

3ikakke Datascience with python course Schedule

3ikakke

January 2022

Course Schedule

Week 01

- (Mon) Introduction to Course - Data, datascience tools, real life examples, careers in data science, house keeping (start times, duration, recording of sessions, shared data repositories), what you will benefit (First timers, Career people, Academics, People giving it a second chance)
- (Fri) Python - **Introduction to Python** - Datatypes (strings, integers, floats, booleans, none, collections), mapping python data types to statistics and data science, Operators (mathematical, assignment, concatenation, comparison, logical, special), creating variables

Week 02

- (Mon) Python - **Control Structure Logic** (if else statments, if elif else, ternary)
- (Fri) Python - **Control Structure Loops** (for loops, while loops)

Week 03

- (Mon) Python - **Functions** (in built and user defined)
- (Fri) Python - **Collections** (structure and functions), tuples, lists, dictionaries, sets, indexing, slicing, looping, some simple functions (print, formatted strings, input, type)

Week 04

- (Mon) Python - **Collections** (lambdas, map, reduce, comprehensions), more functions
- (Fri) Python - **Object Oriented Programing** - classes, objects, methods, attributes

Week 05

- (Mon) Python - **Modules** inbuilt (requests, datetime,) and user defined modules (understanding **init.py**, **name**, **main**), importing, importing from
- (Fri) Python - **Intro Numpy and Pandas** - Numpy arrays, numpy array functions, introduction to pandas and the concept of dataframes (a new type of collection)

Week 06

- (Mon) Python - **Pandas** Importing data and Data Manipulation - selecting columns, filtering through rows, searching, sub setting
- (Fri) Python - **SQL with SQLAlchemy** - Setting up SQLAlchemy, create_engine, DSN, and overview of SQL - SELECT, FROM, WHERE, GROUP BY, IN, BETWEEN, DISTINCT, COUNT, AS, JOIN, ON, HAVING, ORDER BY, DESC, ASC, CASE, WHEN, END

Week 07

- (Mon) Statistics - **Introduction to basic stats** - Probability, Scipy, Data as quantitative, qualitative (binary or categorical), Summaries - Frequencies and Proportions, 5 number summary, mean and standard deviation
- (Fri) Statistics - **Univariate and Bivariate Analysis**, Matplotlib and Seaborn - cross tables (2-way tables), histograms, boxplots, barplots, correlations

Week 08

- (Mon) Statistics - **AB and Hypothesis testing** - TTests, ANOVA, ChiSquared Tests, Introducing regressions (conceptually)
- (Fri) Machine Learning - **Introducing** - SKLearn, Supevised and Unsupervised Learning, Data Pre-processing

Week 09

- (Mon) Machine Learning - **Supervised Learning Regressions** (Simple Linear Regressions, Multi Linear Regressions, SVM Regressions, Decision Tree Regression, RMSE, Regression Plots)
- (Fri) Machine Learning - **Supervised Learning Classification** (Logistic Regression, SVM Classification, Decision Tree Classification, KNN, Confusion matrices)

Week 10

- (Mon) Machine Learning - **Clustering** (K-means clustering)
- (Fri) Machine Learning - **Ensemble learning** (Random Forest Regressions and Classification, XG-Boost), concept of bagging and boosting

Week 11

- (Mon) Machine Learning - **Dimensionality Reduction** - Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA)
- (Fri) Machine Learning - **Recommender Systems** - Apriori and Eclat

Week 12

- (Mon) Machine Learning - **Natural Language Processing (NLP)**
- (Fri) Deep Learning - **Intro, Artificial Neural Networks** -

Week 13

- (Mon) Deep Learning - **Convolutional Neural Networks**
- (Fri) **What next?**

Introduction to Course

- Looking at data, an overview of the course, what we will not cover, expectations, housekeeping, introduction to programming
- Objectives for:
 - First timers
 - Career data users
 - Academics/researchers
 - People giving data science a second chance

Python

1. Introduction - Datatypes, Operators
2. Control Structure Logic - if and else statements and ternaries
3. Control Structure Loops - for and while loops
4. Functions - inbuilt and user-defined functions
5. Collections (structure and functions), tuples, lists, dictionaries and sets
6. Collections (lambdas, map, filter, comprehensions)
6. Object Oriented Programming - classes, objects, attributes, and methods
8. Modules - inbuilt and user-defined
7. Numpy & Pandas Intro
8. Pandas Data Manipulation
11. Pandas SQL (sqlalchemy)

Statistics

1. Introduction - Scipy, probability, review basics of data, the scientific method
2. Univariate and Bivariate - Data types, Summaries, 2way tables, correlation, Matplotlib, Seaborn
3. Hypothesis testing - AB testing, ttests, ANOVA, chi-square tests, regressions

Machine Learning

1. Introduction - SKLearn, algorithms
2. Supervised Learning - Regressions :Simple Linear Regression, MultiLinear Regression, Polynomial, Decision Tree, Support Vector Machine (SVM)
3. Supervised Learning - Classification (Logistic regression, Decision Tree Classification, SVM classification, K-Nearest Neighbor (KNN))
4. Clustering - K-means clustering
5. Ensemble - Random Forest (regression and classification) and XGBoost
6. Dimensionality Reduction - PCA & LDR
7. Recommender Systems - Apriori and Eclat
8. Natural Language Processing (NLP)

Deep Learning

1. Intro, Artificial Neural Networks (ANN)
2. Convolutional Neural Networks (CNN)

What next?

- Where can you apply this knowledge
- Setting up your projects and portfolio
- Opportunities out there