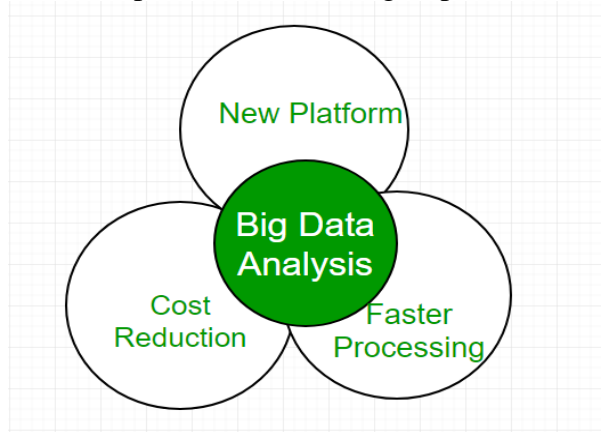# Benefits of Big Data

As per Oxford English Dictionary Big data is "data of a very large size, typically to the extend that its manipulation and management present logistical challenges". Big Data refers to technologies and initiatives that involve data that is too diverse, fast-changing or massive for conventional technologies, skills and infrastructure to address efficiently.

But nowadays with the help of new technologies it is very easy to realize value of Big data, for example to identity shopping behavioural trends of costumers to improve stockage, pricing of the items. Government processes are also get benefits and banking institutions are capturing data and customer interaction to model risk and fraud.

Benefits in IT sectors:

- Many old IT companies are fully depends on big data in order to modernise their outdated mainframes by identifying the root causes of failures and issues in real time and antiquated code bases. Many organisation are replacing their tradition system with open source platforms like Hadoop.
- Most big data solutions are based on hadoop, which is designs to scale up from single machine to thousands of machines, each offering local computation and storage, moreover it is "free" open source platforms, allowing minimizing capital investment of an organisation in acquiring new platforms.
- With the help of big data technologies IT companies are able to processed fast the third-party data that is often hard to understand at once by having inherently high-horsepower and parallelized working of platforms.

Benefits of Big Data in Business:

- Data quality has a direct impact on business process efficiency. In purchase to pay process, poor quality vendors data can cause missing of purchase contracts or pricing information which can lead to delays in procuring vital goods. Many company using big data solutions or algorithms to simply do what they have already been doing , so that there is no data loss moreover if we run the algorithm against the data set ,the result might the list of individual who exhibits attributes of fraudulent behaviour.
- In order to cash process ,incomplete or inaccurate credit limits or pricing information can lead overall customer service lost or reduce revenue or may increase service cost ,with the help of big data technologies and ability to run various algorithm more quickly ,the data can be updated at regular intervals throughout the day.

- The systematic analysis of data or data profiling is used to assess the overall health of the data which lead to proper business decisions in accordance with present situation because sometimes inaccurate data results in incorrect management, means business decisions are based on incorrect information .For example the more one can understand about costumers complex relationships, preference, and interaction history with company, the more relevant and timely once business outreach.

Benefits in Enterprise:

- Big data might allow a company to collect trillions or billions of real-time data points on its products, resources, or costumers- and then repackage the data instantaneously to optimize customer experience.
- Speed at which data updated using big data technologies allow enterprises to more quickly and accurately respond to customer demands. For example MetLife used MongoDB to quickly consolidate customer information over 70 different source and provide single, rapidly-updated view
- Big data can help enterprises to act more nimbly allowing them to adopt to changes faster than their competitors.

Benefits in Other Areas:

- Big data technologies used to predict the 'buy' and 'sell' decisions made on a share of different companies for the costumers.
- Hospitals are analysing medical data and records to predict those patients that are likely to seek readmission within a few months of discharge. The hospital can the avoid costly stay of patients in hospital.
- Search-Engine retrieve lots of Data from different data bases in fractional of seconds using big data technologies. For example Google use MapReduce algorithm to search for a given query. MapReduce divides the task into small parts and assigns those parts to many computers connected over the network, and collect the result to form the final result.

- Financial Services organisation are using big data for data minning about customer interactions to slice and dice their users into finely tuned segments, this will help in creating increasingly relevant and sophisticated offers.
- Insurance companies are using Big Data analysis to see which home insurance application can be immediately processed, and which ones need a validating in-person visit from an agent.
- Web-based business are developing information products that combine data gathered from customers to offer more appealing recommendations and more successful coupon programs.

Big Challenges with Big Data

The challenges in Big Data are the real implementation hurdles. These require immediate attention and need to be handled because if not handled then the failure of the technology may take place which can also lead to some unpleasant result. Big data challenges include the storing, analysing the extremely large and fast-growing data.
Some of the Big Data challenges are:

1. *Sharing and Accessing Data:*
   - Perhaps the most frequent challenge in big data efforts is the inaccessibility of data sets from external sources.
   - Sharing data can cause substantial challenges.
   - It includes the need for inter and intra- institutional legal documents.
   - Accessing data from public repositories leads to multiple difficulties.
   - It is necessary for the data to be available in an accurate, complete and timely manner because if data in the companies information system is to be used to make accurate decisions in time then it becomes necessary for data to be available in this manner.

2. *Privacy and Security:*
   - It is another most important challenge with Big Data. This challenge includes sensitive, conceptual, technical as well as legal significance.
   - Most of the organizations are unable to maintain regular checks due to large amounts of data generation. However, it should be necessary to perform security checks and observation in real time because it is most beneficial.
   - There is some information of a person which when combined with external large data may lead to some facts of a person which may be secretive and he might not want the owner to know this information about that person.
   - Some of the organization collects information of the people in order to add value to their business. This is done by making insights into their lives that they're unaware of.

3. *Analytical Challenges:*

- There are some huge analytical challenges in big data which arise some main challenges questions like how to deal with a problem if data volume gets too large?
- Or how to find out the important data points?
- Or how to use data to the best advantage?
- These large amount of data on which these types of analysisis to be done can be structured (organized data), semi-structured (Semi-organized data) or unstructured (unorganized data). There are two techniques through which decision making can be done:
  - Either incorporate massive data volumes in the analysis.
  - Or determine upfront which Big data is relevant.

4. *Technical challenges:*
   - **Quality of data:**
     - When there is a collection of a large amount of data and storage of this data, it comes at a cost. Big companies, business leaders and IT leaders always want large data storage.
     - For better results and conclusions, Big data rather than having irrelevant data, focuses on quality data storage.
     - This further arise a question that how it can be ensured that data is relevant, how much data would be enough for decision making and whether the stored data is accurate or not.
   - **Fault tolerance:**
     - Fault tolerance is another technical challenge and fault tolerance computing is extremely hard, involving intricate algorithms.
     - Nowadays some of the new technologies like cloud computing and big data always intended that whenever the failure occurs the damage done should be within the acceptable threshold that is the whole task should not begin from the scratch.
   - **Scalability:**
     - Big data projects can grow and evolve rapidly. The scalability issue of Big Data has lead towards cloud computing.
     - It leads to various challenges like how to run and execute various jobs so that goal of each workload can be achieved cost-effectively.
     - It also requires dealing with the system failures in an efficient manner. This leads to a big question again that what kinds of storage devices are to be used.