



APSSDC

Andhra Pradesh State Skill Development Corporation



Skill AP
APSSDC

Data Analysis Using Python



DATA ANALYSIS

Agenda of the Training

Day1

Intro to Data
and Data
Manipulation
with NumPy

Day2

Data
Analysis with
pandas

Day3

Data
Preprocessing
with
Scikit-Learn

Day4

Cleaning Data
in Python

Day5

Introduction
to Data
Visualization
& Matplotlib

Day6

Data
Visualization
using
Seaborn

Data Generation



Introduction

Data Analysis is the process of collecting and organizing data.

DATA



SORTED



ARRANGED



PRESENTED
VISUALLY



Why Data Analysis is Required

Data analysis for B2C marketing plays a critical role in improving your business strategy.

It is important in business to understand problems facing an organisation, and to explore data in meaningful ways.



Process of Analysing the Data

- Import data sets
- Clean and prepare data for analysis
- Manipulate pandas DataFrame
- Visualize the data
- Build machine learning models using scikit-learn



Python Libraries for Analyzing Data



Pandas



matplotlib

Types of Data in Statistics

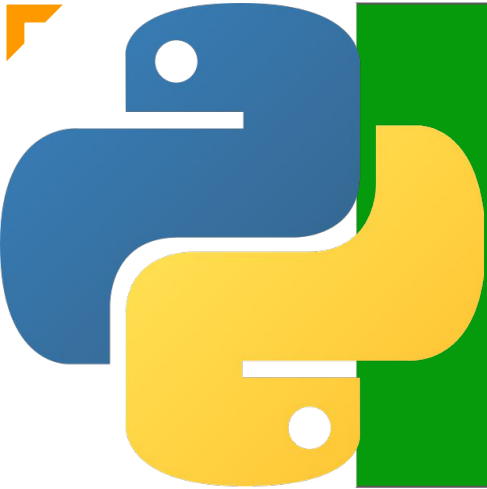
1. Numerical data
 - a. Discrete data
Ex: 0,1 or 0,1,2,3
 - b. Continuous data
Ex: 122,134.6,150,180,232,....
2. Categorical data
Ex: Yes/No, A Grade/ B Grad/ C Grade, Male/Female, etc, ..





APSSDC

Andhra Pradesh State Skill Development Corporation



Print("THANK YOU")

Slides created, maintained and distributed by
Andhra Pradesh State Skill Development Corporation