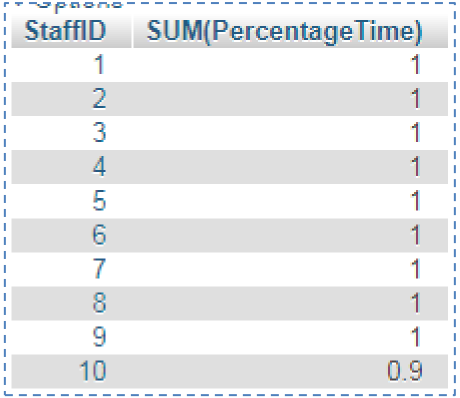
GROUP BY clause

GROUP BY

- The GROUP BY clause allows you to *collapse* multiple records with a common value into groups.
- Use COUNT, MAX, MIN, SUM, AVG as aggregate functions applied to each group separately

- Find the total time allocation of each staff.

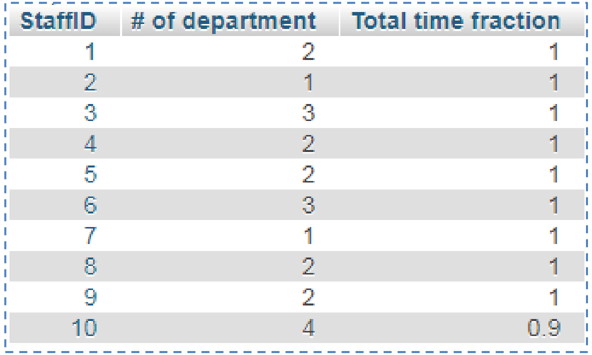
```
SELECT StaffID, SUM(PercentageTime)
FROM workallocation
GROUP BY StaffID;
```



StaffID	SUM(PercentageTime)
1	1
2	1
3	1
4	1
5	1
6	1
7	1
8	1
9	1
10	0.9

- Find the number of departments each staff is working in. Also, show the total time allocation for each staff.

```
SELECT StaffID, COUNT(DepartmentID) AS '# of department',
           SUM(PercentageTime) AS 'Total time fraction'
FROM workallocation
GROUP BY StaffID;
```



StaffID	# of department	Total time fraction
1	2	1
2	1	1
3	3	1
4	2	1
5	2	1
6	3	1
7	1	1
8	2	1
9	2	1
10	4	0.9

GROUP BY

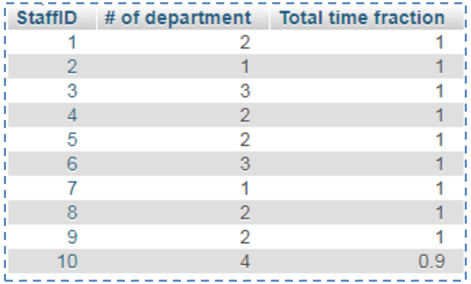
How does it work?

- Make groups based on the column value: StaffID in this case
- If the rows are not ordered, no problem: MySQL still group them
- Apply, any conditions that come in **HAVING clause**!
- Then, apply any aggregate function!

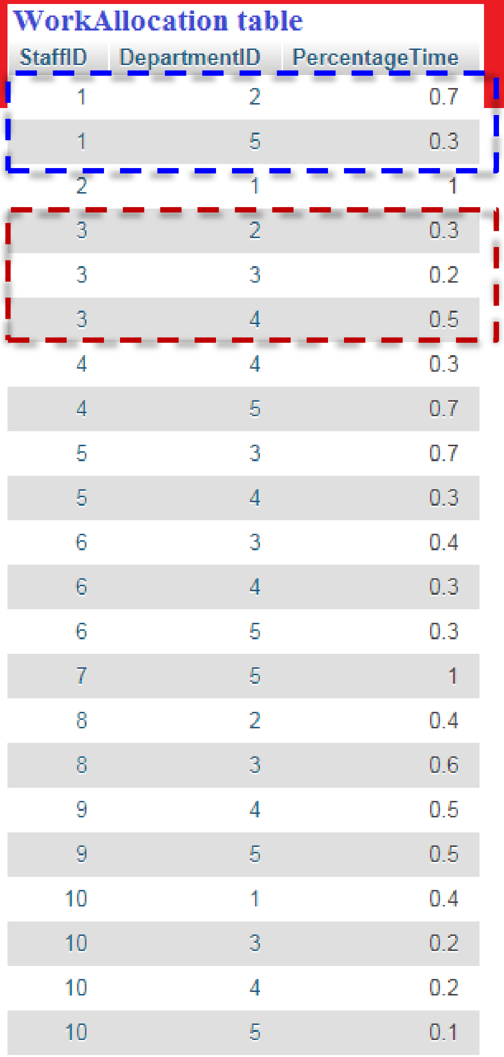
- For example, for this SELECT statement:

```
SELECT StaffID, COUNT(DepartmentID) AS '# of department',
           SUM(PercentageTime) AS 'Total time fraction'
FROM workallocation
GROUP BY StaffID;
```

- First, make groups based on staffID
- No, HAVING clause, so nothing happen
- Finally, apply COUNT() and SUM() functions!
- So, the output looks like:



StaffID	# of department	Total time fraction
1	2	1
2	1	1
3	3	1
4	2	1
5	2	1
6	3	1
7	1	1
8	2	1
9	2	1
10	4	0.9



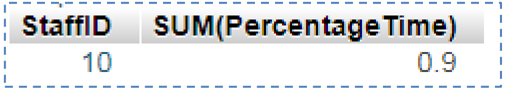
WorkAllocation table

StaffID	DepartmentID	PercentageTime
1	2	0.7
1	5	0.3
2	1	1
3	2	0.3
3	3	0.2
3	4	0.5
4	4	0.3
4	5	0.7
5	3	0.7
5	4	0.3
6	3	0.4
6	4	0.3
6	5	0.3
7	5	1
8	2	0.4
8	3	0.6
9	4	0.5
9	5	0.5
10	1	0.4
10	3	0.2
10	4	0.2
10	5	0.1

GROUP BY with HAVING

- Find the staff who does not work fulltime.

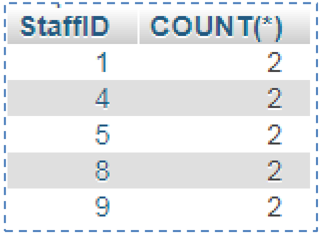
```
SELECT StaffID, SUM(PercentageTime)
FROM workallocation
GROUP BY StaffID
HAVING SUM(PercentageTime) < 1;
```



StaffID	SUM(PercentageTime)
10	0.9

- Find the staff who works in exactly two departments

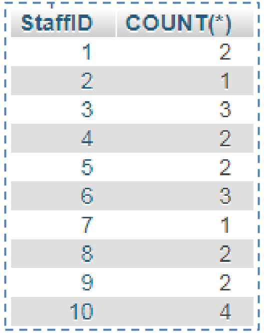
```
SELECT StaffID, COUNT(*)
FROM workallocation
GROUP BY StaffID
HAVING COUNT(*) = 2;
```



StaffID	COUNT(*)
1	2
4	2
5	2
8	2
9	2

  - See without the HAVING condition above

```
SELECT StaffID, COUNT(*)
FROM workallocation
GROUP BY StaffID;
```

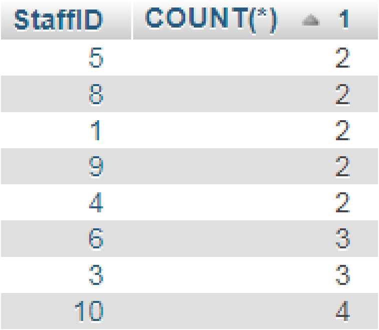


StaffID	COUNT(*)
1	2
2	1
3	3
4	2
5	2
6	3
7	1
8	2
9	2
10	4

GROUP BY & ORDER BY

- Find the staff who works in at least 2 departments. Show your output in ascending order of number of departments each staff work in.

```
SELECT StaffID, COUNT(*)
FROM workallocation
GROUP BY StaffID
HAVING COUNT(*) > 1
ORDER BY COUNT(*) ASC;
```



StaffID	COUNT(*) ▲ 1
5	2
8	2
1	2
9	2
4	2
6	3
3	3
10	4

  - What if you are asked to show staff name?
    - See solution later when joining multiple tables in a query!

Thank you