Babysitter vs Robots

"*Data* is the oil of the 21ˢᵗ century,
and
Analytics is the combustion engine."

- Peter Sondergaard, SVP Gartner Research

Data Sources

GDG DevFest
Baroda 2019

Client facilitation

DB Admin with provide DB
dumps in csv

API access to DB

Come get the data

Someone sold the technology to
client but either the client is not
investing money to manage it *OR*
has no idea how of use it

Make your own data

Simulations

Synthetic Data

Data Scraping

Generic Data Science Pipeline

1

Data procurement, cleaning and preprocessing

2

Understanding the data in the business context

Get your facts straight with 5 W's

3

Exploratory analysis and visualizations

Descriptive Analytics

4

ML Model building, tuning and testing

5

Delivering actionable insights

Productionizing the model

WHY? Its all about making decisions...

"How much price should be quoted to the customer?"	"Is this customer trustworthy?"	"How can I get this customer to spend more money?"
Airline tickets	Loan sanctions	Targeted Advertisements
Cab fares	Fraud detection	Marketing campaigns
Delivery charges	Network security	Recommender systems
	Spam	

FACT

Data Science is like teenage stuff: everyone talks about it, nobody really knows how to do it, everyone thinks everyone else is doing it, so everyone claims they are doing it.

Majority of industry problems can be solved by simple statistical and Machine Learning approaches

THANK YOU
May the data be with you...

Sagar Patel
Consultant Data Scientist

sagarpatel.exe@gmail.com

https://www.linkedin.com/in/codesagar/

https://github.com/codesagar