SUMMARY QUERIES

SUMMARY QUERIES

- There are many situations in which we need to retrieve a value which is a function of all values in a column
- For example we may need to compute average or sum of all values in a column
- We may use aggregate function for this purpose

SUMMARY QUERIES

- Aggregate functions in SQL
  - SUM()
  - AVG()
  - MIN(), MAX()
  - COUNT()
- All these functions returns a single value as a result of a query instead of a table with several rows
- But this is still a table

SUMMARY QUERIES

- **Example:** Find the total of targets and sales of all employees?

```
SELECT SUM(EMP_TARGET), SUM(EMP_SALES)
FROM EMPLOYEES
```

- **Result:** A table with single row and two columns

SUM(EMP_TARGET)	SUM(EMP_SALES)
$3,624,500.00	$4,867,363.00

SUMMARY QUERIES

- What is the average, minimum and maximum of sales made by all offices?

```
SELECT AVG(OFF_SALES), MIN(OFF_SALES),
            MAX(OFF_SALES)
FROM OFFICES
```

- Find the date of the earliest order in the database

```
SELECT MIN(ORDER_DATE)
FROM ORDERS
```

COUNT

- COUNT is used to find the number of elements in a column obtained as a result of a query
- **Example:** How many employees are working in the company?

```
SELECT COUNT(EMP_ID)
FROM EMPLOYEES
```

- **Result:** 10

COUNT

- It does not matter which column you use in COUNT as long as the column does not contain NULL values
- **Example:** How many employees are working in the company?

```
SELECT COUNT(FL_NAME), COUNT(AGE),
            COUNT(TITLE)
FROM EMPLOYEES
```

- **Result:**10, 10, 10

COUNT

- However, if there are NULLS in a column, they will not be counted

```
SELECT COUNT(OFFICE)
FROM EMPLOYEES
```

- **Result:** 9
- NULL values are ignored by other aggregate functions as well

COUNT

- If we want to count number of rows of a table, the best way is to use
  - primary key or
  - *

```
SELECT COUNT(*)
FROM EMPLOYEES
```

- Result: 10

CALCULATED COLUMNS

- Aggregate functions can be applied to calculated columns
- **Example:** Suppose that the performance of an employee is determined by the percentage of sales he/she makes with respect to his/her target.

- What is the best sales performance achieved by an employee?

```
SELECT MAX(100 * (EMP_SALES/EMP_TARGET))
FROM EMPLOYEES
```

CALCULATED COLUMNS

- Compute the average number of years an employee works in the company

CALCULATED COLUMNS

- Compute the average number of years an employee works in the company

```
SELECT AVG(DATEDIFF(year,HIRE_DATE,GETDATE()))
FROM EMPLOYEES
```

COMPLEX QUERIES

- Summary function can be used in complicated queries involving row selections and joins
- **Example:**

```
SELECT AVG(ORD_PRICE)
FROM ORDERS
WHERE CUST_NUM = 108
```

- This give the average amount of orders made by customer with id number 108

COMPLEX QUERIES

- Order of execution for a query involving aggregate function
  - First cartesian product and row selection (also join)
  - Then aggregate function and column selection
- Keep in mind this execution order while writing such queries

COMPLEX QUERIES

- How many offices made sales over their targets?

COMPLEX QUERIES

- How many offices made sales over their targets?

```
SELECT COUNT(OFFICE_ID)
FROM OFFICES
WHERE OFF_SALES > OFF_TARGET
```

COMPLEX QUERIES

- What is the total price of the orders taken by Charles Bass?

COMPLEX QUERIES

- What is the total price of the orders taken by Charles Bass?

```
SELECT SUM(ORD_PRICE)
FROM ORDERS, EMPLOYEES
WHERE FL_NAME = 'Charles Bass'
        AND REP_NUM = EMP_ID
```

DISTINCT VALUES

- In some problems you may want to apply the summary function to distinct values returned by a query
- Use DISTINCT keyword **within** the aggregate function to do this
- **Example:** Find the number of supervisors in the company

```
SELECT COUNT(DISTINCT SUPERVISOR)
FROM EMPLOYEES
```

- **Result:** 4

DISTINCT VALUES

- DINSTINCT keyword should be used inside the aggregate function
- Otherwise, it will not have an effect
- **Example:**

  SELECT DISTINCT COUNT(SUPERVISOR)

  FROM EMPLOYEES
- **Result:** 9