**Group members:**

**(e.g: Username, StudentID, Name, Parts responsible)**

# **Task 1**

Student ID used for publication type selection: davby024

Major type: Article (last digit is even)

Minor type: Book (4%6)+1 = 5

Code used to retrieve your data should be in extract.xql.

# **Task 2**

Start schema definitions (fact table must be the first one).

Fact\_Publication(publication\_id, details\_id, author\_id, publisher\_id, date\_id, links\_id, publication\_type , [publication\_count])

publication\_id (PK)

details\_id (SK)

author\_id (SK)

publisher\_id (SK)

time\_id (SK)

links\_id (SK)

Dim\_Details(details\_id, title, volume, number, pages)

details\_id (SK)

Dim\_Author(author\_id, author\_name, editor\_name)

author\_id (SK)

Dim\_Publisher(publisher\_id, publisher\_name, journal\_name)

publisher\_id (SK)

Dim\_Time(time\_id, year, month)

time\_id (SK)

Dim\_Links(links\_id, url, ee, cdrom, cite)

links\_id (PK)

Descriptions of usefulness dimension tables.

Table creation statements (SQL)

Table schemas (screenshots)

Table constraint summary (screenshot)

# **Task 3**

Summarize transformations and data uploading to oracle.

(note that here, you summarize the steps, detailed code should appear in transform?.\* files.)

transform1.xql:

Aim:

Input:

Processes:

Output:

Data tables (screenshots)

Justification of value and relationship preservation

Examples:

Values:

One-to-many relationships:

Dataset for an analysis.

Analysis:

SQL:

Screenshot:

# **Task 4**

1. Example of data entries in the data source.
2. The data ready for integration (a small table two records)
3. Assumptions and processes to extract and transform data from a) to b).

Assumptions

Processes

1. The purpose of new data. Does it extend an existing dimension or add a new dimension? Present the new table schema.
2. The way in which the new data records will be linked/matched to existing records.
3. Automation challenges.