



Client Overview

- The client is a market leader and a large South East Asian Taxi player, with a combined fleet size of about 23000 spread over 61 depots including major business and commercial districts
- Over 8.5 million passengers per month are served by the client's taxi service



Business Requirements

- The client was capturing geospatial time-series data for every taxi for every second.
- The client was looking at ways to increase the occupancy rate of their taxis by gleaning patterns of the occupancy rates from the time-series data captured



Key Challenges

- Data processing was the key challenge as the data volume was approx. 30GB/month
- The nature of the data was time-series with geospatial information of all taxis, which increased the time to prepare data for analytics
- The task required both Big Data Infrastructure and Advanced Analytics expertise to glean actionable insights



Our Approach

- Calculate the occupancy rates of the taxi's in below scenarios to unearth previously unknown patterns and understand the reasons behind them
 - ✓ At different times of the day
 - ✓ At different places of the city
 - ✓ At various hotspots like airport, city center etc.



Our Solutions

- We created Apache Hive external tables (Hadoop Ecosystem) on top of data files received, pre-processed the data and created derived columns like "half hour of the day", "day of the week" to enable descriptive analytics
- Processed data using Hive queries and created visualizations for consumption by top management
- We leveraged Geospatial analytics concept like Haversine formula to calculate distance between two points and divide the city into grids for further analysis



Results

- It was found that the occupancy rate was high between 10 A.M and 12 P.M, which was counterintuitive. On further analysis it was found that Indonesians visit mosques for prayer in that time period, and hence the surge in occupancy rate
- Based on our grid analysis, we identified upcoming hotspots in the city outskirts, which the management was unaware of



INDIA

Chennai | Bengaluru | Mumbai
Toll-free: 1800-123-1191

USA

Cupertino | Princeton
Toll-free: +1-888-207-5969

UK

London

SINGAPORE

+65 9630 7959