Joining Data in SQL

Introduction to joins

```
inner join : only have things in both table

OuterJoin:
left join/right join: only id in left or right
full join: 两边的都有

cross join: all combinations with left and right

UNION (unique), UNION ALL, INTERSECT and EXCEPT

semi-join/Anti-join: subqueries
```

```
SELECT c.code AS country_code, name, year, inflation_rate
FROM countries AS  c INNER JOIN economies AS e ON c.code = e.code
AND c.xx=e.xxx;

这段代码显示了不需要有前后顺序,AS的东西可以用在一段语句里面。只要在后面出
现的东西其实都可以写在SELECT旁边

SELECT name, continent, code, surface_area,
CASE WHEN surface_area > 2000000 THEN 'large'
WHEN surface_area> 350000  THEN 'medium'
ELSE 'small' END
```

```
AS geosize_group
FROM countries;
这段就是如何用case来进行category的创造
```

命令	解释
SELECT xx.se, yy. dfs FROM xx INNER JOIN yy ON xx.a=yy.b	inner join on a column and finally show 2 columns
很长，放在上面当作例子了	很长，放在上面当作例子了
SELECT xx FROM xx AS xx INNER JOIN yy USING (m)	if same name, we can use USING
In self join, we can let one dataframe to be 2 names	相当于xx AS P1 INNER JOIN xx AS P2
Using case for new categories	在上面
SELECT xx INTO a FROM xx	Save the result

Outer joins and cross joins

命令	解释
SELECT xx.se, yy. dfs FROM xx LEFT JOIN yy ON xx.a=yy.b	same things for join, this time for left join
SELECT xx.se, yy. dfs FROM xx RIGHT JOIN yy ON xx.a=yy.b	same things for join, this time for right join
/* */	多行comment
SELECT xx.se, yy. dfs FROM xx FULL JOIN yy ON xx.a=yy.b	same things for join, this time for Full join with id on double sides
SELECT xx.se, yy. dfs FROM xx CROSS JOIN yy ON xx.a=yy.b	same things for join, this time for cross join with id distribution

Set theory clauses

命令	解释
SELECT xx FROM xx UNION SELECT xx FROM xx	Combine 2 table into one table and remove duplicates

命令	解释
SELECT xx FROM xx UNION ALL SELECT xx FROM xx	Combine 2 table into one table and remains duplicates
SELECT xx FROM xx INTERSECT SELECT xx FROM xx	Combine 2 table into one table and remains things in both tables for both selected column
SELECT xx FROM xx EXCEPT SELECT xx FROM xx	Combine 2 table into one table and remains things in both tables, then only select the column in left
SELECT xx FROM xx WHERE xx IN (SELECT xxx)	WHERE IN statement for subquery
SELECT xx FROM xx WHERE xx NOT IN (SELECT xxx)	WHERE NOT IN statement for subquery

Subqueries

命令	解释
SELECT (SELECT xxx) FROM xxx	Subquery in SELECT
WHERE	上面说到了
SELECT xx FROM (SELECT xxx)	Subquery in FROM