Seventh hands-on exercise (answers)

* **Task 1**

**Answer to Question 3:**

**Create** **table** CFanalysis **AS** (**SELECT** **DISTINCT** username, hotel\_id **FROM** tripadvisor\_review\_sample\_without\_reviewtext **WHERE** username **IS** **NOT** **NULL**);

**ALTER** **TABLE** `CFanalysis` **ADD** **INDEX** `username` (`username`), **ADD** **INDEX** `hotel\_id` (`hotel\_id`);

*# what would happen, if we don’t build index.*

**SELECT** a.username **as** Person1, b.username **as** Person2, **COUNT**(\*) **as** Frequency

**FROM** CFanalysis **as** a **JOIN** CFanalysis **as** b

**ON** a.hotel\_id = b.hotel\_id **AND** a.username > b.username

**GROUP** **BY** a.username, b.username **ORDER** **BY** Frequency **desc**

**Answer to Question 4:** The above results based on association analysis can be applied to an online auction platform that sells used staff. For instance, suppose that user A and user B are a pair who strong similar interests, if user A likes a product (e.g., making a bid for a product), user B is likely to like that product as well. Thus, when a new staff appears on the site when a customer (A) makes a bid to a specific staff, the staff will be recommended to another customer (B), who has exhibited a strong similarity in the preference for products with the customer (A). Please note that A may share similar interests with several other customers, not just B. Also if B also bid for the product, there might be other customers similar to B (but not similar to A) can be recommended.

* **Task 2: please see the course slides for the code**
* **Task 3: please see the course slides for the solution**