



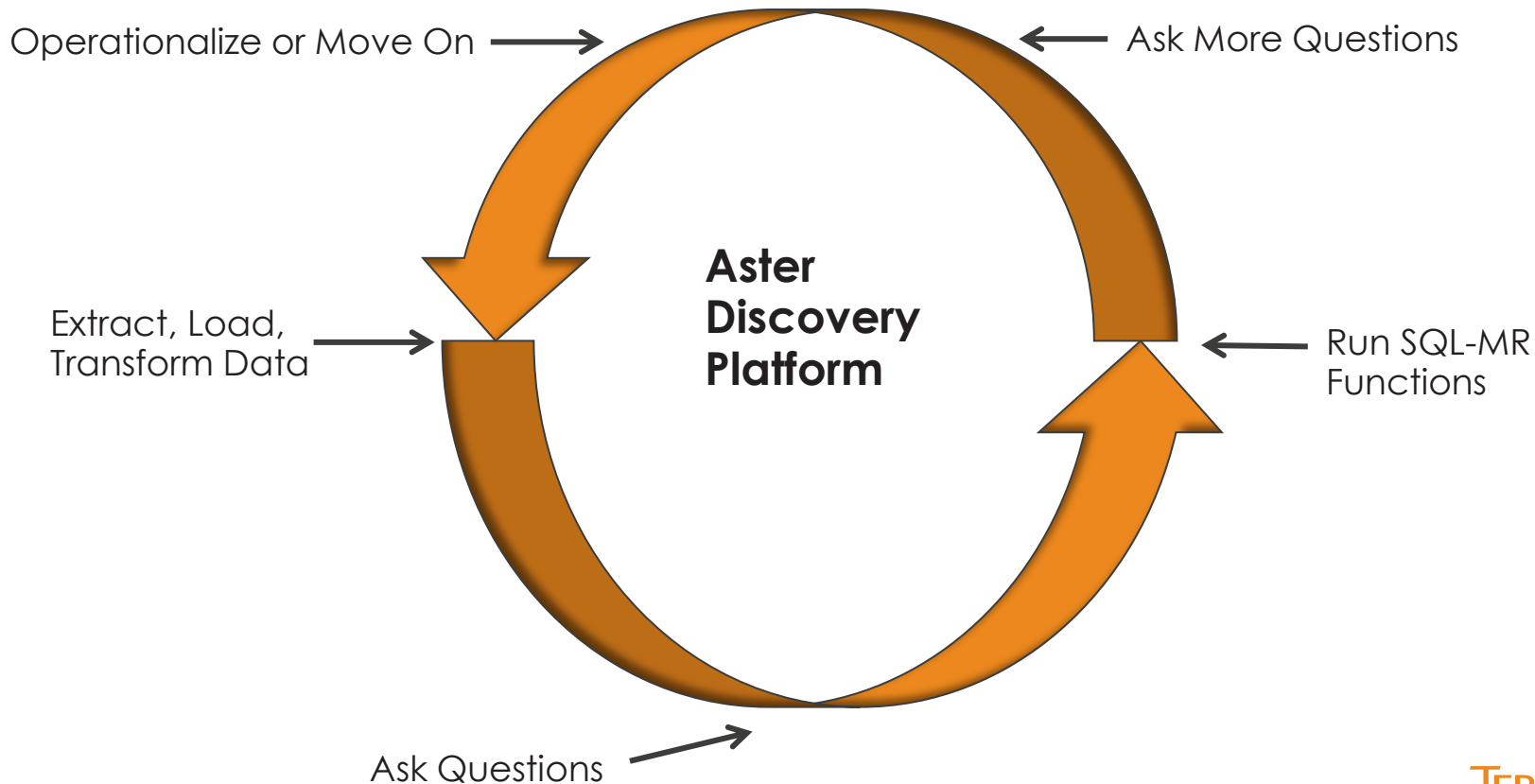
Teradata Aster SkiHackathon Demo

November 7, 2015

Agenda

- Introduction to Aster
- Data Sources for Ski Hackathon Demo
- Unstructured Data Analytics using Trip Advisor Comments
- GLM Model to Predict Traffic
- Predicting Traffic with ARIMA
- Choosing a Route
- Further Development

The Teradata Aster Way



Data Sources

<http://dtdapps.coloradodot.info/otis/TrafficData>



<http://www.tripadvisor.com/>

"Great sites and beautiful weather."

●●●●● Reviewed September 21, 2015 NEW

We rode the gondolas from the parking lot to the main resort area. We purchased tickets to ride the Zephyr to the top of the mountain. The sights on the ski lift were awesome and the lodge at the top was amazing. There was a mountain bike competition the weekend day we were there and it was fun watching the...

More ▾

Helpful? 👍

Report

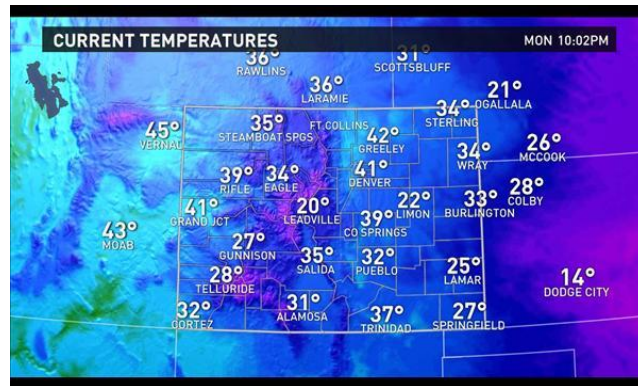
<http://www.onthesnow.com/>



https://www.dpsk12.org/calendars/pdf/calendar_1415_color.pdf



<http://www.wunderground.com/>



TERADATA

Text Mining Trip Advisor

- Steamboat Springs Resort
- Winter Park Resort
- Vail Mountain
- Wolf Creek Resort
- Loveland Ski Resort
- Breckinridge Ski Resort
- Keystone Ski Resort
- Beaver Creek Resort
- Copper Mountain
- Aspen Mountain



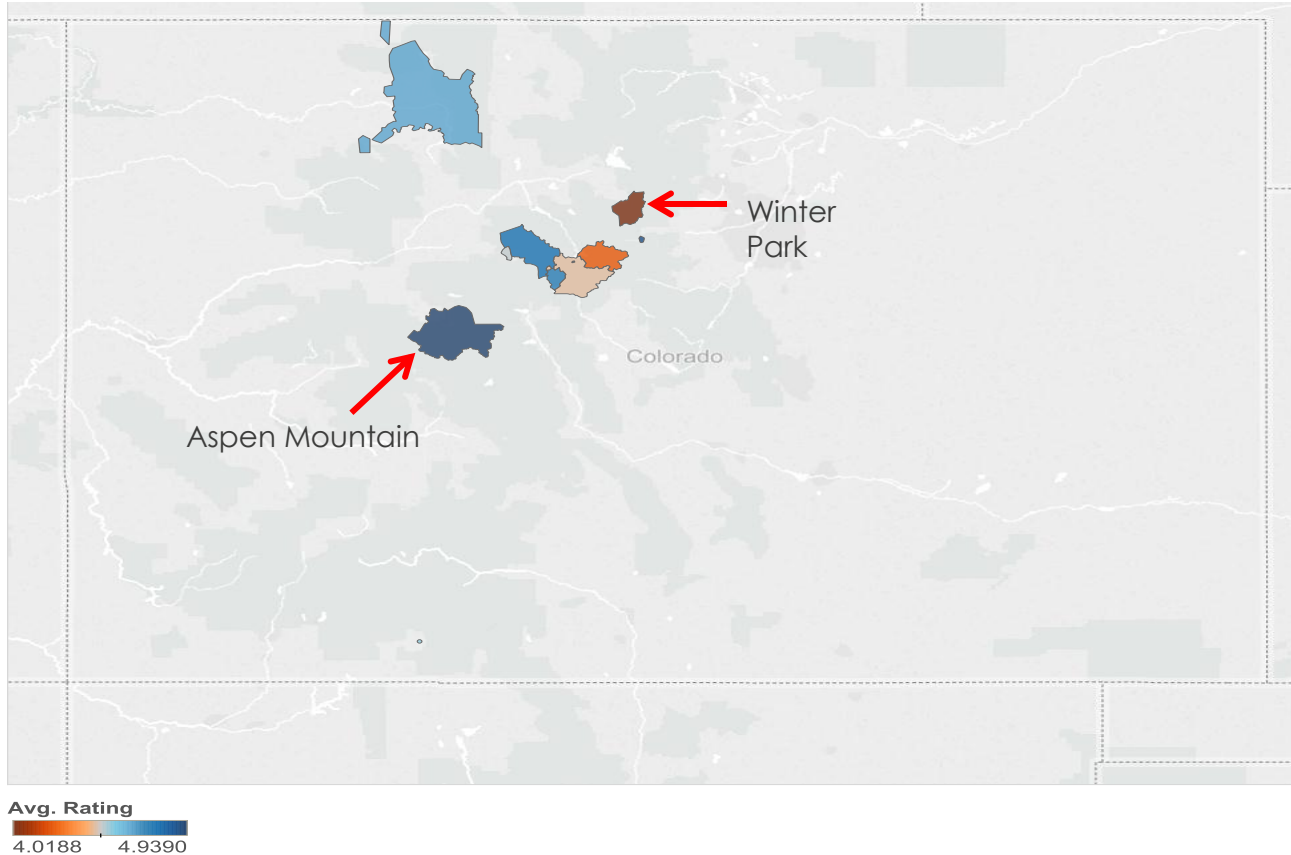
SQL MR For Unstructured Data

```
SELECT * FROM Text_Parser( ON resorttext
TEXT_COLUMN('review')
CASE_INSENSITIVE('true')
STEMMING('false')
PUNCTUATION('\[.,?\\!\]')
REMOVE_STOP_WORDS('true')
ACCUMULATE('id', 'review'))
ORDER BY id;
```

```
SELECT * FROM ExtractSentiment
(
  ON public.resorttext
  TEXT_COLUMN('review')
  MODEL('dictionary')
  LEVEL('sentence')
  ACCUMULATE('id', 'review')
) ORDER BY review;
```

```
SELECT * FROM tf_idf
(
  ON tf
  (
    ON tf_idf_in PARTITION BY term
  ) AS tf PARTITION BY term
  ON (SELECT count(DISTINCT(id)) FROM tf_idf_in)
  AS doccount DIMENSION
) ORDER BY id;
```


Trip Advisor Average Ratings for Resorts



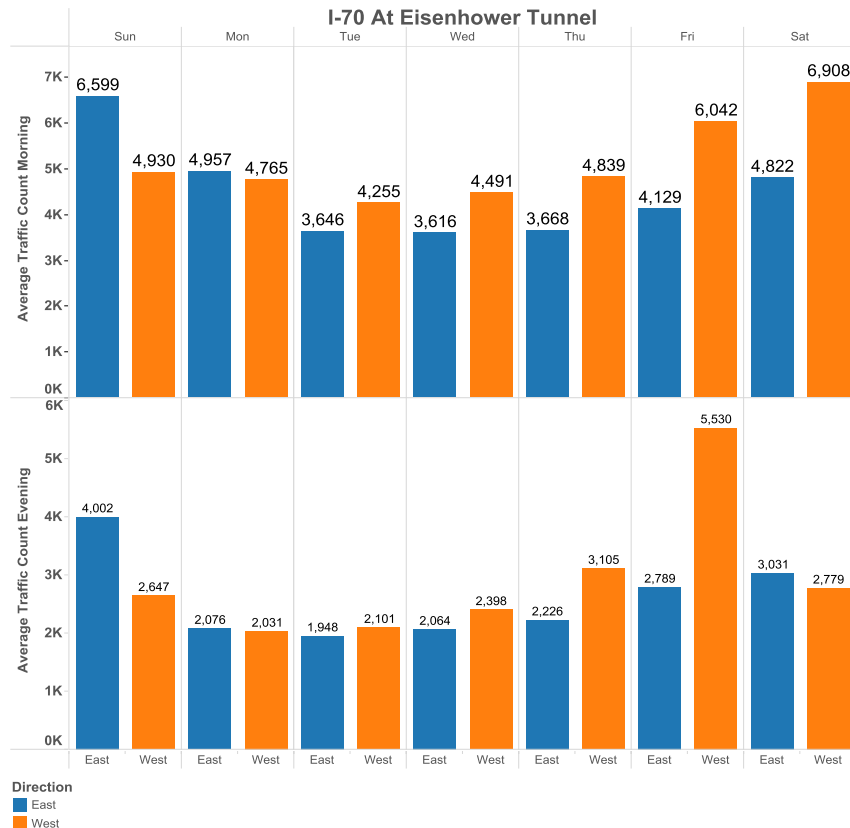
Positive Sentiment for Aspen Mountain



Negative Sentiment for Aspen Mountain



Which Day and Direction to Travel I-70?



3 Busiest Morning Days By Route Location

| locationdata | max(morning) | direction | Date | dayofwk |
|---|--------------|-----------|-----------|---------|
| ON I-70 E/O SH 70 BUS RT, COLORADO BLVD AT IDAHO SPRINGS TWIN TUNNELS | 32466 | West | 1/26/2013 | Sat |
| ON I-70 E/O SH 70 BUS RT, COLORADO BLVD AT IDAHO SPRINGS TWIN TUNNELS | 19239 | West | 1/28/2013 | Sat |
| ON I-70 E/O SH 70 BUS RT, COLORADO BLVD AT IDAHO SPRINGS TWIN TUNNELS | 18824 | West | 1/18/2014 | Sat |
| ON I-70 NE/O SH 9, BLUE RIVER PKWY, AT EISENHOWER TUNNEL | 12993 | East | 7/6/2014 | Sun |
| ON I-70 NE/O SH 9, BLUE RIVER PKWY, AT EISENHOWER TUNNEL | 11626 | East | 9/3/2012 | Mon |
| ON I-70 NE/O SH 9, BLUE RIVER PKWY, AT EISENHOWER TUNNEL | 11573 | West | 1/28/2012 | Sat |
| ON SH 40 E/O HENDERSON MINE RD | 4020 | West | 3/8/2014 | Sat |
| ON SH 40 E/O HENDERSON MINE RD | 3809 | West | 1/11/2014 | Sat |
| ON SH 40 E/O HENDERSON MINE RD | 3631 | West | 1/18/2014 | Sat |
| ON SH 9 S/O TIGER RD, BRECKENRIDGE | 5512 | West | 1/2/2014 | Thu |
| ON SH 9 S/O TIGER RD, BRECKENRIDGE | 5372 | West | 1/18/2014 | Sat |
| ON SH 9 S/O TIGER RD, BRECKENRIDGE | 5362 | West | 1/25/2014 | Sat |
| ON SH 9, BLUE RIVER PKWY N/O HAMILTON CREEK RD, CR 1900, SILVERTHORNE | 5493 | East | 1/11/2014 | Sat |
| ON SH 9, BLUE RIVER PKWY N/O HAMILTON CREEK RD, CR 1900, SILVERTHORNE | 5262 | East | 2/8/2014 | Sat |
| ON SH 9, BLUE RIVER PKWY N/O HAMILTON CREEK RD, CR 1900, SILVERTHORNE | 5025 | East | 2/9/2014 | Sun |

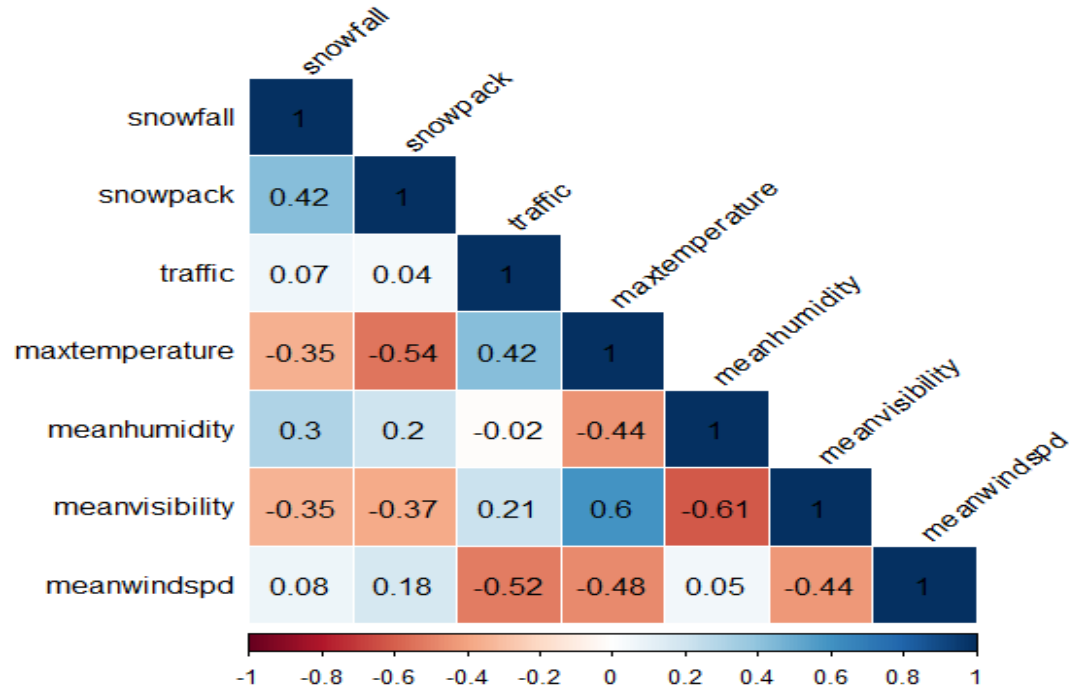
13/15 = 86.7%
Are Ski Days

3 Busiest Evening Days By Route Location

| locationdata | max(evening) | direction | Date | dayofwk |
|---|--------------|-----------|-----------|---------|
| ON I-70 E/O SH 70 BUS RT, COLORADO BLVD AT IDAHO SPRINGS TWIN TUNNELS | 20488 | West | 1/25/2013 | Fri |
| ON I-70 E/O SH 70 BUS RT, COLORADO BLVD AT IDAHO SPRINGS TWIN TUNNELS | 14274 | West | 7/3/2014 | Thu |
| ON I-70 E/O SH 70 BUS RT, COLORADO BLVD AT IDAHO SPRINGS TWIN TUNNELS | 13791 | East | 7/15/2012 | Sun |
| ON I-70 NE/O SH 9, BLUE RIVER PKWY, AT EISENHOWER TUNNEL | 11347 | West | 7/3/2014 | Thu |
| ON I-70 NE/O SH 9, BLUE RIVER PKWY, AT EISENHOWER TUNNEL | 9531 | West | 1/17/2014 | Fri |
| ON I-70 NE/O SH 9, BLUE RIVER PKWY, AT EISENHOWER TUNNEL | 9470 | West | 8/30/2013 | Fri |
| ON SH 40 E/O HENDERSON MINE RD | 3229 | West | 7/3/2014 | Thu |
| ON SH 40 E/O HENDERSON MINE RD | 2878 | West | 8/30/2013 | Fri |
| ON SH 40 E/O HENDERSON MINE RD | 2791 | West | 1/17/2014 | Fri |
| ON SH 9 S/O TIGER RD, BRECKENRIDGE | 3754 | West | 1/17/2014 | Fri |
| ON SH 9 S/O TIGER RD, BRECKENRIDGE | 3718 | West | 7/3/2013 | Wed |
| ON SH 9 S/O TIGER RD, BRECKENRIDGE | 3566 | West | 1/24/2014 | Fri |
| ON SH 9, BLUE RIVER PKWY N/O HAMILTON CREEK RD, CR 1900, SILVERTHORNE | 2962 | West | 8/17/2012 | Fri |
| ON SH 9, BLUE RIVER PKWY N/O HAMILTON CREEK RD, CR 1900, SILVERTHORNE | 2927 | East | 2/14/2014 | Fri |
| ON SH 9, BLUE RIVER PKWY N/O HAMILTON CREEK RD, CR 1900, SILVERTHORNE | 2567 | East | 8/2/2013 | Fri |

6/15 = 40%
Are Ski Days

What's Correlated with Traffic?



GLM Model to Predict Morning Traffic

```
SELECT * FROM glm(  
  ON (SELECT 1)  
  PARTITION BY 1  
  INPUTTABLE('skimasterevent')  
  OUTPUTTABLE('glm_skil')  
  COLUMNNAMES('morning', 'location', 'dayofwk', 'direction', 'snowpack',  
              'maxtemperaturef', 'meanhumidity', 'schoolday')  
  CATEGORICALCOLUMNS('closestcity', 'dayofwk', 'direction', 'schoolday')  
  FAMILY('GAUSSIAN')  
  LINK('IDENTITY'));
```

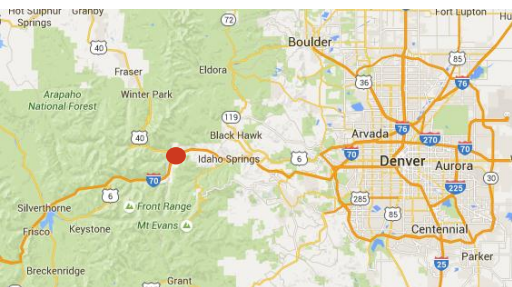
GLM Variables Used in Model

| Variable | +/- Influence on Traffic |
|------------------------|--------------------------|
| Highway Location | Positive and Negative |
| Day of Week | Positive and Negative |
| Direction (East, West) | Positive and Negative |
| Snow Pack | Positive |
| Max Temperature | Positive |
| Mean Humidity | Positive |
| School Day Yes | Negative |

GLM Model Output

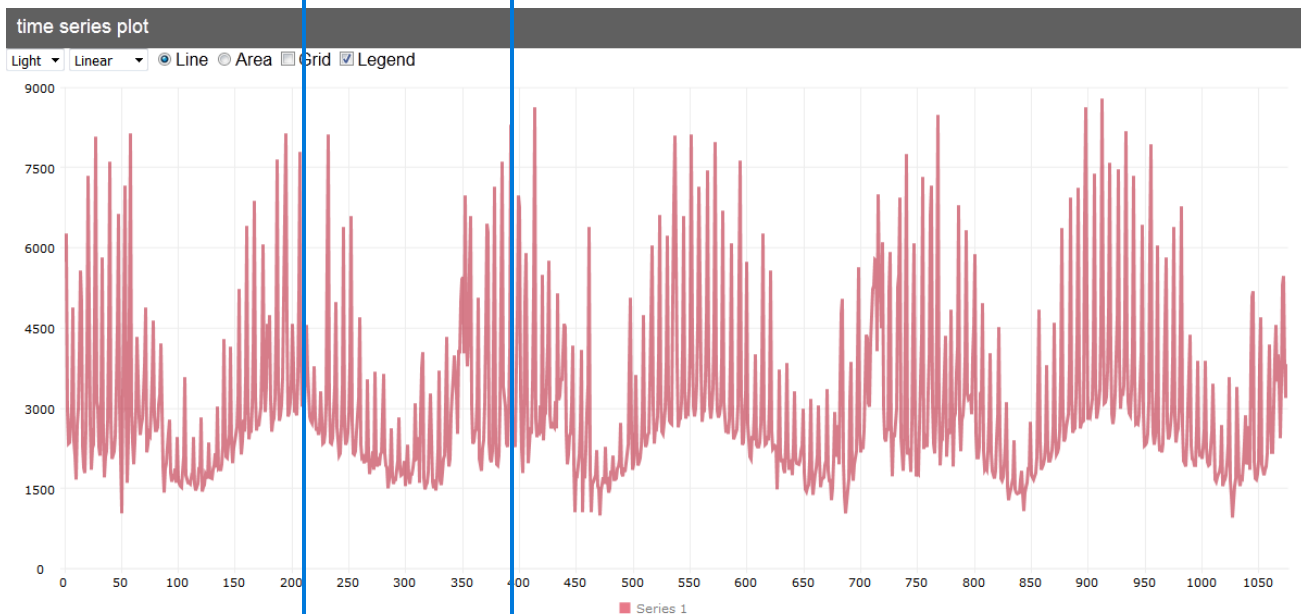
| Model Output | Statistic |
|--------------------|---------------|
| R-Squared | 0.899 |
| Adjusted R-Squared | 0.898 |
| F-Statistic | 5279 |
| Degrees of Freedom | 40 and 23,735 |
| P-Value | <2.2e-16 |

Predicting Traffic – Henderson Rd East



We want to account for peaks in traffic when creating a forecasting model – Seasonal ARIMA

July 28, 2012 February 16, 2013

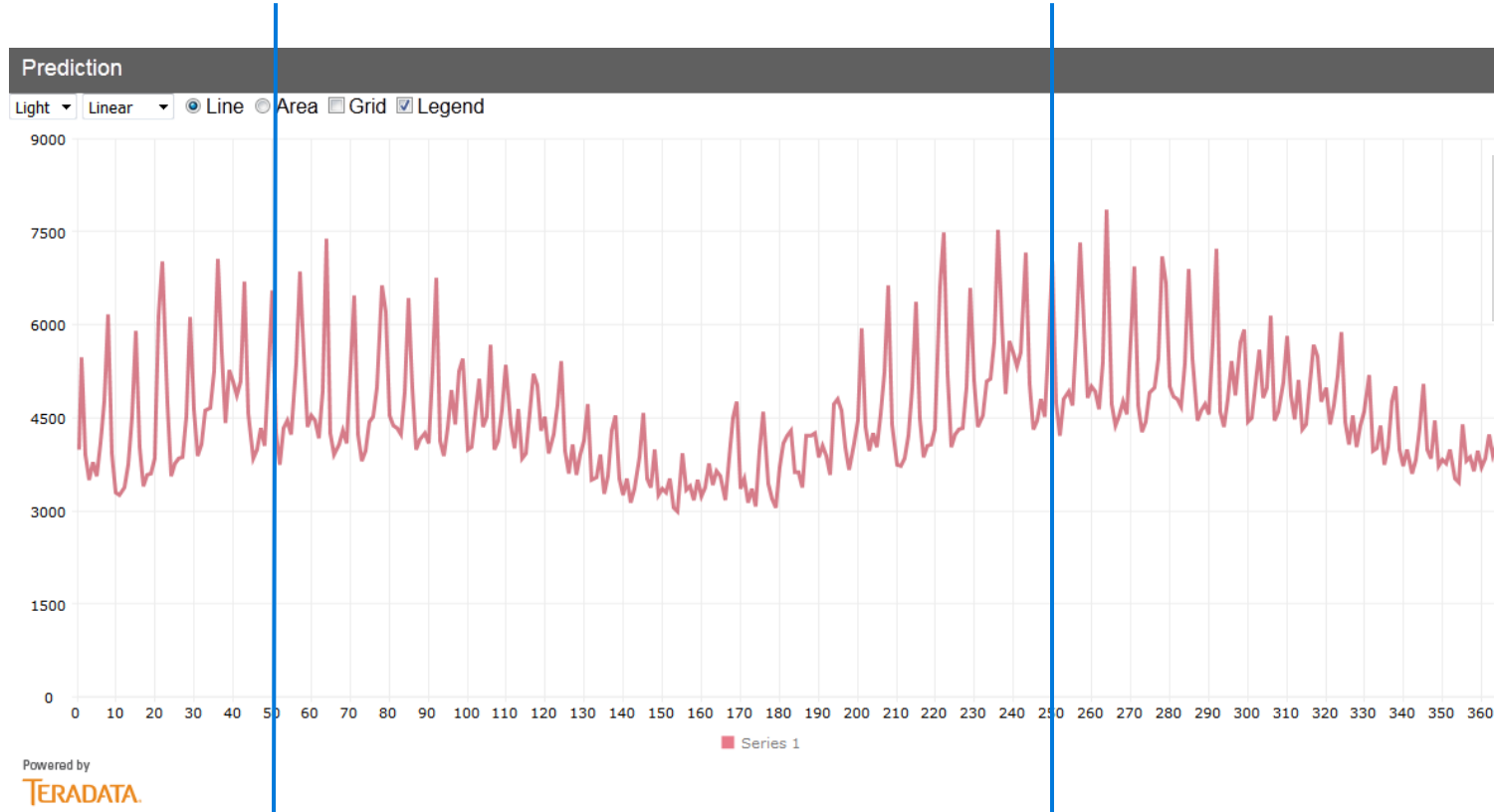


Powered by
TERADATA

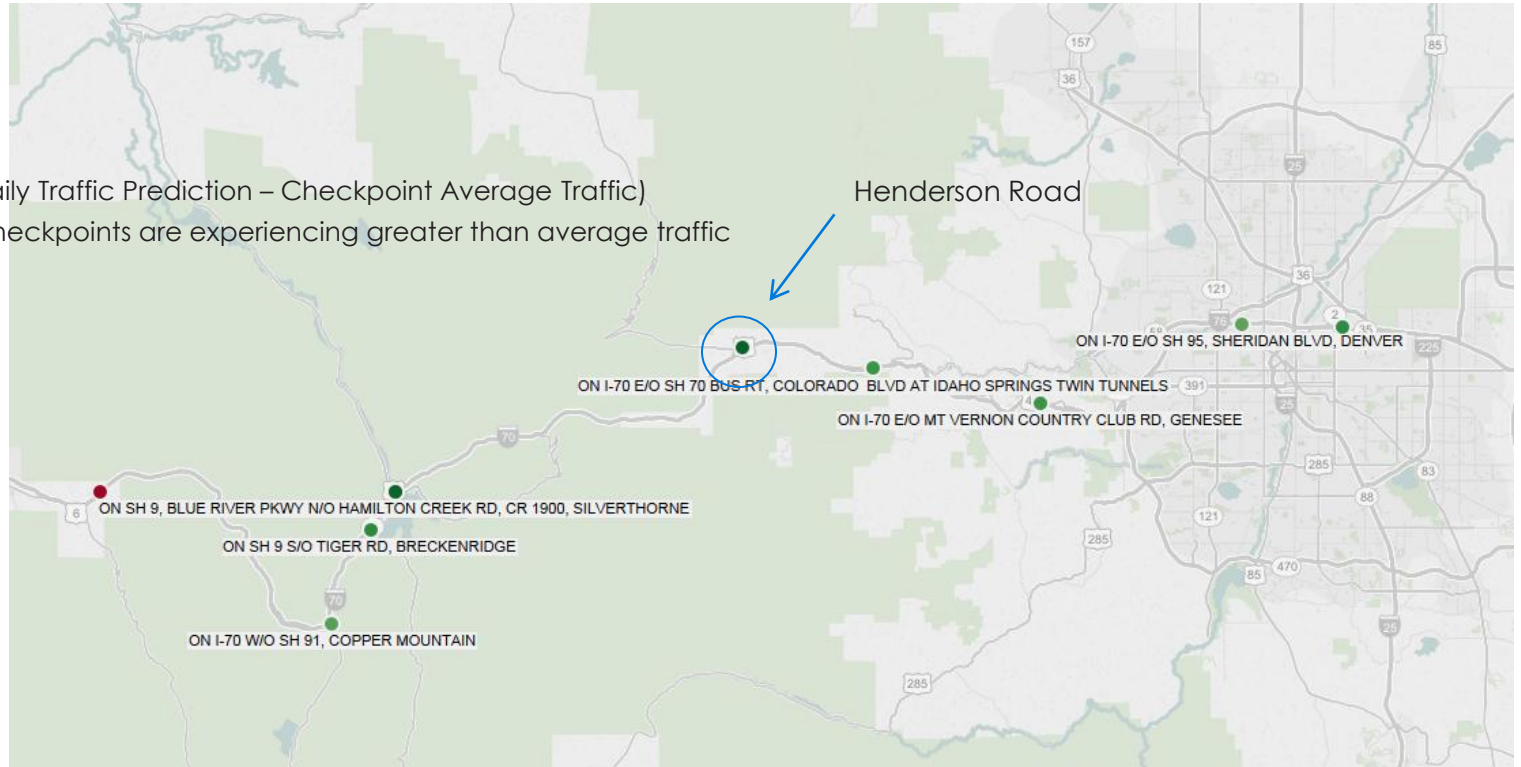
~ 200 days
between peaks

TERADATA

Henderson Road 2015 Traffic Prediction

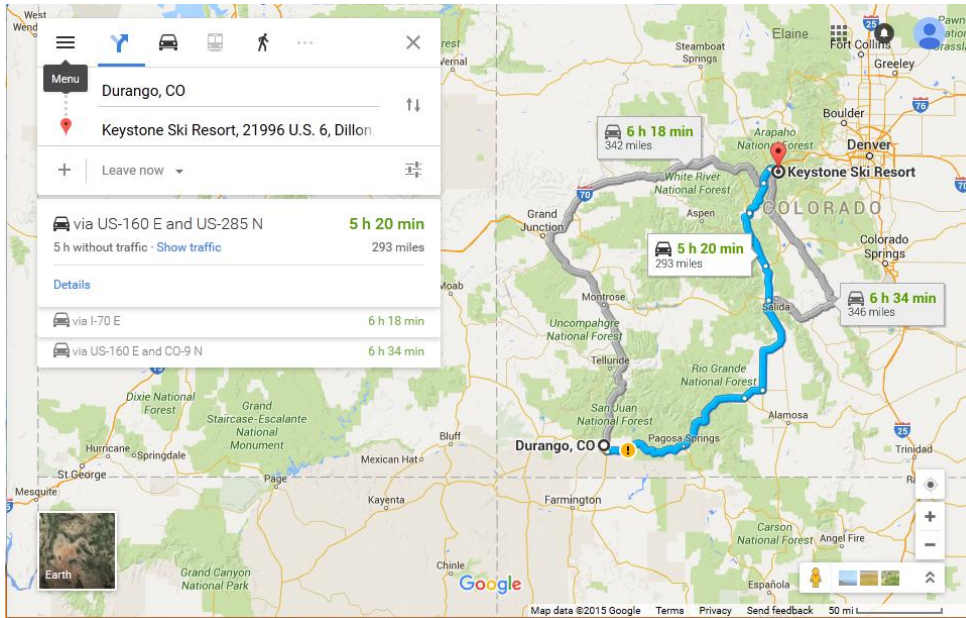


Visualizing Traffic in Tableau



Multiple Routes – Durango to Keystone

- How many people are taking each route?
- Which route is the best in different weather conditions?



Using nPath and AppCenter to Choose a Route

TERADATA AppCenter

Welcome, mp186078

skiapp2 (by Elaine, ver 0.1)

* Title: Durango to Keystone - Mild Weather

Tags: Separate tags with space

* Date: 01/12/2012 8:00 AM

* Starting Location: Durango

* Resort: Keystone

Weather Condition: [Slider]

Run Schedule

Tasks Schedule

Accepts user inputs for date, location, destination, and weather severity

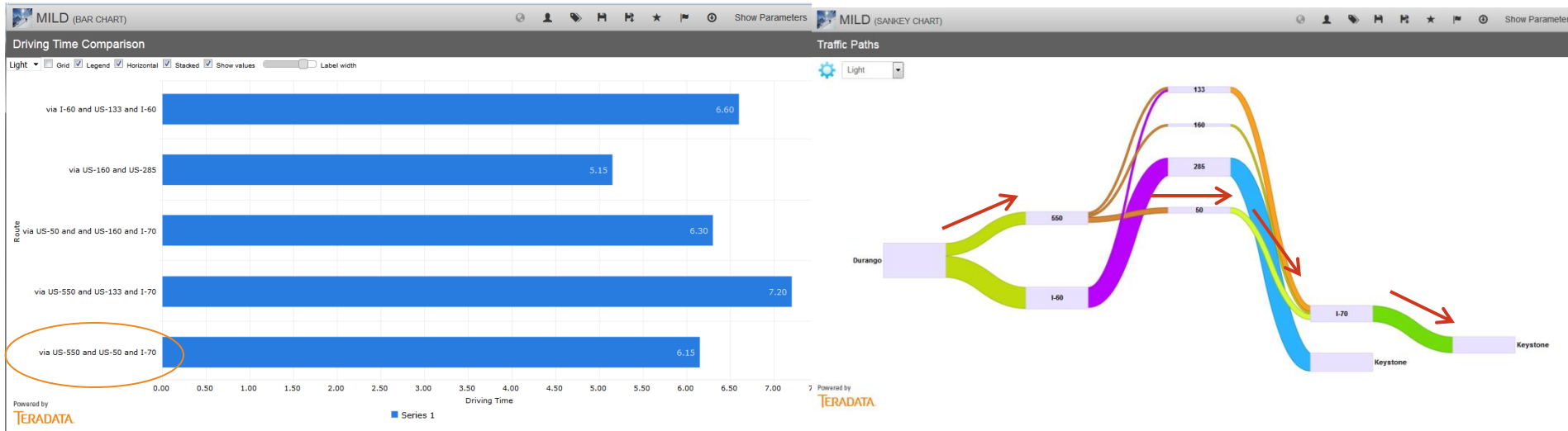
Create Collection

Reports

- All Reports
- My Reports
- Favorites
- Recent
- Shared By Me
- Shared With Me

Create Library

AppCenter Output



Gives the option of adding an hour to the trip, but seeing far fewer cars

Brainstorming an Aster Ski Route App

- If we had real data on how people got to resorts an Aster App could show:
 - Prediction of traffic using factors shown in GLM
 - Optimal Path based on user preferences (e.g. time, number of cars on the road, ease in snow)
 - Advantages/Disadvantages of each resort with Text Analysis
 - Could be done in real time

Aster and AppCenter Resources

➤ Aster and AppCenter Video Tutorials

1. <https://www.youtube.com/channel/UCBfxoagYzPqYIzSbLSHdCoQ>

➤ Teradata Aster User Community

1. <https://aster-community.teradata.com/welcome>

➤ Aster Express Download

1. <http://downloads.teradata.com/download/aster>

➤ Teradata Contacts

1. Elaine Crutchley Elaine.Crutchley@Teradata.com
2. Mark Peterson Mark.Peterson@Teradata.com