

A map of Las Vegas showing AVL trip segments as green lines, parks as green polygons, and AVL 300m proximity data as red and yellow dots. A legend in the top-left corner identifies these elements. A scale bar in the top-right corner shows distances in miles (0, 1250, 2500).

# AVL Data for Descriptive Analytics

Geofencing Using Oracle Spatial and Endeca for the City of Las Vegas

Chris Hughes  
Senior BI Specialist

Oracle Public Sector  
Jan 28, 2015



# Presentation Agenda

- A** ➤ Intros
- B** ➤ Problem Statement
- C** ➤ Solution(s) and Caveat(s)
- D** ➤ Demos
- E** ➤ Questions & Resources



## A. Introductions

# Introductions

- Las Vegas Sponsors

- Deanna Wilson  
IT Business Analyst DE
- Kathleen Fauerbach  
IT Section Manager
- Tim Shattner  
Detention and Enforcement

- Oracle Team

- Chris Hughes  
BI & Spatial Specialist
- Rob Golden  
Account Manager BI



- City of Las Vegas  
Detention and  
Enforcement Dept

- Manages detention facilities for the City and greater Metro
- Marshals Service patrols city facilities like Parks

- Quick Evaluation of Endeca to solve an Analytical problem
- Wanted to avoid the long development cycles involved with OBIEE
- Like the Ease of Use of Endeca for business users






## B. Problem Statement

# Business Problem

- Council and City Management asked the following:
  - How much time is spent patrolling parks?
  - How much time as a proportion to other activity is spent patrolling parks?
  - Which parks are patrolled the most?
- Detention and Enforcement have an AVL/GPS system on their patrol vehicles
  - Position, speed, other telematics collected every 5 secs
  - Each record has a vehicle id and a call sign id (officer id)
  - 1 month of data is about 70MB of data



- How to Approach the problem:
  - Traditional GIS 
  - Spatial Data Marts 
  - Data Discovery
  - Traditional BI 
  - Big Data  
is it hadoopable?

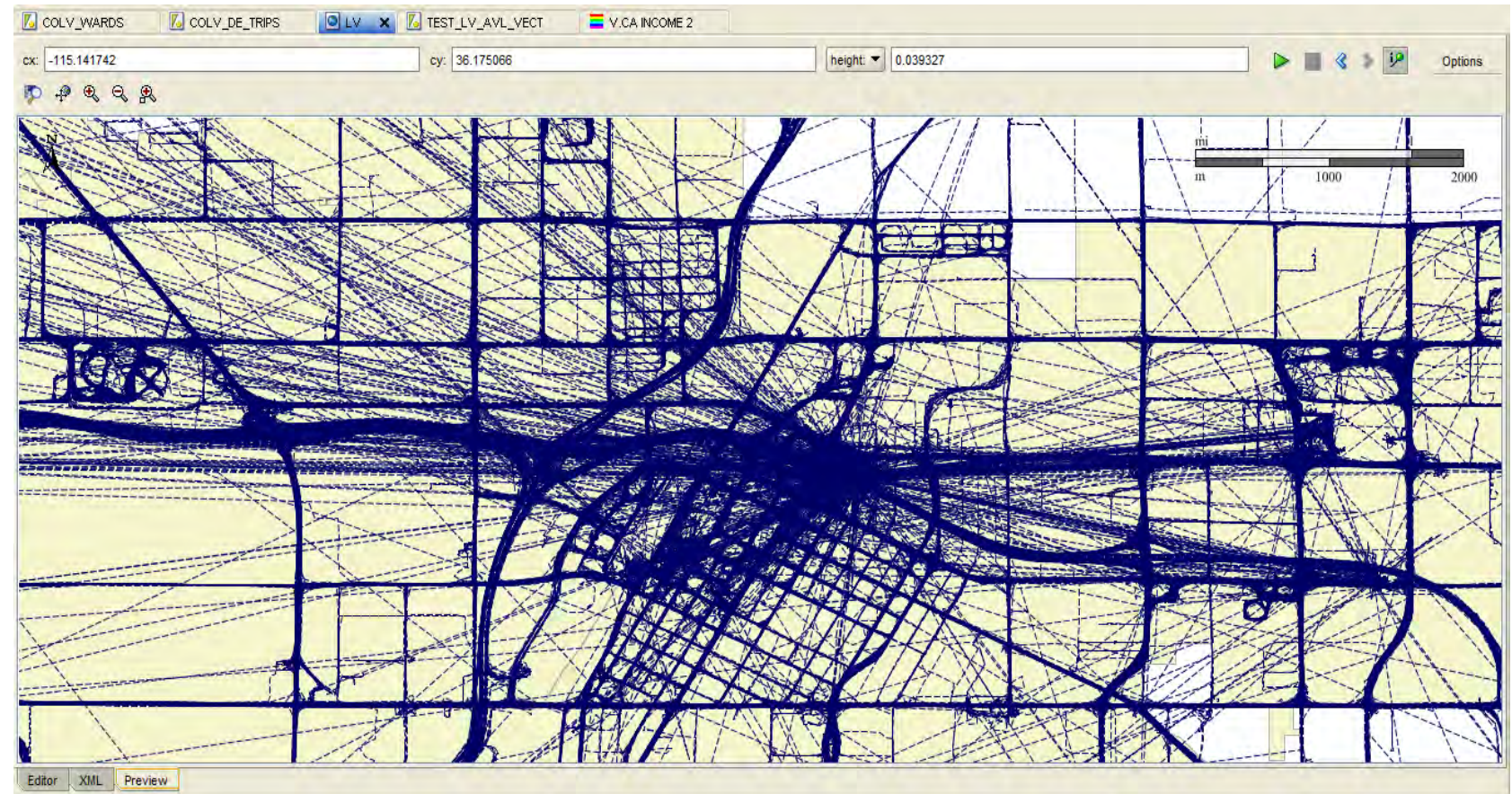




## B. Problem Statement

# Technical Problem(s)

- Data is dirty
- Unsure of aggregation rules (shift changes etc)
- A lot of data to handle
- Unsure if the questions DE is asking will change
  - is it only parks? What about wards?
  - What about micro questions? Where was call sign DM2 at 13:34 May 5?
- Spatial Enrichment problem
- GeoFencing definition
  - some parks you cannot drive in – what constitutes a patrol





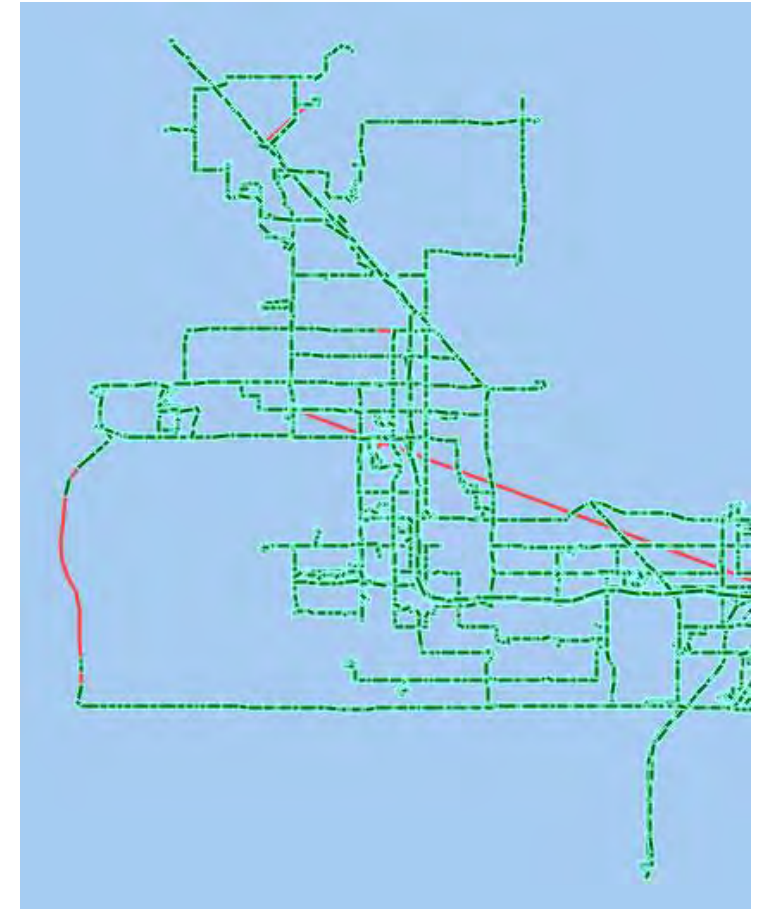
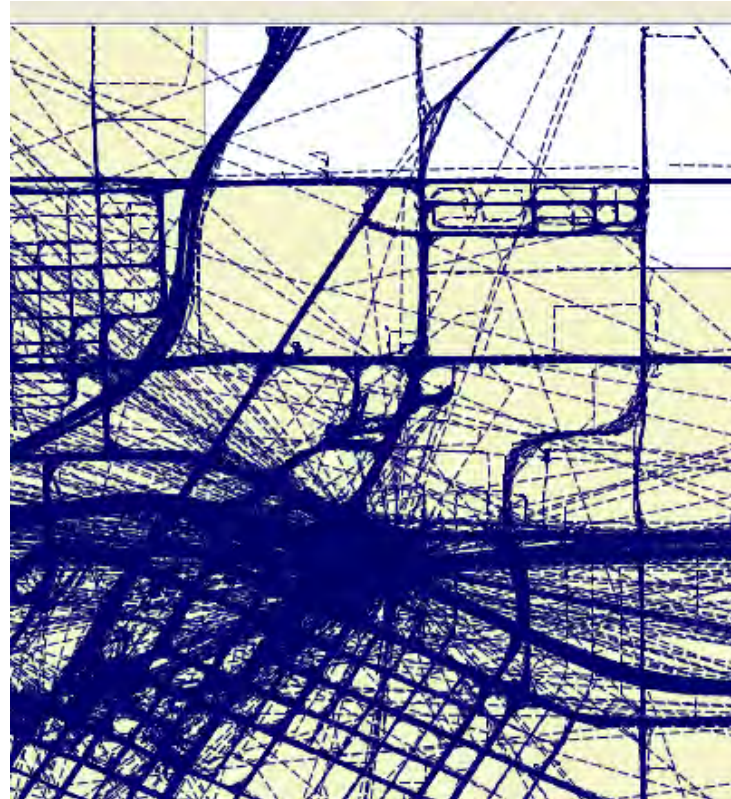
## C. Solution and Caveats

# Las Vegas DE | GeoFence Solution Cookbook

- 1. Clean data**  
filter out or delete incorrect data
- 2. Create value-add data structures**  
monotonic index, trips and trip segments
- 3. Filter out or create breaks in trips**  
when time or distance > threshold
- 4. Create spatial processing procedures like nearest neighbor**  
use spatial operators SDO\_NN and SDO\_RELATE
- 5. Create GeoFence Datamart**  
capture nearest X and the distance to GPS point in a table that can be partitioned
- 5. Process AVL data**  
we may want to process this in near realtime in the future so ensure procs are fast and robust
- 6. Deploy JDBC queries in Endeca**  
use self service for flexibility for business – let them discover the data
- 7. Visualize in tool of your choice**  
keep in mind ~1.5M points isnt useful  
I chose Endeca
- 8. Utilize EQL in Endeca to aggregate group locations in a pseudo-tile**

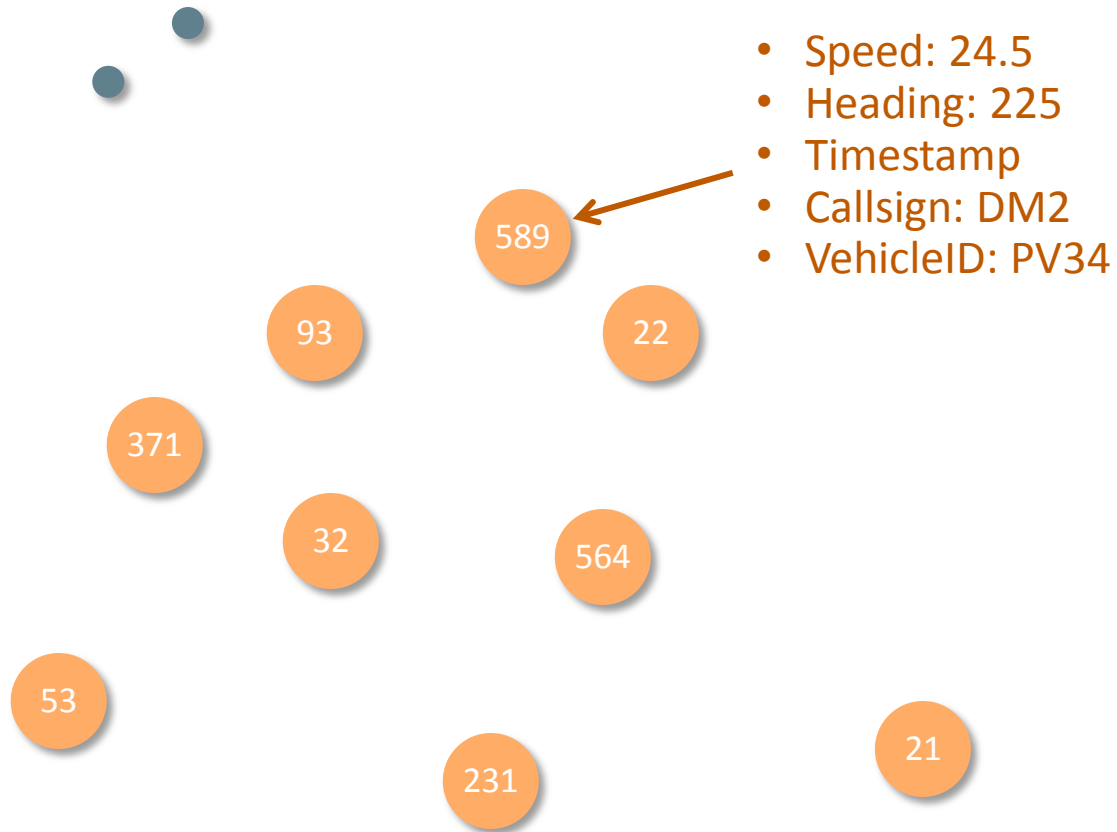


# 1) Clean the AVL data

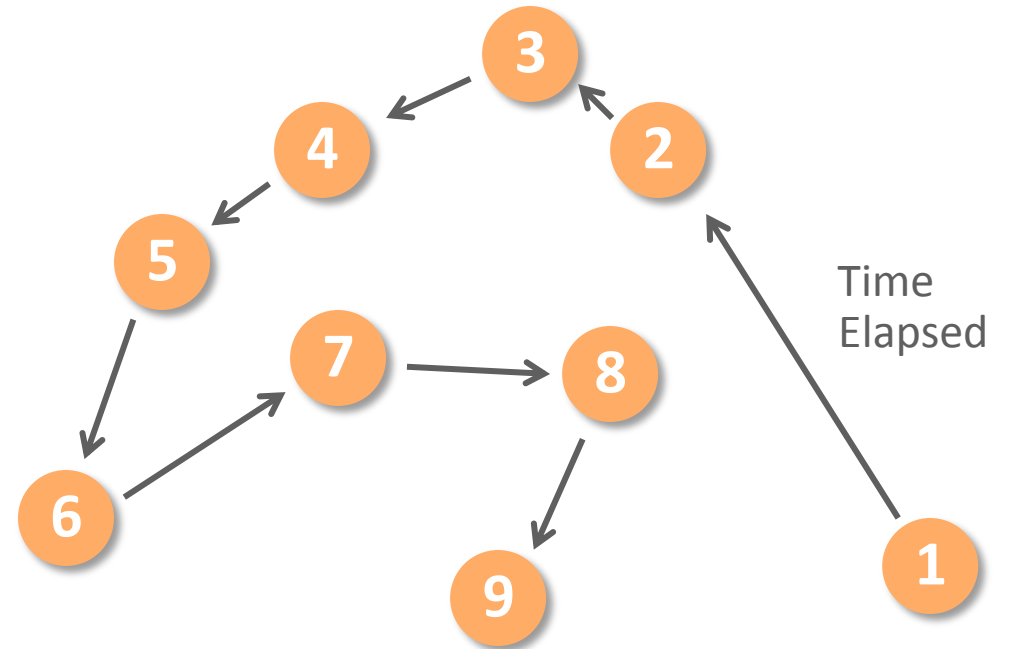


## C. Solution and Caveats

# 2) Create value added data structures



Clean Up

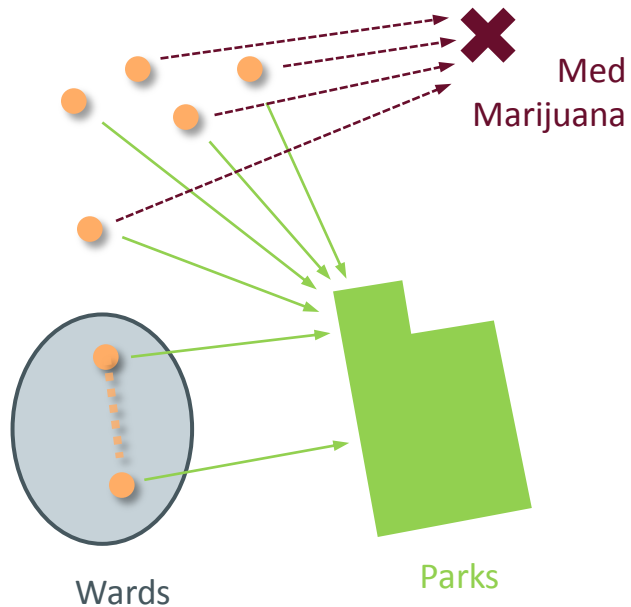




## C. Solution and Caveats

# 4) Develop a Spatial Processing Engine

- Ranking Analytical SQL
- Nearest neighbor and point in poly operations on each position
- Trip calculations



```
PROCEDURE process_overlay(p_geometry IN SDO_GEOMETRY, p_avl_id IN NUMBER);
```

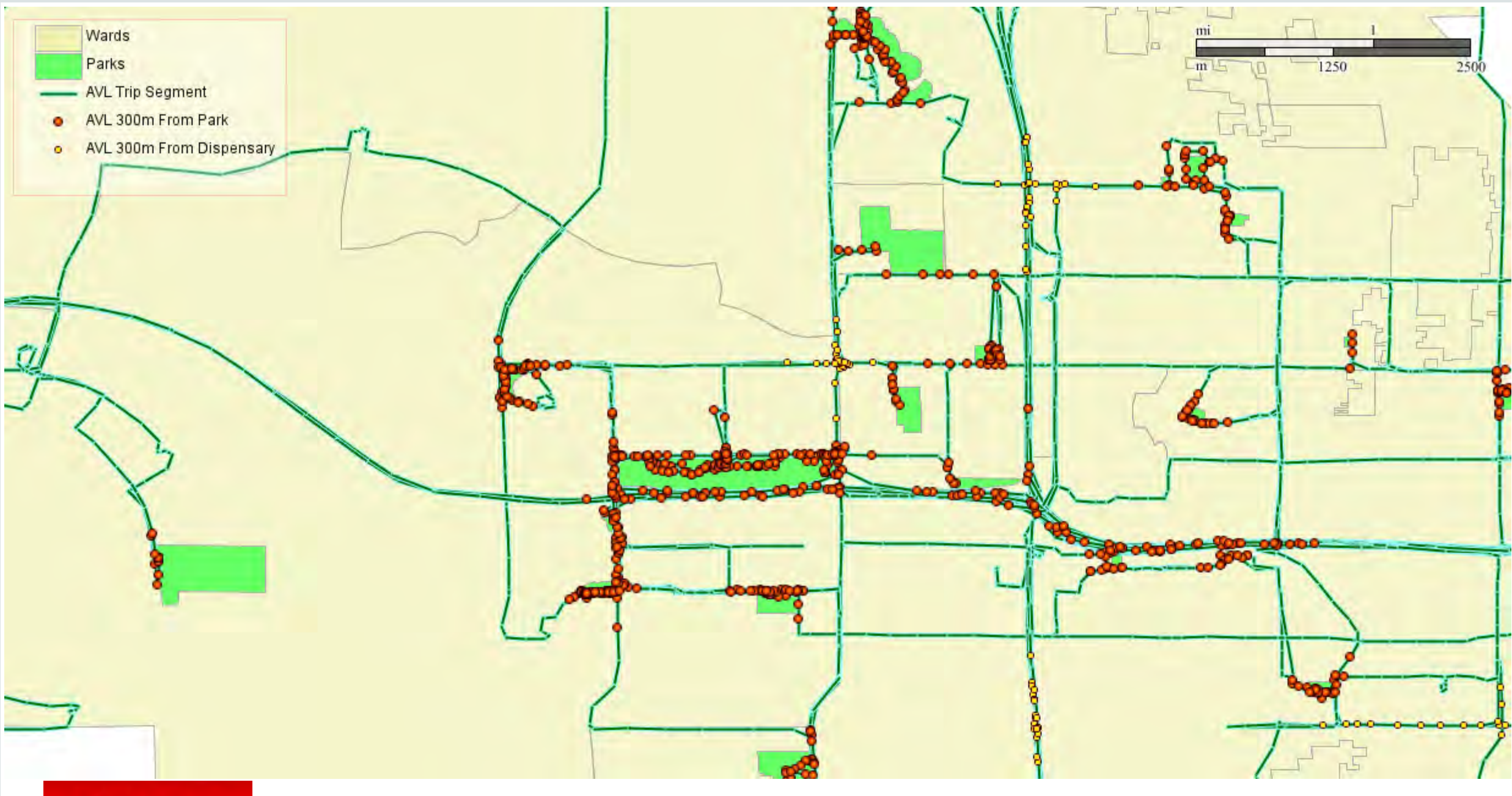
```
PROCEDURE do_overlay(p_geometry IN SDO_GEOMETRY, p_table_name IN VARCHAR2,  
                    p_name_col IN VARCHAR2 := 'NAME', p_result_rowid OUT VARCHAR2,  
                    p_result_name OUT VARCHAR2, p_result_dist_m OUT NUMBER);
```

```
PROCEDURE update_vectors(p_callsign IN VARCHAR2, p_vehid IN VARCHAR2,  
                        p_date IN DATE);
```

```
PROCEDURE update_nodes(p_vectors IN vect_tab, p_trip_id IN NUMBER);
```

```
PROCEDURE update_trip(p_trip_id IN NUMBER);
```

```
PROCEDURE insert_trip(p_trip_date IN DATE, p_vehicleid IN VARCHAR2,  
                    p_callsign IN VARCHAR2, p_avg_speed IN NUMBER, p_geometry IN SDO_GEOMETRY,  
                    p_trip_id IN NUMBER);
```





## C. Solution and Caveats

# 6) Deploy JDBC queries in Endeca

- Use SQL to create geographic objects that Endeca can use

```
SELECT DATETIMESTAMP,  
VEHICLEID, CALLSIGN,  
UNITSTATUS, ASSIGNEDINCIDENT,  
LATITUDE, LONGITUDE,  
LATITUDE || ' ' || LONGITUDE as GEO,  
SPEED, HEADING, ALTITUDE, GPSTIME,  
AVL_ID, SECS_ELAPSED,  
TRIP_RANK, VECTOR_LEN_M, TRIP_ID,  
PARK_NAME, PARK_DIST, PARK_ID,  
STREET, STREET_ID, STREET_DIST,  
ZIP, WARD  
FROM LV_AVL_JOINED_OVERLAYS
```







## C. Solution and Caveats

# 8) Use EQL to cluster points

- The Map View cannot really sustain more than ~5k points
- A heatmap is a better representation
- Create a pseudo tile based on 1/1000's of decimal degrees
- Generates a cluster and summation of various metrics

```
DEFINE DE_AVL_Cluster AS SELECT
TO_GEOCODE(TRUNC("DE_AVL_Overlays.Latitude",3) +
.0005,TRUNC("DE_AVL_Overlays.Longitude",3)+.0005) AS
"ClusterPoint",
count(1) AS "CountAVLPoints",
MAX("DE_AVL_Overlays.eid-reserved-spec") as
"DE_AVL_Overlays.eid-reserved-spec",
SUM("DE_AVL_Overlays.Secs_Elapsed")/360 AS "HoursPatrol",
SUM("DE_AVL_Overlays.Secs_Elapsed") WHERE
("DE_AVL_Overlays.Park_Distance" <= 100)/360 AS
"HoursPatrolPark",
AVG("DE_AVL_Overlays.Park_Distance") AS "Park_Distance",
MAX(TRUNC("DE_AVL_Overlays.Date_Timestamp",DATE)) AS
"DateLastVisited"
FROM DE_AVL_Overlays
GROUP BY ClusterPoint
```



▶ Demonstration





**Search Box**

Search within

**Selected Refinements**

No refinements have been selected.

- Available Refinements**
- Patrol Info
    - Call Sign
    - Unit Status
    - Assigned Incident
    - Date Timestamp
    - Vehicle Id
  - Geography
    - Park Name
 
      - Centennial Plaza (171559)
      - Rafael Rivera Park (142503)
      - Neon Boneyard Park (135806)
      - Gary Reese Freedom Park (71357)
      - Justice Myron E. Leavitt a... (45287)
    - Park Distance
    - Park Uid
    - Street Name
    - Street Id
    - Street Dist Offset
    - Ward

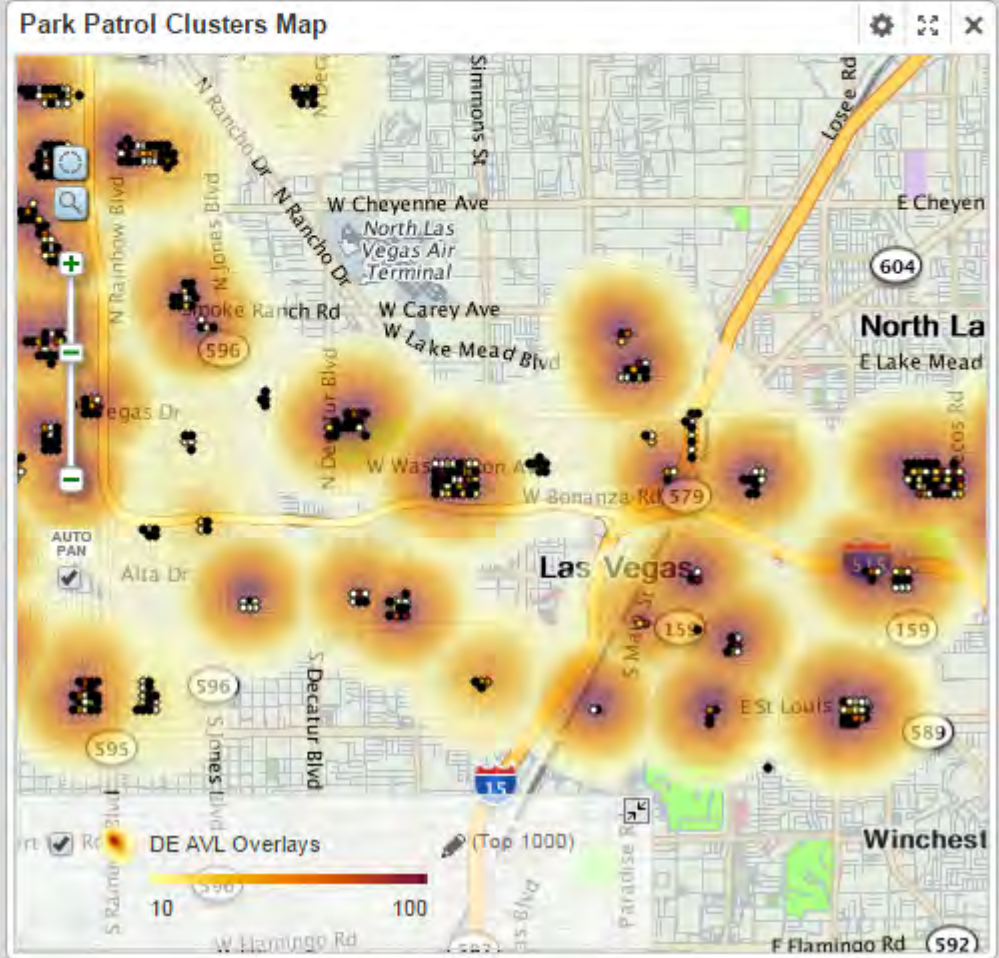
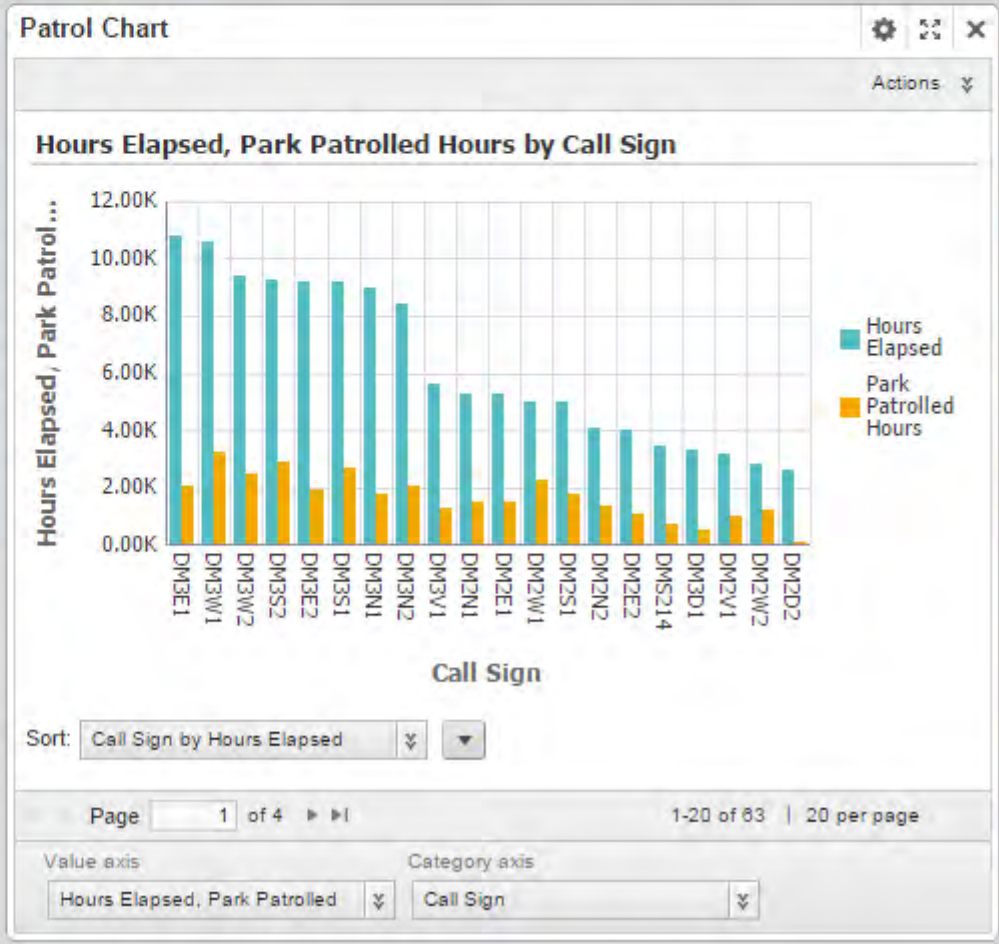
Component Container - Configuration - Remove

**AVL Detail Data Summarization**

1,460,899 AVL Positions	135,556.10 Time Patrolled (Hrs)	DM3E1 Call Sign Most Hours	11.61 Speed (avg)
26.00% % Park Patrolled	DM3W1 Call Sign Most Park Hours	2,317.32 Park Distance (avg)	

**Park Cluster Summarization (patrols within 100M)**

1,021 Num of Park Clusters	33,890.54 Total Cluster Patrol Hours
Lorenzi Park Top ParkNames by HoursPatrol (sum)	Ronemus Dr Top Street Name by Variance

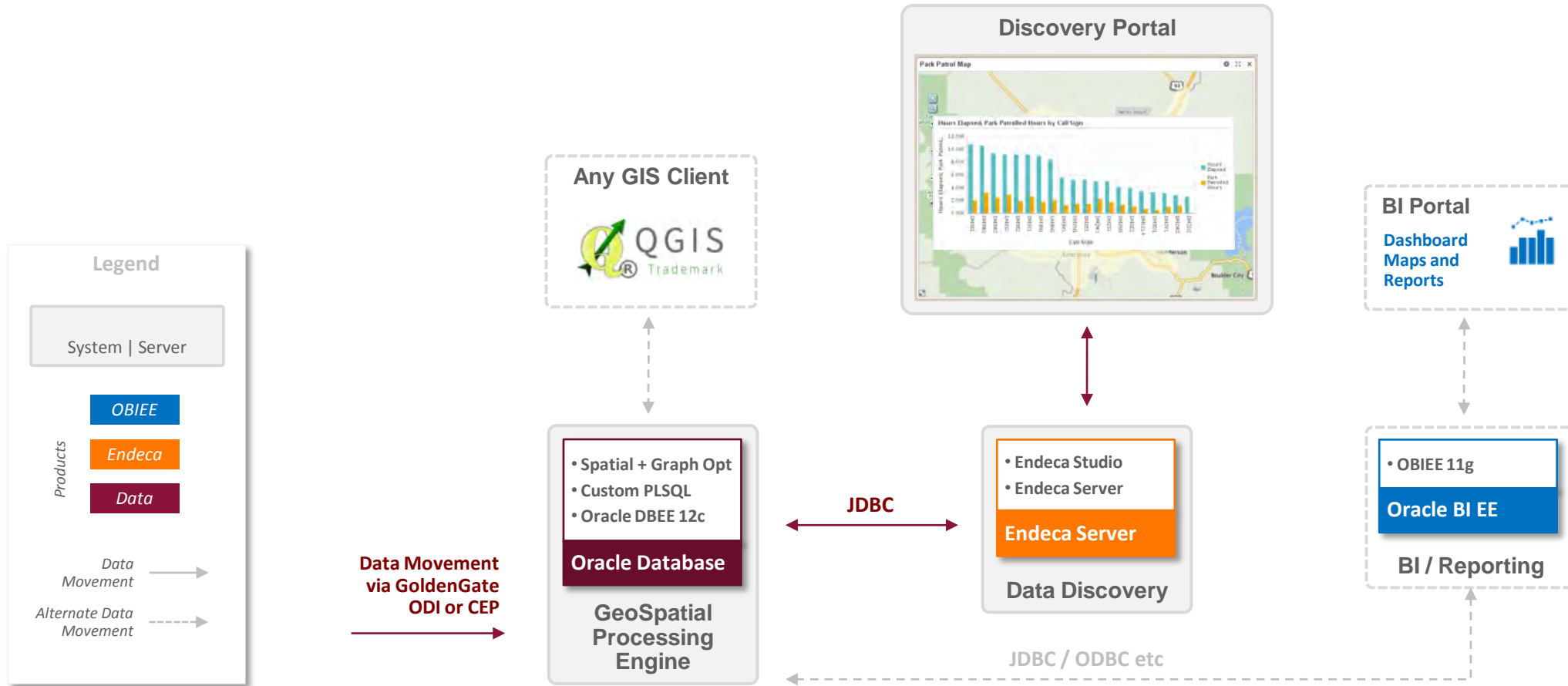






## E. Questions and Resources

# Logical Components to Solution





# Discussion / Q&A



- For this Use Case:
  - How best to enrich the AVL data?
  - How best to manage the AVL Data?
  - How best to analyze the AVL Data?
- What is the best visualization and analysis platform?
- Contact:
  - Chris Hughes  
[chris.m.hughes@oracle.com](mailto:chris.m.hughes@oracle.com)  
619-933-3336



# Appendix A

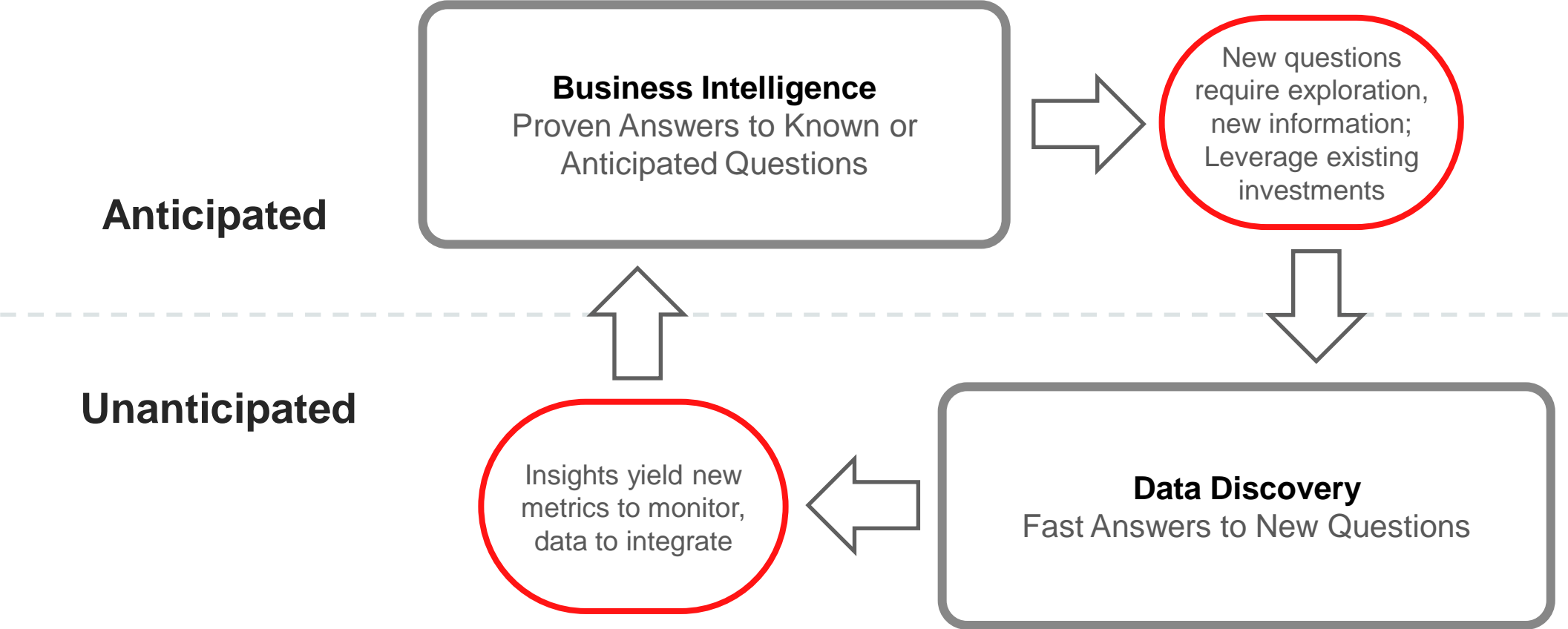
## Extra Slides | Endeca Capabilities





# Business Intelligence and Data Discovery

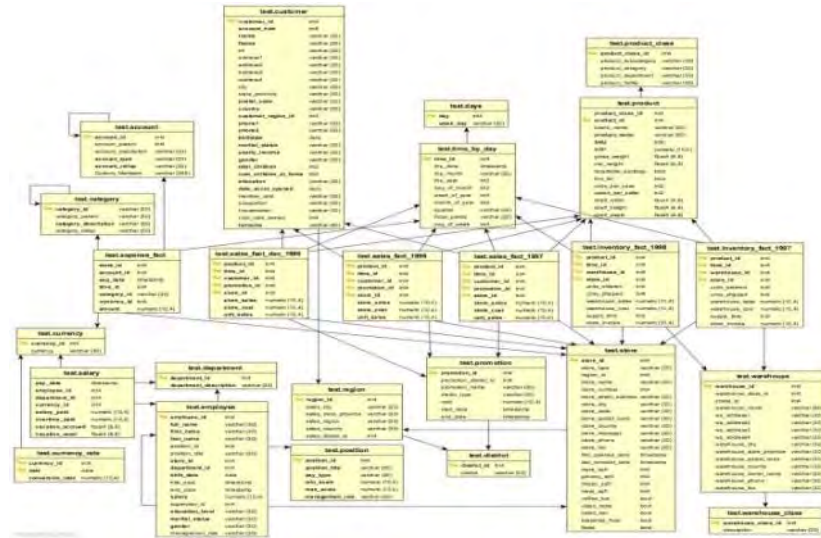
Complimentary Concepts – Technology is More Powerful Together



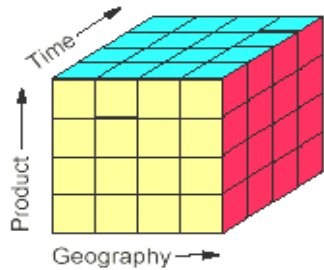
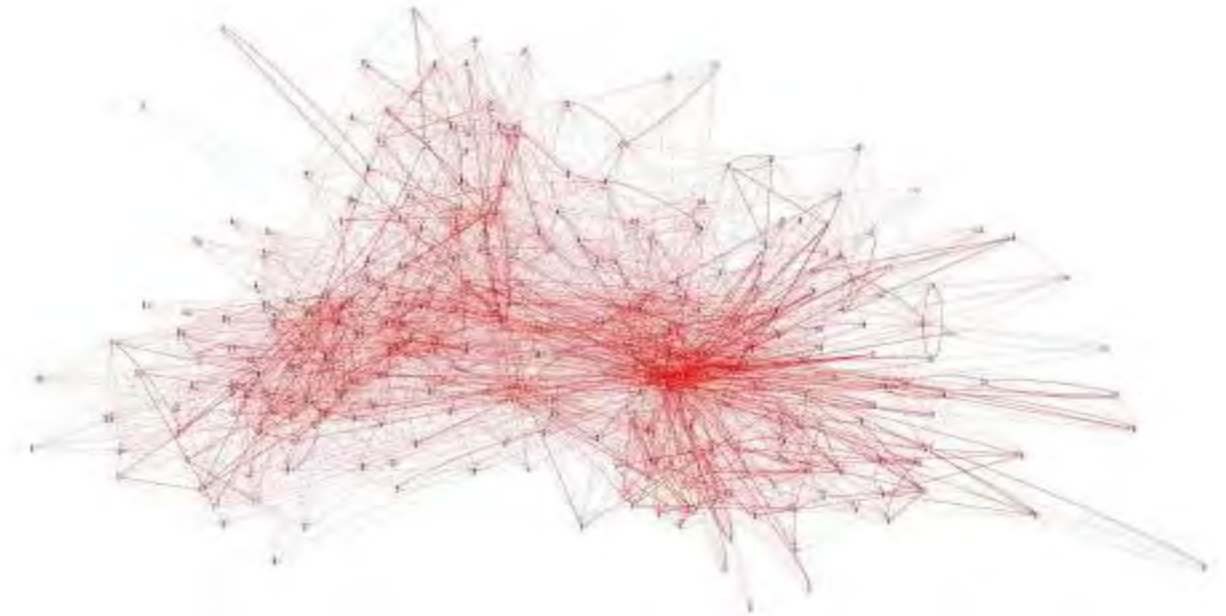
# Endeca's Brain – MDEX Data Store

Columns, Rows, Tables & Cubes

Semantic network of objects



VS





# Oracle Endeca Server: A Flexible 'NoSQL-like' Data Model

