Retail Marketing Analysis: Hydrogen Refueling Stations

Project ID#: SA053

P.I. Name: Kent Schlesselman
Presented by: Ian Thompson
Kalibrate Technologies plc
June 11, 2015


This presentation does not contain any proprietary, confidential, or otherwise restricted information.
Overview

• **Timeline:**
  - Start - December 17, 2013
  - Finish – December 17, 2014
  - 100% Complete

• **Barriers**
  - Systems Analysis – 4.5
  - Technical Challenges:
    A. Future Market Behavior
    D. Insufficient Suite of Models and Tools
    E. Unplanned Studies and Analysis

• **Budget:**
  - $90,000

• **Partners:**
  - National Renewable Energy Laboratories (NREL)
  - California Fuel Cell Partnership
Process: Retail Marketing Analysis for Hydrogen Refueling Stations

Analysis Framework
- Discovery Session
- Global Network Planning Expertise
- Data: (Kalibrate site surveys, NREL, Nielsen, InfoUSA)

Models & Tools
- Location Insite – Ranking Model
- Network Planning Site Placement

Studies & Analysis
- Hydrogen Refueling Station: Location Ranking, Network Plan

Outputs & Deliverables
- Ranking of Points, Recommendations, Report, Future Work

Internal & External Reviews:
- Two External Reviewers
- Fuel Pathway Integration Technical Team (FPITT) Presentation
- California Fuel Cell Partnership Presentation
- Internal reviews at DOE and NREL

NREL
- Participated in Discovery Session and provided data
Relevance: Objectives

- Develop an analytic approach to prioritize and identify the best locations for hydrogen refueling stations.
- Apply this framework to California to prioritize station network expansion beyond existing and planned locations.

Objectives have been achieved and this project is complete.

Where?

How many?

The result of this approach will be a geographic representation of a hydrogen refueling station infrastructure for the State of California.
Approach

• Objectively evaluated locations
  • Determined supply and demand variables
  • Determined dependent and independent variables
  • Identified competitive area or trading area size for aggregating data
  • Established weights on variables
  • Ranked 30,620 traffic points

• As a complement to Spatially & Temporally Resolved Energy & Environmental Tool (STREET) and cluster studies, proposed locales for refueling stations based on rankings:
  • Identified attractive areas outside a competitive distance from current/existing and planned hydrogen refueling stations (Coverage Network)
  • Identified attractive areas in the major urban areas of Los Angeles and San Francisco, offering minimal competition with current/existing and planned hydrogen refueling stations (Urban Incremental Network)
Approach: Locations Selected for Ranking

- 30,620 Traffic Count locations
  - Published Counts – 21%
  - Current Year Estimates – 79%

Criteria
- Traffic Counts < 10,000 cars per day were eliminated
- Eliminated Published Counts older than 2011 and replaced with Current Year Estimates when available
# Accomplishments: Variables, Weights and Thermal Map of Traffic Point Ranking

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Households with income &gt; $100,000</td>
<td>7.20</td>
</tr>
<tr>
<td>No. of gas stations</td>
<td>6.74</td>
</tr>
<tr>
<td>No. of hydrogen vehicles</td>
<td>6.67</td>
</tr>
<tr>
<td>Total no. of employees</td>
<td>6.09</td>
</tr>
<tr>
<td>Hydrogen permitting constraints</td>
<td>5.55</td>
</tr>
<tr>
<td>Gas stations with sufficient lot size</td>
<td>4.08</td>
</tr>
<tr>
<td><strong>Hydrogen stations</strong> - existing/planned</td>
<td>2.32</td>
</tr>
<tr>
<td>Households with solar panels</td>
<td>2.31</td>
</tr>
<tr>
<td>Average commute time</td>
<td>1.01</td>
</tr>
<tr>
<td>Distance to hydrogen auto dealership</td>
<td>-2.40</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Gasoline Expenditure</td>
<td></td>
</tr>
</tbody>
</table>

Weighting of independent variables vs. dependent variable allows ranking of locations. Thermal map gives graphic representation of ranking.

1) Residential San Francisco
2) West Los Angeles
3) Silicon Valley
4) Orange County
5) San Diego
Accomplishments: Ranking of Points-Most Attractive Area of San Francisco

Applied ranking to census shapes in San Francisco. Red blocks are most attractive.
Accomplishments: Ranking of Points - Most Attractive Area of Los Angeles

Applied ranking to census shapes in Los Angeles. Red blocks are most attractive.
## Accomplishments: Comparison of Best Point in San Francisco vs. Best Point in Los Angeles

<table>
<thead>
<tr>
<th>Variable</th>
<th>No. 1 in Los Angeles County</th>
<th>No. 1 in San Francisco County</th>
<th>Weights for 2 Mile Radius</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location Insite Ranking</td>
<td>69.46</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Gasoline Expenditure ($000,000 per year)</td>
<td>222</td>
<td>346</td>
<td>Dependent Variable</td>
</tr>
<tr>
<td>Distance to Nearest Hydrogen Dealership (miles)</td>
<td>1.1</td>
<td>0.3</td>
<td>-2.40</td>
</tr>
<tr>
<td>Hydrogen Permit Constraints</td>
<td>5</td>
<td>5</td>
<td>5.55</td>
</tr>
<tr>
<td>Thousands of Households with Income over $100,000</td>
<td>10</td>
<td>67</td>
<td>7.20</td>
</tr>
<tr>
<td>Average Commuting Time (minutes)</td>
<td>31</td>
<td>29</td>
<td>1.01</td>
</tr>
<tr>
<td>Thousand Employees</td>
<td>217</td>
<td>430</td>
<td>6.09</td>
</tr>
<tr>
<td>Number of Gas Stations</td>
<td>45</td>
<td>35</td>
<td>6.74</td>
</tr>
<tr>
<td>Minimum Size Requirement Gas Stations</td>
<td>16</td>
<td>6</td>
<td>4.08</td>
</tr>
<tr>
<td>Households with Solar Energy in Use</td>
<td>0</td>
<td>1,293</td>
<td>2.31</td>
</tr>
<tr>
<td>Anticipated Fuel Cell Vehicles</td>
<td>143</td>
<td>125</td>
<td>6.67</td>
</tr>
<tr>
<td>Anticipated Fuel Cell Refueling Stations</td>
<td>0</td>
<td>1</td>
<td>2.32</td>
</tr>
</tbody>
</table>

Scoring higher in heavily weighted items leads to higher overall rank for S.F. point
Accomplishments: Identification of Attractive Areas for Hydrogen Refueling Stations

• Building on the STREET model results and cluster strategies, use the rankings of 30,620 points:
  • To identify attractive areas that are unlikely to compete with current/existing and planned hydrogen refueling stations (Coverage Network)
  • To identify attractive areas in the major urban areas of Los Angeles and San Francisco, offering minimal competition with current/existing and planned hydrogen refueling stations (Urban Incremental Network)

Using rankings, choose locations that are unlikely to compete with stations currently planned.
Accomplishments: Proposed locales for hydrogen refueling stations (Coverage Network)

Proposed locales (green points) are in attractive areas, but will not compete with current hydrogen refueling stations (blue points) and do not compete with planned stations (red points).
Accomplishments: Greater San Francisco - Proposed locales for hydrogen refueling stations (Urban Incremental Network)

Given the blue polygons representing six minute drive times around existing and planned stations, attractive areas remain in census shapes highlighted in red.
Accomplishments: Greater Los Angeles - Proposed locales for hydrogen refueling stations (Urban Incremental Network)

The Greater Los Angeles area shows a high concentration of existing and planned stations in blue polygons, but additional opportunities remain in census shapes highlighted in red.
Accomplishments: Proposed Hydrogen Refueling Infrastructure – State of California

Infrastructure to provide California with geographic coverage of hydrogen refueling stations.

<table>
<thead>
<tr>
<th>No. of FCEV Refueling Stations</th>
<th>Type of Hydrogen Refueling Stations</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>Permanent (aka existing or current)</td>
</tr>
<tr>
<td>42</td>
<td>Planned</td>
</tr>
<tr>
<td>29</td>
<td>Proposed Urban Incremental</td>
</tr>
<tr>
<td>52</td>
<td>Proposed Coverage Network</td>
</tr>
<tr>
<td>134</td>
<td>California Total</td>
</tr>
</tbody>
</table>
Response to Previous Year Reviewer’s Comments; Remaining Challenges and Barriers

• Response to Previous Year Reviewer’s Comments: This project was not reviewed last year

• Remaining Challenges and Barriers: This project has been completed on time and on budget
Collaborations


  - NREL/Kalibrate Collaboration:
    - To determine list of candidate independent and dependent variables for ranking of points
    - Guidelines for determining and naming coverage network and urban incremental network
    - Assistance with results presentation content

- California Fuel Cell Partnership (CaFCP)

  - CaFCP/Kalibrate Collaboration:
    - Presented preliminary results
    - Collected feedback
Proposed Future Work

• This specific project, Retail Marketing Analysis: Hydrogen Refueling Stations for the State of California, is complete.
• Given the opportunity for additional work, Kalibrate can assist with:
  • Similar work in other geographies: Other US States or regions
  • Develop a tool to provide instant analysis of ranked map points.
  • Given a sufficiently developed infrastructure of hydrogen refueling stations, create of a full featured network planning tool for the industry.
  • Contribute to NREL infrastructure simulation capabilities by improving on station sizing, footprint restrictions, market saturation, and inter-station demand shifting algorithms.
Summary

• Objectives:
  • Prioritize and identify the best locations for hydrogen refueling stations
  • Apply this framework to prioritize station network expansion

• Approach:
  • Objectively evaluated locations
  • Proposed locales for refueling stations based on rankings:

• Collaborations:
  • National Renewable Energy Laboratory (NREL)
  • California Fuel Cell Partnership

• Accomplishments:
  • Ranked 30,620 geographic points
  • Identified attractive areas with distance from planned hydrogen refueling stations
  • Identified attractive areas in the major urban areas of Los Angeles and San Francisco
  • Defined a geographic representation of a hydrogen refueling station infrastructure
  • On time and on budget

• Future Work:
  • Similar work in other US States
  • Develop a tool to provide instant analysis
  • Create of a full featured network planning tool
  • Enhance simulation capabilities
Acronyms and Definitions

• Nielson: The Nielson Company – source for demographic information

• InfoUSA: Business and consumer data resource. Gasoline retail outlet location information used to supplement Kalibrate data

• Location InSite: Kalibrate tool used to rank locations. Uses statistical concepts of normalization, correlation and significance to determine the relationship between independent and dependent variables.

• STREET model: Spatially & Temporally Resolved Energy & Environmental Tool. Results of STREET model commonly referenced for early hydrogen refueling station development plans.

• California Fuel Cell Partnership (CaFCP): A collaboration of organizations, including auto manufacturers, energy providers, government agencies and fuel cell technology companies, that work together to promote the commercialization of hydrogen fuel cell vehicles.

• FCEV: Fuel Cell Electric Vehicles
Technical Backup Slides
# Statistical Analysis

## Objectives

- Analyze the key factors for the introduction of hydrogen fueling infrastructure and fuel cell vehicles
- Identify the features of the information gathered
- Understand the importance of variables for the hydrogen fueling infrastructure
- Provide data correlations to pass through to later stages of predictive algorithms
- Calculate weights to rank the traffic points to identify the best locations for hydrogen refueling stations
Statistical Analysis - Methodology

**Statistical analyses**

- Descriptive statistics
- Correlation Analysis
- T-test for independent groups
- Collinearity and multicollinearity analyses
- Multiple linear regression
- Location Insite weight calculation

**Descriptive Statistics**

- Mean
- Median
- Standard deviation
- Skewness
- Kurtosis

**Graphical Methods**

- Density plot
- Histogram
- Normal Q-Q plot
**Statistical Analysis - Findings**

**Multiple Linear Regression**
- Measure the contribution and impact of the key factors on gasoline expenditure

**Select Best Model**
- Best model variables and adjusted r-squared are reported for all subsets
- 10 fold cross validation is applied to find the best minimum mean squared error

**Collinearity and Multicollinearity**
- Applied variance inflation factor (VIF) to best subsets
- Excluded the variables with problematic VIF values from the model
Multiple Linear Regression

- Evaluated the model accuracy reporting r-squared value
- Lowest r-squared reported is 0.85
- Diagnosed the model reporting
  - Detected outliers using Bonferroni outlier test
  - Used diagnostic tests of linearity, global validation test, non-constant error variance test to measure asymmetry, peakedness, linearity, and heteroscedasticity
  - Used diagnostic graphs residuals versus fitted, normal Q-Q, scale-location and residuals versus leverage graphs
### Statistical Analysis - Findings

**Location Insite**

All these analyses are utilized to calculate Location Insite weights.

<table>
<thead>
<tr>
<th>INDEPENDENT VARIABLES</th>
<th>1 Mile</th>
<th>2 Miles</th>
<th>3 Miles</th>
<th>5 Miles</th>
<th>7 Miles</th>
<th>10 Miles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traffic Count</td>
<td>8.4846</td>
<td>0.7044</td>
<td>0.7930</td>
<td>0.8521</td>
<td>0.8313</td>
<td>0.7846</td>
</tr>
<tr>
<td>Distance to nearest highway ramp (miles)</td>
<td>-1.9367</td>
<td>-1.7747</td>
<td>-1.6866</td>
<td>-1.5693</td>
<td>-1.4648</td>
<td>-1.3365</td>
</tr>
<tr>
<td>Distance to nearest dealer (miles)</td>
<td>-2.6679</td>
<td>-2.4844</td>
<td>-2.2546</td>
<td>-2.0285</td>
<td>-1.8468</td>
<td>-1.6688</td>
</tr>
<tr>
<td>Average commute time (minutes)</td>
<td>1.1151</td>
<td>1.0086</td>
<td>0.9437</td>
<td>0.9541</td>
<td>1.1054</td>
<td>1.4963</td>
</tr>
<tr>
<td>Total white collar employees</td>
<td>5.5913</td>
<td>6.8935</td>
<td>6.4488</td>
<td>6.7655</td>
<td>6.7587</td>
<td>6.6125</td>
</tr>
<tr>
<td>Gas stations</td>
<td>5.7060</td>
<td>6.7356</td>
<td>6.9196</td>
<td>6.9435</td>
<td>6.8728</td>
<td>6.6818</td>
</tr>
<tr>
<td>Gas stations meeting the min size req.</td>
<td>2.5667</td>
<td>4.0783</td>
<td>5.0248</td>
<td>5.8705</td>
<td>6.1599</td>
<td>6.2276</td>
</tr>
<tr>
<td>Vehicles per household</td>
<td>-0.0072</td>
<td>0.5829</td>
<td>1.0324</td>
<td>1.7418</td>
<td>2.2922</td>
<td>2.5922</td>
</tr>
<tr>
<td>Luxury vehicles</td>
<td>7.2034</td>
<td>6.4288</td>
<td>6.0942</td>
<td>5.8886</td>
<td>5.8764</td>
<td>5.9550</td>
</tr>
<tr>
<td>Households with solar panels</td>
<td>2.4338</td>
<td>2.3098</td>
<td>2.8799</td>
<td>1.7853</td>
<td>1.7424</td>
<td>1.9772</td>
</tr>
<tr>
<td>Permit constraints</td>
<td>6.9550</td>
<td>5.5507</td>
<td>4.7762</td>
<td>3.9024</td>
<td>3.4169</td>
<td>2.9985</td>
</tr>
<tr>
<td>Incentive Influence</td>
<td>7.8448</td>
<td>6.8636</td>
<td>6.4282</td>
<td>6.0088</td>
<td>5.8692</td>
<td>5.8537</td>
</tr>
<tr>
<td>Hydrogen stations</td>
<td>1.8169</td>
<td>2.3198</td>
<td>2.7484</td>
<td>3.2346</td>
<td>3.6676</td>
<td>4.4462</td>
</tr>
</tbody>
</table>

**Yellow and orange highlighted variables together represent the initial models with significant variables. Orange variables only represent the suggested models with the significant variables.**