



# Data - Lifeline of the Modern World

---

Ishaan Shrivastav  
Redlands High School



# Data - Lifeline of the Modern World

---

## **Introduction - Data is an integral part of modern world**

Role big data plays in today's world

How big data can make an impact in the future

Tools to store, analyze, and visualize big data

My big data project to analyze social media data

# Types of Data

**Communication:** phones-calls  
emails satellite-transmission  
internet

**Transportation:** Flight-status  
vehicle/plane-location airport-  
queue traffic-condition  
passenger-count parking-meters

**Public Safety/Health:**  
Police ambulance crime-data diseases-incidences  
hospital-location vaccination-records



**Public Records:** census parcels  
population tax-records  
birth/death/marriage  
records

**Environment:** noise, CO2 humidity  
temperature air-pressure air-quality  
water-quality rain-gauges ozone-  
concentration

# Who is producing Big Data?

## Social media

Twitter, Instagram, Facebook, WhatsApp, Youtube, ...

## Scientific Data

Satellite Imagery, Health metrics, Space observations



# Who is producing Big Data?

## Modern 'Live' Devices (Internet of Things)

Mobile phones, Smart meters, Weather sensors, In-built car sensors

## Corporations

Financial transactions, Sales

# Data Explosion in Social Media

What happens  
**online**  
in 60 seconds?  
(2012 - 2014)



Picture by Centre for Learning and Teaching  
3<sup>rd</sup> November 2014

## By 2020!!

30 Billion Mobile Devices

40 Trillion Gigabytes of  
Data

10 Million mobile apps

All for 8 Billion people

# Data - Lifeline of the Modern World

---

Introduction - Data is an integral part of modern world

**Role big data plays in today's world**

What Big Data can be used for in the Future

Tools to Store, Analyze, and Visualize data

My Big Data project to analyze social media interaction

# Role Big Data Plays in Today's World

---

- Transportation

- Checking flight status, identifying plane location, finding traffic conditions

- Public Safety

- Police Stations, Fire Stations, and Surveillance all use data coming from different sources and make real-time decisions for public safety



# Role Big Data Plays in Today's World

---

- Energy Usage

- Electricity, Gas, and Smart meters produce large volume of data that is analyzed by utility companies for best utilization of resources

- Weather

- Prediction for weather (temperature, precipitation etc.) and natural disasters (tornadoes, hurricanes, tsunami) are made by analyzing large volume of scientific data coming from sensors around the world.

# Data - Lifeline of the Modern World

---

Introduction - Data is an integral part of modern world

Role big data plays in today's world

**How Big Data can be used in the Future**

Tools to Store, Analyze, and Visualize data

My Big Data project to analyze social media interaction

# How Big Data can be used in the Future

---

Self driven cars will communicate with each other about road conditions, traffic and gas prices data and making intelligent decision based on them.

Dating sites will find the perfect soulmate based on data available about members interest, movies they like, donation they make to charities, association with different organization etc. :)

## Better healthcare management

Chronic disease like diabetes and high blood pressure patients will be wearing devices that will be constantly sending their health related data to doctors office so they can make appropriate adjustments

# Data - Lifeline of the Modern World

---

Introduction - Data is an integral part of modern world

Role big data plays in today's world

What Big Data can be used for in the Future

**Tools to Store, Analyze, and Visualize data**

My Big Data project to analyze social media interaction

—

# Tools to Store, Analyze and Visualize Big Data

# Storing Big Data

---

Distributed file systems

Hadoop, cloud data storage (S3, azure blob storage)

Distributed Data Management systems

Elasticsearch, CouchDB, DynamoDB



# Analyzing Big Data

Elasticsearch

Microsoft Power BI

SAP HANA, Teradata

An example of analyzing big data stored in elasticsearch using different techniques

Police department wants to improve public safety by taking citizen petrol help for non-violent crimes such as vandalism

# Analyzing Big Data

---

## **Crime\_incidents**

```
{  
  "Type" : "vandalism",  
  "Time" : "15:30",  
  "Date" : "12/31/2016"  
  "Location" : {  
    "Lat" : 34.0656  
    "Long" : 117.16645  
  }  
  "Case_solved" : "No",  
  "Officer_assigned" : "James Smith"  
}
```



# Spatial analysis

- Find statistics for number of vandalism within 10 miles of city hall

- {

```
"query": {  
  "bool" : {  
    "must" : {  
      "match_all" : {}  
    },  
    "filter" : {  
      "geo_distance" : {  
        "distance" : "10miles",  
        "pin.location" : {  
          "lat" : 40,  
          "lon" : -70  
        }  
      }  
    }  
  }  
}
```

# Analyzing Big Data

- Filter data

- Find all vandalism type crimes that occurs

```
“query”: {  
    “term” : {“Type” : “vandalism”} }
```

# Analyzing Big Data

- Aggregation

- Find when do vandalism crime happens most

GET /\_search

```
{
  "query": {
    "term" : {"Type" : "vandalism"}
  },
  "aggs": {
    "tweets_over_time": {
      "date_histogram": {
        "field": "Time",
        "interval": "60m"
      }
    }
  }
}
```

# Total Query

GET /\_search

```
{
  "query": {
    "bool" : {
      "must" : {
        "match_all" : {
          "query": {
            "term" : {"Type" : "vandalism"}  }
          },
          "filter" : {
            "geo_distance" : {
              "distance" : "10mi",
              "pin.location" : {
                "lat" : 40,
                "lon" : -70
              }
            }
          }
        }
      }
    }
  }
}
```

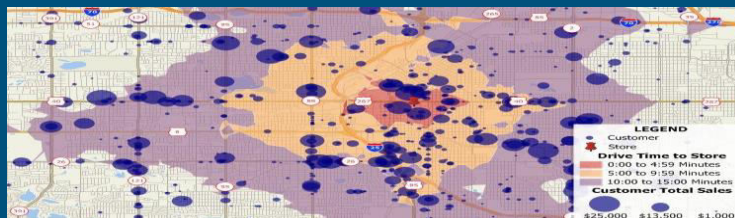
# Visualizing Big Data

## ArcGIS

Mapping tool that can use data to show trends across space

## Tableau

Tool that can create charts and diagrams to display trends in data



# Data - Lifeline of the Modern World

---

Introduction - Data is an integral part of modern world

Role big data plays in today's world

What Big Data can be used for in the Future

Tools to Store, Analyze, and Visualize data

**My Big Data project to analyze social media interaction**

# My Big Data Project:

—Goal: To discover social pattern based on big data available in social media

Project: To see if there was a correlation between the number of tweets and scoreline during the SuperBowl game

Software used:

Storage and analysis: Elasticsearch including logstash extension

# My Big Data Project:

---

## Storage

### Twitter feeds

Use Logstash pluggin in Elasticsearch to store live twitter feeds into elasticsearch during super bowl

Filter twitter feeds based on keywords for NFL teams (patriots and falcons)

1.7 million tweets



# My Big Data Project:

## Analysis

Used Elasticsearch search and aggregation queries to find tweets for different teams during different time of the game

Search Query to find certain tweets  
team

Search query to find number of tweets per quarter per

```
GET /_search
```

```
{  
  "query": {  
    "term" : {"text": "patriots"}  
  }  
}
```

# My Big Data Project:

