**Introduction To Capstone Project**

You have recently been hired as a Data Analyst by a global IT and business consulting services firm that is known for their expertise in IT solutions and their team of highly experienced IT consultants. In order to keep pace with changing technologies and remain competitive, your organization regularly analyzes data to help identify future skill requirements.

As a Data Analyst, you will be assisting with this initiative and have been tasked with collecting data from various sources and identifying trends for this year's report on emerging skills.

Your first task is to collect the top programming skills that are most in demand from various sources including:

* Job postings
* Training portals
* Surveys

Once you have collected enough data, you will begin analyzing the data and identify insights and trends that may include the following:

* What are the top programming languages in demand?
* What are the top database skills in demand?
* What are the popular IDEs?

You will begin by scraping internet web sites and accessing APIs to collect data in various formats like .csv files, excel sheets, and databases.

Once this is completed, you will make that data ready for analysis using data wrangling techniques.

When the data is ready you will then want to apply statistical techniques to analyze the data. Then bring all of your information together by using IBM Cognos Analytics to create your dashboard. And finally, show off your storytelling skills by sharing your findings in a presentation.

You will be evaluated using quizzes in each module as well as the final project presentation.

# Syllabus

### **Module 1: Data Collection**

* Collecting Data Using APIs
* Collecting Data Using Webscraping
* Exploring Data

### **Module 2: Data Wrangling**

* Finding Missing Values
* Determine Missing Values
* Finding Duplicates
* Removing Duplicates
* Normalizing Data

### **Module 3: Exploratory Data Analysis**

* Distribution
* Outliers
* Correlation

**Module 4: Data Visualization**

* Visualizing Distribution of Data
* Relationship
* Composition
* Comparison

### **Module 5: Dashboard Creation**

* Dashboards

### **Module 6: Presentation of Findings**

* Final Presentation