



**Course:** MSc in Computer Engineering / MSc in Bioinformatics  
**U.C.:** NoSQL Databases

| Exercise Sheet PL02     |                                   |
|-------------------------|-----------------------------------|
| Teacher:                | António Abelha / Cristiana Neto   |
| Theme:                  | Data Migration: Oracle to MongoDB |
| Class:                  | Laboratory Practice               |
| Academic Year:          | 2023-2024 – 2nd Semester          |
| Duration of the lesson: | 2 hours                           |

## Exercise

The objective of this exercise is to develop a Python script to migrate data from an Oracle database to a MongoDB collection. You must extract data from various Oracle tables and insert it into MongoDB documents according to a pre-defined structure.

Consider the tables from the Oracle database in sheet FE01:

AUTOR {ID\_AUTOR, NOME}  
EDITORA {ID\_EDITORA, NOME}  
GENERO {ID\_GENERO, NOME}  
SUPORTE {ID\_SUPORTE, NOME}  
TITULO {ID\_TITULO, TITULO, PRECO, DTA\_COMPRA, ID\_EDITORA, ID\_SUPORTE, ID\_GENERO, ID\_AUTOR}  
MUSICA {ID\_MUSICA, NOME, ID\_AUTOR, ID\_TITULO}  
REVIEW {ID\_REVIEW, ID\_TITULO, DTA\_REVIEW, CONTEUDO}

You must develop a Python script that connects to Oracle, extracts data from these tables, and formats it to match the desired document structure in MongoDB. Then, the formatted documents must be inserted into a MongoDB collection.

### Tasks:

1. Write a Python script that connects to the Oracle database and extracts data from all the tables mentioned above.
2. Format the extracted data to match the desired document structure.
3. Connect to MongoDB and insert the formatted documents into the MongoDB collection.

The desired document structure for MongoDB is attached.

### Technical requirements:

- Python 3.x
- Oracledb packages for connecting to Oracle and pymongo for connecting to MongoDB.

Submit the developed Python script along with any documentation required to understand and run the script.



University of Minho  
Department of Computer Science

# FE02

Documents structure:

```
{
  "id_titulo": int,
  "titulo": string,
  "preco": float,
  "dta_compra": date,
  "editora": {
    "id_editora": int,
    "nome": string
  },
  "suporte": {
    "id_suporte": int,
    "nome": string
  },
  "genero": {
    "id_genero": int,
    "nome": string
  },
  "autor": {
    "id_autor": int,
    "nome": string
  },
  "musicas": [
    {
      "id_musica": int,
      "nome": string,
      "autor": {
        "id_autor": int,
        "nome": string
      }
    }
  ],
  "reviews": [
    {
      "id_review": int,
      "dta_review": date,
      "conteudo": string
    }
  ]
}
```