MBTA PARATRANSIT SERVICES: Demand + Cancellation Forecasting & Automated Address Geocoding

**Project Overview**

**The RIDE**
- MBTA's door-to-door shared ride paratransit service
- Responsible for 1.8 million trips in 2019

**Predictions**
- Forecast demand and cancellations for each hour of next service day
- Understand factors that drive both

**Geocoding**
- Automate geocoding of requested pick-up and drop-off addresses
- Refine coordinates of most visited locations

---

**Clustering Analysis**
- Performed k-means clustering on pick-up and drop-off positions
- Predictions to be made for each pick-up and drop-off cluster pair per hour to capture geographical patterns (16 total pairs)

**Methodology**
- Regression Trees, Random Forest, Gradient Boosting, Optimal Regression Trees, Optima Regression Trees w/ Linear Predictions & Ensemble Methods

---

**Predictions**

**Prediction Framework**

<table>
<thead>
<tr>
<th>Features</th>
<th>Pick-up Cluster</th>
<th>Drop-off Cluster</th>
<th>Time of Day</th>
<th>Day of Week</th>
<th>Holidays</th>
<th>Daily Lags</th>
<th>Hourly Lags</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weather Factors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weather Factors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Methods**

- Regression Trees, Random Forest, Gradient Boosting, Optimal Regression Trees, Optima Regression Trees w/ Linear Predictions & Ensemble Methods

**Takeaways**

- Critical Factors: daily/hourly lags (seasonality), geography & time of day
- DEMAND PREDICTION IMPROVEMENT

---

**Automated Geocoding**

**Geocoder**
- "Geocode address: " to trigger slackbot
- Only returns coordinates in The RIDE service area
- Tuning: obtained list of top addresses for manual recoding based on discrepancies in trip GPS data, up to discretion of booking agent

**Examples**

Geocode address: 100 Main St, Cambridge, MA 02142

Geocoder IP: 192.3.0.22
Latitude: 42.36193215
Longitude: -71.0825089725973
Address: E62 - MIT Sloan School of Management, 100, Main Street, East Cambridge, Cambridge, Middlesex County, Massachusetts, 02142, United States

Geocode address: 1623 Main St, Agawam, MA

Geocoder IP: 192.3.0.22
Unable to find address with first geocoder. Trying second geocoder...
Address does not appear to be in the service area.
Please check that the address is valid or manually geocode this address.