**Exploratory Data Analysis – Ames House Price Dataset**

A city street

Description automatically generated

Figure 1 – Ames Iowa Main Street

In this coursework you will utilize Google Colab to generate a Jupyter notebook for analysing the Ames Housing dataset. The dataset describes the sale of individual residential property from 2006 to 2010 in Ames (a small town in Iowa, USA). It was compiled by Dean De Cock who wrote a paper on this (which you should read). More details of this dataset are described in the paper: [*Ames, Iowa: Alternative to the Boston Housing Data as an End of Semester Regression Project*](https://ww2.amstat.org/publications/jse/v19n3/decock.pdf)*.[*[[1]](#footnote-1)*]*

This CW will require you to read in the data, perform any pre-processing required on the data, and utilise the Python Data Science Software Ecosystem to undertake and Exploratory Data Analysis of the House prices detailing and appropriately documenting any information you can extract from the data.

The dataset is included as a text file within the Coursework section on GCULearn.

**Jupyter Notebook**

This notebook should provide a discussion and demonstration of the steps you have undertaken to perform exploratory data analysis on the dataset detailing both the software required and the reasoning for undertaking each step.

You should discuss:

* The formatting of the data
* Data cleaning
* Exploratory Data Analysis techniques to describe the data
* Data visualization.

Coursework reports should be submitted to GCULearn via Turnitin no later than **23.59 on Monday 25th July 2022**.

Please note that only one submission can be made so please make sure you are happy with your notebook before you submit.

A rubric for this component of the coursework is provided below:

Table

Description automatically generated

Figure 2 - Coursework 1 Report RUBRIC

1. De Cock, D., 2011. Ames, Iowa: Alternative to the Boston housing data as an end of semester regression project. Journal of Statistics Education, 19(3). [↑](#footnote-ref-1)