**Ontology of Generated SQL Errors  
The model generates synthetically valid SQL variations of a given sampled query. A total of 100 queries are sampled from a pool of 1534 queries. The samples are selected randomly making sure that all the 11 databases (the queries are sourced from) are represented equally. Furthermore, the original query has Database names at the end of the SQL statement. The generated variation query omits this database name (so that it does not throw errors upon execution).**

**1. Query Structure Errors**

* **1.1 Modify Filter Clause**
  + Description: Changes made to the **WHERE** clause, such as adding, removing, or modifying conditions.
  + Example: Adding **AND T2.City = 'Prague'** to the **WHERE** clause.
  + Count: 55
* **1.2 Modify Join Type**
  + Description: Changes made to the type of join (e.g., **INNER JOIN** to **LEFT JOIN**).
  + Example: Changing **INNER JOIN** to **LEFT JOIN**.
  + Count: 33
* **1.3 Modify Group by Clause**
  + Description: Changes made to the **GROUP BY** clause, such as adding or removing columns.
  + Example: Adding **GROUP BY T3.Description**.
  + Count: 7
* **1.4 Modify Having Clause**
  + Description: Changes made to the **HAVING** clause, such as adding or modifying conditions.
  + Example: Adding **HAVING COUNT(\*) > 10**.
  + Count: 2

**2. Operator Errors**

* **2.1 Change Comparison Operator**
  + Description: Changes made to comparison operators in conditions (e.g., **=** to **!=**, **>** to **<**).
  + Example: Changing **WHERE T2.Country = 'CZE'** to **WHERE T2.Country != 'CZE'**.
  + Count: 54
* **2.2 Swapping Logical Operators**
  + Description: Changes made to logical operators (e.g., **AND** to **OR**).
  + Example: Changing **WHERE T2.Country = 'CZE' AND T2.City = 'Prague'** to **WHERE T2.Country = 'CZE' OR T2.City = 'Prague'**.
  + Count: 18

**3. Clause Errors**

* **3.1 Add/Modify Limit Clause**
  + Description: Changes made to the **LIMIT** clause, such as adding or modifying the limit value.
  + Example: Adding **LIMIT 10** to the query.
  + Count: 32
* **3.2 Add/Remove Distinct Clause**
  + Description: Adding or removing the **DISTINCT** keyword from the **SELECT** clause.
  + Example: Changing **SELECT DISTINCT T3.Description** to **SELECT T3.Description**.
  + Count: 26
* **3.3 Modify Order by Clause**
  + Description: Changes made to the **ORDER BY** clause, such as modifying the column or direction.
  + Example: Changing **ORDER BY T3.Description ASC** to **ORDER BY T3.Description DESC**.
  + Count: 19

**4. Function Errors**

* **4.1 Modify Aggregate Function**
  + Description: Changes made to aggregate functions (e.g., **COUNT** to **MAX**).
  + Example: Changing **COUNT(\*)** to **MAX(T1.id)**.
  + Count: 23

A screenshot of a computer program

AI-generated content may be incorrect.