

# HADOOP

1. [Big Data Basics](#)
2. [Hadoop](#)
3. [Hadoop Analytical Tools](#)
4. [Internal working process of Hadoop](#)
  - 4.1. [Hadoop Daemons](#)
  - 4.2. [Hadoop cluster overview](#)
5. [Hadoop Ecosystem](#)
6. [Hadoop Security](#)
7. [Hadoop installation steps - Ubuntu](#)
8. [HDFS](#)
  - 8.1. [HDFS operation using CLI](#)
  - 8.2. [Running MapReduce using HDFS](#)
  - 8.3. [HDFS- site setup](#)
9. [MapReduce](#)
  - 9.1. [MapReduce Template](#)
  - 9.2. [Running MapReduce using CLI](#)
  - 9.3. [MapReduce-site setup](#)
10. [YARN](#)
  - 10.1. [YARN-site setup](#)
11. [HIVE](#)
12. [HIVE Installation steps](#)
13. [AWS-EMR](#)
14. [Challenges in installation and running Hadoop](#)
15. Case Studies
  - 15.1. [Running word count program on local and AWS service](#)
    - MapReduce code
    - Input data
    - Results
    - Jar file
    - Word count program without Hadoop
  - 15.2. [Web log analysis](#)
    - Problem Statement
    - Data pre-processing and analysis MapReduce code
    - Input data
    - Results
    - Jar file
    - Output results for running 1 million web log records using AWS
  - 15.3. [Retail data analysis](#)
    - Problem Statement
    - MapReduce code for analysis
    - Input data

- Results

#### 15.3. Retail data analysis

- Jar file

16. [Comparing performance of word count program with MapReduce on local Hadoop and AWS vs without MapReduce on local machine.](#)