



Real World Examples of Big Data

Anurag Mantripragada
Vikas Hanumegowda
Khyati Ganatra
Arathi Mary Sunder



Online Ticket Booking Data

- The **volume** of data collected is increasing with the increase in usage of “The Internet” to make lazy and informed decisions.
- A large **variety** of data is collected from around the world on which “values” can be filtered out.
- Data-**velocity** peaks during holidays and characteristic events like NYEs and exotic indigenous events.
- **Variability** exists in the collection due to the above events and values from this variation can be used by other data moguls like Ad, insurance and other companies for profits.
- Hubs like LA, NYC, Dubai, etc. have huge data flow throughout the year whereas lesser used airports like Cape Cod have seasonal volumes.
- **Unintended consequences** are ever existent! - ranging from making profits by varying ticket prices (:-}) to calculating numbers for acts of terror (:-{).



Fitness Trackers

→ **How is it collected?:** Smart devices for health collect huge data for an individual and there are millions of users. So the data has **volume**. These devices have inbuilt GPS trackers and heart rate monitors that collect the heart rate every minute. GPS trackers collect your movement through GPS location changes.

→ **Temporal and Spatial Coverage:** Different aspects of your fitness are collected like heart rate, running speed, geolocation etc. Hence there is variety in the data.

Also, data is collected everyday that we use, so it tracks the steps you walk on a particular day(Apple health).

→ Since all this data can be used for better user experiences plus health of the user, this adds value to the data. These three Vs make this kind of data as big data.

-**Unintended consequences:** Your fitness details can be joined with other information like your visits to a healthcare locations and sold to health insurance agencies.



Uber/Lyft Users data

- This is an example of Big Data as millions of users are generating data on daily basis.
- With more people using on-demand transportation, more data is obtained. This is **volume**
- **Velocity** is when there are thousands of users using the service at a given point of time
- As information like ride history of individuals, credit card number, name, etc are collected, we can say there is **variety** in data
- Given that there are peak hours, with more users, busy cities and favourite spots, there is also **variability** in the data
- **Unintended use**: It is a great loss if sensitive information like credit card information is misused. Also, ride history can be used to get information like Home address, work address, etc which can be misused to plan a theft or any bad activity



Conclusion

- Data is the fuel of 21st century. Due to the constantly increasing use of Internet and smart, IoT devices, **Big Data** comes into picture.
- No matter what we think of in our everyday life, it is information and that can be tapped for data.
- With data, comes power both good and evil; and we hope to learn about it (a lot!) in this course.



Thank you