Conditional logic

CASE

Syntax form

**CASE** expression

**WHEN** test **THEN** **result**

…

**ELSE** otherResult

**END**

Lets create a new field based on cost

**SELECT** aircraft, airline, cost,

**CASE**

**WHEN** cost = 0

**THEN** 'NO COST'

**WHEN** cost >0 **AND** cost < 100000

**THEN** 'MEDIUM COST'

**ELSE**

'HIGH COST'

**END**

**AS** cost\_category

**FROM** birdstrikes

**ORDER** **BY** cost\_category;

Exercise1

Do the same with speed. If speed is NULL or speed < 100 create a “LOW SPEED” category, otherwise, mark as “HIGH SPEED”. Use IF instead of CASE!

Aggregations

**COUNT**

**Counting the number of records**

COUNT(\*) - **counts the number of records**

**SELECT** **COUNT**(\*) **FROM** birdstrikes;

COUNT(column) - **counts the number of not NULL records for the given column**

**SELECT** **COUNT**(reported\_date) **FROM** birdstrikes;

**DISTINCT**

**How do we list all distinct states? (Remember last seminar!)**

**SELECT** **DISTINCT** **state** **FROM** birdstrikes;

**Count number of distinct states**

**SELECT** **COUNT**(**DISTINCT** **state**) **FROM** birdstrikes;

Exercise2

How many distinct ‘aircraft’ we have in the database?

**MAX, AVG, SUM**

**The sum of all repair costs of birdstrikes accidents**

**SELECT** **SUM**(cost) **FROM** birdstrikes;

**Speed in this database is measured in KNOTS. Let’s transform to KMH. 1 KNOT = 1.852 KMH**

**SELECT** (**AVG**(speed)\*1.852) **as** avg\_kmh **FROM** birdstrikes;

**How many observation days we have in birdstrikes**

**SELECT** DATEDIFF(**MAX**(reported\_date),**MIN**(reported\_date)) **from** birdstrikes;

Exercise3

What was the lowest speed of aircrafts starting with ‘H’

Grouping

**GROUP BY**

**What is the highest speed by aircraft type?**

**SELECT** **MIN**(speed), aircraft **FROM** birdstrikes **GROUP** **BY** aircraft;

**Which state for which aircraft type paid the most repair cost?**

SELECT state, aircraft, SUM(cost) AS sum FROM birdstrikes WHERE state !='' GROUP BY state, aircraft ORDER BY sum DESC;

Exercise4

Which phase\_of\_flight has the least of incidents?

Exercise5

What is the rounded highest average cost by phase\_of\_flight?

**HAVING**

**We would like to filter the result of the aggregation. In this case we want only the results where the avg speed is equal to 50.**

**SELECT** **AVG**(speed) **AS** avg\_speed,**state** **FROM** birdstrikes **GROUP** **BY** **state** **WHERE** ROUND(avg\_speed) = 50;

**Crashbummbang! The correct keyword after GROUP BY is HAVING**

**SELECT** **AVG**(speed) **AS** avg\_speed,**state** **FROM** birdstrikes **GROUP** **BY** **state** **HAVING** ROUND(avg\_speed) = 50;

Exercise6

What the highest AVG speed of the states with names less than 5 characters?

Homework 3

* Upload the solution of exercise 1-6 to your GitHub repo in a folder called HW3
* Make sure to submit both the SQL statements and answers to the questions
* The required data format for submission is a .sql file
* Submit GitHub repo link to moodle when you are ready