



## Big Data Use Case Group Discussion Activity



# Today's Agenda....

## (First Slot)

1. Group Exercise on a use case

## (Second Slot)

1. Linux and HDFS commands



# Group Exercise





# Problem Background

- Any e-commerce industry requires customer retainment.
- Retaining customers is a key metric for retailers, given the cost of acquiring new customers.
- To serve better we should know our **customer better.**





# Problem Statement

A online retail store wants to improve their customer satisfaction and sales opportunities by integrating information of product data and the customer data from various data sources that are available online.

Arrive a solution that can help the client to improve sales by providing the recommendations as well as improving on his products.

# Group Exercise



- Understand the Problem
- Justify whether it is a single problem
- Analyze the data required for the problem.
- Justify if it as a Big Data Problem.
- What type of problems are involved.

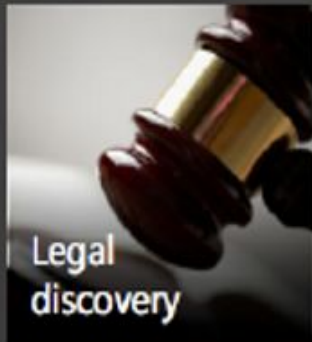
# BIG DATA Application Examples



IT infrastructure optimization



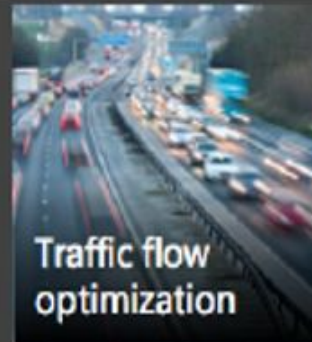
Legal discovery



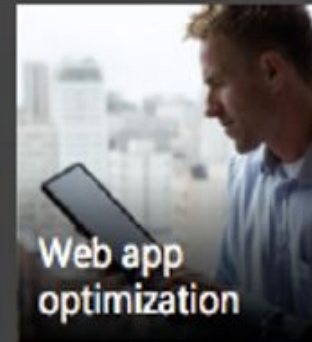
Social network analysis



Traffic flow optimization



Web app optimization



Churn analysis



Natural resource exploration



Weather forecasting



Healthcare outcomes



Fraud detection

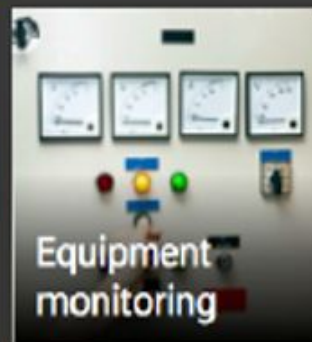


Life sciences research

Advertising analysis



Equipment monitoring



Smart meter monitoring







# Group Activity

## – Case Study – Group Activity

- Pick one domain (i.e. Retail, Telecom etc.) And will work on the following.
- Identify different type and sources of data.
- Identify the engineering problems they may have solving.
- Classify the problem.
- How this can be solved using regular ML/Analytics tools.
- What is the type of data – Volume, Velocity and Variety aspect?
- Extend the same to big data analytics.