About this course

Analyzing data with Python is an essential skill for Data Scientists and Data Analysts. This course will take you from the basics of data analysis with Python to building and evaluating data models.

Topics covered include: - collecting and importing data - cleaning, preparing & formatting data - data frame manipulation - summarizing data - building machine learning regression models - model refinement - creating data pipelines You will learn how to import data from multiple sources, clean and wrangle data, perform exploratory data analysis (EDA), and create meaningful data visualizations. You will then predict future trends from data by developing linear, multiple, polynomial regression models & pipelines and learn how to evaluate them. In addition to video lectures you will learn and practice using hands-on labs and projects. You will work with several open source Python libraries, including Pandas and Numpy to load, manipulate, analyze, and visualize cool datasets. You will also work with scipy and scikit-learn, to build machine learning models and make predictions. If you choose to take this course and earn the Coursera course certificate, you will also earn an IBM digital badge.

**This course belongs to several programs**

This course can be applied to multiple Specialization or Professional Certificate programs. Completion of this course will count toward your studies in any of these programs:

* [Професійний сертифікат Наука про дані IBM](file:////specializations/ibm-data-science)
* [Професійний сертифікат Аналітик даних від IBM](file:////specializations/ibm-data-analyst)
* [Спеціалізація прикладна наука про дані](file:////specializations/applied-data-science)

**WHAT YOU WILL LEARN**

* + Develop Python code for cleaning and preparing data for analysis - including handling missing values, formatting, normalizing, and binning data
  + Perform exploratory data analysis and apply analytical techniques to real-word datasets using libraries such as Pandas, Numpy and Scipy
  + Manipulate data using dataframes, summarize data, understand data distribution, perform correlation and create data pipelines
  + Build and evaluate regression models using machine learning scikit-learn library and use them for prediction and decision making

**SKILLS YOU WILL GAIN**

* + Predictive Modelling
  + Python Programming
  + Data Analysis
  + Data Visualization (DataViz)
  + Model Selection