# Business Intelligence Capstone

A capstone might be called a culmination project, senior thesis, or a final exhibition. ... The capstone exercise is intended to apply all of the knowledge and skills you've gained over the course in one assignment. This capstone exercise will focus on the following:

* Analysis of raw data
* Visualization of the raw data
* Driving insights out of pre-existing dashboards

The capstone will not be a paper or thesis, but a live demonstration of the knowledge gained using such tools as python, Excel and Tableau. The student will be able to demonstrate their level of expertise in a manner that is befitting a budding Data Analyst/Scientist.

## Analysis of Data

The student will pick and chose from many different publicly available datasets or also have the option to illustrate a particularly important step in the unit testing and user acceptance testing phase of an initiative by generating raw data in a programmatic fashion.

### Programmatic Fashion

Using python, generate up to 100,000 rows of data at various dimensionality in order to satisfy several use cases (both positive and negative) for inclusion in the tasks to come as part of the complete capstone exercise. The student should be able to walk through the code, providing a straightforward explanation of the various use cases and specifically point out initial inputs and expected results. Be creative with this as your audience will not want to look at a record level, tabular file (hint – a bit of programmatic visualization never hurt anybody).

### Internet Available

There are many websites that supply various datasets that have been used throughout time by students as well as professionals in the field to develop techniques for analysis of raw data. The student will need to select a prepared dataset, give a short statement about the contents of the dataset and define several use cases (positive and negative) that will make use of the information to generate actionable insights. Here are a few of my favorites:

<https://community.tableau.com/s/question/0D54T00000G557ESAR/data-sets>

<https://www.kaggle.com/datasets>

<http://statweb.stanford.edu/~sabatti/data.html>

<https://opendatainception.io/>

<https://www.stats.govt.nz/large-datasets/csv-files-for-download/>

Again, the student should be able to walk through the code, providing a straightforward explanation of the various use cases and specifically point out initial inputs and expected results. Be creative with this as your audience will not want to look at a record level, tabular file (hint – a bit of programmatic visualization never hurt anybody).

## Visualization of Data

Now that the student has selected a dataset that is of interest to them, the next step in this exercise is to become truly familiar with the details that are used to describe the data in a more statistical manner.

Descriptive statistics can be useful for two purposes: 1) to provide basic information about variables in a dataset and 2) to highlight potential relationships between variables. The three most common descriptive statistics can be displayed graphically or pictorially. It is recommended that the student use the more common to describe or summarize the characteristics of a sample or data set, such as a variable's mean, standard deviation, or frequency. The inclusion of others will gain a better understanding of these three basic stats and is highly recommended.

Here is a hint as to how they should be used on your sample dataset - <https://www.investopedia.com/terms/d/descriptive_statistics.asp> or <https://corporatefinanceinstitute.com/resources/knowledge/other/descriptive-statistics/>.

Now that the student has a better understanding of the dataset based on descriptive statistics it is time to generate a few charts, graphs and even reports to enlighten us on some of the insightful aspects of the data. Feel free to use Tableau to generate sheets or even a quite simple dashboard with three or four graphics to tell us a story. Remember, you are not looking for perfection in your graphics at this point in the capstone. The purpose of this exercise is to think fast and act quickly to generate a couple of interesting artifacts.

## Driving Actionable Insights

This is the final step in the process of illustrating that the student has grasped the concepts associated with data analytics/business intelligence.

Tableau provides a stunning set of data visualization examples from around the world created with Tableau Public (the licensed software used for this class). There is a Viz of the Day as well as many Featured vizs that the student will be able to chose from for the exercise. As discussed in the final week of the class, driving out actionable insights involves a methodology. Use the concepts of this methodology to peruse the Gallery and find one or two vizs that allow you to demonstrate your prowess. Remember to keep it simple and be sure it is engaging!

<https://public.tableau.com/en-us/gallery/?tab=viz-of-the-day&type=viz-of-the-day>

The final step in the capstone exercise will be to walk the talk and talk the walk. If you are not tired of my humor yet, then you are certain to make a big splash in the world of Data Analytics.

You worked very smart (anybody can work hard) on putting together a first-class story from scratch and now it is time to shout it from the roof tops. Use your final project to tell everybody a story!

This might seem a bit daunting, but I assure you that these are the typical types of activities I have performed on a daily basis in order to advance through the ranks of Data Analytics/Business Intelligence.

**CONGRATULATIONS IS IN ORDER!**