Module 1: Data Science Introduction

Lecture 1: Introduction to Data Science

* What is Data Science
* Goals of Data Science
* Methods of Data Science
* Structure of Data Science
* Characteristics of Big Data
* Technologies for Big Data
* Data Science projects
* Where to learn more

Lecture 2: Installation of Environment

* Anaconda distribution
* Github
* Installation on Windows
* Installation on Mac
* Installation on Linux

Lecture 3: Python

* Python programming language
* Comments in Python
* Constants
* Variables
* Type conversion
* Variable names
* Reserved words
* Assignment
* Numeric operations
* Collections
* Lists
* Strings
* Strings Operations
* Sets
* Dictionaries
* Tuples
* Conditionals
* Comparison operations
* Indentation
* Errors and Exceptions
* Functions
* Loops
* List/tuples/dictionaries comprehensions
* Loop over elements, search for elements
* I/O

Module 2: Collecting data

Lecture 1: Data Formats

* Types of data
* HTML format
* CSV format
* XML format
* XML schema
* JSON format
* Graphics formats
* Video formats

Lecture 2: Web scrapping

* Regular expressions in Python
* HTTP requests
* URLLIB
* Web crawler
* Parsing HTML with Beautiful Soup
* API
* SOAP API
* REST API
* Examples of API services
* API Security and Rate Limits

Lecture 3: SQL

* Relational Database (RDB)
* RDB terminology
* SQL
* RDB systems
* ER Diagram
* Relational Schema
* SQLite
* CREATE Tables
* INSERT
* DELETE
* UPDATE
* Retrieve records
* ORDER BY
* GROUP BY
* Database design
* Keys
* Relationships
* Join Tables
* Python interacting with SQLite

Module 3: Processing and analyzing data

Lecture 1: Data Manipulation

* Numpy module
* Create arrays
* Describe arrays
* Reshape arrays
* Indexing and slicing of arrays
* Iterating through arrays
* Operations with arrays
* Pandas module
* Timestamps
* Series
* Series creation
* Query series
* Iterate through series
* Function on series
* DataFrames
* DataFrames creation
* Query DataFrames
* Operations on DataFrames
* Missing values
* Pandoric way to process DataFrames
* Join DataFrames
* GroupBy
* Aggregate
* Pivot Tables

Lecture 2: Descriptive statistics

* Statistics in DataFrames
* Visualization in DataFrames
* Scipy module
* Distributions
* Hypothesis testing
* A/B testing
* Outliers/Inliers
* Anomaly detection (static data)
* Anomaly detection (time series)

Lecture 3: Data Science pipeline

* Elements of Data Science Pipeline
* Find
* Collect
* Store
* Verify
* Clean, transform
* Combine
* Analyze
* Visualize

Module 4: Data visualization and presentation

Lecture 1: Data Visualization

* Visualization wheel
* Effect of the audience
* Qualities of great visualization
* Truthful
* Functionality
* Beauty
* Insightful
* Enlightening
* Matplotlib
* Matplotlib architecture
* Scatter plots
* Line plots
* Barcharts
* Subplots
* Histograms
* Box plots
* Heatmaps
* Animation
* Interactivity

Lecture 2: Infographics

* Seaborn
* Interactive infographics with d3.js
* Business dashboards with Tableau

Lecture 3: Data Science, where to go from here

* Stay current on new developments
* Sharpen your skills
* Learn more
* Defense of projects