SQL for Data Analysis 101

EP01 - ER Diagram

> **ER Diagram** ย่อมาจาก Entity-Relationship Diagram เป็นเหมือน แผนที่ ที่บอกเราว่าข้อมูลใน database มีอะไรบ้าง และความสัมพันธ์ของ แต่ละ table เป็นอย่างไร

- รูปแบบความสัมพันธ์ที่เราใช้ใน database มีสามแบบหลักๆคือ

  1. One-to-One
  2. One-to-Many (Hit!)
  3. Many-to-Many

EP02 - Basic Select

```
-- select all columns from a table
SELECT * FROM customers;

-- we can use select to do basic calculation
SELECT 10 + 20;
SELECT 10 * 2;
SELECT 10 / 2;
SELECT 10 - 5;

-- use ROUND to determine decimal places
SELECT  name, milliseconds
        ROUND(milliseconds/6000.0,2) AS minute,
        ROUND(bytes/(1024*1024.0),4) AS MBs
FROM tracks;
```

EP03 - Select Columns

```
-- select columns from table
SELECT
  firstname,
  lastname,
  country
FROM customers;
```

EP04 - Create New Columns

```
-- create/ transform columns
SELECT
  firstname,
  lastname,
  firstname || ' ' || lastname AS fullname,
  LOWER(firstname) || '@company.com' AS email
FROM customers;
```

EP05 - CASE When

🌵 `CASE WHEN` ใช้ในการเขียนเงื่อนไข สร้างคอลัมน์ใหม่ เราจะเขียน `CASE` `WHEN` ใน `SELECT` clause

เราสามารถเขียนได้มากกว่าหนึ่งเงื่อนไข ได้ผลลัพธ์เหมือนกับฟังก์ชัน `=IFS()` ของ Google Sheets/ Excel

```
-- CASE is like IF-ELSE in Google Sheets
SELECT
  company,
  CASE
    WHEN company IS NOT NULL THEN 'Corporate'
    ELSE 'End Customer'
  END AS segment
FROM customers;

-- ตัวอย่างการเขียน CASE WHEN เพื่อสร้างคอลัมน์ region
SELECT
  country,
  CASE
    WHEN country IN ('Canada', 'USA') THEN 'America'
    WHEN country IN ('Belgium', 'France', 'Italy') THEN 'Europe'
    ELSE 'Other'
  END AS region
FROM customers
```

```
-- Method 1
SELECT  firstname,
    company,
        COALESCE(company,"No Information") AS clean_company,
        CASE
          WHEN company IS NULL   THEN 'B2C'
            WHEN company IS NOT NULL THEN 'B2B'
        END
FROM customers

-- Method 2
SELECT  firstname,
    company,
        COALESCE(company,"No Information") AS clean_company,
```

```
        CASE
          WHEN company IS NULL   THEN 'B2C'
            ELSE 'B2B'
        END
FROM customers


=IF(ISBLANK(company),"B2C","B2B") ** in google sheet
```

EP06 - Date Time

```
-- STRFTIME is a specific function for SQLite
SELECT
    STRFTIME("%Y",invoicedate) AS YEAR,
    STRFTIME("%m",invoicedate) AS MONTH,
    STRFTIME("%d",invoicedate) AS DAY,
    billingaddress
FROM invoices
WHERE year='2009' AND month='01' AND day='01'


SELECT  invoicedate,
        STRFTIME("%Y-%m",invoicedate) AS monthid,
        billingaddress
FROM invoices
WHERE monthid  ='2011-05'
```

EP07 - Select from Multiple Tables

```
-- review join syntax
SELECT
  A.column,
  B.column
FROM tableA AS A
JOIN tableB AS B
  ON A.primary_key = B.foreign_key;
```

```
SELECT * FROM artists
JOIN albums
ON artists.artistid = albums.artistidd
WHERE artists.artistid =100;

-- Join Syntax
SELECT
    A.artistid,
    B.name,
    C.title
FROM artists AS A
```

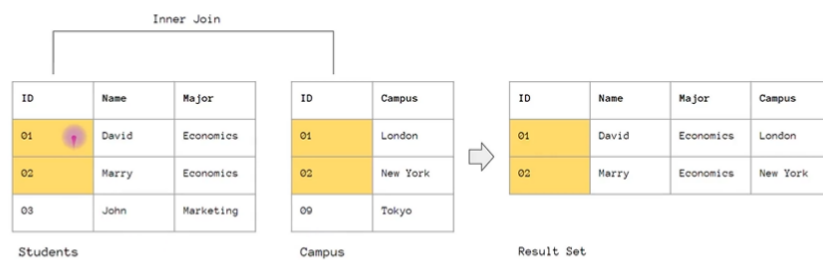```
JOIN albums AS B
ON A.artistid = B.albumid
```

## EP08 - Inner Vs. Left join

💡 เวลาสัมภาษณ์งานตำแหน่ง Data Analyst คำถามนี้เจอบ่อยมาก "จงอธิบายความ
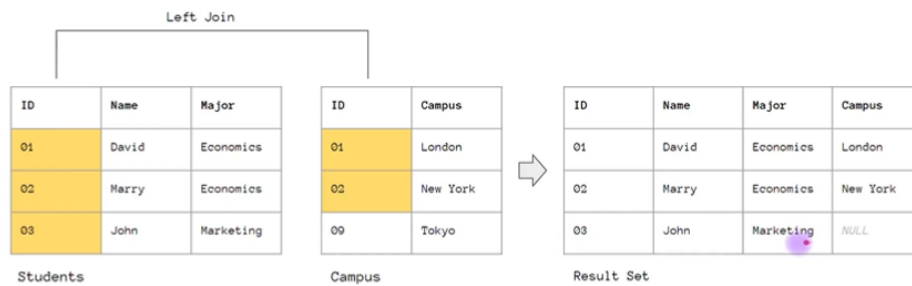แตกต่างระหว่าง inner vs. left join"

```
-- Inner join
SELECT
    A.artistid,
    B.name,
    C.title
FROM artists AS A
INNER JOIN albums AS B
ON A.artistid = B.artistid
```

Inner Join Illustration


```
-- Left join
SELECT
    A.artistid,
    B.name,
    C.title
FROM artists AS A
LEFT JOIN albums AS B
ON A.artistid = B.artistid
```

Left Join Illustration

EP09 - Practicing Joins example

```
-- Example Join in the video
SELECT
    art.artistsid,
    art.name,
    alb.title,
    tra.name,
    tra.composer
FROM artists AS art
JOIN albums AS alb
ON art.artistid = alb.artistid
JOIN tracks AS tra
ON alb.albumid = tra.albumid
WHERE art.name = 'Aerosmith';  -- WHERE is used after JOINs completed
```

EP10 - Random Sampling Rows

```
-- Random Sampling Example
SELECT name FROM tracks
ORDER BY RANDOM()
LIMIT 10;
```