

Group By and Calculaton Examples

Aggregating Rows versus Aggregating Columns

Spatial Database Management

GEP 664 / GEP 380

Review of GROUP BY and Calculated Fields

Frank Donnelly

Dept of EECS, Lehman College CUNY

Spring 2019

Examples use annual passenger data for the three international airports in the NYC metro area

```
CREATE TABLE airports (  
  air_id varchar(3) PRIMARY KEY,  
  state varchar(2),  
  y2015 numeric (5,1),  
  y2014 numeric (5,1));
```

```
INSERT INTO airports  
VALUES ('EWR', 'NJ', 37.5, 35.6),  
('LGA', 'NY', 28.4, 27.0),  
('JFK', 'NY', 56.8, 53.2);
```

1: GROUP BY - Correct

Sums across rows for columns that can be grouped

Summarize passenger travel by state

```
SELECT state,  
SUM(y2015) AS y2015t, SUM(y2014) AS y2014t  
FROM airports  
GROUP BY state;
```

Table: original

air_id	state	y2015	y2014
EWR	NJ	37.5	35.6
JFK	NY	56.8	53.2
LGA	NY	28.4	27.0

Table: result

state	y2015t	y2014t
NJ	37.5	35.6
NY	85.2	80.2

1: GROUP BY - Wrong

Includes columns that cannot be grouped

Summarize passenger travel by state

```
SELECT air_id, state,  
SUM(y2015) AS y2015t, SUM(y2014) AS y2014t  
FROM airports  
GROUP BY air_id, state;
```

Table: original

air_id	state	y2015	y2014
EWR	NJ	37.5	35.6
JFK	NY	56.8	53.2
LGA	NY	28.4	27.0

Table: result

air_id	state	y2015t	y2014t
EWR	NJ	37.5	35.6
JFK	NY	56.8	53.2
LGA	NY	28.4	27.0

2: Calculation - Correct

Sums across columns

Calculate total passenger travel for airports for all years

```
SELECT air_id, state, (y2015+y2014) AS total
FROM airports;
```

Table: original

air_id	state	y2015	y2014
EWR	NJ	37.5	35.6
JFK	NY	56.8	53.2
LGA	NY	28.4	27.0

Table: result

air_id	state	total
EWR	NJ	73.1
JFK	NY	110.0
LGA	NY	55.4



2: Calculation - Wrong

Unnecessary use of aggregate function and group by

Calculate total passenger travel for airports for all years

```
SELECT air_id, state, SUM(y2015+y2014) AS total
FROM airports
GROUP BY airport, state;
```

Table: original

air_id	state	y2015	y2014
EWR	NJ	37.5	35.6
JFK	NY	56.8	53.2
LGA	NY	28.4	27.0

Table: result

air_id	state	total
EWR	NJ	73.1
JFK	NY	110.0
LGA	NY	55.4



3: GROUP BY and Calculate - Correct

Sums across rows for columns that can be grouped

Summarize travel by state and calculate total for all years

```
SELECT state,
SUM(y2015) AS y2015t, SUM(y2014) AS y2014t,
(SUM(y2015)+SUM(y2014)) AS total
FROM airports
GROUP BY state;
```

Table: original

air_id	state	y2015	y2014
EWR	NJ	37.5	35.6
JFK	NY	56.8	53.2
LGA	NY	28.4	27.0

Table: result

state	y2015t	y2014t	total
NJ	37.5	35.6	73.1
NY	85.2	80.2	165.4



3: GROUP BY and Calculate - Wrong

Includes columns that cannot be grouped, incorrect use of SUM

Summarize travel by state and calculate total for all years

```
SELECT air_id, state,
SUM(y2015) AS y2015t, SUM(y2014) AS y2014t,
SUM(y2015+y2014) AS total
FROM airports
GROUP BY air_id, state;
```

Table: original

air_id	state	y2015	y2014
EWR	NJ	37.5	35.6
JFK	NY	56.8	53.2
LGA	NY	28.4	27.0

Table: result

air_id	state	y2015t	y2014t	total
EWR	NJ	37.5	35.6	73.1
JFK	NY	56.8	53.2	110.0
LGA	NY	28.4	27.0	55.4

