## Analysing COVID-19 data for India

## **Business Problem:**

countries, India too is severely impacted: 34 states / UT affected. First Case Reported on: 30 January In the current pandemic situation where COVID-19 has spread across the world impacting 150+ lower than in the worst affected countries. 134,843 recoveries and 7,751 deaths the infection rate of COVID-19 in India is: 1.7 (significantly 2020 as of 10th June, Ministry of Health and Family Welfare reported a total of 276,804 cases,

## **Business Objective:**

and predict the saturation point of the disease spread state and district level spread. To be able to foresee the local transmission rate, top affected districts To analyse the covid-19 India data and find key insights on patient demographics, patient clusters,

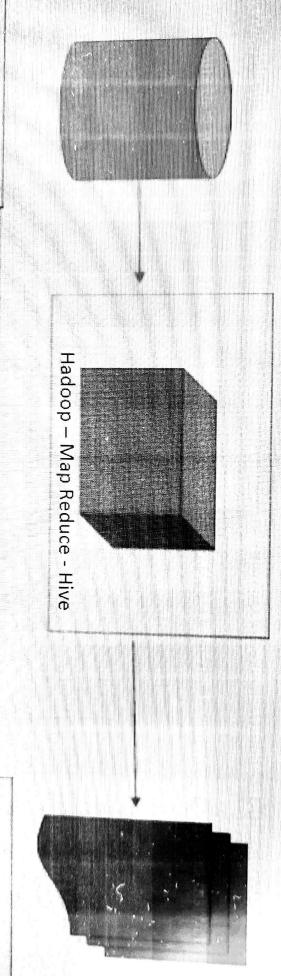
Reg. JD	Reg_ID Requirements	Priority
1	Data preparation: Loading into Hadoop	14
	Quality data preparation	
	Identifying tool for BI	
2	Identifies the top 6 affected districts in India Improves planning and	먼
	strategy to provide	
ప	Identifies patient demographics to understand the spread and trend in	P1
	daíly cases treatment	
4	Provides key insights for new policy formation like lockdown extension	P1
5	Predicts the saturation point for the spread	P1
6	Average time taken by patients to recover, average time a patient stays in	P1
	hospital - hospital occupancy	
7	To track load on healthcare facility and predict number of ventilators	P2
	required in future	

## Descriptive Analysis:

Level View Visualize the data to have a view for different levels: India Level, Patient Demographics and District

**Predictive Analysis:** 

Analysis to predict saturation point for new cases at district level (top 6 affected districts).



Data from different

segruos

Big data Environment on Hadoop

BI tool reports and document