

# Industry session on Analytics

**Shiv Kumar**

Department Head : Data  
Sciences

Mahindra and Mahindra

*Session : Series 01 (Not for circulation- Academic purpose only )*

*Approach represents industry view on exploring career in analytics Industry & setting advanced analytics labs*

# AGENDA

Managing thoughts as a Data Scientist

Analytics problem statement Overview

Approach discussion

Skill development techniques

# Experienced analytics professional in diversified analytics application

## Current Role –

Department Head : Mahindra and Mahindra

- ❖ *Leading the Data Sciences Practice at Information Insight Center – Automotive, Farm and Agri in year -1(16 Bn USD business), Analytics set up initiative for across the M&M group analytics services*

## Application areas

- ❖ *Commercial credit risk, retail analytics, Marketing and Sales Analytics, reliability engineering, Demand forecasting, social media analytics, coupon redemption / promotional effectiveness analysis for retailers, drug discovery analytics, Lead prioritization modeling*

## Analytics career

- ❖ *Data science and Strategy Consulting practitioner with ~11 years of experience with management consulting and predictive analytics firms*
- ❖ *Past Associations with boutique advanced analytics companies :*

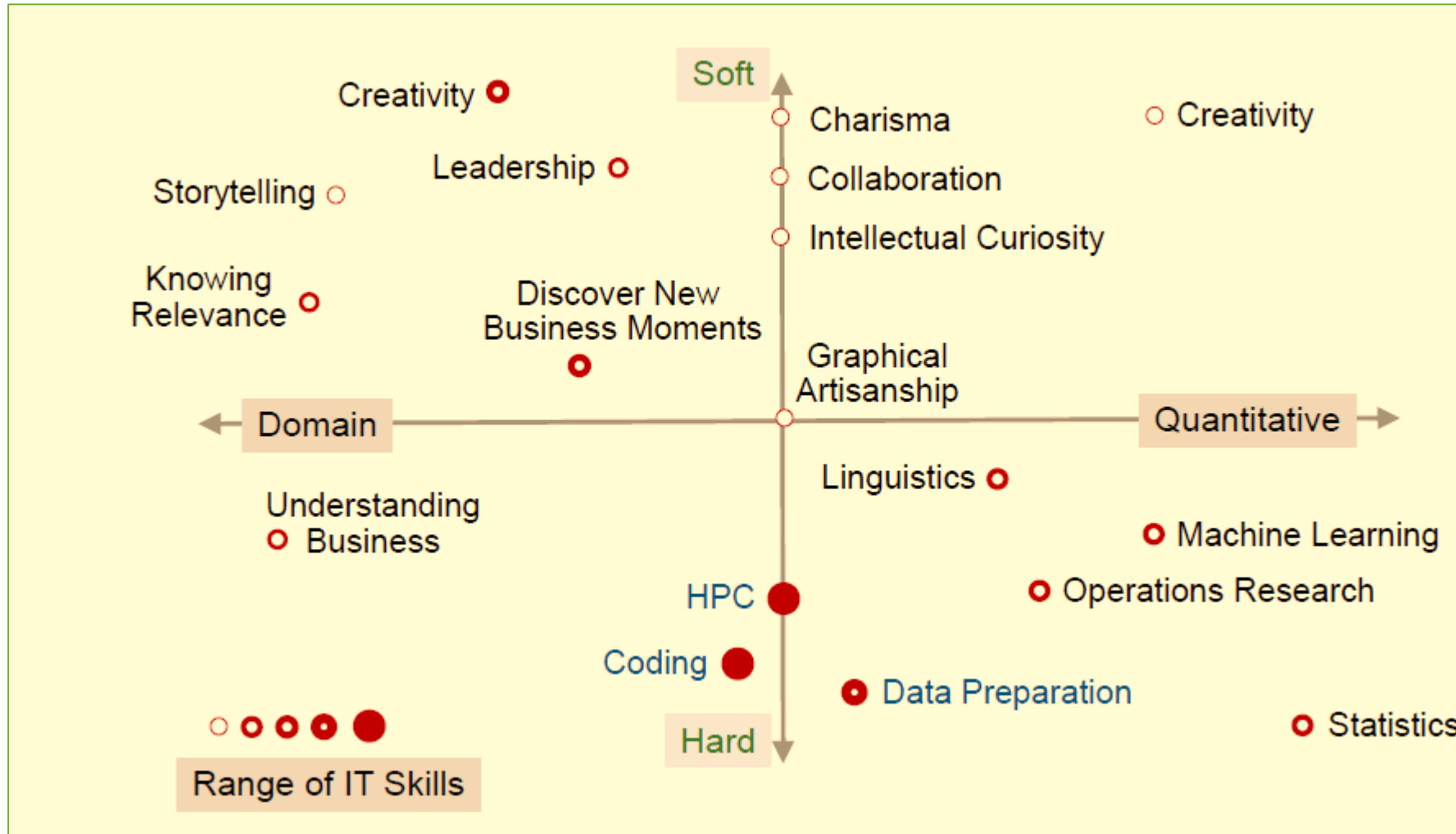


## Academics

- ❖ *MBA – Indian Institute of Management Kozhikode*
- ❖ *BS – Mathematical Statistics Institute of Science – BHU*



# Analytics leadership skill quadrant ...List your key strength, Locate your quadrant



Securing internal and external collaboration to build the talent pool: (Analytics partner companies, Academic institutions to resources and R&D pool, Tools and technology partner companies )

*Lets look at a problem from data scientist lenses ... A wine tasting experience ..*



LOOK

SWIRL

SMELL

TASTE

Wine Quality = 12.145 + 0.00117 Winter Rainfall + 0.0614 Average growing season temperature - 0.00386 Harvest rainfall

# Discussion topics: Framework on career moves | opportunities | Leadership in Analytics Practices | Industry Practice

Understanding Different Roles in  
analytics Industry

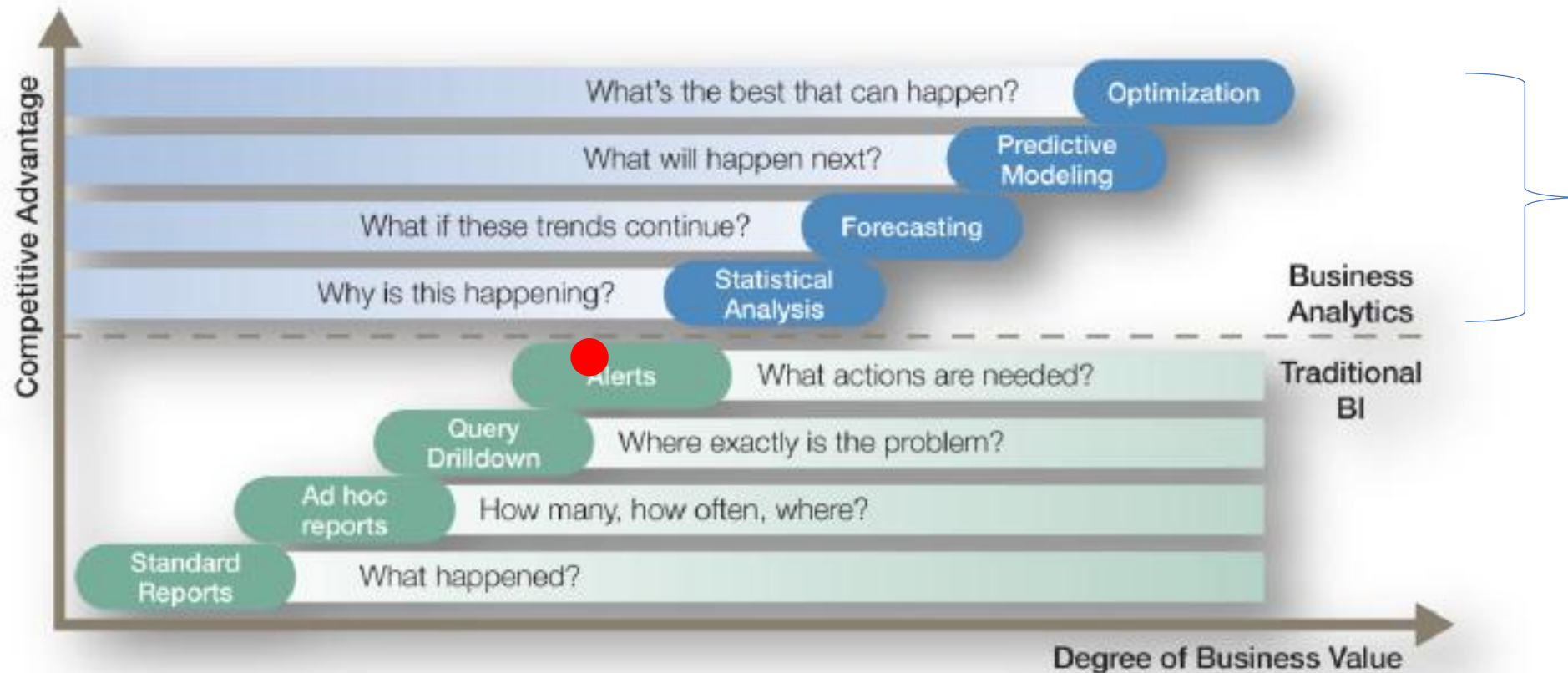
Ways and means to Crack the  
Analytics opportunity  
(Internal/External )

**In-house analytics  
Incubation practices for  
leadership**

Learning the Hard and Soft skills for  
Analytics professionals – HPC ,  
Storytelling , Graphical Artisanship

Analytics adoption maturity in  
various Industry

# Maturity Scale of Analytics....



Source: Developing an Analytics Center of Excellence (Or The Care and Feeding of Magical Creatures) - Charles D. Kincaid, Experis Business Analytics Practice, Portage, MI

# *\*Become Analytics Evangelist.. Evaluate internal opportunity..*

## Four pillars to assess analytics maturity of the business unit / Department

**1 People**

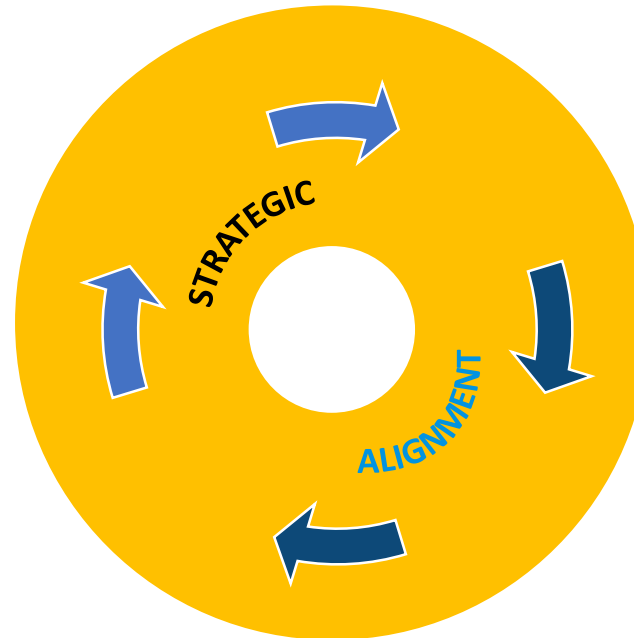
**Mix of data science, business acumen and technical expertise**

- Able to identify the most relevant type of insight to the business
- A team with different but overlapping skills that knows collaboration

**3 Data**

**Quality, consistent data shared in a manner that is easy to access?**

- Right systems to capture the required data
- Central repository of data
- Collecting useful data
- Right technology to access and store data
- Leverage external data sources



**2 Intent**

**Resolving to be data driven ; creating structures, processes and incentives to support analytical decision making**

- Top leadership define how analytics will shape the business performance
- Organization makes data driven decisions very frequently and much faster than competition

**4 Tools**

**State of the art tools like unstructured databases, scale out compute clusters and heuristic instrumentation**

- Advanced analytics and statistical tools used
- Open source vs high end tools
- Ecosystem of mature analytics, visualization and data management

# *Analytics Evangelist.. Needs to research around*

- Firms' internal R&D is largely focused on improving the firms' own performance
- Relatively slow pace of digital and analytics transformation of the economy
- Many business leaders are uncertain about
  - what exactly AI can do for them,
  - where to obtain AI-powered applications,
  - how to integrate them into their companies,
  - how to assess the return on an investment in the technology

# Setting up data science labs...

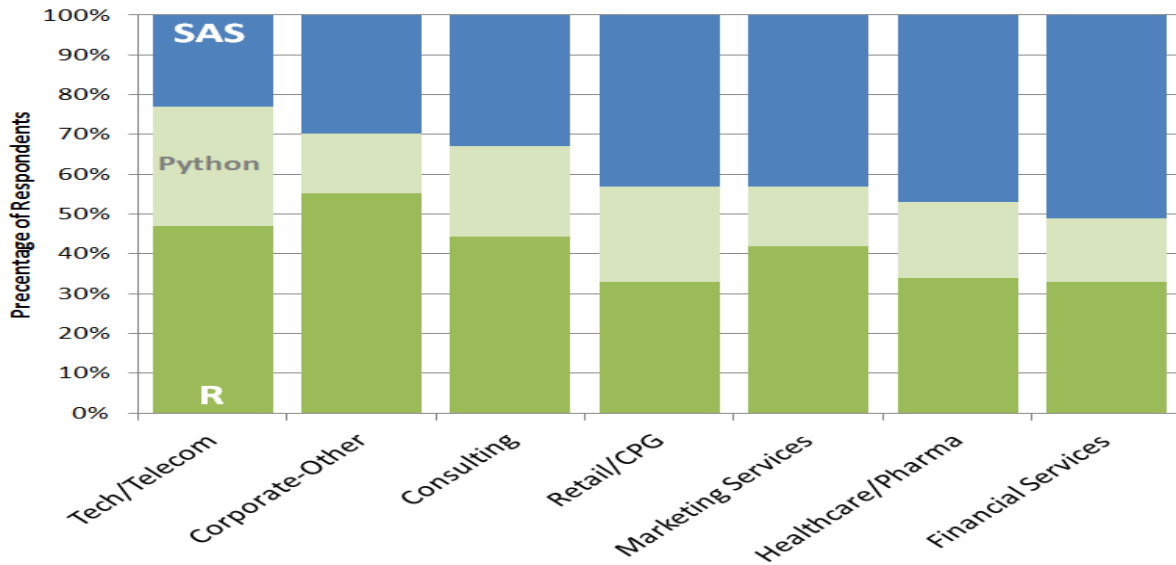
- **Data and analytics** have altered the dynamics in many industries, can fuel discovery and innovation
- Combining analytical tools with business insights for decision makers **multiplies the business value**
- **Data-Driven Thinking** and **Innovation**, Big data exploration, Process refinement, Cross-Functional View etc. are important principles of **Data Science Labs**
- Massive data integration capabilities can break down organizational silos

## Modelling Play Store

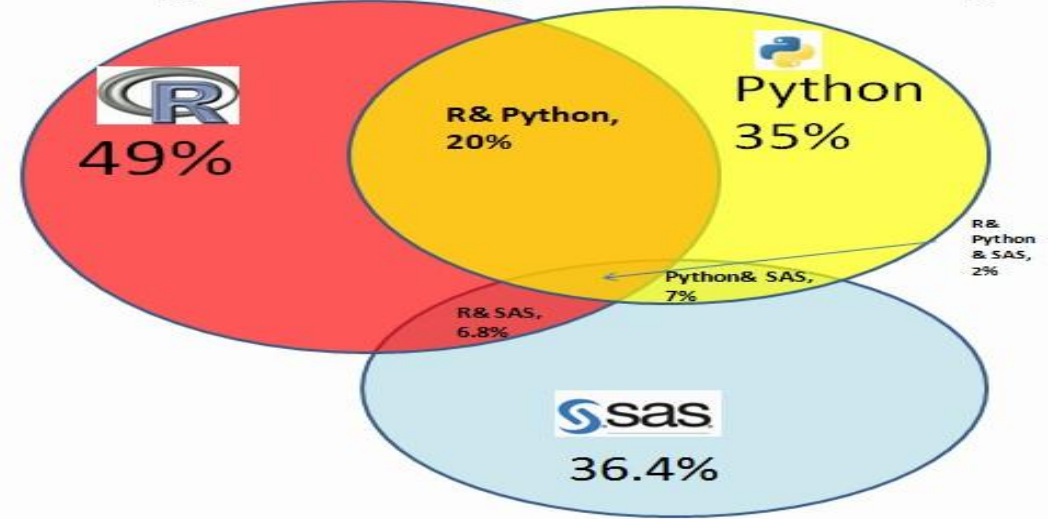
- Comprehensive module aiding audience to understand **statistical techniques** to solve **complex predictive analytics** problem
- Multiple domains like Warranty, Retail, Ecommerce, Banking etc. can be undertaken
- Uncover business insights from **structured and unstructured data** to have better understanding of the current business scenario
- Overview about most frequently used ML techniques for solving business problem like **Customer Segmentation, Clustering** and **Business Forecasting**
- The modelling Playstore module will be followed up by a small problem solving activity with the participants consisting predictive models such as
  - Decision Tree
  - Linear Regression
  - Logistic Regression
  - Clustering (k-Means)

# Comparison among tools and Techniques

## Tool Preference by Industry



## KDNuggets 2014 Poll: Languages used for Analytics/Data Mining, 2



	SAS	SPSS	R	Python
Advantages	<ol style="list-style-type: none"> <li>High adoption rate in major industries</li> <li>Flow based interface with drag and drop</li> <li>Official support</li> <li>Handling large datasets</li> <li>'PROC SQL'</li> </ol>	<ol style="list-style-type: none"> <li>Used a lot in universities</li> <li>Good user interface with extensive documentation</li> <li>Click &amp; Play functionality</li> <li>Writing code made easy using the 'paste' button.</li> <li>Official support</li> </ol>	<ol style="list-style-type: none"> <li>Big community who creates libraries</li> <li>Free</li> <li>Early adopter in explanatory and predictive modeling.</li> <li>Easy to connect to data sources, including NoSQL and webscraping.</li> </ol>	<ol style="list-style-type: none"> <li>Scalability</li> <li>General purpose language</li> <li>Easy to learn</li> <li>Good in machine learning</li> <li>Big community</li> <li>Free</li> </ol>
Disadvantages	<ol style="list-style-type: none"> <li>Relatively high cost</li> <li>For not-standard options not in interface, you'll need to write the code</li> <li>Slow adapting to new techniques</li> <li>Different programs for visualization or Data Mining</li> </ol>	<ol style="list-style-type: none"> <li>Relatively high cost</li> <li>different licenses for different functionalities.</li> <li>Syntax limited</li> <li>Slow adapting to new techniques</li> <li>Slow in handling large datasets</li> </ol>	<ol style="list-style-type: none"> <li>Can be slow with big datasets</li> <li>Steep learning curve</li> <li>No official support</li> <li>No user interface</li> </ol>	<ol style="list-style-type: none"> <li>Not as strong in explanatory modeling</li> <li>Choice of version: 2.7 or 3.5?</li> <li>No user interface</li> <li>No official support</li> </ol>

# Approach Creation -

**BUSINESS  
BASELINING**

**Think Data**

**Model Selection**

**Model  
Validation**

- Business Validation
- Insights Creation
- Business adoption
- Post analytics intervention measurement

# Project 1 - Driving sales transformation using predictive algorithms to enhance conversion rates

“~60% of the Sales happening from classified High Priority Enquiries, it’s conversion rate being ~3.0 times the average conversion rate of low & medium priority enquiries ”

## Project Outline

### 1 Background

- **M&M** is a key player in Indian Automotive sector & receives **huge volume of enquiries** for purchase of personal & commercial segment vehicles.
- Dealerships across Pan-India need to **prioritize** such high volume of enquiries to **maximize the sale**

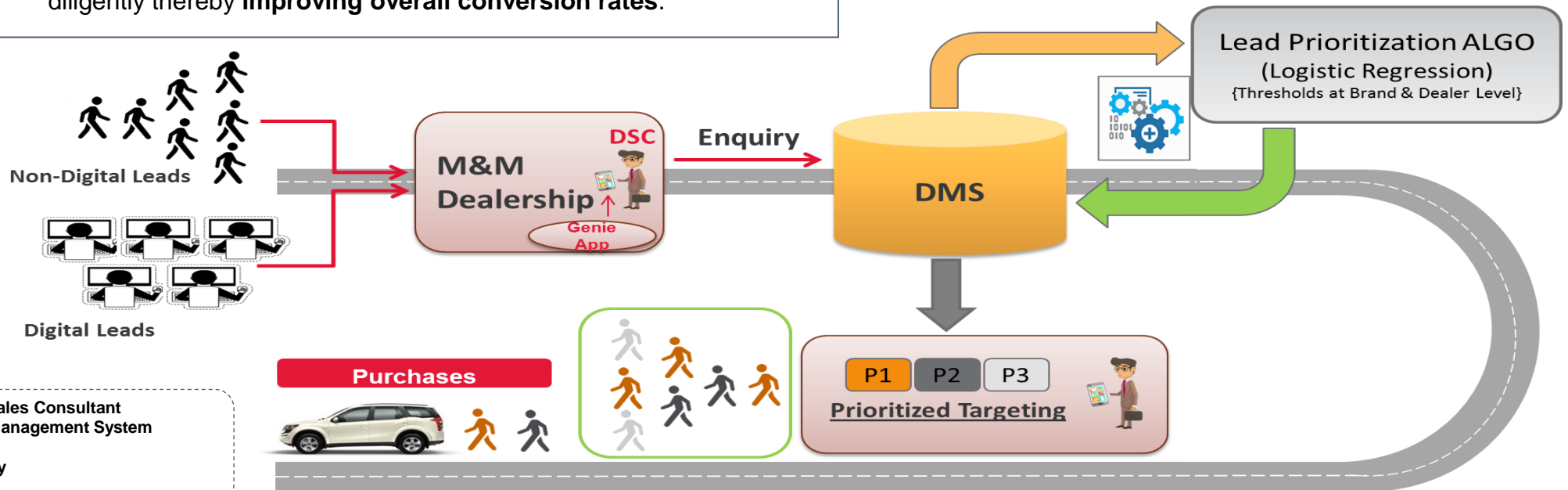
### 2 Objective

- The project is aimed at building and sustaining a robust **predictive algorithm** which can predict & identify high quality enquiries.
- The algorithm **prioritizes incoming enquiries** in a hierarchy based on buying propensity post which the sales consultant targets them diligently thereby **improving overall conversion rates**.

## Benefits

- Significantly reduce the **time/effort of DSCs** and margin of error in targeting/following-up high quality enquiries
- Enable **decision support system** leading to improved conversion rate with prioritized targeting
- Substantial **uplift in sales** from baseline period for dealerships where the program was rolled-out

### 3 Approach



- DSC – Dealership Sales Consultant
- DMS – Dealership Management System
- P1 – High Priority
- P2 – Medium Priority
- P3 – Low Priority

# Project 2 – Enhancing product quality using Predictive Warranty Analytics

Improving Quality using predictive and prescriptive analytical solutions like Reliability analysis, Text mining, Concern Tracking and Sustenance, Early Warning System for Strategic decision making



## Project Outline

### 1 Background

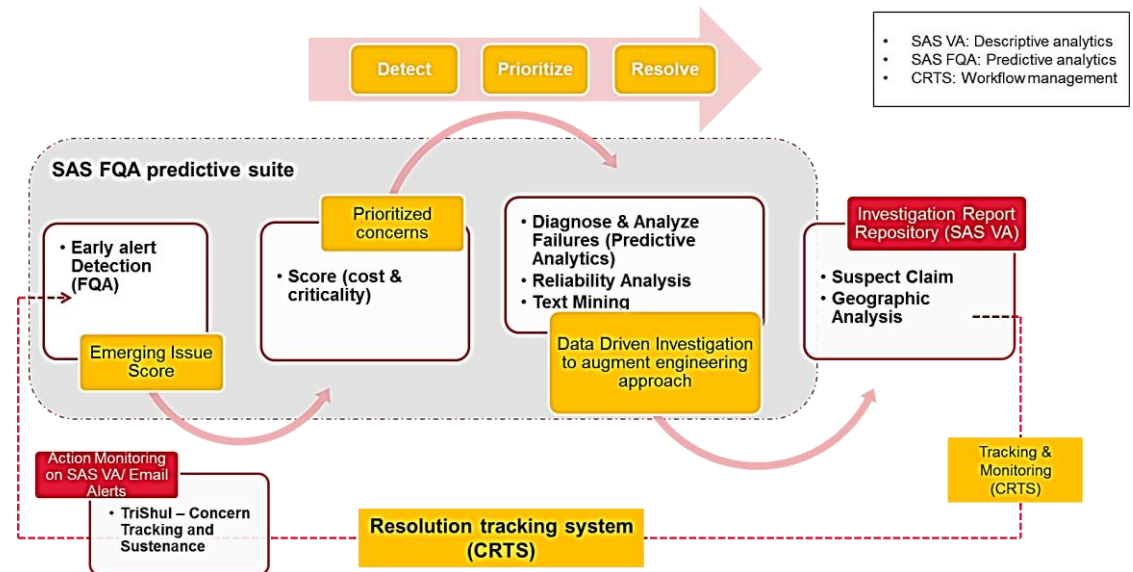
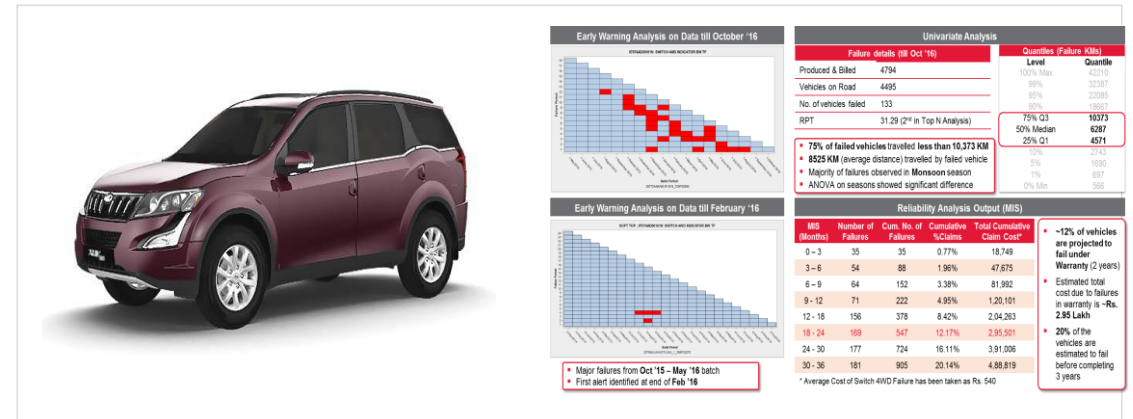
- M&M incurs high warranty cost due to failures in the vehicles on road
- Claims data analysis indicates high failure rates (RPT) on some of the parts for different models
- Impact on decision making in absence of reliable claims forecast

### 2 Objective

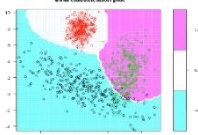
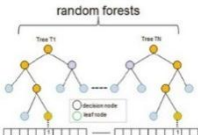
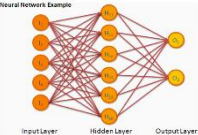
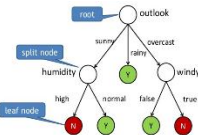
- Reliability predictive analysis on parts for all the models of AD and FD (High priority cases)
- Customized **Tracking & Sustenance** System Development for AD - **TriShul**
- Automated solution to predict 12 and 24 MIS RPT/CPU for immature mfg. batches: **AUTO-PRED**
- Early Warning Reports for AD & FD**– Top claims based on incidences and claims cost identified and prioritized

### 3 Benefits

- Identification of **emerging issues at early stage**
- Claim frequency and warranty cost forecasting to initiate proactive maintenance and **decision support for recall**
- Efficient customers' & technicians' comments analysis using **Text mining**
- Quality team can **prioritize the efforts** on warranty issues based on severity and scores generated in early warnings



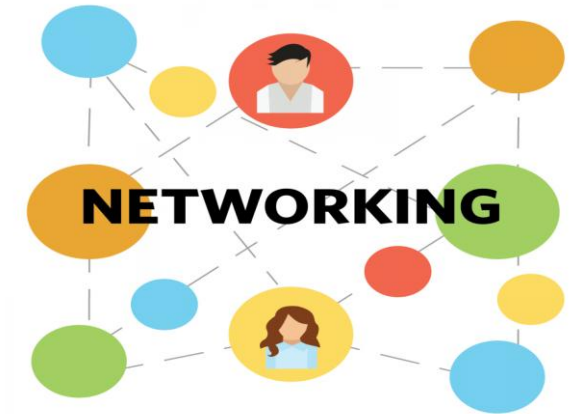
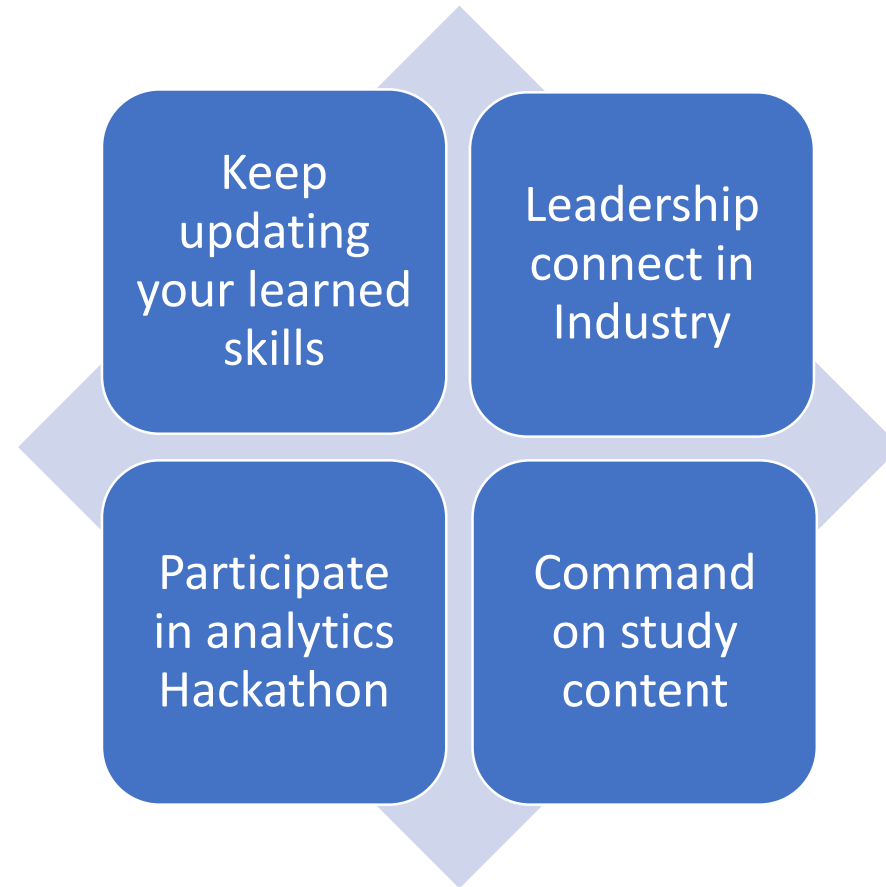
# Additional Machine Learning Techniques..

Techniques	Use	Advantages
 <p><b>Support Vector Machines</b></p>	Used in non-linearly separable problem	<ol style="list-style-type: none"> <li>1. SVMs handles highly dimensional space efficiently</li> <li>2. Works better for text classification</li> </ol>
 <p><b>Random Forest</b></p>	Used in Complex Relationship among variables	<ol style="list-style-type: none"> <li>1. Reduces variances among different samples</li> <li>2. Highly compatible for large number of samples</li> </ol>
 <p><b>Neural Network</b></p>	Used in Image Classification and sound recognition	<ol style="list-style-type: none"> <li>1. Can approximate any function, regardless of its linearity</li> <li>2. Great for complex/abstract problems</li> </ol>
 <p><b>CART</b></p>	Used for forecasting and in financial sectors	<ol style="list-style-type: none"> <li>1. CART has no assumptions and computationally fast</li> <li>2. Results are invariant to monotone transformations of response variables</li> </ol>

# Keeping in touch with opportunity ...



**DATA**  
HACKATHON



# Lets Analyze some of Job Description...

## ABOUT THE POSITION

---

Provide high quality and proficient support for our work across the data and analytics spectrum including data collection, analysis, modeling and visualization.

For this role, we are looking specifically for candidates with the following technical capabilities:

- Ability to work quickly and effectively with mathematical concepts, modeling, and statistics/finance
- Advanced skills in Microsoft Excel (Pivot Tables, comfort with advanced formulas, VBA a plus)
- Ability to express oneself clearly and concisely, both verbally and in writing
- Attention to detail and the ability to prioritize and work on multiple projects
- Team player with the ability to adjust work pace to match client priorities
- Experience with Tableau, @Risk, R or a programming language is a plus

Successful Linksbridge consultants generally have the following traits:

- A demonstrated commitment to making their world a better place
- A breadth of knowledge and interests in various areas of focus, especially social issues

- Strong professional and/or academic track record and experience
- Outstanding analytical and creative problem-solving skills
- Strong interpersonal, collaboration, leadership, and communication skills
- A healthy sense of humor, quirkiness and humility
- Business and financial acumen
- Intellectual curiosity and an appetite for personal and professional development
- A penchant for non-traditional consulting strategies
- Skilled in and willing to embrace new technologies and tools

We are looking for professionals with 2-5 years of work experience from various backgrounds,

# Lets Analyze some of Job Description...

## Job Description

We are looking for a data scientist that will help us discover the information hidden in vast amounts of data, and help us make smarter decisions to deliver even better products. Your primary focus will be in applying data mining techniques, doing statistical analysis, and building high quality prediction systems integrated with our products. **{{Depending on your needs, you can write very specific requirements here, like: “automate scoring using machine learning techniques”, “build recommendation systems”, “improve and extend the features used by our existing classifier”, “develop internal A/B testing procedures”, “build system for automated fraud detection”, etc.}}**

## Responsibilities

- Selecting features, building and optimizing classifiers using machine learning techniques
- Data mining using state-of-the-art methods
- Extending company’s data with third party sources of information when needed
- Enhancing data collection procedures to include information that is relevant for building analytic



# Case study presentation | Logical Aptitude

## FEW INTERVIEW FORMATS

- 1) First round was written - Subjective SAS questions
- 2) Two rounds of interview where they discussed about:
  - a) My past projects
  - b) Some statistical questions
  - c) Some case studies

### Interview Questions

Detailed working of Cluster Analysis

# Case study presentation | Logical Aptitude

## FEW INTERVIEW FORMATS

- 1) First round was written - Subjective SAS questions
- 2) Two rounds of interview where they discussed about:
  - a) My past projects
  - b) Some statistical questions
  - c) Some case studies

### Interview Questions

Detailed working of Cluster Analysis



Use mathematical operators  
like +, -, x, and Division -----

7,7,7,7,1 to make it 100

*Example like 2,2,1 make it 5*

$$2+2+1=5$$

*Make it 2*

$$2/2 + 1 = 2$$

# Middle level manager :Case study presentation | Logical Aptitude

## Sample case studies



Coca-Cola's profits have been declining in recent years.

Can you help find the root-cause of the profit decline and turn the situation around?



Nike wants to launch a new line of sneakers in Europe.

How would you go about advising the CEO on what new product to launch?



Foreign &  
Commonwealth  
Office

UK Foreign Office wants to restructure its workforce

What are the different elements you would consider to make a recommendation

Q&A

Thank You !