

Sessions

```
graph TD; Sessions[Sessions] --- S1[Basic Data Structure x 1]; Sessions --- S2[Basic Data Structure & Data Manipulation x 1]; Sessions --- S3[Data Exploration - Sanity Checks and Quality Reports x 1]; Sessions --- S4[Outliers, Missing Value Treatment, Variable Transformation and Information Value x 2];
```

Basic Data Structure x 1

Basic Data Structure & Data Manipulation x 1

Data Exploration - Sanity Checks and Quality Reports x 1

**Outliers, Missing Value Treatment, Variable
Transformation and Information Value x 2**

What you should know

```
graph TD; A[What you should know] --- B[Import and Export Data]; A --- C[Filter, Partition and Sort Data]; A --- D[Identify Data Types and convert to required format]; A --- E[Identify Features with less Variance and drop them]; A --- F[Identify and treat Missing Values]; A --- G[Identify and treat outliers]; A --- H[Transform Numerical Features]; A --- I[Transform Categorical Features];
```

Import and Export Data

Filter, Partition and Sort Data

Identify Data Types and convert to required format

Identify Features with less Variance and drop them

Identify and treat Missing Values

Identify and treat outliers

Transform Numerical Features

Transform Categorical Features

Python for Machine Learning

Import & Export Data

Data Types

Primitive

Number

Integer

Float

Boolean

String (also Object)

Containers

Lists

Tuples

Sets

Dictionary

Array

Numpy Array

Datetime

How to convert data types

Number to String

String to Number

String to Datetime

Basic Data Manipulation

Level 1

Filter

Select

Sort

Create Basic Features

Summary

Level 2

Grouping

Merge

Concatenate

Create Intermediate Features

Transition to Machine Learning

Feature Selection

Identify Low Variance Features and Drop them

Numerical

Missing Value - Identify and Impute

Outlier - Identify and Impute

Transformation

Non Normal to Normal or Scale to same range

Binning

Categorical

Missing Value - Identify and Impute

Outlier - Identify and Impute

Dummy Variable Creation

One Hot Encoding

Encoding

Ordinal Encoding

Mean Encoding

For Modelling

Train Test Split

Create DataFrames

Evaluate Models

Merge Data