**GOVERNMENT POLYTECHNIC, NASHIK**

**(An Autonomous Institute of Govt. Of Maharashtra)**

****

A

Seminar Report

On

**“Big Data & Hadoop”**

For The Course

Third Year Diploma In Computer Technology

Submitted By

Sudarshan Sharad Gawale

( 176114 )

Guided By

Mr. P. B. Mali

Submitted To

Government Polytechnic

Nashik

**GOVERNMENT POLYTECHNIC, NASHIK**

**(An Autonomous Institute of Govt. Of Maharashtra)**

****

**CERTIFICATE**

This is to Certify that Seminar Report on **Big Data & Hadoop** has been successfully completed by, **Sudarshan Sharad Gawale** In the fulfillment of requirement of Diploma in “Computer Technology” from “Government Polytechnic, Nashik” during the Academic Year 2019-2020 is record of his own work carried in my guidance. He has satisfactorily completed this seminar.

|  |  |
| --- | --- |
| **Guided By** | **H.O.D** |
| Mr. P. B. Mali | Mr. Y. B. Sanap |

**Principal**

Prof. D. P. Nathe

Government Polytechnic, Nashik

**Acknowledgement**

I express our deep sense of gratitude and respect to our Guide Mr. P. B. Mali. His valuable Guidance and inflicting help during the report work.

His esteemed suggestions and encouragements during the entire period have been valuable for us in achieving these goals.

I also wish to express our deepest gratitude to all staff members of Computer Technology Department for their valuable support.

I also want to thank all my colleagues for contribution I making this seminar a success.

Sudarshan Sharad Gawale

Roll No: 176114

(Third Year Computer Technology)

Govt. Polytechnic, Nashik

**Abstract**

Everyday enormous amount of data is being produced worldwide. Big Data analytics has brought a big opportunity for organizations. Companies capture trillions of bytes of information about their customers, suppliers, and operations. IT organizations are exploring the analytics technologies to explore web-based data sources and extract value from the social networking boom. In the western world, organizations are wondering about the kind of business intelligence they could derive from all the information they have at their disposal.? The organizations are trying to leverage Big Data by trying to make sense from the data that they have and by securing it. In the next three to five years, there will be a widening gap between companies that understand and exploit Big Data and companies that are aware of it but do not know what to do with it. Already the forward thinking players of the banking, insurance, manufacturing, retail, wholesale, healthcare, communications, transportation, construction, utilities, and education are successfully using big data by exploiting meaningful information from all the Data they have and using those information in formulating their strategic moves. Those companies who will be able to use Big Data successfully will be clearly ahead of those who will react slowly to capitalize on Big Data.

**Index**

|  |  |  |
| --- | --- | --- |
| **Sr. No.** | **Chapter Name** | **Page No** |
|  |  |  |
| 1 | What is Big Data? | 1 |
|  |  |  |
| 2 | Characteristics of Big Data | 3 |
|  |  |  |
| 3 | What is Hadoop? | 6 |
|  |  |  |
| 4 | Hadoop : Big Data & Hadoop – Restaurant Analogy | 9 |
|  |  |  |
| 5 | Hadoop Features | 15 |
|  |  |  |
| 6 | Conclusion | 18 |
|  |  |
|  |  |  |
| 7 | Image References | 19 |
|  |  |
|  |  |  |
| 8 | References | 20 |
|  |  |
|  |  |  |