Data Management Wrap-up
Wed. December 2 - Wrapup

Selected Project Presentations:

1. WDB: Morris Tirey Hurt Joshua Wu SongTing D1  
   https://www.github.com/tireymorris/F15D1
2. WDB: Deng Qicong Chee Raymond D2  
   https://github.com/alvin319/F15D2

Wrap-up:

- Class Notes
CS411: Intro to Upper Division Electives

April 13, 2015

Dr. Philip E. Cannata - Oracle and UT Adjunct Professor
CS347 – Data Management

CS347 will prepare you for this.
CS347 – Data Management

Growth of Data vs. Growth of Data Analysts

Stored Data accumulating at 28% annual growth rate
Data Analysts in workforce growing at 5.7% growth rate

CS347 will prepare you for this.
CS347 – Data Management

Talent Remains Key Big Data Challenge

Sixty-one percent of IT leaders expect spending on big data initiatives to increase, while only 5% expect decreases. The challenge: Finding the right big data talent to fulfill those initiatives, according to a recent survey by TEKsystems.

Nearly 60% of the respondents are confident that their IT department can satisfy big data demands of the business and 14% are not confident, according to TEKsystems, a provider of IT staffing and talent management services.

"The data indicates current expectations of big data are still somewhat unrealistic due to market hype," the report states. "Despite IT leaders spending to increase, the confidence level in their department's ability to meet big data demands in comparison to broader IT initiatives is lower."

About two thirds of the IT executives rank big data architects as the most difficult role to fill. Data scientists (48%) and data modelers (43%) round out the top three most difficult positions to fill. More technical big data positions are ranked less difficult to fill.

CS347 will prepare you for these.
You will learn – SQL, Data Modeling, Web/Database Application Development, the Oracle DBMS Architecture, Transaction Processing, Indexing, NoSQL and Graph Databases, Normalization, and SQL Query Optimization.
Got this in the email last October

Dear Phil,

Please register today for our October online Oracle APEX classes...

<table>
<thead>
<tr>
<th>Course</th>
<th>Comments</th>
<th>Start Date</th>
<th>Start Time</th>
<th>End Time</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction to APEX</td>
<td>78</td>
<td>Oct 6 - 10</td>
<td>1 PM - 5 PM EDT</td>
<td></td>
<td>$1,500</td>
</tr>
<tr>
<td>Advanced Oracle APEX</td>
<td>46</td>
<td>Oct 20 - 24</td>
<td>1 PM - 5 PM EDT</td>
<td></td>
<td>$1,500</td>
</tr>
<tr>
<td>HTML 5, CSS &amp; JavaScript (German)</td>
<td>(new)</td>
<td>Oct 6 - 10</td>
<td>12 PM - 5 PM GMT</td>
<td></td>
<td>$1,500</td>
</tr>
<tr>
<td>Introduction to SQL</td>
<td>38</td>
<td>Oct 6 - 10</td>
<td>1 PM - 5 PM EDT</td>
<td></td>
<td>$1,500</td>
</tr>
<tr>
<td>Advanced SQL Queries</td>
<td>74</td>
<td>Oct 20 - 24</td>
<td>1 PM - 5 PM EDT</td>
<td></td>
<td>$1,500</td>
</tr>
</tbody>
</table>

Online, instructor-led, hands-on class taught by APEX and Oracle experts Tyson Jouget, Christian Rokitta and Geoff Wiland.

Register online or call 1.888.803.5607. Call for group and alumni discounts.

Remember: Call us if you need an APEX DBA or Developer!
Continuing Data Management Research

Possible Opportunities:

- WDB Open Source Project
- SQL to SPARQL translation,
  see next 2 pages for examples
- SIM Query Language to SPARQL translation

  RETRIEVE Project-Title of Sub-Projects of Project, Name of Project-Team of Sub-Projects WHERE Project-Title of Project = “Annual Report Preparation”

- ReL Open Source Project
- ARL RFE in SIM and Apex

All of these are probably publishable and would make very good cs370 projects, also.
5.1 Basic Example

SQL:

select deptno, sal from emp where SAL > 1000

SPARQL:

SELECT v1 "DEPTNO", v2 "SAL"
FROM TABLE(SEM_MATCH('SELECT ?v1 ?v2 WHERE
   ?s1 rdf:type :EMP .
   OPTIONAL { ?s1 :DEPTNO ?v1 } /* Select them */
   OPTIONAL { ?s1 :SAL ?v2 }
   OPTIONAL { ?s1 :SAL ?v3 }
   FILTER(?f1 > 1000)
)

5.2 Join Example

SQL:

select dname, sal from emp e, dept d where e.deptno = d.deptno where sal > 1000

SPARQL:

SELECT v1 "DNAME", v2 "SAL"
FROM TABLE(SEM_MATCH('SELECT * WHERE {
   ?s1 rdf:type :EMP .
   ?s2 rdf:type :DEPT .
   OPTIONAL { ?s2 :DNAME ?v1 }
   OPTIONAL { ?s1 :SAL ?v2 }
   FILTER(?f1 = d.deptno && ?f2 > 1000) }

Dr. Philip Cannata
5.3 Complete Example

**SQL:**

```sql
select deptno, avg(sal) from emp
group by deptno
having avg(sal) > 1000
order by avg(sal)
```

**SPARQL:**

```sparql
SELECT v1 "DEPTNO", n2 "AVG(EMP.SAL)"
FROM TABLE(SEM_MATCH('SELECT ?v1 (avg(?v2) as ?n1) WHERE {
  ?s1 rdf:type :EMP .
  OPTIONAL { ?s1 :DEPTNO ?v1 }
  OPTIONAL { ?s1 :SAL ?v2 }
  OPTIONAL { ?s1 :SAL ?v3 }
  OPTIONAL { ?s1 :SAL ?v4 }
  OPTIONAL { ?s1 :DEPTNO ?v5 }
} GROUP BY ?v5
HAVING( avg(?v3) > 1000)
ORDER BY avg(?v4)
```
Final Review

- Review the Midterm, Homework (especially homework 8 and 9) and Quizzes (especially Quizzes 5, 6, and 7)
- SQL on the class website
- Normalization
- Query Execution Plans

\[ \approx 30\% \]
Thank You!

Chia-Chen Hsu and Chris Timaeus
Midterm Exam Grade = \text{max}(\text{Midterm Grade, Final Exam Grade})

\[
AB3 = (25 \times (\max((V3+8)/140, AE3))) + (20 \times (Y3/10)) + (15 \times (AA3/10)) + (15 \times (W3/100)) + \text{Round}
\]
Final Grade = max(Current Grade, Current Grade + Final Exam)

(i.e., the Final is optional, it can’t hurt you, it can only help – Thank you for your effort this semester especially on the projects.)

<table>
<thead>
<tr>
<th>Homework 1</th>
<th>Quiz 1</th>
<th>Quiz 2</th>
<th>Homework 2</th>
<th>Quiz 3</th>
<th>Homework 3</th>
<th>Quiz 4</th>
<th>Homework 4</th>
<th>Quiz 5</th>
<th>Homework 5</th>
<th>Quiz 6</th>
<th>Homework 6</th>
<th>Test 1</th>
<th>Quiz 7</th>
<th>Project</th>
<th>Current Numberic Grade</th>
<th>Current Letter Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>9</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>8.75</td>
<td>9.5</td>
<td>10</td>
<td>8</td>
<td>10</td>
<td>6.8</td>
<td>7</td>
<td>92</td>
<td>10</td>
<td>7.5</td>
<td>82</td>
<td>91.35</td>
</tr>
</tbody>
</table>

Have a Great Christmas!