**Homework-02-Report**

**Introduction**

This report summarizes the hypothesis and results found on weather data that includes hourly/daily information about the weather for Szeged, Hungary between 2006-2016. The data was obtained from Kaggle API, Click on the handy link for more information [Weather in Szeged 2006-2016](https://www.kaggle.com/budincsevity/szeged-weather)

**Hypothesis**

We want to study the relationship between the Humidity and Apparent Temperature (in Celsius) using regression. We want to see if we can predict the Apparent Temperature(C) given Humidity.

**Formulas**

We have used the formula of the linear regression that consists of the dependent variable Y and the independent variable X.

Where X is a matrix of the independent variable that represent the Humidity values and Y is a vector of the dependent variable that represent the Apparent Temperature(C).

We have used the following formula to estimate the coefficients :

We have used three separate files: **Regression.py** for the function the take care of the regression, **RegressionTest.py** to test the regression function when it is given bad inputs, and **Data-Export.ipynb** to handle data export and run the regression function on them. A sample of 250 points were drawn from the original data to serve the goal of our hypothesis.

**Result**

The result shows that there is a linear regression relationship between Humidity and Apparent Temperature. This can be shown from the plots of the data after running the regression function, showing the regression line through the data.

**Links**

[Credible Intervals](https://www.econometrics-with-r.org/5-2-cifrc.html)