

```
-- Recap:
-- We learned that when we GROUP BY, there will ultimately be a single
-- row reported for each group, and that this has implications for what
-- can go in the SELECT clause: Everything we SELECT must either be
-- an attribute according to which we grouped (so it will have one value per group),
-- or an attribute that we have aggregated (so that SQL can compute one value
-- per group).
```

```
-- And to demonstrate that once again with an example, here is the Student table:
```

```
csc343h-diane> select *
csc343h-diane> from student;
  sid | firstname | surname | campus | email | cgpa
-----+-----+-----+-----+-----+-----
 99132 | Avery     | Marchmount | StG    | avery@cs | 3.13
 98000 | William   | Fairgrieve | StG    | will@cs  | 4.00
 99999 | Afsaneh   | Ali        | UTSC   | aali@cs  | 2.98
   157 | Leilani   | Lakemeyer  | UTM    | lani@cs  | 3.42
 11111 | Homer     | Simpson    | StG    | doh@gmail | 0.40
(5 rows)
```

```
-- and here is what happens when we break the rule summarized above:
```

```
csc343h-diane> select * from student group by campus;
ERROR: column "student.sid" must appear in the GROUP BY clause or be used in an aggregate function
LINE 1: select * from student group by campus;
          ^
```

```
-- Here we have the same query, but with a SELECT clause that respects the rule:
```

```
csc343h-diane> select campus, count(sid)
csc343h-diane> from student
csc343h-diane> group by campus;
 campus | count
-----+-----
   UTSC |      1
    StG |      3
    UTM |      1
(3 rows)
```

```
-- Here's the new fact: PostgreSQL will overlook this violation of the rule:
```

```
csc343h-diane> select *
csc343h-diane> from student
csc343h-diane> group by sid;
  sid | firstname | surname | campus | email | cgpa
-----+-----+-----+-----+-----+-----
 11111 | Homer     | Simpson | StG    | doh@gmail | 0.40
 98000 | William   | Fairgrieve | StG    | will@cs  | 4.00
 99132 | Avery     | Marchmount | StG    | avery@cs  | 3.13
 99999 | Afsaneh   | Ali        | UTSC   | aali@cs  | 2.98
   157 | Leilani   | Lakemeyer | UTM    | lani@cs  | 3.42
(5 rows)
```

```
-- We shouldn't be able to include surname without aggregating it, because
-- it is not in the GROUP BY. But it's allowed because there cannot be more
-- than one surname per sid. Or more than one firstname, campus, email or cgpa.
-- Why? Because sid is a key in this table, so no sid can occur in more than one
-- row, so none of these other attributes can have more than one value.
```

```
-- PostgreSQL is willing to make this inference and let us off the hook.
-- I wouldn't count on all other DBMSs doing the same.
```