

# Big Data Analytics

## EVENT DATA

Don't Think Records...Think Events!  
Events are happening all the time:

- >Senors ping every second
- >Websites--search, scroll, each twitter feed = EVENT

## INDUSTRIES

- >Banking/Financial Service
- >Bio-Tech/Pharmaceuticals
- >Entertainment

## POSSIBLE BUSINESS PROBLEMS

- >BI/DW taking long to process with high turnaround time?
- >Are there needs for Real time Access to enable faster decision making?
- >Like to improve Customer Experience
- >What other data related pain points are there?
- >Have a need to create a business case or compete better with

## SKILLS NEEDED

- >Data Engineering
- >Data Science (Stats and Math background)
- >Data Modelers
- >Solution Architecture

## TOOLS/TECHNOLOGY STACK

- >Real-time data processing
- >Queuing Mechanism (Kafa/Flume)
- >Big Data on Cloud (Amazon AWS)

- >Hadoop/Spark
- >MapReduce
- >PIG
- >HIVE
- >Cloudera
- >HortonWorks
- >MapR
- >Amazon

## NO SQL STORES

- >CouchDB -- KeyValue
- >Cassandra -- Column Store
- >BynomoDB --Cloud Based
- >MongoDB -- Document Based
- >Hbase -- Column Based
- >Beo4J -- Graph DB

## VISUALIZE/ANALYTICS

Recommendation  
Customer Experience  
Email Targeting  
Predict Customer Behavior

Build Analytics Models  
NumPy/SciKit<  
R<  
Matlab<  
SAS<

Tableau<  
D3<  
Qlik<

## HOW TO PREVENT DATA LOSS, KEEP DATA IN SYNC AND SERVE ACROSS THE WORLD?

