1.1 Lesson Summary - The Zen of Data

The proliferation of data and data analysis tools in recent years offers us an opportunity to discover meaningful, actionable information that would have been previously inaccessible. This course offers you an introduction to the tools and methods used to extract meaning from data.

Concept: The field of **Data Science** explores the scientific processes used in extracting knowledge from data.

Concept: **Data Analytics** is the process of collecting, manipulating, and analyzing data to extract useful information.

Concept: Throughout this course you will be introduced to a variety of tools commonly used in data analytics. **Microsoft Excel** is a spreadsheet management software with numerous tools for manipulating and visualizing data. **Python** is a portable versatile coding language useful in software development and data analysis. **Jupyter** is a free open-source editor for Python. **pandas** is a coding library used for Python based numerical analysis. **Matplotlib** is a coding library for Python used for visualizing data. When using Python the coding library **Beautiful Soup** will allow you to get data from websites. Structured query language (**SQL**) is a language and specification for structured relational databases. **MongoDB** is used for managing unstructured document databases. Hypertext markup language (**HTML**) is used to specify the contents of a webpage. Formatting webpages is most often done using cascading style sheets (**CSS**). **JavaScript** is a programming language used for making websites interactive. You can create web-based data visualizations using the JavaScript library **D3.js**. Visualizing map-based data on the web can be accomplished using **Leaflet.js** or **Google Maps**. **Tableau** is a spreadsheet-based software used for organizing and visualizing data. **Hadoop** is open source software that allows the user to utilize networks of computers to process data. You only need to be generally familiar with these tools at the outset of the course as they will be discussed in greater detail as they are needed.

Concept: The process of data analytics usually involves the following common pattern. Deconstruct the questions being asked. Consider what information might be useful in answering the question. Find data sources and how to get them. Transform the data into something useful. Analyze the data. Address any shortcomings in the results of your analysis. Illustrate and present your data.

* Activity: 01-Ins\_GreatDebate

This first lesson offers a brief overview of the material that will be covered in this course and an illustration of how what you learn in this course can be used.