Département Informatique et Statistique, ICOM, Université Lumière Lyon 2
M1 Informatique – Year 2015-2016
Semi-structured data & XML – Labwork #6: Complex XQueries
J. Darmont (*http://eric.univ-lyon2.fr/~jdarmont/*), 16/09/16

Preliminaries

1. Download the following XML documents into a new directory:
   - http://eric.univ-lyon2.fr/~jdarmont/docs/customers.xml
   - http://eric.univ-lyon2.fr/~jdarmont/docs/orders.xml
   - http://eric.univ-lyon2.fr/~jdarmont/docs/books.xml.

2. Within BaseX, create a new database (menu Database/New); indicate the directory you created in question #1 as input.

Exercise #1: Joins

Formulate the following XQueries.

1. List of customers.
2. Names of customers who ordered books (formulate join condition in a `where` clause).
3. Distinct names of customers who ordered books (use function `distinct-values` and formulate join condition in a path predicate).
4. Distinct names of customers who ordered books (use function `exists`). Conclusion?
5. Names of customers who did not pass any order.
6. Distinct names of customers who ordered any book in quantity greater than 1 (formulate condition in a `where` clause).
7. List of orders indicating the name of customer and the titles of ordered books (formulate join condition in a `where` clause). Return result under the following format.

```
<order>
    <customer>CUSTOMER_NAME</customer>
    <date>ORDER_DATE</date>
    <book>BOOK_TITLE
         <qty>QUANTITY_ORDERED</qty>
    </book>
</order>
```

8. List of orders indicating the name of customer and the titles of ordered books (formulate join condition in a path predicate).
9. Names of customers who ordered Science Fiction books.

Exercise #2: Aggregation and grouping

Formulate the following queries in XQuery 3.0.

1. Total number of books in the catalog.
2. Average price of books of genre “Computer”.
3. Minimum, average and maximum price of books.
4. Number of books per author.
5. Number and average price of books per genre, ordered by genre. From now on, always order the result with respect to the grouping element(s).
6. Average length of description per genre.
7. Maximum of average lengths of description per genre.
8. Number of books per author and per genre

(on both `books.xml` and `orders.xml`)

9. Total quantity ordered per book (indicate book title).
10. Total quantity ordered per genre of books.