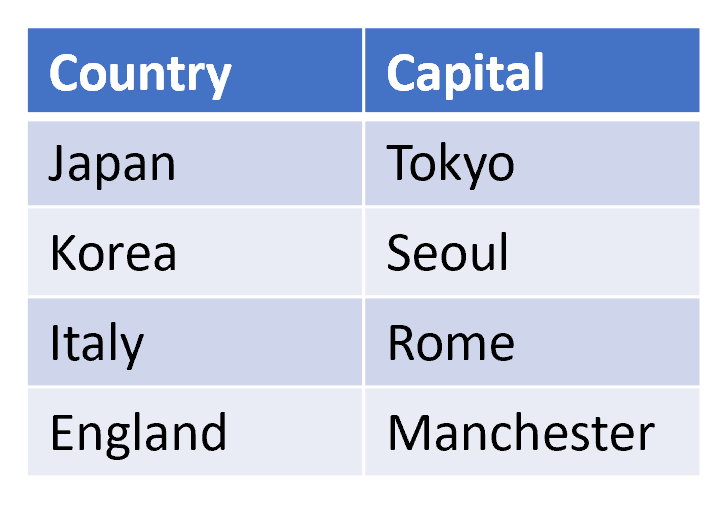
DSCI 558: Building Knowledge Graphs

Quiz 16 (5 minutes)

**Question 1:**

Map the following data table with the correct type of errors it shows: wrong values, format inconsistencies, violated attribute dependency (3pts)



**Question 2: Randomly choose 7 out of 9**

True/False

1. KATARA (KG-based error detection system) uses SPARQL to handle partial KB coverage cases. F
2. KATARA uses crowdsourcing to help produce the source models of tables. T
3. ML-based error detection systems can use a wide range of features to capture different types of errors. T
4. ML-based error detection systems can extract generalized features from the data and thus do not suffer from the rarity of errors. F
5. In programming-by-example systems, clustering before transforming helps users detect transformation errors easier. T
6. In programming-by-example systems, the final output is a combination of input substrings and constant strings. T
7. ML-based error detection systems can leverage outlier detection methods and thus require no training data. F
8. KATARA uses the sum of TF-IDF and PMI scores to rank the set of candidate table patterns (subject-predicate-object in source models) T
9. In error detection, active learning and few-shot learning techniques can improve the system’s performance when compared to normal supervised methods. F