**DSA 8203 - Homework 1**

Notes:

* Make a copy of this document in your drive and share it with me
  + (File, Make a copy)
  + ([cnjunge@strathmore.edu](mailto:cnjunge@strathmore.edu) )
* Change the file name to include your first initial and last name
* Your responses should be about a page in length, be concise.

Look for a dataset you will be using to analyze during the course of the semester (based on your country). The syllabus has some links to ‘Dataset options’, ‘Kaggle’, and you may also search on Google with the suffix ‘csv’ to find a dataset of interest. I suggest you select something of interest to you professionally or personally.

1. Where is the data found? (Add link)
2. Is it in a format that pandas can read? (Structured, CSV/Excel/SQL)
3. What do you hope to discover in the data?
4. What challenges do you foresee with using this data? Perhaps the data might not have enough information about each sample. Maybe there is too much information. Maybe you will need to read up on Python and or Pandas.
5. What information would you expect your final presentation to include? (Consider the CRISP-DM framework? Will outputs be charts, model, a dashboard, app etc.)
6. Is there an example online of an analysis and report that you like?

This homework assignment is due week four before class starts.