Data Science

## Competency development areas in data science

* Foundations for Data Science:
  + Understanding the role of data science in extracting insights from data and making data-driven decisions.
  + Python Foundations (Libraries: Pandas, NumPy, Arrays and Matrix handling, visualization)
  + Statistics Foundations: Basic/Descriptive Statistics, Distributions (Binomial, Poisson, etc.), Bayes, Inferential Statistics.
  + Exploratory Data Analysis (EDA).
* Data science lifecycle:
  + Data Collection and Cleaning: The basics of data collection, data sources, and data formats; techniques for cleaning and preprocessing data to ensure data quality.
  + Exploratory Data Analysis (EDA): Exploring and visualizing data to understand its characteristics, identify patterns, and detect anomalies. Descriptive statistics and data visualization techniques.
  + Feature Engineering: Creating meaningful features from raw data to improve model performance; techniques for feature extraction, transformation, and selection.
  + Foundational Machine Learning: Fundamentals of machine learning and basic concepts (features, labels, and models). Introduction to supervised learning (regression, classification). Introduction to unsupervised learning. Model evaluation (cross validation and bootstrapping).
  + Data Visualization: Creating visual representations of data to effectively communicate insights. Different types of visualizations, tools like Matplotlib or Tableau, and best practices for visualizing data.

## Online training resources

* Introduction to Data Science Specialization by Coursera and IBM: This online training program is designed to provide learners with a comprehensive introduction to the field of data science. The specialization consists of a set of courses that cover various topics and skills related to data science. (coursera.org/specializations/introduction-data-science).
* Coursera: Offers a wide range of data science courses and specializations from top universities and institutions. (coursera.org/browse/data-science).
* edX: Provides online courses in data science from leading universities and institutions. (edx.org/learn/data-science).
* Udacity: Offers nanodegree programs and courses in data science, machine learning, and related fields.
* IBM Data Science: Provides a range of data science courses and learning paths. (ibm.com/training/search?query=data science).
* The National Consortium for Data Science (NCDS): NCDS helps members take advantage of data in ways that result in new jobs and transformative discoveries. (https://datascienceconsortium.org/).