---- SQL Subqueries

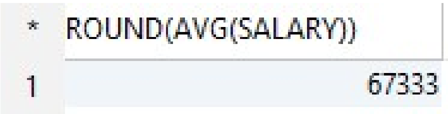
Prof. Dr. Ingo Claßen

Mittwoch, 26. Mai 2021 16:56

Average Salary as Condition

Mittwoch, 2. Juni 2021 10:38

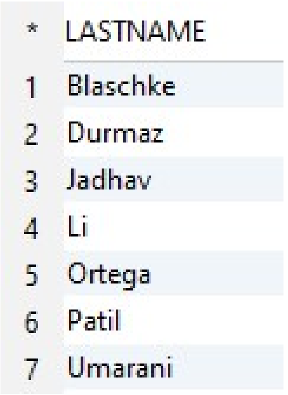
```
select
  round(avg(salary))
from employee;
```



*	ROUND(AVG(SALARY))
1	67333

problem: manual calculation of average value and insertion in second query

```
select lastname
from employee
where salary> 67333
order by lastname;
```



*	LASTNAME
1	Blaschke
2	Durmaz
3	Jadhav
4	Li
5	Ortega
6	Patil
7	Umarani

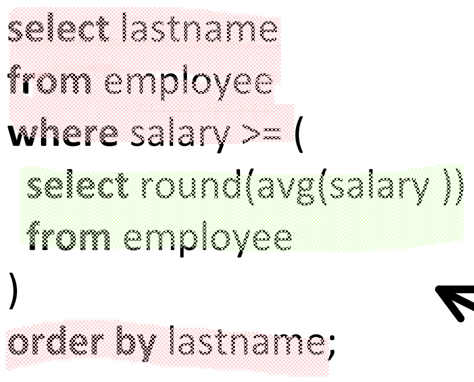
notice: query needs values on two aggreagtion levels:

1. average salary (aggregated level)
2. output of lastname (detail level)

Nesting of Queries

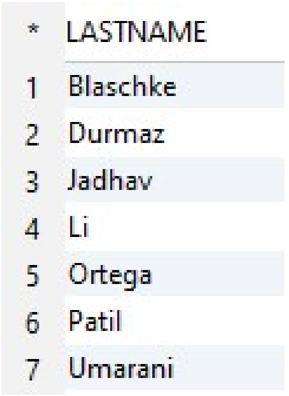
Mittwoch, 2. Juni 2021 10:38

main query



```
select lastname
from employee
where salary >= (
 select round(avg(salary ))
 from employee
)
order by lastname;
```

subquery



*	LASTNAME
1	Blaschke
2	Durmaz
3	Jadhav
4	Li
5	Ortega
6	Patil
7	Umarani

aggregated value in subquery is regarded on detail level in main query

Problem on Conceptual Level

Mittwoch, 2. Juni 2021 10:38

```
select lastname
from employee
where salary >= (
  select round(avg(salary ))
  from employee
)
order by lastname;
```

salary is a number

result of a query
is a table

Comparison "**>=**" requires left-hand side and right-hand side to be of compatible type

Is satisfied in this case, because the subquery delivers a table with one row and one column

Data Cells

Mittwoch, 2. Juni 2021 09:48

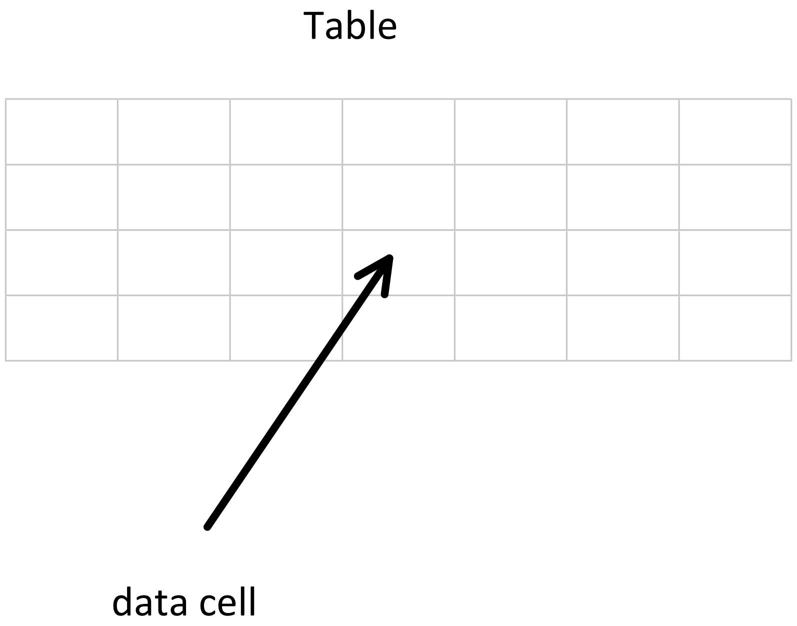


table with one row and one column

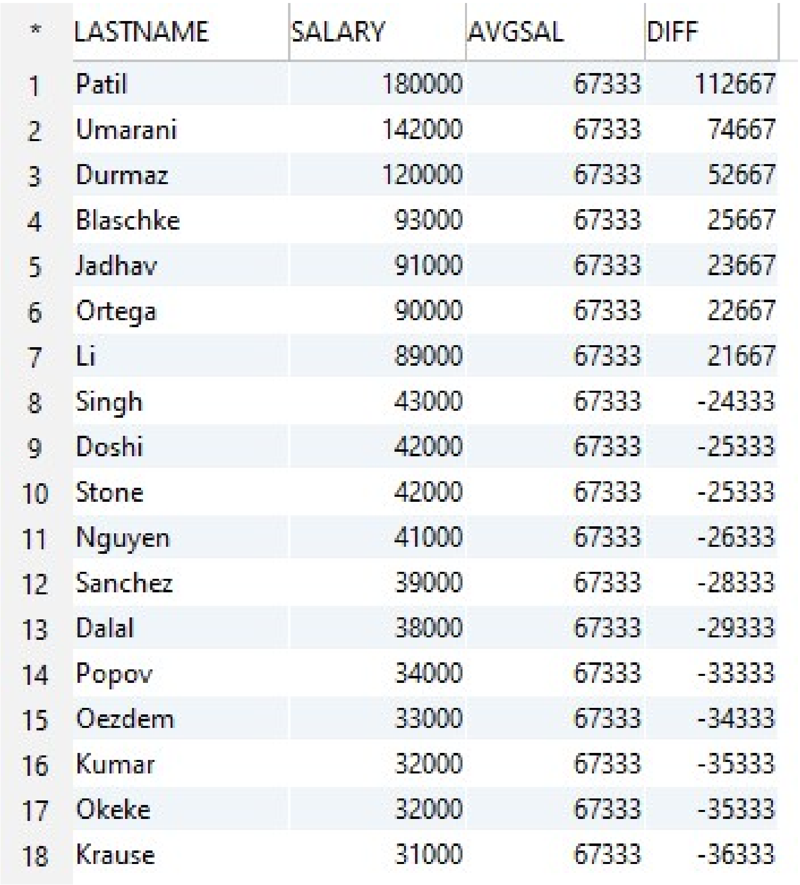
$$\boxed{7} \longrightarrow 7$$

table can be regarded as value

Calculations with Subqueries

Mittwoch, 2. Juni 2021 10:38

```
select
  lastname,
  salary,
  (select round(avg(salary)) from employee) as avgsal,
  salary- (select round(avg(salary)) from employee) as diff
from employee
order by salary- (select round(avg(salary)) from employee) desc;
```



*	LASTNAME	SALARY	AVGSAL	DIFF
1	Patil	180000	67333	112667
2	Umarani	142000	67333	74667
3	Durmaz	120000	67333	52667
4	Blaschke	93000	67333	25667
5	Jadhav	91000	67333	23667
6	Ortega	90000	67333	22667
7	Li	89000	67333	21667
8	Singh	43000	67333	-24333
9	Doshi	42000	67333	-25333
10	Stone	42000	67333	-25333
11	Nguyen	41000	67333	-26333
12	Sanchez	39000	67333	-28333
13	Dalal	38000	67333	-29333
14	Popov	34000	67333	-33333
15	Oezdem	33000	67333	-34333
16	Kumar	32000	67333	-35333
17	Okeke	32000	67333	-35333
18	Krause	31000	67333	-36333

subqueries enable calculations with values on different aggregation levels

notice: lot of code duplication

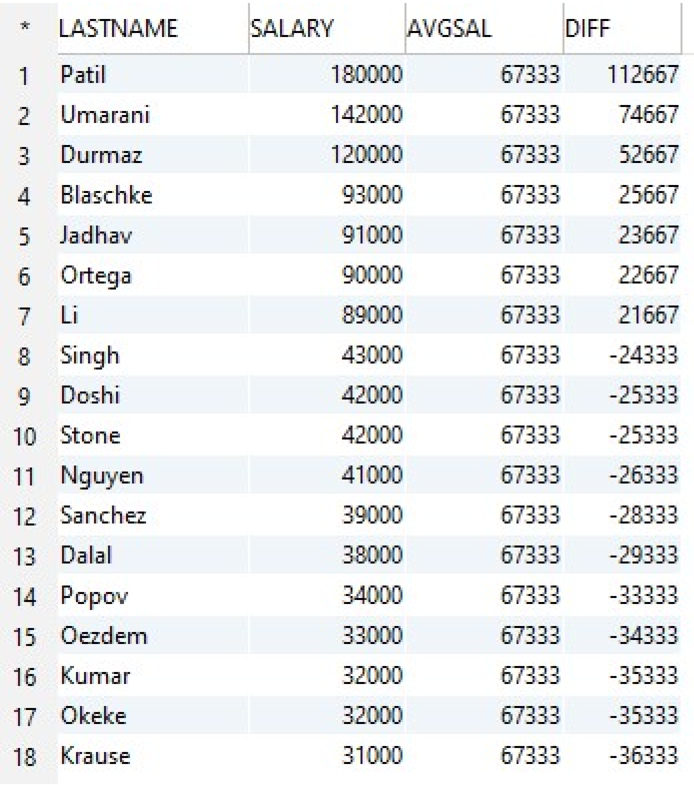
Common Table Expression (CTE)

Mittwoch, 2. Juni 2021 10:38

Remove code duplication by extracting common subqueries and define them with keyword **with**

```
with
  asal as (select round(avg(salary)) as avgsal from employee)
select
  lastname,
  salary,
  avgsal,
  salary-avgsal as diff
from employee cross join asal
order by salary-avgsal desc;
```

add column avgsal to employee table
 -> **cross join** with table asal



*	LASTNAME	SALARY	AVGSAL	DIFF
1	Patil	180000	67333	112667
2	Umarani	142000	67333	74667
3	Durmaz	120000	67333	52667
4	Blaschke	93000	67333	25667
5	Jadhav	91000	67333	23667
6	Ortega	90000	67333	22667
7	Li	89000	67333	21667
8	Singh	43000	67333	-24333
9	Doshi	42000	67333	-25333
10	Stone	42000	67333	-25333
11	Nguyen	41000	67333	-26333
12	Sanchez	39000	67333	-28333
13	Dalal	38000	67333	-29333
14	Popov	34000	67333	-33333
15	Oezdem	33000	67333	-34333
16	Kumar	32000	67333	-35333
17	Okeke	32000	67333	-35333
18	Krause	31000	67333	-36333

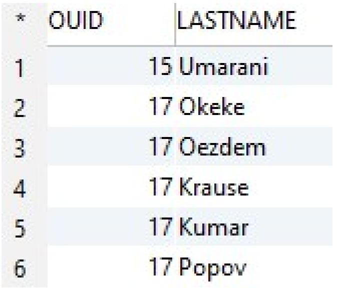
Subquery with List of Values

Mittwoch, 2. Juni 2021 10:38

all employees that work in org units whose head is Umarani (109)

```
select ouid, lastname
from employee
where ouid in (
  select ouid
  from orgunit
  where head=109
);
```

Subquery delivers:
15
17



*	OUID	LASTNAME
1	15	Umarani
2	17	Okeke
3	17	Oezdem
4	17	Krause
5	17	Kumar
6	17	Popov

after keyword **in** there must be a list of values, i.e. subquery with one column and many rows

important: join according to column head

- Same query with use of lastname 'Umarani'
- Requires a join inside subquery with table employee

```
select ouid, lastname
from employee
where ouid in (
  select ou.ouid
  from orgunit ou
     join employee e on e.eid=ou.head
  where lastname='Umarani'
);
```

Average Salary per Orgunit

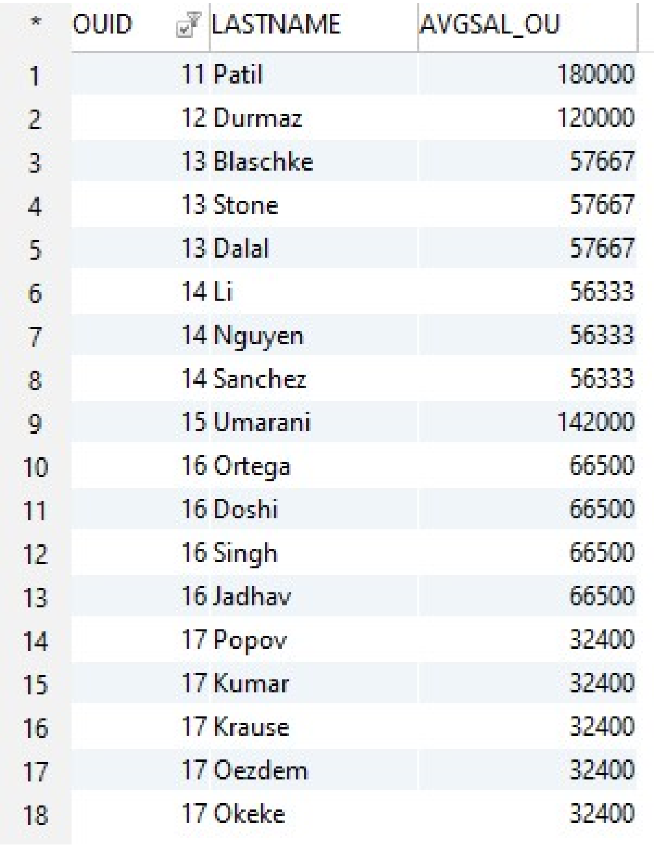
Mittwoch, 2. Juni 2021 16:22

```
select
  ouid,
  lastname,
   (select round(avg(salary))
    from employee asal
    where asal.ouid=e.ouid) as avgsal_ou
from employee e
order by ouid, eid;
```

reference from subquery to main query, called correlated subquery

Correlated subqueries can lead to long runtimes:

- rewrite with subquery in *from part* of main query
- or use window function



*	OUID	LASTNAME	AVGSAL_OU
1	11	Patil	180000
2	12	Durmaz	120000
3	13	Blaschke	57667
4	13	Stone	57667
5	13	Dalal	57667
6	14	Li	56333
7	14	Nguyen	56333
8	14	Sanchez	56333
9	15	Umarani	142000
10	16	Ortega	66500
11	16	Doshi	66500
12	16	Singh	66500
13	16	Jadhav	66500
14	17	Popov	32400
15	17	Kumar	32400
16	17	Krause	32400
17	17	Oezdem	32400
18	17	Okeke	32400

different values of average salaries for each orgunit

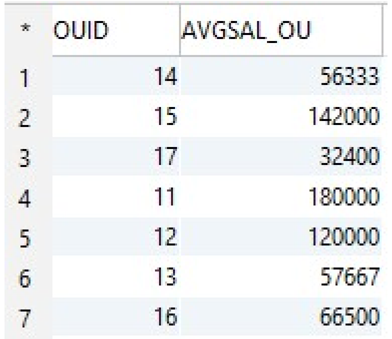
Subquery in From Part

Mittwoch, 2. Juni 2021 10:38

compare with previous slide

```
select
  e.ouid,
  lastname,
  salary,
  avgsal_ou,
  salary- avgsal_ou as diff
from employee e
     join
        (select ouid, round(avg(salary)) as avgsal_ou
         from employee
         group by ouid) asal
     on e.ouid=asal.ouid
order by e.ouid, e.eid;
```

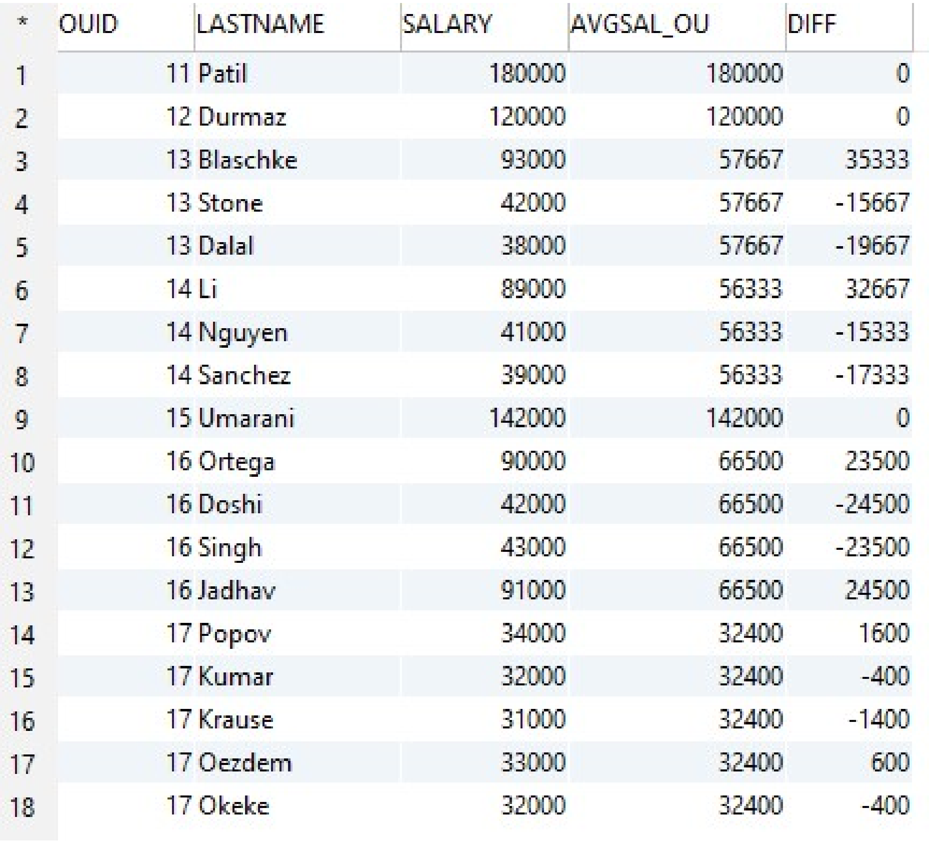
result of subquery



*	OUID	AVGSAL_OU
1	14	56333
2	15	142000
3	17	32400
4	11	180000
5	12	120000
6	13	57667
7	16	66500

subquery can deliver a real table

result of main query

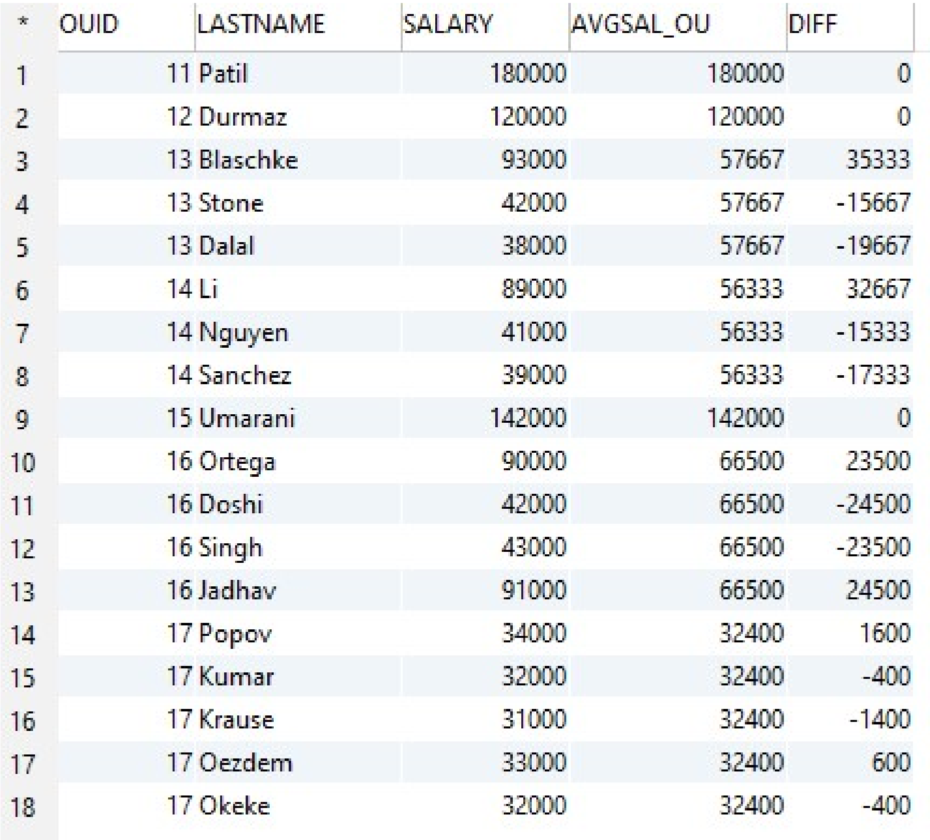


*	OUID	LASTNAME	SALARY	AVGSAL_OU	DIFF
1	11	Patil	180000	180000	0
2	12	Durmaz	120000	120000	0
3	13	Blaschke	93000	57667	35333
4	13	Stone	42000	57667	-15667
5	13	Dalal	38000	57667	-19667
6	14	Li	89000	56333	32667
7	14	Nguyen	41000	56333	-15333
8	14	Sanchez	39000	56333	-17333
9	15	Umarani	142000	142000	0
10	16	Ortega	90000	66500	23500
11	16	Doshi	42000	66500	-24500
12	16	Singh	43000	66500	-23500
13	16	Jadhav	91000	66500	24500
14	17	Popov	34000	32400	1600
15	17	Kumar	32000	32400	-400
16	17	Krause	31000	32400	-1400
17	17	Oezdem	33000	32400	600
18	17	Okeke	32000	32400	-400

With CTE

Mittwoch, 2. Juni 2021 10:38

```
with
  asal as (
    select ouid, round(avg(salary)) as avgsal_ou
    from employee
    group by ouid)
select
  e.ouid,
  lastname,
  salary,
  avgsal_ou,
  salary-avgsal_ou as diff
from employee e
      join asal on asal.ouid=e.ouid;
```



*	OUID	LASTNAME	SALARY	AVGSAL_OU	DIFF
1	11	Patil	180000	180000	0
2	12	Durmaz	120000	120000	0
3	13	Blaschke	93000	57667	35333
4	13	Stone	42000	57667	-15667
5	13	Dalal	38000	57667	-19667
6	14	Li	89000	56333	32667
7	14	Nguyen	41000	56333	-15333
8	14	Sanchez	39000	56333	-17333
9	15	Umarani	142000	142000	0
10	16	Ortega	90000	66500	23500
11	16	Doshi	42000	66500	-24500
12	16	Singh	43000	66500	-23500
13	16	Jadhav	91000	66500	24500
14	17	Popov	34000	32400	1600
15	17	Kumar	32000	32400	-400
16	17	Krause	31000	32400	-1400
17	17	Oezdem	33000	32400	600
18	17	Okeke	32000	32400	-400

Positioning of Subqueries

Prof. Dr. Ingo Claßen

Mittwoch, 2. Juni 2021 09:53

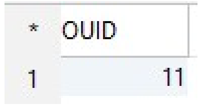
select	(select …) - one row, one column
from	(select …) - no restrictions
where	(select …) - one row, one column - many rows, one column (**in** predicate)
group by	Not allowed
having	(select …) - one row, one column
order by	(select …) - one row, one column

All

Donnerstag, 3. Juni 2021 09:11

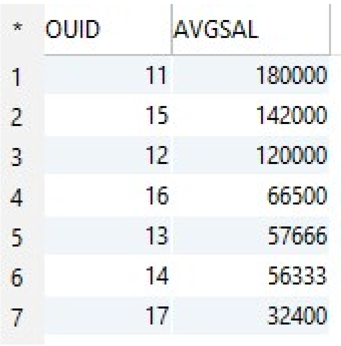
orgunits with largest average salary

```
with
  asal as (
    select ouid, avg(salary) as avgsal
    from employee
    group by ouid
  )
select ouid
from asal
where avgsal >= all (select avgsal from asal);
```



*	OUID
1	11

average salaries



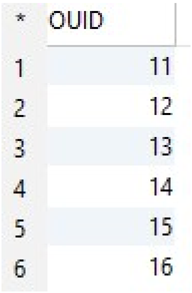
*	OUID	AVGSAL
1	11	180000
2	15	142000
3	12	120000
4	16	66500
5	13	57666
6	14	56333
7	17	32400

Some

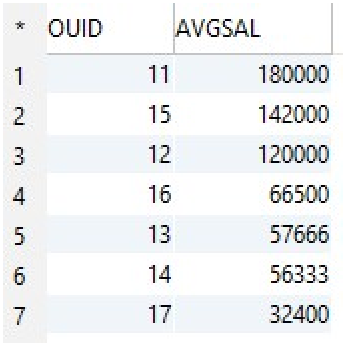
Donnerstag, 3. Juni 2021 09:19

orgunit with non-smallest average salary

```
with
  asal as (
    select ouid, avg(salary) as avgsal
    from employee
    group by ouid
  )
select ouid
from asal
where avgsal > some (select avgsal from asal)
order by ouid;
```



*	OUID
1	11
2	12
3	13
4	14
5	15
6	16



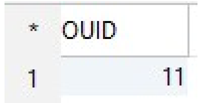
*	OUID	AVGSAL
1	11	180000
2	15	142000
3	12	120000
4	16	66500
5	13	57666
6	14	56333
7	17	32400

Exists

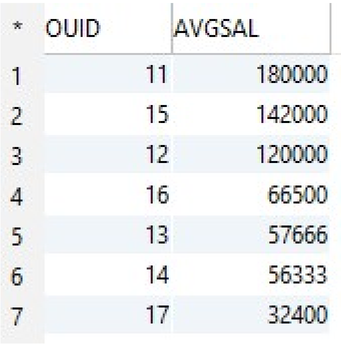
Donnerstag, 3. Juni 2021 09:11

orgunits with largest average salary

```
with
  asal as (
    select ouid, avg(salary) as avgsal
    from employee
    group by ouid
  )
select ouid
from asal asal1
where not exists (
   select avgsal
   from asal asal2
   where asal2.avgsal > asal1.avgsal
);
```



*	OUID
1	11



*	OUID	AVGSAL
1	11	180000
2	15	142000
3	12	120000
4	16	66500
5	13	57666
6	14	56333
7	17	32400