Analysis of Airport Data using Hadoop-Hive: A Case Study

Abstract

In the contemporary world, Data analysis is a challenge in the era of varied inter-disciplines though there is a specialization in the respective disciplines. In other words, effective data analytics helps in analyzing the data of any business system. But it is the big data which helps and axialrates the process of analysis of data paving way for a success of any business intelligence system. With the expansion of the industry, the data of the industry also expands. Then, it is increasingly difficult to handle huge amount of data that gets generated no matter...
what's the business is like, range of fields from social media to finance, flight data, environment and health. Big Data can be used to assess risk in the insurance industry and to track reactions to products in real time. Big Data is also used to monitor things as diverse as wave movements, flight data, traffic data, financial transactions, health and crime. The challenge of Big Data is how to use it to create something that is value to the user. How can it be gathered, stored, processed and analyzed it to turn the raw data information to support decision making. In this paper Big Data is depicted in a form of case study for Airline data based on hive tools.

References

- Data set is taken from edureka http://www.edureka.co/my-course/big-data-and-hadoop
- Manjunath T N et. al, The Descriptive Study of Knowledge Discovery from Web Usage Mining, IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 5, No 1, September 2011 ISSN
- HiveQL Language Manual
- Apache Tez
- Working with Students to Improve Indexing in Apache Hive

Index Terms

Computer Science

Distributed Systems

Keywords

Airline Data Set

Hive Tools.