaussian function at a point x based on the parameters amplitude (k) and standard deviation (σ), relative to the mean $\overline{x}$. The function is bell-shaped and describes the probability density of a data set.

The amplitude a determines the peak height, σ controls the width of the curve, and the term $x - \overline{x}$ measures the deviation from the mean. The exponent $\frac{-(x-\overline{x})^2}{2\sigma^2}$ governs the rate of decrease as you move away from the mean.

It is for the reasons above that the Gaussian fit was a theoretically suitable fit for the data characterising the fourth wave. This fit was coded and is shown below:

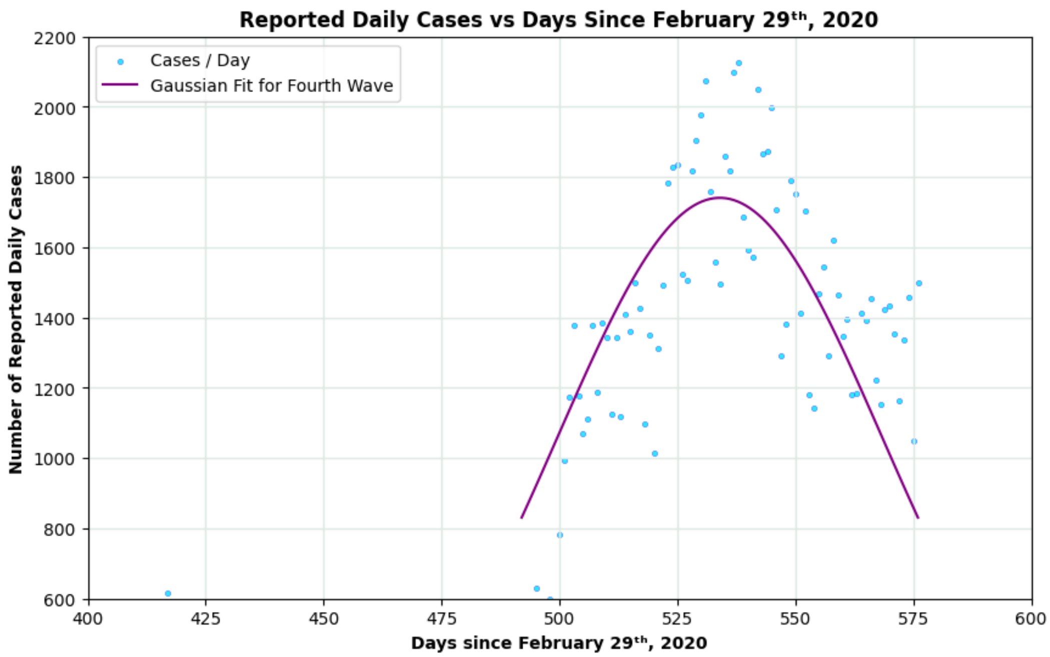

Figure 5: *Gaussian Fit for Reported Daily Cases for the 4^{th} wave of Covid-19 in Ireland*