Business Intelligence
Forecasting and Modeling
with Oracle10g SQL

Anthony D. Noriega
MSCS, MBA, BSSE, OCP-DBA
Speaker Qualifications

- Independent Consultant, ADN
- Over 20 years of IT experience, 16 with Oracle
- Former OCP Instructor
- Speaker at NYOUG and IOUG events
- MS Computer Science, NJIT, 1993
- PhD CIS candidate, NJIT, 1997
- MBA MIS, Montclair State University, 2006
- Forecasting projects at M&M Mars, AT&T
- Analytical/Statistical projects with FMC and AT&T
- Modeling Projects with Bowne, DB, eWayDirect
Objective

- Emphasize Oracle10g SQL Modeling and Forecasting capabilities.
- Discuss areas of current and future applications, and relevant studies.
- Introduce and expand on a frame of reference to approach BI modeling and forecasting.
Business Framework

Traditional Methodologies

- Statistical Analysis
  - Linear Regression
  - Multiple Regression
  - Factorial Design
  - Latin Quarters

- Exponential Smoothing
  - Simple Exponential Smoothing
  - Double Exponential Smoothing
Business Framework

Traditional Methodologies

- Time Series Analysis
  - Moving Average
  - Moving Average with Seasonal Adjustments
  - ARMA (Autoregressive Moving Average)
    - Box, Jenkins (1976)
  - ARIMA (Autoregressive Integrated Moving Average)
Business Framework

Traditional Methodologies
- Operating Research
  - Markov Chains (Stochastic Processes)
  - Bayesian Models (Probabilistic Models)
- Game Theory Approaches
- Mathematical Models
  - Linear Algebra
  - Structured Matrices and Polynomials Models
BI Software Market

Oracle in the Leader’s Quadrant
Data Mining Software Market

Oracle in the Leader’s Quadrant
# Oracle’s BI Approach

<table>
<thead>
<tr>
<th>Time Frame</th>
<th>Time Level</th>
<th>Typical Forecasting Horizon</th>
<th>Best Approach</th>
<th>Product Level</th>
<th>Other Dimension Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short</td>
<td>Week, Biweek, or Month</td>
<td>Up to 18 months</td>
<td>Time Series</td>
<td>UPC, SKU, NDC, ISBN</td>
<td>Level of interest</td>
</tr>
<tr>
<td>Medium</td>
<td>Month or Quarter</td>
<td>6 to 36 months</td>
<td>Causal Analysis</td>
<td>Brand</td>
<td>Level of interest</td>
</tr>
<tr>
<td>Long</td>
<td>Quarter or higher</td>
<td>19 months to 5 years</td>
<td>Expert Opinion</td>
<td>Brand, Company, Market</td>
<td>Level of interest</td>
</tr>
</tbody>
</table>

*Table 1. Oracle Corporation Vision to Forecasting Strategies*
What to Forecast

World Population: 1950-2050

- 3 Billion in 1950
- 4 Billion in 1970
- 5 Billion in 1980
- 6 Billion in 2000
- 7 Billion in 2020
- 8 Billion in 2040
- 9 Billion in 2050

Source: U.S. Census Bureau, International Data Base, August 2006 version.
# What to Forecast

## People Living with HIV 2000-2004 and Forecast Through 2050

**Human Population Statistics Given in Millions**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>34.00</td>
<td>31.00</td>
<td>38.00</td>
<td>4.00</td>
<td>3.52</td>
<td>5.00</td>
<td>0.93</td>
<td>0.95</td>
<td>0.94</td>
<td>0.94</td>
</tr>
<tr>
<td>2001</td>
<td>35.00</td>
<td>32.00</td>
<td>39.20</td>
<td>4.20</td>
<td>3.80</td>
<td>6.04</td>
<td>0.97</td>
<td>0.97</td>
<td>0.97</td>
<td>0.97</td>
</tr>
<tr>
<td>2002</td>
<td>36.60</td>
<td>33.30</td>
<td>41.10</td>
<td>4.40</td>
<td>3.90</td>
<td>6.20</td>
<td>1.02</td>
<td>1.02</td>
<td>1.02</td>
<td>1.02</td>
</tr>
<tr>
<td>2003</td>
<td>38.00</td>
<td>35.00</td>
<td>42.65</td>
<td>4.65</td>
<td>4.21</td>
<td>6.59</td>
<td>1.08</td>
<td>1.06</td>
<td>1.07</td>
<td>1.06</td>
</tr>
<tr>
<td>2004</td>
<td>39.40</td>
<td>35.90</td>
<td>44.30</td>
<td>4.90</td>
<td>4.30</td>
<td>6.30</td>
<td>1.11</td>
<td>1.05</td>
<td>1.05</td>
<td>1.06</td>
</tr>
<tr>
<td>2005</td>
<td>41.67</td>
<td>37.85</td>
<td>46.94</td>
<td>5.42</td>
<td>4.78</td>
<td>6.97</td>
<td>1.18</td>
<td>1.07</td>
<td>1.07</td>
<td>1.07</td>
</tr>
<tr>
<td>2006</td>
<td>44.62</td>
<td>40.34</td>
<td>50.36</td>
<td>6.39</td>
<td>5.61</td>
<td>8.22</td>
<td>1.32</td>
<td>1.09</td>
<td>1.08</td>
<td>1.09</td>
</tr>
<tr>
<td>2007</td>
<td>48.66</td>
<td>43.66</td>
<td>55.05</td>
<td>8.42</td>
<td>7.39</td>
<td>10.83</td>
<td>1.59</td>
<td>1.12</td>
<td>1.11</td>
<td>1.12</td>
</tr>
<tr>
<td>2008</td>
<td>54.32</td>
<td>48.34</td>
<td>61.65</td>
<td>13.39</td>
<td>11.75</td>
<td>17.22</td>
<td>2.16</td>
<td>1.14</td>
<td>1.14</td>
<td>1.15</td>
</tr>
<tr>
<td>2009</td>
<td>62.35</td>
<td>54.92</td>
<td>71.05</td>
<td>28.94</td>
<td>25.39</td>
<td>37.20</td>
<td>3.42</td>
<td>1.17</td>
<td>1.17</td>
<td>1.19</td>
</tr>
<tr>
<td>2010</td>
<td>74.07</td>
<td>64.43</td>
<td>84.80</td>
<td>98.84</td>
<td>86.74</td>
<td>127.08</td>
<td>5.92</td>
<td>1.24</td>
<td>1.22</td>
<td>1.24</td>
</tr>
</tbody>
</table>
What to Model

- Finance
- Economics
- NASDAQ
- Dow
- S&P
Technical Framework

Analytical Workspace

- Tables
- Dimensions
  - Levels
  - Hierarchies
    - Level-Based
    - Value-Based
- Measures
Technical Framework

Analytical Workspace
- Logical Cubes
- Rules
- Attributes
Analytical Workspace

- OLAP Registry
- Analytical Workspace Manager
Analytical SQL Syntax

Analytical SQL DML Syntax Highlights

## A Simple Regression Model

```sql
SQL> get d:\batch\man_hist.sql
1 SELECT backup_type, db_name, avg_db_size, max_duration,  
2 REGR_SLOPE(avg_db_size, max_duration) slope,  
3 REGR_INTERCEPT(avg_db_size, max_duration) INTcpt,  
4 REGR_R2(avg_db_size, max_duration) r2,  
5 OVER (PARTITION BY backup_type) slope,  
6 OVER (PARTITION BY backup_type) intcpt,  
7 OVER (PARTITION BY backup_type) r2,  
8 OVER (PARTITION BY backup_type) count,  
9 OVER (PARTITION BY backup_type) avgx,  
10 OVER (PARTITION BY backup_type) avgy,  
11 OVER (PARTITION BY backup_type) avgx,  
12 OVER (PARTITION BY backup_type) avgy  
13 FROM man_hist  
14 ORDER BY regr_desc  
SQL> /  
```

<table>
<thead>
<tr>
<th>BACKUP_TYPE</th>
<th>DB_NAME</th>
<th>AVG_DB_SIZE</th>
<th>MAX_DURATION</th>
<th>SLOPE</th>
<th>INTcpt</th>
<th>R2R</th>
<th>COUNT</th>
<th>AVGX</th>
<th>AVGY</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAT</td>
<td>Uat19</td>
<td>878</td>
<td>44.66</td>
<td>27.2728662</td>
<td>-71.451668</td>
<td>.915574947</td>
<td>15</td>
<td>31.78</td>
<td>795.28</td>
</tr>
<tr>
<td>CAT</td>
<td>Uat19</td>
<td>878</td>
<td>44.66</td>
<td>27.2728662</td>
<td>-71.451668</td>
<td>.915574947</td>
<td>15</td>
<td>31.78</td>
<td>795.28</td>
</tr>
<tr>
<td>CAT</td>
<td>Risk18</td>
<td>716.8</td>
<td>38.08</td>
<td>27.2728662</td>
<td>-71.451668</td>
<td>.915574947</td>
<td>15</td>
<td>31.78</td>
<td>795.28</td>
</tr>
<tr>
<td>CAT</td>
<td>Prd16</td>
<td>819.2</td>
<td>28.16</td>
<td>27.2728662</td>
<td>-71.451668</td>
<td>.915574947</td>
<td>15</td>
<td>31.78</td>
<td>795.28</td>
</tr>
<tr>
<td>CAT</td>
<td>Dev17</td>
<td>870.8</td>
<td>30.56</td>
<td>27.2728662</td>
<td>-71.451668</td>
<td>.915574947</td>
<td>15</td>
<td>31.78</td>
<td>795.28</td>
</tr>
<tr>
<td>CAT</td>
<td>Risk18</td>
<td>460.8</td>
<td>24.48</td>
<td>27.2728662</td>
<td>-71.451668</td>
<td>.915574947</td>
<td>15</td>
<td>31.78</td>
<td>795.28</td>
</tr>
<tr>
<td>CAT</td>
<td>Uat19</td>
<td>579.6</td>
<td>29.26</td>
<td>27.2728662</td>
<td>-71.451668</td>
<td>.915574947</td>
<td>15</td>
<td>31.78</td>
<td>795.28</td>
</tr>
<tr>
<td>CAT</td>
<td>Risk1210</td>
<td>1254</td>
<td>46.4</td>
<td>27.2728662</td>
<td>-71.451668</td>
<td>.915574947</td>
<td>15</td>
<td>31.78</td>
<td>795.28</td>
</tr>
<tr>
<td>CAT</td>
<td>Risk1210</td>
<td>628</td>
<td>23.2</td>
<td>27.2728662</td>
<td>-71.451668</td>
<td>.915574947</td>
<td>15</td>
<td>31.78</td>
<td>795.28</td>
</tr>
<tr>
<td>CAT</td>
<td>Risk1210</td>
<td>1884</td>
<td>69.6</td>
<td>27.2728662</td>
<td>-71.451668</td>
<td>.915574947</td>
<td>15</td>
<td>31.78</td>
<td>795.28</td>
</tr>
<tr>
<td>CAT</td>
<td>Prd16</td>
<td>3517.2</td>
<td>10.56</td>
<td>27.2728662</td>
<td>-71.451668</td>
<td>.915574947</td>
<td>15</td>
<td>31.78</td>
<td>795.28</td>
</tr>
<tr>
<td>CAT</td>
<td>Dev17</td>
<td>358.4</td>
<td>12.6</td>
<td>27.2728662</td>
<td>-71.451668</td>
<td>.915574947</td>
<td>15</td>
<td>31.78</td>
<td>795.28</td>
</tr>
<tr>
<td>CAT</td>
<td>Risk18</td>
<td>2084.8</td>
<td>10.88</td>
<td>27.2728662</td>
<td>-71.451668</td>
<td>.915574947</td>
<td>15</td>
<td>31.78</td>
<td>795.28</td>
</tr>
<tr>
<td>CAT</td>
<td>Prd16</td>
<td>1331.2</td>
<td>45.76</td>
<td>27.2728662</td>
<td>-71.451668</td>
<td>.915574947</td>
<td>15</td>
<td>31.78</td>
<td>795.28</td>
</tr>
<tr>
<td>CAT</td>
<td>Dev17</td>
<td>1382.4</td>
<td>48.6</td>
<td>27.2728662</td>
<td>-71.451668</td>
<td>.915574947</td>
<td>15</td>
<td>31.78</td>
<td>795.28</td>
</tr>
<tr>
<td>NOCAT</td>
<td>Q015</td>
<td>758</td>
<td>14.5</td>
<td>44.4942561</td>
<td>167.5266572</td>
<td>.610395486</td>
<td>15</td>
<td>2.9</td>
<td>296.56</td>
</tr>
<tr>
<td>NOCAT</td>
<td>prod12</td>
<td>39.6</td>
<td>1.6</td>
<td>44.4942561</td>
<td>167.5266572</td>
<td>.610395486</td>
<td>15</td>
<td>2.9</td>
<td>296.56</td>
</tr>
<tr>
<td>NOCAT</td>
<td>dev13</td>
<td>38.4</td>
<td>1.18</td>
<td>44.4942561</td>
<td>167.5266572</td>
<td>.610395486</td>
<td>15</td>
<td>2.9</td>
<td>296.56</td>
</tr>
<tr>
<td>NOCAT</td>
<td>dev14</td>
<td>99.2</td>
<td>1.72</td>
<td>44.4942561</td>
<td>167.5266572</td>
<td>.610395486</td>
<td>15</td>
<td>2.9</td>
<td>296.56</td>
</tr>
<tr>
<td>NOCAT</td>
<td>Q015</td>
<td>158</td>
<td>2.9</td>
<td>44.4942561</td>
<td>167.5266572</td>
<td>.610395486</td>
<td>15</td>
<td>2.9</td>
<td>296.56</td>
</tr>
<tr>
<td>NOCAT</td>
<td>prod11</td>
<td>281.6</td>
<td>1.54</td>
<td>44.4942561</td>
<td>167.5266572</td>
<td>.610395486</td>
<td>15</td>
<td>2.9</td>
<td>296.56</td>
</tr>
<tr>
<td>NOCAT</td>
<td>prod12</td>
<td>237.6</td>
<td>1.96</td>
<td>44.4942561</td>
<td>167.5266572</td>
<td>.610395486</td>
<td>15</td>
<td>2.9</td>
<td>296.56</td>
</tr>
<tr>
<td>NOCAT</td>
<td>dev13</td>
<td>166.4</td>
<td>1.78</td>
<td>44.4942561</td>
<td>167.5266572</td>
<td>.610395486</td>
<td>15</td>
<td>2.9</td>
<td>296.56</td>
</tr>
<tr>
<td>NOCAT</td>
<td>dev14</td>
<td>347.2</td>
<td>2.52</td>
<td>44.4942561</td>
<td>167.5266572</td>
<td>.610395486</td>
<td>15</td>
<td>2.9</td>
<td>296.56</td>
</tr>
<tr>
<td>NOCAT</td>
<td>Q015</td>
<td>458</td>
<td>8.7</td>
<td>44.4942561</td>
<td>167.5266572</td>
<td>.610395486</td>
<td>15</td>
<td>2.9</td>
<td>296.56</td>
</tr>
<tr>
<td>NOCAT</td>
<td>prod11</td>
<td>537.6</td>
<td>2.94</td>
<td>44.4942561</td>
<td>167.5266572</td>
<td>.610395486</td>
<td>15</td>
<td>2.9</td>
<td>296.56</td>
</tr>
<tr>
<td>NOCAT</td>
<td>prod12</td>
<td>435.6</td>
<td>1.76</td>
<td>44.4942561</td>
<td>167.5266572</td>
<td>.610395486</td>
<td>15</td>
<td>2.9</td>
<td>296.56</td>
</tr>
<tr>
<td>NOCAT</td>
<td>dev13</td>
<td>294.4</td>
<td>1.38</td>
<td>44.4942561</td>
<td>167.5266572</td>
<td>.610395486</td>
<td>15</td>
<td>2.9</td>
<td>296.56</td>
</tr>
<tr>
<td>NOCAT</td>
<td>dev14</td>
<td>595.2</td>
<td>4.32</td>
<td>44.4942561</td>
<td>167.5266572</td>
<td>.610395486</td>
<td>15</td>
<td>2.9</td>
<td>296.56</td>
</tr>
<tr>
<td>NOCAT</td>
<td>prod11</td>
<td>25.6</td>
<td>.14</td>
<td>44.4942561</td>
<td>167.5266572</td>
<td>.610395486</td>
<td>15</td>
<td>2.9</td>
<td>296.56</td>
</tr>
</tbody>
</table>

30 rows selected.
## Analytical Functions (SQL)

<table>
<thead>
<tr>
<th>Statistical Analysis</th>
<th>Statistical Analysis</th>
<th>Object Reference</th>
<th>Model Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>AVG *</td>
<td>RANK</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CORR *</td>
<td>RATIO_TO_REPORT</td>
<td>DEREF</td>
<td>ITERATION_NUMBER</td>
</tr>
<tr>
<td>COVAR_POP *</td>
<td>REGR_SLOPE</td>
<td>MAKE_REF</td>
<td>PRESENTNINV</td>
</tr>
<tr>
<td>COVAR_SAMP *</td>
<td>REGR_INTERCEPT</td>
<td>REF</td>
<td>PRESENTV</td>
</tr>
<tr>
<td>COUNT *</td>
<td>REGR_COUNT</td>
<td>REFTOHEX</td>
<td>PREVIOUS</td>
</tr>
<tr>
<td>CUME_DIST</td>
<td>REGR_R2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DENSE_RANK</td>
<td>REGR_AVGX</td>
<td>Model Functions</td>
<td></td>
</tr>
<tr>
<td>FIRST</td>
<td>REGR_AVGY</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FIRST_VALUE *</td>
<td>REGR_SXX</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LAG</td>
<td>REGR_SYY</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LAST</td>
<td>REGR_SKY</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LAST_VALUE *</td>
<td>ROW_NUMBER</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LEAD</td>
<td>STDDEV*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAX *</td>
<td>STDDEV_POP *</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MIN *</td>
<td>STDDEV_SAMP *</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NTILE</td>
<td>SUM *</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PERCENT_RANK</td>
<td>VAR_POP *</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PERCENTILE_CONT</td>
<td>VAR_SAMP *</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PERCENTILE_DISC</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Statistical Analysis Functions

<table>
<thead>
<tr>
<th>Statement</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CATEGORIZE</td>
<td>Groups the values of a numeric expression into categories.</td>
</tr>
<tr>
<td>CORRELATION</td>
<td>Returns the correlation coefficients for the pairs of data values in two expressions.</td>
</tr>
<tr>
<td>NORMAL</td>
<td>Returns a random value from a normal distribution with a specified mean and standard deviation. The result returned by NORMAL accounts for all the dimensions of the mean and standard deviation</td>
</tr>
<tr>
<td>RANDOM</td>
<td>Produces a number that is randomly distributed between specified low and high boundaries.</td>
</tr>
<tr>
<td>STDDEV</td>
<td>Calculates the standard deviation of the values of an expression.</td>
</tr>
</tbody>
</table>

*Table 5. Statistical Analysis Functions*
<table>
<thead>
<tr>
<th>Function</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEPRDECL</td>
<td>Computes the depreciation expenses for a series of assets using the declining balance method</td>
</tr>
<tr>
<td>DEPRDECLS</td>
<td>Calculates the depreciation expenses for a series of assets using a variation on the declining balance method to depreciate assets over the specified lifetime of the assets.</td>
</tr>
<tr>
<td>DEPRSL</td>
<td>Calculates the depreciation expenses for a series of assets. DEPRSL uses the straight line method to depreciate the assets over the specified lifetime of the assets.</td>
</tr>
<tr>
<td>DEPRSOYD</td>
<td>Calculates the depreciation expenses for a series of assets. DEPRSOYD uses the sum-of-years’-digits method to depreciate the assets over the specified lifetime of the assets.</td>
</tr>
<tr>
<td>FINTSCHED</td>
<td>Calculates the interest portion of the payments on a series of fixed-rate installment loans that are paid off over a specified number of time periods.</td>
</tr>
<tr>
<td>FPMTSCHED</td>
<td>Calculates a payment schedule (principal plus interest) for paying off a series of fixed-rate installment loans over a specified number of time periods.</td>
</tr>
<tr>
<td>GROWRATE</td>
<td>Calculates the growth rate of a time-series expression, based on the first and last values of the series.</td>
</tr>
<tr>
<td>IRR</td>
<td>Computes the internal rate of return associated with a series of cash flow values. Each value of the result is calculated to be the discount rate for a period that makes the net present value of the corresponding cash flows equal to zero.</td>
</tr>
<tr>
<td>NPY</td>
<td>Computes the net present value of a series of cash flow values.</td>
</tr>
<tr>
<td>VINTSCHED</td>
<td>Calculates the interest portion of the payments on a series of variable-rate installment loans that are paid off over a specified number of time periods.</td>
</tr>
<tr>
<td>VPMTSCHED</td>
<td>Calculates a payment schedule (principal plus interest) for paying off a series of variable-rate installment loans over a specified number of time periods.</td>
</tr>
</tbody>
</table>
Forecast / Regression Statements

- Simple Forecast and Regressions
  - REGRESS.report
  - SMOOTH

- Forecast and Regression Statements using Context
Analytical Operators

- Rollup
- Cube
Group-by Aggregation

- Aggregation using Grouping Sets
- Aggregation using Operators
- Aggregation with the Group Id and Grouping Id Differentiators.
**Dimensions (MOLAP)**

```sql
CREATE DIMENSION RMAN_REC_DIM
  LEVEL DB_NAME IS
    (RMAN_HIST,DB_NAME)
  LEVEL CITY IS
    (RMAN_HIST,CITY)
  LEVEL STATE IS
    (RMAN_HIST,STATE)
  LEVEL REGION IS
    (RMAN_HIST,REGION)
HIERARCHY INST_ROLLUP
  (DB_NAME CHILD OF
   CITY CHILD OF
   STATE CHILD OF
   REGION)
ATTRIBUTE DB_NAME DETERMINES
  (RMAN_HIST,HOSTNAME ,
   RMAN_HIST,MAX_DURATION ,
   RMAN_HIST,AVG_DB_SIZE ,
   RMAN_HIST,RMAN_BKP_MAXSIZE ,
   RMAN_HIST,RMAN_BKP_MINSIZE ,
   RMAN_HIST,RECOVERY_TYPE ,
   RMAN_HIST,MIN_DURATION );
```

1. BEGIN
2.  dbms_olap.validate_dimension(
3.    dimension_name => 'RMAN_REC_DIM',
4.    dimension_owner => 'ANTHONY');
5.  END;

SQL> /
PL/SQL procedure successfully completed.

---

**Exhibit 5A. Create Dimension Statement (OLAP SQL DDL)**

**Exhibit 5B. Dimension Validation with DBMS_OLAP**
MODEL Clause

MODEL
[<global reference options>]
[<reference models>]
[MAIN <main-name>]
[PARTITION BY (<cols>)]
DIMENSION BY (<cols>)
MEASURES (<cols>)
[<reference options>]
[RULES] <rule options>
(<rule>, <rule>,.., <rule>)
<global reference options> ::= <reference options> <ret-opt>
<ret-opt> ::= RETURN {ALL|UPDATED} ROWS
<reference options> ::= [IGNORE NAV | [KEEP NAV]
[UNIQUE DIMENSION | UNIQUE SINGLE REFERENCE]
<brule options> ::= [UPDATE | UPSERT | UPSERT ALL]
[AUTOMATIC ORDER | SEQUENTIAL ORDER]
[ITERATE (<number>) [UNTIL <condition>]]
<reference models> ::= REFERENCE ON <ref-name> ON (<query>)
DIMENSION BY (<cols>) MEASURES (<cols>) <reference options>
MODEL-Related Clauses

- Partition by
- Dimension by
- Measures
- Rules
  - Rules Options
    - Update, Upsert, Upsert all
In addition to matching rules these conditions can filter the rule application:

- IS ANY
- IS PRESENT
Unified Perspective

- SQL Functions and API
  - Functions
  - Statements
- OLAP DML API
- J2EE
- XML
- SOA
Current Benefits

- Analytic Workspace Support
- OLAP Registry
- Oracle’s OLAP DML J 2EE/ Java and PL/ SQL Integration
- E-Business Support
- Hot Plug-in
- Enhanced OLAP DBA, DBA Capabilities.
SELECT  borough,  Book_title,
        To_char(Due_date,'MM') Month_mm,
        fines, total_dues
FROM library_fines
WHERE borough IN ('Brooklyn', 'Manhattan')
MODEL
  RETURN UPDATED ROWS
  PARTITION BY (SUBSTR(borough,1,12) AS borough)
  DIMENSION BY (Book_title AS b, Month_mm)
  MEASURES (fines, 0 AS total_dues)
  RULES
  (total_dues['Madame Bovary', '03'] = fines['Madame Bovary','0 2'] * 0.5,
   fines['Madame Bovary', '03'] = fines['Madame Bovary', '01'] +
   fines['Madame Bovary', '02']
   total_dues['Madame Bovary', '02'] = ['Madame Bovary', '01' * 0.5])
ORDER BY borough, month_mm;
SELECT RTRIM(country) country, SUBSTR(product,1,16) product, year, sales
FROM sales_view
WHERE country in ('France', 'China')
MODEL
  RETURN UPDATED ROWS
  MAIN simple_model
  PARTITION BY (country)
  DIMENSION BY (product, year)
  MEASURES (s)
  RULES
    (s['iPod', 2002] = 2000,
     s['iPod', 2003] = s['iPod', 2001] + s['iPod', 2000],
     s['PlayStation2', 2002] = s['PlayStation2', 2001])
ORDER BY country, product, year;
SELECT country,  
    product,  
    year,  
    sales,  
    profits  
FROM sales_view  
WHERE country IN ('France', 'China')  
MODEL  
    RETURN UPDATED ROWS  
    PARTITION BY (SUBSTR(country,1,20) AS country)  
    DIMENSION BY (product AS p, year)  
    MEASURES (sales, 0 AS profits)  
    RULES  
        (profits['iPod', 2003] = sales['iPod', 2002] * 0.4,  
        profits['iPod', 2003] = sales['iPod', 2003] * 0.5)  
ORDER BY country, year;
SELECT country,
    year,
    sales,
    usd_sales
FROM sales_view
GROUP BY country, year
MODEL
    REFERENCE conv_ref ON
        (SELECT country, exchange_rate FROM dollar_conv_tbl)
    DIMENSION BY (country c) MEASURES (exchange_rate)
IGNORE NAV
    REFERENCE growth_ref ON
        (SELECT country, year, growth_rate FROM growth_rate_tbl)
    DIMENSION BY (country c, year y) MEASURES (growth_rate)
IGNORE NAV
    MAIN projection
    DIMENSION BY (country, year) MEASURES (SUM(sales) sales, 0
usd_sales)
IGNORE NAV
    RULES
    (usd_sales[ANY, 2004] = sales[CV(country), 2003] *
growth_rate[CV(country), CV(year)] *
exchange_rate[CV(country)]);
SQL Examples

SELECT country,
       year,
       sales,
       usd_sales
FROM sales_view
GROUP BY country, year
MODEL
  REFERENCE conv_ref ON
    ( SELECT country, exchange_rate
      FROM dollar_conv_tbl
    )
  DIMENSION BY (country) MEASURES (exchange_rate) IGNORE NAV
  MAIN conversion
    DIMENSION BY (country, year)
    MEASURES (SUM(sales) sales, SUM(sales) usd_sales) IGNORE NAV
  RULES
    (usd_sales['France', 2004] = sales[CV(country), 2003] *.95 *
      conv_ref.exchange_rate['France'],
    usd_sales['Japan', 2004] =
      sales['Japan', 2003] *.105 * exchange_rate['Japan']);
SELECT DISTINCT R,C,H,DB,AVG_S,rs,max_dur,dur_diff,slope,r2
FROM
(
SELECT region r ,
    city c,
    hostname h ,
    db_name db,
    avg(avg_db_size) over (partition by region,city order by region,city)
    avg_s,
    avg(rman_bkp_maxsize) over (partition by region,city order by region,city)
    rs ,
    avg(max_duration) over (partition by region,city order by region,city)
    max_dur,
    avg(max_duration-min_duration) over (partition by region,city order by region,city)
    dur_diff,
    regr_slope(avg_db_size,max_duration) over (partition by region,city order by region,city)
    slope,
    regr_r2(avg_db_size,max_duration) over (partition by region,city order by region,city)
    r2
from rman_hist
) A
order by 1,2,3;
Concluding Remarks

Oracle SQL analytical framework has attained an outstanding level of maturity deserving a leadership positioning.

Oracle SQL and OLAP API can integrate with custom J2EE, XML, and .net applications, and for a key reporting Data Warehousing capability with Text and related tools.
Concluding Remarks

SQL Model capabilities have achieved a transactional and reporting level which allows the visualization of modeling and forecasting in query-like mode, with greater flexibility than conventional statistical analysis and modeling tools.

Oracle10g support through both OLAP DML and API and SQL is essentially solid for conventional forecasting methods such as Time Series (Moving Average and ARIMA) and Exponential Smoothing, among others.
Future Forecasting Models

- Who
- Where
- When
- Extent
- How
- How much
Case Studies

- Health (HIV, AIDS)
- Demographics (Census)
- Financial Markets (Nasdaq)
- Econometrics (Income/Gender)
- Marketing Research (4Ps)
Concluding Remarks

Integration
MODEL
Forecasting