ECEN 321 Lab 5

Regression

Niels Clayton : 300437590

Introduction

Regression is a commonly used tool for assessing the relationship between variables. It is often used to compute the line of best fit of a dataset, and can even be used to help predict future values of a dataset. In this Lab we will be performing a regression on temperature anomaly data from a collection of locations around the globe. We will attempt to use this regression to come to a conclusion on if the average global temperature is increasing, decreasing, or remaining static.

Theory

For a simple linear regression, the first step is to calculate the correlations coefficient of the data sets, often denoted r. This is done with the following equation:

$$r = \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{x_i - \bar{x}}{s_x}\right)\left(\frac{y_i - \bar{y}}{s_y}\right)$$

$$= \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2(y_i - \bar{y})^2}}$$

Using the above equation we are able to calculate the correlation coefficient between variable x and variable y. From the correlation coefficient, we are then able to calculate the slope of the line of best fit, and pot it against the data using the following equation:

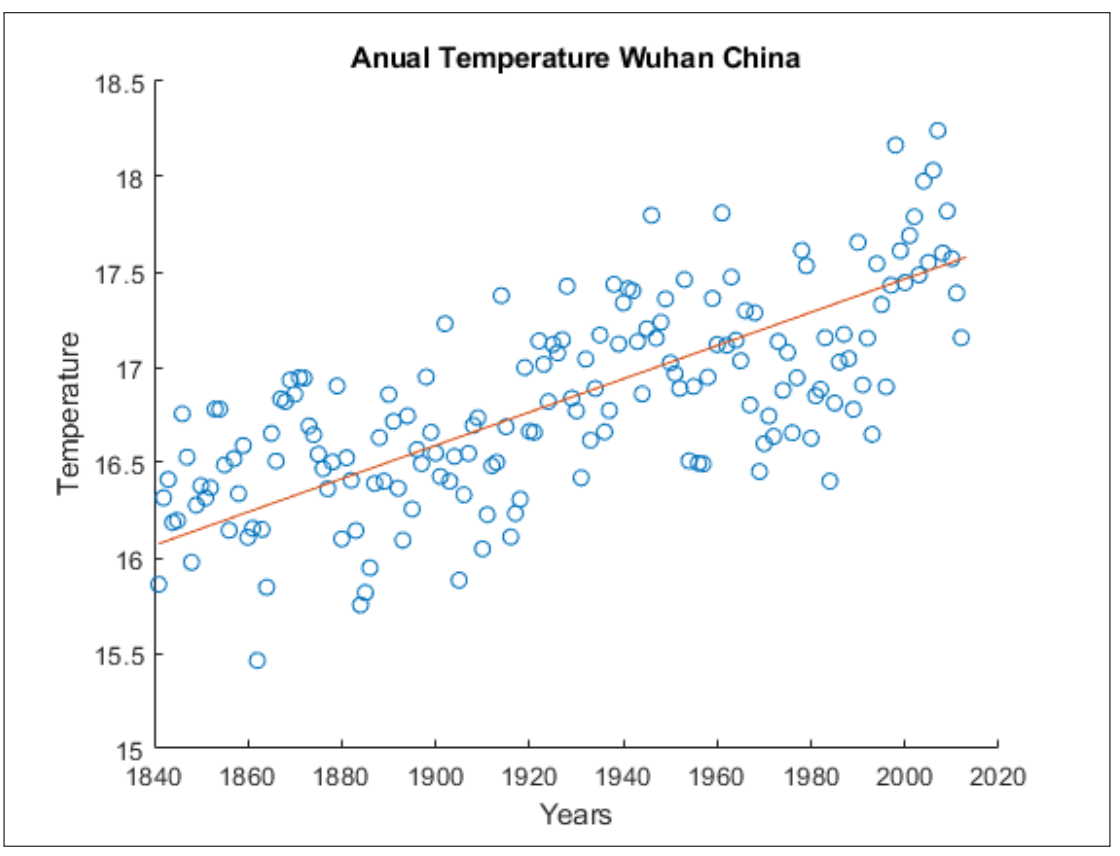
$$\beta_1 = r\frac{s_y}{s_x}$$

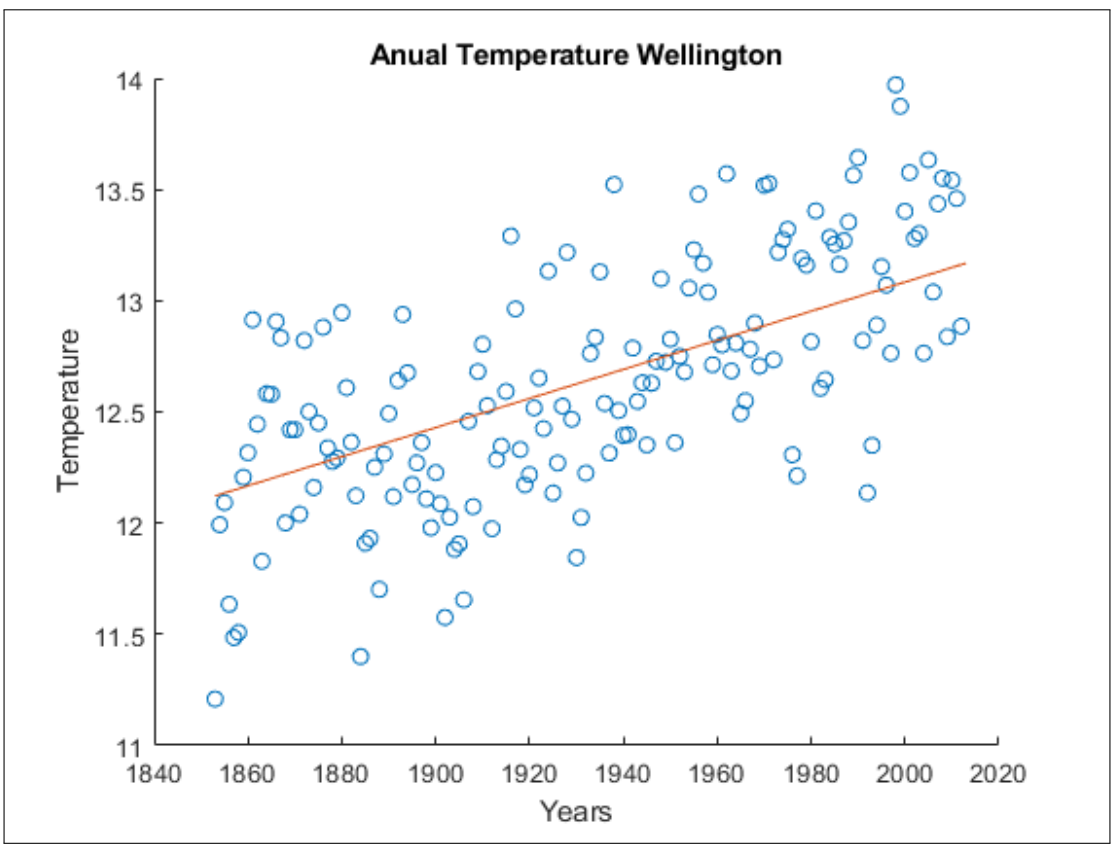
Once this correlation coefficient has been computed, we can compute the p-statistic of the correlation coefficient using the following equation:

$$U = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \sim t(n-2)$$

Results

This regression was performed on data from Wuhan China, and Wellington New Zealand, the results can be seen below in figures 1, and 2





Using the null hypotheses that the correlation coefficient is zero the following was calculated:
$H_o : r = 0$
$H_1 : r > 0$

using the equation above, the p-static was calculated to be:

Wuhan China - $p = 0$
Wellington New Zealand - $p = 5.9 \times 10^{-15}$

Due to the the insignificance of the p statistic for both Wuhan China, and Wellington New Zealand, we can confidently disregard the null hypotheses that $H_o : r = 0$. This would suggest that there is an increase in the average temperature in these locations.