Assignment 3

1. Explain the difference between outer join and inner join with suitable illustrations.
2. Describe a method for resolving mismatched keys in two datasets.
3. Provide a detailed example of how to merge two datasets using a common key, highlighting the role of indices.
4. Using Python's Pandas library, demonstrate how to perform each type of join with code examples.
5. Discuss the implications of using different types of joins on data completeness and accuracy.
6. Discuss the implications of ignoring duplicate keys in the merging process. What potential issues could arise?
7. Discuss the algorithms commonly used for fuzzy matching. Provide examples of their applications in real-world data scenarios.
8. Explain the potential pitfalls of using fuzzy matching in data merging. How can these be addressed?