Grouping and Having

INTRODUCTION TO SQL SERVER

John MacKintosh
Instructor

A simple SELECT

```sql
SELECT
    SUM(demand_loss_mw) AS lost_demand
FROM grid;
```

```
+-------------+
| lost_demand |
+-------------+
| 177888      |
+-------------+
```

Grouping error

Can we break this down by adding an additional column?

```sql
SELECT
  SUM(demand_loss_mw) AS lost_demand,
  description
FROM grid;
```

```
Msg 8120, Level 16, State 1, Line 1
Column 'grid.description' is invalid in the select list because it is not contained in
either an aggregate function or the GROUP BY clause.
```

```sql
SELECT
  SUM(demand_loss_mw) AS lost_demand,
  description
FROM grid
GROUP BY description;
```

```
+----------- -+-----------------------------------------------------+
| lost_demand | description                                         |
+-------------+-----------------------------------------------------+
| NULL        | Actual Physical Attack                              |
| NULL        | Cold Weather Event                                  |
| NULL        | Cyber Event with Potential to Cause Impact          |
| 40          | Distribution Interruption                           |
| 2           | Distribution System Interruption                    |
| NULL        | Earthquake                                          |
| NULL        | Electrical Fault at Generator                       |
| 338         | Electrical System Islanding                         |
| 24514       | Electrical System Separation Islanding              |
| 15          | Electrical System Separation Islanding Severe Weather |
+-------------+-----------------------------------------------------+
```

```sql
SELECT
  SUM(demand_loss_mw) AS lost_demand,
  description
FROM grid
WHERE
  description LIKE '%storm'
  AND demand_loss_mw IS NOT NULL
GROUP BY description;
```

```
+----------- -+---------------------------------------------+
| lost_demand | description                                 |
|-------------+---------------------------------------------|
| 996         | Ice Storm                                   |
| 420         | Load Shed Severe Weather   Lightning Storm |
| 332         | Major Storm                                 |
| 3           | Severe Weather  Thunderstorm                |
| 413         | Severe Weather  Wind Storm                  |
| 4171        | Severe Weather  Winter Storm                |
| 1352        | Winter Storm                                |
+------------+----------------------------------------------+
```

HAVING

- Can use aggregate functions in `SELECT`

- Filter data using `WHERE`

- Split data into groups using `GROUP BY`

- What if we want to sum values based on groups?

- ... and then filter on those sums?

```sql
SELECT
    SUM(demand_loss_mw) AS lost_demand,
    description
FROM grid
WHERE
    description LIKE '%storm'
    AND demand_loss_mw IS NOT NULL
GROUP BY description;
```

```
+----------- -+-------------------------------------------+
| lost_demand | description                               |
+------------+-------------------------------------------+
| 996        | Ice Storm                                 |
| 420        | Load Shed Severe Weather   Lightning Storm |
| 332        | Major Storm                               |
| 3          | Severe Weather  Thunderstorm              |
| 413        | Severe Weather  Wind Storm                |
| 4171       | Severe Weather  Winter Storm              |
| 1352       | Winter Storm                              |
+------------+-------------------------------------------+
```

```sql
SELECT
  SUM(demand_loss_mw) AS lost_demand,
  description
FROM grid
WHERE
  description LIKE '%storm'
  AND demand_loss_mw IS NOT NULL
GROUP BY description
HAVING SUM(demand_loss_mw) > 1000;
```

```
+------------ -+-------------------------------------------+
| lost_demand | description                               |
|-------------+-------------------------------------------|
| 4171        | Severe Weather  Winter Storm              |
| 1352        | Winter Storm                              |
+-------------+-------------------------------------------+
```

Summary

- `GROUP BY` splits the data up into combinations of one or more values

- `WHERE` filters on row values

- `HAVING` appears after the `GROUP BY` clause and filters on groups or aggregates