CSCD 327: RELATIONAL DATABASE SYSTEMS

WORKING WITH DIFFERENT DATA TYPES

Instructor: Dr. Dan Li

Working with Strings

- Traverse a string
- Count the occurrences of a character
- Remove unwanted characters
- Order by parts of a string
- ...

What does the following code do?

```
SELECT substr( e.ename, iter.pos, 1 ) AS C
FROM
      (SELECT ename
      FROM emp
      WHERE ename = 'KING') e,
       (SELECT id AS pos
      FROM t10
      ) iter
WHERE iter.pos <= length( e.ename )
```

What does this code return?

```
SELECT (length( '10,CLARK,MANAGER' ) - length( replace(
'10,CLARK,MANAGER', ',', '' ) ) ) / length( ',' )  AS cnt
FROM t1
```

Removing Unwanted Characters
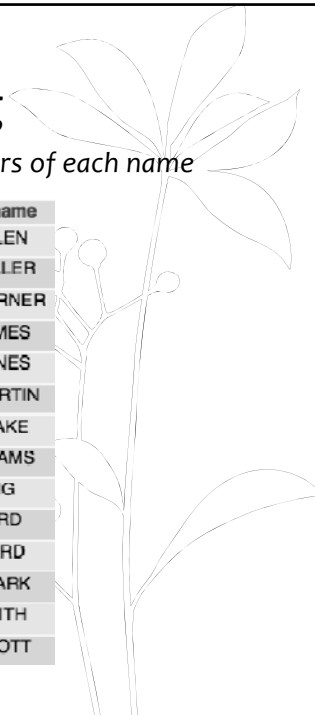
• Remove vowels from ENAME

ename	stripped
SMITH	SMTH
ALLEN	LLN
WARD	WRD
JONES	JNS
MARTIN	MRTN
BLAKE	BLK
CLARK	CLRK
SCOTT	SCTT
KING	KNG
TURNER	TRNR
ADAMS	DMS
JAMES	JMS
FORD	FRD
MILLER	MLLR

Ordering by Parts of a String

• *Order ENAME based on the last two characters of each name*

ename
ALLEN
MILLER
TURNER
JAMES
JONES
MARTIN
BLAKE
ADAMS
KING
FORD
WARD
CLARK
SMITH
SCOTT

Creating a Delimited List from Table Rows

- Return table rows as values in a delimited list, perhaps delimited by commas, rather than in vertical columns as they normally appear.

- E.g. convert a result set from this:

```
DEPTNO EMPS
------ ----------
    10 CLARK
    10 KING
    10 MILLER
    20 SMITH
    20 ADAMS
    20 FORD
    20 SCOTT
    20 JONES
    30 ALLEN
    30 BLAKE
    30 MARTIN
    30 JAMES
    30 TURNER
    30 WARD
```

to this:

```
DEPTNO EMPS
------ ----------------------------------
    10 CLARK,KING,MILLER
    20 SMITH,JONES,SCOTT,ADAMS,FORD
    30 ALLEN,WARD,MARTIN,BLAKE,TURNER,JAMES
```

What does this code do?

SUBSTRING_INDEX(*str*,*delim*,*count*) Returns the substring from string *str* before *count* occurrences of the delimiter *delim*. If *count* is positive, everything to the left of the final delimiter (counting from the left) is returned. If *count* is negative, everything to the right of the final delimiter (counting from the right) is returned.

```
select substring_index(substring_index(y.ip,'.',1),'.',-1) a,
       substring_index(substring_index(y.ip,'.',2),'.',-1) b,
       substring_index(substring_index(y.ip,'.',3),'.',-1) c,
       substring_index(substring_index(y.ip,'.',4),'.',-1) d
  from (select '92.111.8.2' as ip from t1) y
```

More String Functions

- CONCAT_WS(): return concatenate with separator
- FORMAT(): return a number formatted to specified number of decimal places
- INSERT(): insert a substring at the specified position up to the specified number of characters
- LEFT(): return the leftmost number of characters as specified
- LOCATE(): return the position of the first occurrence of substring
- REVERSE(): reverse the characters in a string
- TRIM(): remove leading and trailing spaces
- LOWER()/UPPER()
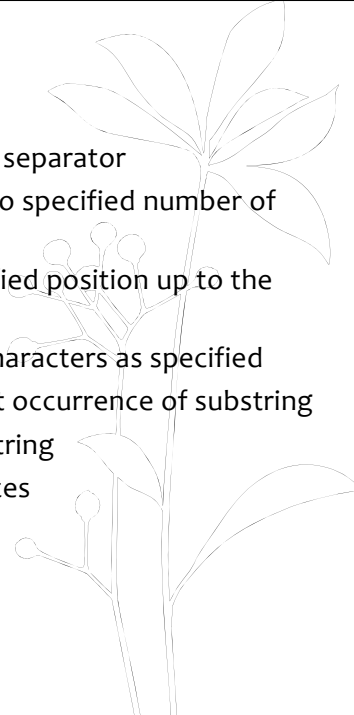
Working with Numbers

- AVG(), COUNT(), SUM(), MIN(), MAX()
- Generate a running total
- Calculate a mode
- Calculate a median
- Compute averages without high and low values
- ...

Generating a Running Total

| empno	ename	sal	running_total
7369	SMITH	800	800
7499	ALLEN	1600	2400
7521	WARD	1250	3650
7566	JONES	2975	6625
7654	MARTIN	1250	7875
7698	BLAKE	2850	10725
7782	CLARK	2450	13175
7788	SCOTT	3000	16175
7839	KING	5000	21175
7844	TURNER	1500	22675
7876	ADAMS	1100	23775
7900	JAMES	950	24725
7902	FORD	3000	27725
7934	MILLER	1300	29025

Calculating a Mode

- Find the mode (the *mode* in mathematics is the element that appears most frequently for a given set of data) of the values in a column.
- E.g. find mode of the salaries in DEPTNO 20. Based on the following salaries, the mode is 3000.

```
select sal
   from emp
  where deptno = 20
  order by sal

    SAL
----------
    800
   1100
   2975
   3000
   3000
```
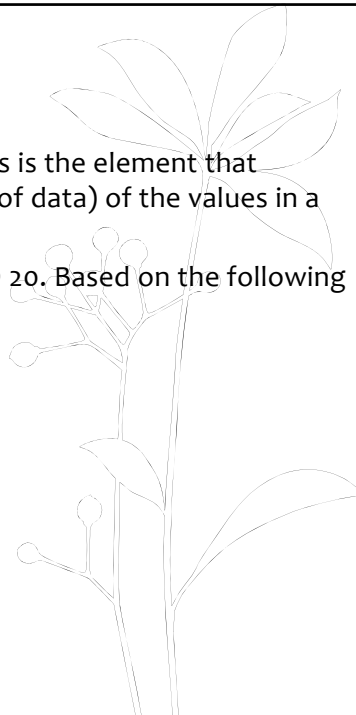
Calculating a Median

- *Find the median of the salaries in DEPTNO 20.*

```
1 select median(sal)
2   from emp
3  where deptno=20

1 select percentile_cont(0.5)
2          within group(order by sal)
3   from emp
4  where deptno=20
```

- *Here shows Oracle solution. MySQL doesn't support median() and percentile_cont().*

Determining the Percentage of a Total

- Determine the percentage that values in a specific column represent against a total
- E.g. determine the percentage that DEPTNO 10 salaries contribute to the total salary.

Including Null Values in Aggregation

• *Determine the average commission for employees in DEPTNO 30.*

• The **COALESCE** function in SQL returns the first non-NULL expression among its arguments.

• It is the same as the following **CASE** statement:

• **SELECT CASE**
  WHEN "expression 1 is not NULL" THEN "expression 1"
  WHEN "expression 2 is not NULL" THEN "expression 2"
  ...
  [ELSE "NULL"]
  END
  FROM "table_name"

Computing Averages Without High and Low Values

• *Compute the average salary of all employees excluding the highest and lowest salaries.*

Working with Date

• Add and subtract days, months, years
• Determine the number of days
• Determine the number of months/years
• Determining the first and last day of a month
• …

Adding and Subtracting Days, Months, and Years

• Add or subtract some number of days, months, or years from a date.
• E.g., using the HIREDATE for employee CLARK you want to return six different dates: five days before and after CLARK was hired, five months before and after CLARK was hired, and, finally, five years before and after CLARK was hired.

ename	hiredate	hd_minus_5D	hd_plus_5D	hd_minus_5M	hd_plus_5M	hd_minus_5Y	hd_plus_5Y
CLARK	1981-06-09	1981-06-04	1981-06-14	1981-01-09	1981-11-09	1976-06-09	1986-06-09

INTERVAL keyword specifying the unit of time to add or subtract

Determining the Number of Days Between Two Dates

- *Find the difference in days between the HIREDATEs of employee ALLEN and employee WARD.*

ward_hd	allen_hd	datediff(allen_hd,ward_hd)
1981-02-22	1981-02-20	-2

Determining the Number of Months or Years Between Two Dates

- *Find the number of months between the first and last employees hired, and express that value as some number of years.*

```
select min(hiredate) as min_hd,
       max(hiredate) as max_hd
   from emp

MIN_HD       MAX_HD
-----------  -----------
17-DEC-1980  12-JAN-1983
```

| mnth	mnth/12
25	2.0833

Searching on Specific Units of Time

- *Find all employees hired in February or December, as well as employees hired on a Tuesday.*

Determining the Number of Days in a Year

- The number of days in the current year is the difference between the first day of the next year and the first day of the current year (in days).
  - Find the first day of the current year.
  - Add one year to that date (to get the first day of the next year).
  - Subtract the current year from the result of Step 2.

```
1 select datediff((curr_year + interval 1 year),curr_year)
2   from (
3 select adddate(current_date,-dayofyear(current_date)+1) curr_year
4   from t1
5       ) x
```
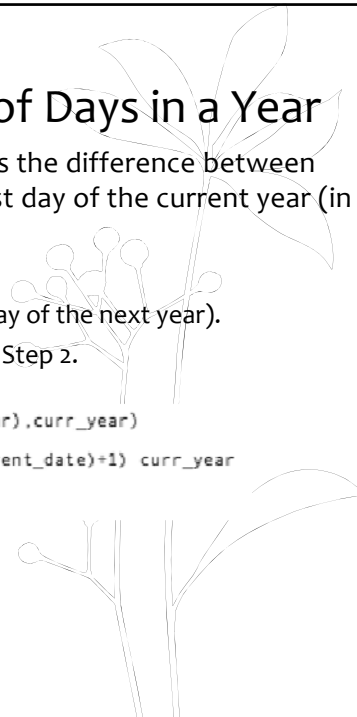
Determining the First and Last Day of Current Month

- To find the first day of the month, use the DAY function. The DAY function conveniently returns the day of the month for the date passed. If you subtract the value returned by DAY(CURRENT_DATE) from the current date, you get the last day of the prior month; add one day to get the first day of the current month. To find the last day of the current month, simply use the LAST_DAY function.

```
select date_add(current_date,
        interval -day(current_date)+1 day) firstday,
        last_day(current_date) lastday
    from t1
```