Big Data or not Big Data?
Opportunities and Strategies in Transportation
What is Big Data?

How big is big?
   No right answer

3Vs (Gartner)
   Volume
   Velocity
   Variety

Volume
Big Data
Velocity
Variety
What is Big Data?

My answer
Whatever you currently cannot do
Teradata appliances
eBay has 40 Petabytes
Petabyte = 32,000 iPhone’s
What about a query that takes 24 hours?
This is also big data
What is Big Data?

Velocity
Sensors
Aircraft
Locomotives
Trucks
GPS

Variety/Complexity

<table>
<thead>
<tr>
<th>Account</th>
<th>Miles</th>
<th>Member since</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diego Klabjan</td>
<td>60,000</td>
<td>2/1/2011</td>
</tr>
<tr>
<td>John Doe</td>
<td>40,000</td>
<td>4/5/2005</td>
</tr>
</tbody>
</table>

"His size required both armrests to be raised, and for his body to cover half of my seat," he says.

I was on a flight from Toronto to Las Vegas, So I paid $25 for an upgrade, and they only let me in the club. Granted most of it was costume work, but...
Variety and Complexity
Frequent Flier Programs

Holy Grail

360 view of customers

Each component in a data warehouse

Has my Gold customer being delayed on a flight?

It is all about customer service

Big Data Solution

Frequent Flier Program

Partner Programs

OTA

3rd Party

Facebook

Flight Status
Why Hot?

Commodity hardware

Thousands of PCs

Becoming ubiquitous

MPP=massively parallel processing

Cray or other MPP
Yesterday and Today

Yesterday
- Sensor data
  - Collect
  - Store
  - Analyze off-line
  Problems detected later

Today
- Sensor data
  - Stream
  - Analyze in real-time
  Real time detection

Yesterday
- Dumb humans

Today
- Mr. Watson

Yesterday
- Sensor data
  - Collect
  - Store
  - Analyze off-line

Today
- Customer segmentation
  - Macro-analytics
  - Know your segment
    - Targeted marketing
    - Product design

Yesterday
- Customer

Today
- Micro-analytics
  - Know your customer
    - One-on-one marketing
    - Produce design

Yesterday
- Read
  - Documents
  - Text

Today
- Automatically analyze
  - Documents
  - Text

Yesterday
- Human resources
B2C

- 360 view of customers
- Customer sentiment
- Machine data
- HR
- Real-time analytics

B2B

- Machine data
- HR
- Real-time analytics
- Location
Center of Excellence in Analytics

Support day-to-day analytics
  Apply known techniques
  New problems
  New business value

Innovations in analytics
  Risky for everyday analytics
  Uncertain outcome and return
Center for Excellence in Analytics

Establish an internal group/center

Objective
   Innovate with analytics
   Ahead of competition

Composition
   Part time data analysts
   Full time data analysts

Corporate support and buy-in
Real-time Analytics

The biggest impact on the customer are immediate actions

Enabler

Fast queries
Results back in seconds
Not in hours or minutes

Thought leaders see this as a big opportunity
Final Thoughts

Big data is a new project
Treat as such
Scope, evaluate, assess risks
ROI
Start small
Ten nodes
Expand

Do not jump on the bandwagon
Hurdles

Trained workforce
   Strong in IT
Innovations
   Center for excellence
Years before ubiquitous as SQL
Difference with traditional analytics
   ETL becomes ELT
   Low level programming