

The background of the slide is a photograph of a cityscape, likely Lisbon, Portugal. It features numerous buildings with characteristic red-tiled roofs. In the distance, a tall, white church tower with a dome is visible against a blue sky with scattered clouds. The overall scene is a high-angle view looking down on the rooftops.

# Data Science methodologies for Forecasting

Anirudh Prabhu

Strategic Finance - Data Science

Uber Technologies Inc

# About me



E-commerce development  
DataBase modeling  
Tableau analytics  
SAP reporting



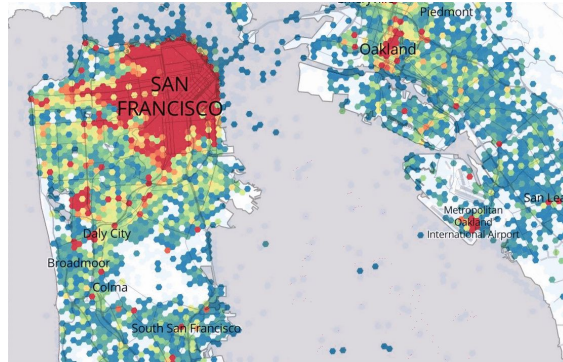
Payment Risk  
Anomaly Detection  
Forecasting  
Strategic Finance  
Fintech

# Data Science is the key to Uber's problems

Scale and Growth



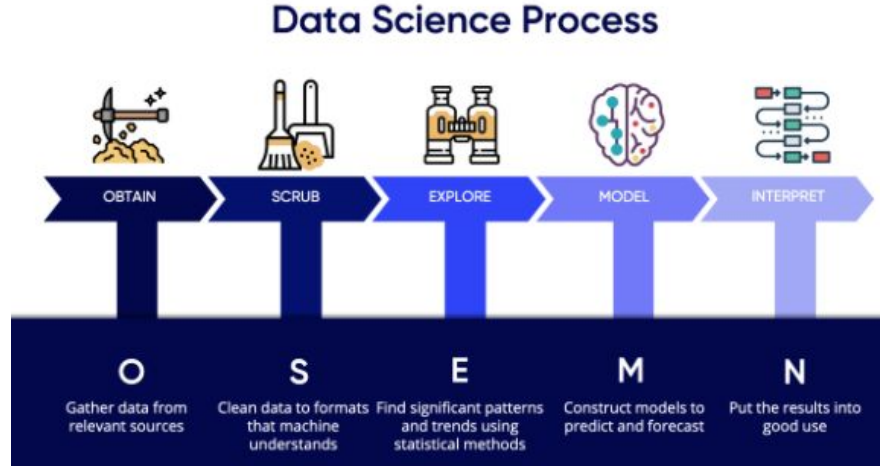
Real-time, complex decisions



Uniqueness



# Stages in Data Science Lifecycle



# Forecasting Methodologies

## Classical, statistical

- Autoregressive Integrated Moving Average (ARIMA)
- Exponential Smoothing Methods, e.g. Holt-Winters, Theta
- ...

## Machine Learning

- Quantile Regression Forest (QRF)
- Support Vector Regression (SVR)
- Recurrent Neural Networks (RNNs)
- ...

Best model depends on, e.g.

- Amount of historical data
- Correlation with explanatory variables
- How the forecast will be used, e.g. interpretability needs, computational complexity

*Compare multiple approaches.*

# 3 of Many Forecasting Use Cases at Uber

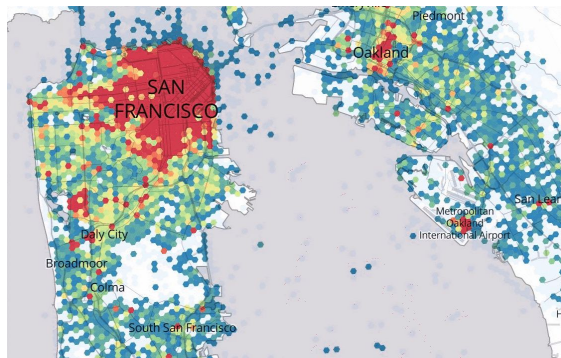
## Marketplace Forecasting

Supply

open cars /  
available driver partners

Demand

app views / sessions

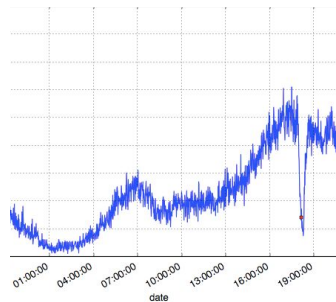


~ 5B+ forecasts / min

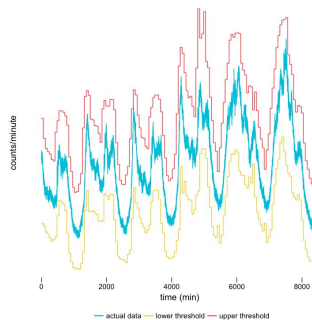
## Real-Time System Outage Detection

Anomaly detection in time-series based on prediction intervals.

~ 500M+ metrics



Example outage



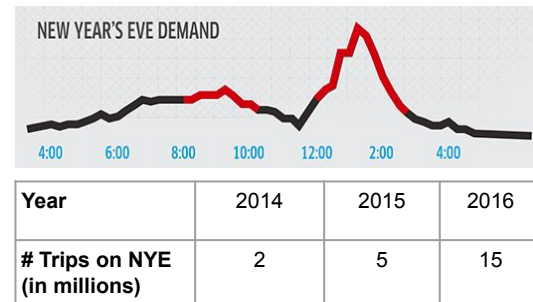
Dynamic thresholds

**Goal:** 99.99 percent availability → 1.01 min / week

## Hardware Capacity Planning

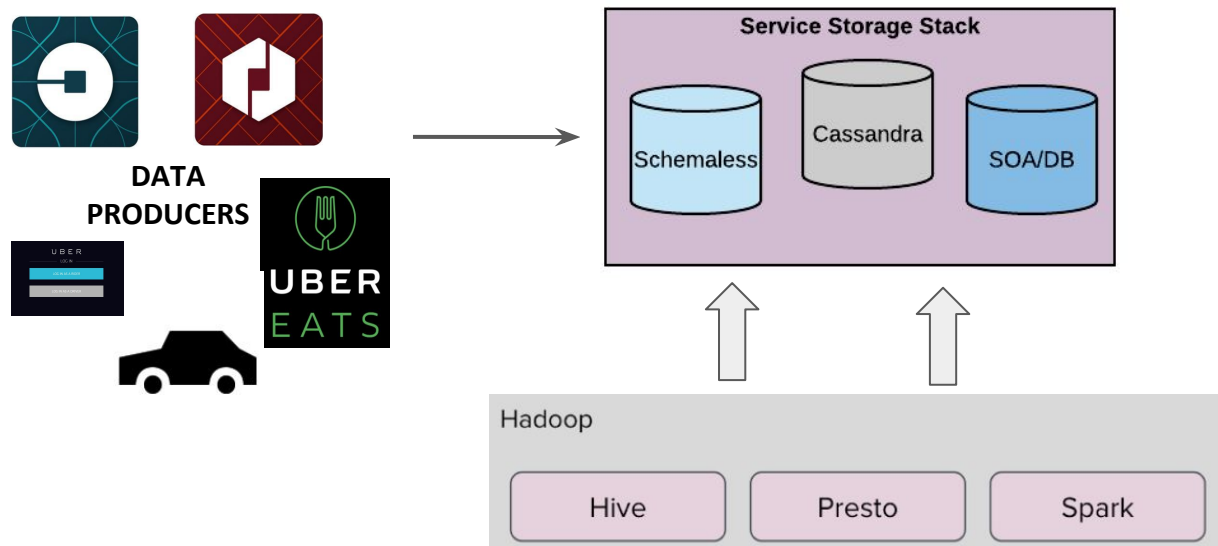
Accurately predict

- minute-by-minute
- concurrent trips
- for high-demand events
- months in advance



| Year                         | 2014 | 2015 | 2016 |
|------------------------------|------|------|------|
| # Trips on NYE (in millions) | 2    | 5    | 15   |

# Data Platform at Uber



Ways to read data for DS/DA/Ops

