Data Cube: A Relational Aggregation Operator Generalizing Group-By, Cross-Tab, and Sub-Totals

Jim Gray  Surajit Chaudhuri  Adam Bosworth  Andrew Layman  Don Reichart  Murali Venkatrao  Frank Pellow  Hamid Pirahesh

October 31+1, 2006

Presented by Michael Lawrence
Outline

1. Data Analysis

2. Analysis Operations
   - Histograms
   - Roll-Up/Drill-Down
   - Cross-Tab

3. The Data Cube
   - The ALL Dummy Value
   - The CUBE Operator

Gray et al.
Data Cube: A Relational Aggregation Operator Generalizing Group-By, Cross-Tab, and Sub-Totals
Data Analysis

Key features

- Dimensionality reduction
- Aggregation

Gray et al.

Data Cube: A Relational Aggregation Operator Generalizing Group-By, Cross-Tab, and Sub-Totals
## Example Data

<table>
<thead>
<tr>
<th>Model</th>
<th>Year</th>
<th>Colour</th>
<th>Num_Sold</th>
</tr>
</thead>
<tbody>
<tr>
<td>Golf</td>
<td>2001</td>
<td>Black</td>
<td>40</td>
</tr>
<tr>
<td>Golf</td>
<td>2001</td>
<td>Silver</td>
<td>65</td>
</tr>
<tr>
<td>Golf</td>
<td>2002</td>
<td>Black</td>
<td>70</td>
</tr>
<tr>
<td>Golf</td>
<td>2002</td>
<td>Silver</td>
<td>100</td>
</tr>
<tr>
<td>Jetta</td>
<td>2001</td>
<td>Black</td>
<td>70</td>
</tr>
<tr>
<td>Jetta</td>
<td>2001</td>
<td>Silver</td>
<td>50</td>
</tr>
<tr>
<td>Jetta</td>
<td>2002</td>
<td>Black</td>
<td>60</td>
</tr>
<tr>
<td>Jetta</td>
<td>2002</td>
<td>Silver</td>
<td>80</td>
</tr>
</tbody>
</table>

Gray et al.

Data Cube: A Relational Aggregation Operator Generalizing Group-By, Cross-Tab, and Sub-Totals
Discussion 1: Multi-dimensional aggregation

- How useful is multi-dimensional aggregation?

- Besides the data warehousing applications mentioned in the paper, can you think of any other applications?
Problems with SQL GROUP BY

Common analysis operations awkward

- Histograms
- Roll-Up/Drill-Down
- Cross-Tabulations
### Histograms

Categorize models by type

<table>
<thead>
<tr>
<th>Model</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Golf</td>
<td></td>
</tr>
<tr>
<td>Rabbit</td>
<td>Hatchback</td>
</tr>
<tr>
<td>:</td>
<td></td>
</tr>
<tr>
<td>Jetta</td>
<td></td>
</tr>
<tr>
<td>Passat</td>
<td>Sedan</td>
</tr>
<tr>
<td>:</td>
<td></td>
</tr>
</tbody>
</table>
# Histograms

- **Type** | **Year** | **Colour** | **Num_Sold**
- Hatchback | 2001 | Black | 40
- Hatchback | 2001 | Silver | 65
- Sedan | 2001 | Black | 70
- Sedan | 2001 | Silver | 50

**Problem:** not directly supported by SQL

---

*Gray et al.*

Data Cube: A Relational Aggregation Operator Generalizing Group-By, Cross-Tab, and Sub-Totals
### Histograms

<table>
<thead>
<tr>
<th>Type</th>
<th>Year</th>
<th>Colour</th>
<th>Num_Sold</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hatchback</td>
<td>2001</td>
<td>Black</td>
<td>40</td>
</tr>
<tr>
<td>Hatchback</td>
<td>2001</td>
<td>Silver</td>
<td>65</td>
</tr>
<tr>
<td>Sedan</td>
<td>2001</td>
<td>Black</td>
<td>70</td>
</tr>
<tr>
<td>Sedan</td>
<td>2001</td>
<td>Silver</td>
<td>50</td>
</tr>
</tbody>
</table>

**Problem:** not directly supported by SQL
Roll-Up/Drill-Down

View data at decreasing/increasing levels of granularity

Model

Model, Year

Model, Year, Colour

Gray et al.

Data Cube: A Relational Aggregation Operator Generalizing Group-By, Cross-Tab, and Sub-Totals
## Roll-Up/Drill-Down (1)

<table>
<thead>
<tr>
<th>Model</th>
<th>Year</th>
<th>Colour</th>
<th>Num_MCY</th>
<th>Num_MY</th>
<th>Num_M</th>
</tr>
</thead>
<tbody>
<tr>
<td>Golf</td>
<td>2001</td>
<td>Black</td>
<td>40</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Silver</td>
<td>65</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2002</td>
<td>Black</td>
<td></td>
<td>70</td>
<td>105</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Silver</td>
<td></td>
<td>100</td>
<td>170</td>
</tr>
</tbody>
</table>

Problem: not relational

Gray et al.

Data Cube: A Relational Aggregation Operator Generalizing Group-By, Cross-Tab, and Sub-Totals
### Roll-Up/Drill-Down (1)

<table>
<thead>
<tr>
<th>Model</th>
<th>Year</th>
<th>Colour</th>
<th>Num_MCY</th>
<th>Num_MY</th>
<th>Num_M</th>
</tr>
</thead>
<tbody>
<tr>
<td>Golf</td>
<td>2001</td>
<td>Black</td>
<td>40</td>
<td></td>
<td>105</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Silver</td>
<td>65</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2002</td>
<td>Black</td>
<td></td>
<td>70</td>
<td>170</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Silver</td>
<td></td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>

Problem: not relational

---

Gray et al.

Data Cube: A Relational Aggregation Operator Generalizing Group-By, Cross-Tab, and Sub-Totals
### Roll-Up/Drill-Down (2)

Chris Date (1996)

<table>
<thead>
<tr>
<th>Model</th>
<th>Year</th>
<th>Colour</th>
<th>Num_Sold</th>
<th>Num_MY</th>
<th>Num_M</th>
</tr>
</thead>
<tbody>
<tr>
<td>Golf</td>
<td>2001</td>
<td>Black</td>
<td>40</td>
<td>105</td>
<td>275</td>
</tr>
<tr>
<td>Golf</td>
<td>2001</td>
<td>Silver</td>
<td>65</td>
<td>105</td>
<td>275</td>
</tr>
<tr>
<td>Golf</td>
<td>2002</td>
<td>Black</td>
<td>70</td>
<td>170</td>
<td>275</td>
</tr>
<tr>
<td>Golf</td>
<td>2002</td>
<td>Silver</td>
<td>100</td>
<td>170</td>
<td>275</td>
</tr>
</tbody>
</table>

Problem: 2D columns
## Roll-Up/Drill-Down (2)

Chris Date (1996)

<table>
<thead>
<tr>
<th>Model</th>
<th>Year</th>
<th>Colour</th>
<th>Num_Sold</th>
<th>Num_MY</th>
<th>Num_M</th>
</tr>
</thead>
<tbody>
<tr>
<td>Golf</td>
<td>2001</td>
<td>Black</td>
<td>40</td>
<td>105</td>
<td>275</td>
</tr>
<tr>
<td>Golf</td>
<td>2001</td>
<td>Silver</td>
<td>65</td>
<td>105</td>
<td>275</td>
</tr>
<tr>
<td>Golf</td>
<td>2002</td>
<td>Black</td>
<td>70</td>
<td>170</td>
<td>275</td>
</tr>
<tr>
<td>Golf</td>
<td>2002</td>
<td>Silver</td>
<td>100</td>
<td>170</td>
<td>275</td>
</tr>
</tbody>
</table>

Problem: **2D columns**
Cross-Tab

Roll-Up

Model

Model, Year

Model, Year, Colour

Cross Tab

Model

Model, Year

Model, Year, Colour

Model, Colour

Model, Year, Colour

Model, Year, Colour

Data Analysis

Analysis Operations

The Data Cube

Cross-Tab

Data Cube: A Relational Aggregation Operator Generalizing Group-By, Cross-Tab, and Sub-Totals
## Pivot Table (Excel)

<table>
<thead>
<tr>
<th>Model</th>
<th>Black</th>
<th>Silver</th>
<th>2001 Total</th>
<th>2002 Black</th>
<th>2002 Silver</th>
<th>2002 Total</th>
<th>Grand Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Golf</td>
<td>40</td>
<td>65</td>
<td>105</td>
<td>70</td>
<td>100</td>
<td>170</td>
<td>275</td>
</tr>
<tr>
<td>Jetta</td>
<td>70</td>
<td>50</td>
<td>120</td>
<td>60</td>
<td>80</td>
<td>140</td>
<td>260</td>
</tr>
<tr>
<td>Total</td>
<td>110</td>
<td>115</td>
<td>225</td>
<td>130</td>
<td>180</td>
<td>310</td>
<td>535</td>
</tr>
</tbody>
</table>

Problem: $N \times M$ values

Gray et al.

Data Cube: A Relational Aggregation Operator Generalizing Group-By, Cross-Tab, and Sub-Totals
### Pivot Table (Excel)

#### Model, Colour, Year

<table>
<thead>
<tr>
<th></th>
<th>2001</th>
<th>2002</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Black</td>
<td>Silver</td>
<td>Total</td>
</tr>
<tr>
<td>Golf</td>
<td>40</td>
<td>65</td>
<td>105</td>
</tr>
<tr>
<td>Jetta</td>
<td>70</td>
<td>50</td>
<td>120</td>
</tr>
<tr>
<td>Total</td>
<td>110</td>
<td>115</td>
<td>225</td>
</tr>
</tbody>
</table>

Problem: $N \times M$ values

---

Gray et al.

Data Cube: A Relational Aggregation Operator Generalizing Group-By, Cross-Tab, and Sub-Totals
### Pivot Table (Excel)

#### Model, Year

<table>
<thead>
<tr>
<th>Model</th>
<th>Black</th>
<th>Silver</th>
<th>2001 Total</th>
<th>2002</th>
<th>Grand Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Golf</td>
<td>40</td>
<td>65</td>
<td>105</td>
<td>70</td>
<td>170</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Jetta</td>
<td>70</td>
<td>50</td>
<td>120</td>
<td>60</td>
<td>140</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>80</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>110</td>
<td>115</td>
<td>225</td>
<td>130</td>
<td>310</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>2002</th>
<th>2002 Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>535</td>
<td>260</td>
</tr>
</tbody>
</table>

Problem: $N \times M$ values
## Pivot Table (Excel)

<table>
<thead>
<tr>
<th></th>
<th>2001</th>
<th></th>
<th>2002</th>
<th></th>
<th></th>
<th>2002</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td></td>
<td>Black</td>
<td>Silver</td>
<td>2001</td>
<td>Black</td>
<td>Silver</td>
<td>2002</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td></td>
<td></td>
<td>Total</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Golf</td>
<td>40</td>
<td>65</td>
<td>105</td>
<td>70</td>
<td>100</td>
<td>170</td>
<td>275</td>
</tr>
<tr>
<td>Jetta</td>
<td>70</td>
<td>50</td>
<td>120</td>
<td>60</td>
<td>80</td>
<td>140</td>
<td>260</td>
</tr>
<tr>
<td>Total</td>
<td>110</td>
<td>115</td>
<td>225</td>
<td>130</td>
<td>180</td>
<td>310</td>
<td>535</td>
</tr>
</tbody>
</table>

**Problem:** $N \times M$ values
## Pivot Table (Excel)

### Model

<table>
<thead>
<tr>
<th></th>
<th>2001</th>
<th></th>
<th>2002</th>
<th></th>
<th>2002</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Black</td>
<td>Silver</td>
<td>Total</td>
<td>Black</td>
<td>Silver</td>
<td>Total</td>
</tr>
<tr>
<td>Golf</td>
<td>40</td>
<td>65</td>
<td>105</td>
<td>70</td>
<td>100</td>
<td>170</td>
</tr>
<tr>
<td>Jetta</td>
<td>70</td>
<td>50</td>
<td>120</td>
<td>60</td>
<td>80</td>
<td>140</td>
</tr>
<tr>
<td>Total</td>
<td>110</td>
<td>115</td>
<td>225</td>
<td>130</td>
<td>180</td>
<td>310</td>
</tr>
</tbody>
</table>

**Grand Total**:
- 2001: 315
- 2002: 500
- Total: 815

**Problem**: $N \times M$ values
### Pivot Table (Excel)

<table>
<thead>
<tr>
<th>Model</th>
<th>Black</th>
<th>Silver</th>
<th>2001 Total</th>
<th>2002 Black</th>
<th>2002 Silver</th>
<th>2002 Total</th>
<th>Grand Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Golf</td>
<td>40</td>
<td>65</td>
<td>105</td>
<td>70</td>
<td>100</td>
<td>170</td>
<td>275</td>
</tr>
<tr>
<td>Jetta</td>
<td>70</td>
<td>50</td>
<td>120</td>
<td>60</td>
<td>80</td>
<td>140</td>
<td>260</td>
</tr>
<tr>
<td>Total</td>
<td>110</td>
<td>115</td>
<td>225</td>
<td>130</td>
<td>180</td>
<td>310</td>
<td>535</td>
</tr>
</tbody>
</table>

**Problem:** $N \times M$ values

---

Gray et al.

*Data Cube: A Relational Aggregation Operator Generalizing Group-By, Cross-Tab, and Sub-Totals*
## Pivot Table (Excel)

<table>
<thead>
<tr>
<th>Model</th>
<th>Black</th>
<th>Silver</th>
<th>2001 Total</th>
<th>2002 Black</th>
<th>2002 Silver</th>
<th>2002 Total</th>
<th>Grand Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Golf</td>
<td>40</td>
<td>65</td>
<td>105</td>
<td>70</td>
<td>100</td>
<td>170</td>
<td>275</td>
</tr>
<tr>
<td>Jetta</td>
<td>70</td>
<td>50</td>
<td>120</td>
<td>60</td>
<td>80</td>
<td>140</td>
<td>260</td>
</tr>
<tr>
<td>Total</td>
<td>110</td>
<td>115</td>
<td>225</td>
<td>130</td>
<td>180</td>
<td>310</td>
<td>535</td>
</tr>
</tbody>
</table>

**Problem:** $N \times M$ values

---

*Gray et al.*

Data Cube: A Relational Aggregation Operator Generalizing Group-By, Cross-Tab, and Sub-Totals
### Pivot Table (Excel)

<table>
<thead>
<tr>
<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></td>
<td>Black</td>
<td>Silver</td>
<td>Total</td>
<td>Black</td>
<td>Silver</td>
<td>Total</td>
<td>Total</td>
<td>Total</td>
<td>Total</td>
</tr>
<tr>
<td>Golf</td>
<td>40</td>
<td>65</td>
<td>105</td>
<td>70</td>
<td>100</td>
<td>170</td>
<td>225</td>
<td>130</td>
<td>310</td>
</tr>
<tr>
<td>Jetta</td>
<td>70</td>
<td>50</td>
<td>120</td>
<td>60</td>
<td>80</td>
<td>140</td>
<td>110</td>
<td>115</td>
<td>225</td>
</tr>
<tr>
<td>Total</td>
<td>110</td>
<td>115</td>
<td>225</td>
<td>130</td>
<td>180</td>
<td>310</td>
<td>535</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Problem:** $N \times M$ values

Gray et al.

Data Cube: A Relational Aggregation Operator Generalizing Group-By, Cross-Tab, and Sub-Totals
Add special `ALL` value

<table>
<thead>
<tr>
<th>Model</th>
<th>Year</th>
<th>Colour</th>
<th>Num_Sold</th>
</tr>
</thead>
<tbody>
<tr>
<td>Golf</td>
<td>2001</td>
<td>Black</td>
<td>40</td>
</tr>
<tr>
<td>Golf</td>
<td>2001</td>
<td>Silver</td>
<td>65</td>
</tr>
<tr>
<td>Golf</td>
<td>2001</td>
<td>ALL</td>
<td>105</td>
</tr>
<tr>
<td>Golf</td>
<td>2002</td>
<td>Black</td>
<td>70</td>
</tr>
<tr>
<td>Golf</td>
<td>2002</td>
<td>Silver</td>
<td>100</td>
</tr>
<tr>
<td>Golf</td>
<td>2002</td>
<td>ALL</td>
<td>175</td>
</tr>
<tr>
<td>Golf</td>
<td>ALL</td>
<td>ALL</td>
<td>275</td>
</tr>
</tbody>
</table>

**SQL:** `UNION of GROUP-BYS on ALL`
Solution: Overload Column Values

Add special **ALL** value

<table>
<thead>
<tr>
<th>Model</th>
<th>Year</th>
<th>Colour</th>
<th>Num_Sold</th>
</tr>
</thead>
<tbody>
<tr>
<td>Golf</td>
<td>2001</td>
<td>Black</td>
<td>40</td>
</tr>
<tr>
<td>Golf</td>
<td>2001</td>
<td>Silver</td>
<td>65</td>
</tr>
<tr>
<td>Golf</td>
<td>2001</td>
<td><strong>ALL</strong></td>
<td>105</td>
</tr>
<tr>
<td>Golf</td>
<td>2002</td>
<td>Black</td>
<td>70</td>
</tr>
<tr>
<td>Golf</td>
<td>2002</td>
<td>Silver</td>
<td>100</td>
</tr>
<tr>
<td>Golf</td>
<td>2002</td>
<td><strong>ALL</strong></td>
<td>175</td>
</tr>
<tr>
<td>Golf</td>
<td><strong>ALL</strong></td>
<td><strong>ALL</strong></td>
<td>275</td>
</tr>
</tbody>
</table>

**SQL:** `UNION of GROUP-BYS on ALL`
Solution: Overload Column Values

Add special \texttt{ALL} value

<table>
<thead>
<tr>
<th>Model</th>
<th>Year</th>
<th>Colour</th>
<th>Num_Sold</th>
</tr>
</thead>
<tbody>
<tr>
<td>Golf</td>
<td>2001</td>
<td>Black</td>
<td>40</td>
</tr>
<tr>
<td>Golf</td>
<td>2001</td>
<td>Silver</td>
<td>65</td>
</tr>
<tr>
<td>Golf</td>
<td>2001</td>
<td>\texttt{ALL}</td>
<td>105</td>
</tr>
<tr>
<td>Golf</td>
<td>2002</td>
<td>Black</td>
<td>70</td>
</tr>
<tr>
<td>Golf</td>
<td>2002</td>
<td>Silver</td>
<td>100</td>
</tr>
<tr>
<td>Golf</td>
<td>2002</td>
<td>\texttt{ALL}</td>
<td>175</td>
</tr>
<tr>
<td>Golf</td>
<td>\texttt{ALL}</td>
<td>\texttt{ALL}</td>
<td>275</td>
</tr>
</tbody>
</table>

\textbf{SQL:} \texttt{UNION of GROUP-BYS on ALL}
The **ALL** Dummy Value

**Solution: Overload Column Values**

Add special **ALL** value

<table>
<thead>
<tr>
<th>Model</th>
<th>Year</th>
<th>Colour</th>
<th>Num_Sold</th>
</tr>
</thead>
<tbody>
<tr>
<td>Golf</td>
<td>2001</td>
<td>Black</td>
<td>40</td>
</tr>
<tr>
<td>Golf</td>
<td>2001</td>
<td>Silver</td>
<td>65</td>
</tr>
<tr>
<td>Golf</td>
<td>2001</td>
<td><strong>ALL</strong></td>
<td>105</td>
</tr>
<tr>
<td>Golf</td>
<td>2002</td>
<td>Black</td>
<td>70</td>
</tr>
<tr>
<td>Golf</td>
<td>2002</td>
<td>Silver</td>
<td>100</td>
</tr>
<tr>
<td>Golf</td>
<td>2002</td>
<td><strong>ALL</strong></td>
<td>175</td>
</tr>
<tr>
<td>Golf</td>
<td><strong>ALL</strong></td>
<td><strong>ALL</strong></td>
<td>275</td>
</tr>
</tbody>
</table>

**SQL:** `UNION of GROUP-BYS on ALL`
### Solution: Overload Column Values

Add special `ALL` value

<table>
<thead>
<tr>
<th>Model</th>
<th>Year</th>
<th>Colour</th>
<th>Num_Sold</th>
</tr>
</thead>
<tbody>
<tr>
<td>Golf</td>
<td>2001</td>
<td>Black</td>
<td>40</td>
</tr>
<tr>
<td>Golf</td>
<td>2001</td>
<td>Silver</td>
<td>65</td>
</tr>
<tr>
<td>Golf</td>
<td>2001</td>
<td>ALL</td>
<td>105</td>
</tr>
<tr>
<td>Golf</td>
<td>2002</td>
<td>Black</td>
<td>70</td>
</tr>
<tr>
<td>Golf</td>
<td>2002</td>
<td>Silver</td>
<td>100</td>
</tr>
<tr>
<td>Golf</td>
<td>2002</td>
<td>ALL</td>
<td>175</td>
</tr>
<tr>
<td>Golf</td>
<td>ALL</td>
<td>ALL</td>
<td>275</td>
</tr>
</tbody>
</table>

**SQL:** `UNION of GROUP-BYS on ALL`
Is **ALL** Enough?

- $2D$ UNIONS for Roll-Up/Drill-Down
- $2D$ UNIONS for Cross-Tab
- Awkward
- Inefficient

Gray et al.

Data Cube: A Relational Aggregation Operator Generalizing Group-By, Cross-Tab, and Sub-Totals
Is \textbf{ALL} Enough?

- $2D$ \textsc{Unions} for Roll-Up/Drill-Down
- $2D$ \textsc{Unions} for Cross-Tab
- Awkward
- Inefficient
Is ALL Enough?

- 2D UNIONS for Roll-Up/Drill-Down

- 2D UNIONS for Cross-Tab

- Awkward

- Inefficient

Gray et al.
Data Cube: A Relational Aggregation Operator Generalizing Group-By, Cross-Tab, and Sub-Totals

The ALL Dummy Value
Is ALL Enough?

- $2D$ UNIONS for Roll-Up/Drill-Down
- $2^D$ UNIONS for Cross-Tab
- Awkward
- Inefficient
Is **ALL** Enough?

- **2D UNIONs** for Roll-Up/Drill-Down
- **2D UNIONs** for Cross-Tab
- Awkward
- Inefficient
Discussion 2: The CUBE Operator

- How hard did you find it to understand the CUBE operator?

- As a query writer, would you feel comfortable using it? Or, would you rather use the "solutions" described in the previous slides?
The **CUBE** Operator

**The CUBE Operator**

```sql
SELECT <attribute-list> <aggregated-measurements>
FROM <table>
WHERE <selection-cons>
GROUP BY CUBE <attribute-list>
```

Generates equivalent result as **UNION** of **GROUP BY** on all subsets of `<attribute-list>` using **ALL**.

Gray et al.

Data Cube: A Relational Aggregation Operator Generalizing Group-By, Cross-Tab, and Sub-Totals
SELECT <attribute-list> <aggregated-measurements>
FROM <table>
WHERE <selection-cons>
GROUP BY CUBE <attribute-list>

Generates equivalent result as UNION of GROUP BY on all subsets of <attribute-list> using ALL.

Gray et al.
Data Cube: A Relational Aggregation Operator Generalizing Group-By, Cross-Tab, and Sub-Totals
The CUBE Operator

Data Cube: A Relational Aggregation Operator Generalizing Group-By, Cross-Tab, and Sub-Totals
Discussion 3: Overloading NULL

The authors intended to reduce special-case handling code by “overloading” NULL with the “ALL” value.

- Do you think overloading NULL is a good idea?

- Does this actually reduce special-case handling code, or is special-case handling code still required, but even more difficult to deal with because of the overloading?