

REPORT ON DATA SCIENCE TRAINING

Topic: Data Science

Organised by: CBSE in partnership with Microsoft

Date, Time: 17th May – 24th May 2022, 4PM to 5PM

Speaker/Resource Person: Mr. Anip Sharma

Attendees: Nidhi Arora & Jyoti Khurana

The training on the topic “Data Science” was conducted by CBSE in partnership with Microsoft. The training on Data Science curriculum by CBSE from classes VIII to XII was covered in 7 different sessions:

Session 1: 17 May 2022

Topics discussed and explained:

- Introduction to Data and Data Science
- Ethics in Data Science and Data Governance
- DIKW Model
- Careers in Data Science

Session 2: 18 May 2022

Discussed about data collection and data visualisation:

- Arranging and Collecting Data
- Data Visualisation
- Use of Statistics in Data Science
- Distributions in Data Science

Session 3: 19 May 2022

Topics discussed and explained:

- Identify Patterns, Manipulation of Data
- Data Merging
- Assessing Data
- Forecasting on Data
- Data Randomisation

Session 4: 20 May 2022

Discussed different elements and structure of codes using R Studio

- Introduction to R Studio
- Explanation and demo of vectors, list, matrices, arrays and data frames

Session 5: 21 May 2022

Discussed Supervised Learning algorithms in detail.

- Classification Algorithms
- Regression Algorithms
- Introduction to Linear regression, Multiple regression and Non-Linear regression

Session 6: 23 May 2022

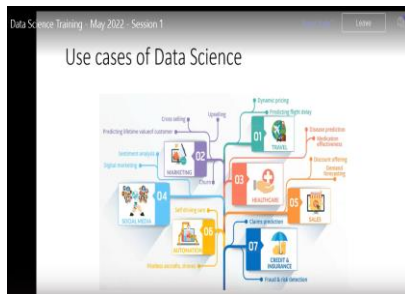
Discussed Unsupervised Learning algorithms in detail.

- Unsupervised Learning
- Real world applications of Unsupervised learning
- Introduction to Clustering
- K-Means Clustering

Session 7: 24 May 2022

Detailed discussion and demonstration of projects in data science.

- Projects in Data Sciences- Meteor-Showers, Space jam- A New Legacy



Big Data

- Data volumes exceeding the processing capacities of traditional databases
- Example: Data from millions of social media users
- Big Data Systems can extract statistical insights from a huge amount of data

Illustration of people interacting with data visualizations.

```
25 # u = [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15]
26
27 # u vector with sequence 1 to 15
28 u = np.arange(1, 16)
29 print(u)
30
31 # v vector with sequence 2.2 to 10.2
32 v = 2.2 + np.arange(1, 16)
33 print(v)
34
35 # Print out the dot product of u and v
36 # Use np.dot() to calculate the dot product of u and v
37 # Print out the result
38 # In [ ]:
39 #
```

