



UNIVERSIDAD DE MÁLAGA



Graduados en Ingeniería de la Salud

## Biological Databases

Nutritional Database

Made by

Susana Rocío Fernández Giacomassi  
Pablo Moreno García-Espina

Supervised by

Ismael Navas Delgado

Department

Lenguajes y Ciencias de la Computación

MÁLAGA, June 2023



UNIVERSIDAD  
DE MÁLAGA



ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA INFORMÁTICA  
GRADUADO EN INGENIERÍA DEL SOFTWARE

Bases de Datos Biológicas

**Nutritional Database**

Made by  
**Susana Rocío Fernández Giacomassi,  
Pablo Moreno García-Espina**

Supervised by  
**Ismael Navas Delgado**

Department  
**Lenguajes y Ciencias de la Computación**

UNIVERSIDAD DE MÁLAGA  
MÁLAGA, JUNE 2023

# Abstract

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

**Keywords:** A, B, C

# Resumen

## 0.1 Initial project description

We are going to create a database which use is going to be directly related to aliments. The goal of our database is to provide a determined patient with a diet following its main characteristics (age, height, weight) Therefore, we will need a group of nutritionists. They are the ones who will be in charge of the development of the diets. We will storage data for aliments mainly, and we will use this data in order to create new diets for our patients in the nutritional center.

We will use a relational database, a type of database that stores and provides access to data points that are related to one another.

In order to generate data for our database, we will use csv files for some of our tables. For the rest of the tables, we will try to generate data automatically with the tools we will see in class (Python)

## 0.2 RDBMS Deployment

In the schema of our database, we will have the following tables: Food, Patient, Nutritionist, Diet and List Of Food. One Diet will consist of different kinds of Food, a Diet will be created by a Nutritionist (although a Nutrionist can create more than one) and a Nutritionist will take care of one or more Patients. The table List Of Food is created as a consequence of the relation between Diet and Food (a relation M:N creates a new table including the primary keys of both tables included in the relation).

Firstly, we have introduced the table Food, which has the food ID as the primary key. The rest of columns of the table are related with information such as the energy in kcal and KJ that provides the food, the quantity of Protein and Fiber in food in every 100 grams... Then the table to register the information about the Patient has been created and fulfilled with relevant information about it in order

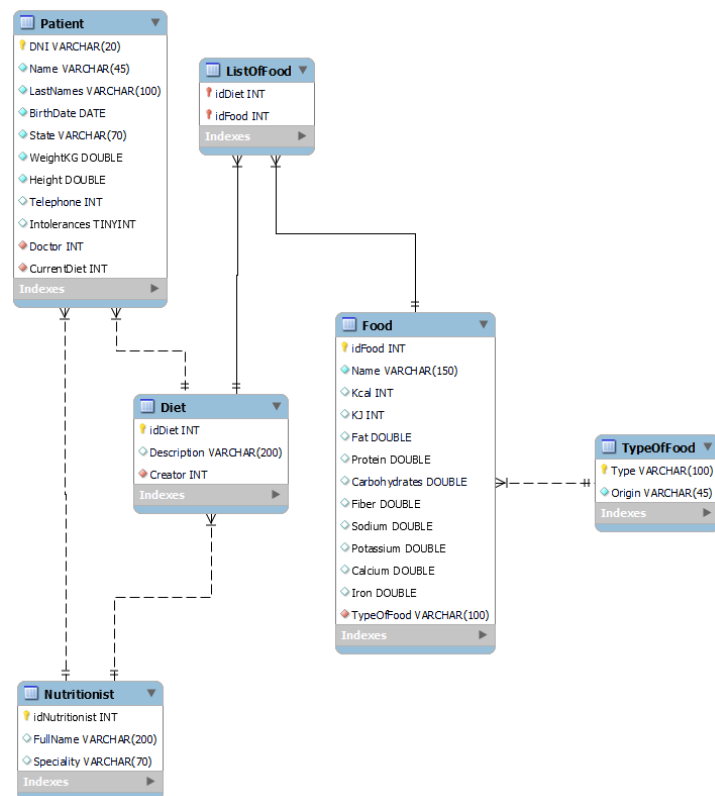


Figure 1: EER model

to organize the new diet. It also stores the current diet that the person is following and the intolerances, information that can be quite useful in order to design a new diet.

The table of the Nutritionist is created also with an ID as the primary key. It also contains information about the name of the nutritionist and the specialty that he or she has (obesity, for example)

For the diets we have another table. As an identifier, we have an ID for the diet (primary key) Also, a description and a Creator are columns for this table. The Description will contain the different steps to follow in order to do the diet. The data which has been introduced in the table Food has been obtained by a csv file in <https://www.fao.org/infoods/infoods/tablas-y-bases-de-datos/es/>

### 0.3 Query Design

· Food with more than 20 grams of protein (Fig. 2):

```
SELECT NutritionalDB.Food.Name
FROM NutritionalDB.Food WHERE Food.Protein > 20;
```

· All the names of patients treated by the doctor Pablo Moreno Garcia-Espina (Fig. 3):

```
SELECT NutritionalDB.Patient.Name
FROM NutritionalDB.Patient JOIN NutritionalDB.
    Nutritionist
ON Patient.Doctor = Nutritionist.idNutritionist
WHERE Nutritionist.FullName = 'Pablo Moreno Garcia-
    Espina';
```

· Description of the diets created by the doctor Susana Rocio Fernandez Giacomassi (Fig. 4)

```
SELECT NutritionalDB.Diet.Description
```

Name
► Soy flour
Wheat germ
Anchovy
Barracuda
Bluefish
Pollock
Rainbow trout
Salmon
Sardines
Shark
Skipjack tuna
Sole
Swordfish
Trout
Tuna
Yellowfin tuna
Yellowtail
Cheese cheddar
Mozzarella cheese
Blue cheese
Goat cheese
Beef round bottom r...
Beef round bottom r...
Beef round top roun...
Beef round top roun...

Figure 2: First query output

Name
► Juan
Ines
Pedro
Ines
Isabel
Antonio
Sofia
Pablo

Figure 3: Second query output

	Description
►	Dieta sin lactosa para veganos
	Dieta rica en fibra para mejorar la digestion
	Dieta sin sal para mejorar la circulacion
	Dieta para mejorar la salud intestinal
	Dieta para reducir el estres
	Dieta para mejorar la salud ocular
	Dieta para mejorar la circulacion sanguinea
	Dieta para mejorar la salud del higado
	Dieta para mejorar la salud de las articulaciones

Figure 4: Third query output

```
FROM NutritionalDB.Diet JOIN NutritionalDB.
    Nutritionist
ON Diet.Creator = Nutritionist.idNutritionist
WHERE Nutritionist.FullName = 'Susana Rocio Fernandez
    Giacomassi';
```

· Obtain the description and the name of their creator of all diets with more than 10 foods (Fig. 5):

```
SELECT NutritionalDB.Diet.Description as 'Diets with
    more than ten foods', NutritionalDB.Nutritionist.
    FullName as 'Creator'
FROM NutritionalDB.Diet JOIN NutritionalDB.
    Nutritionist
ON Diet.Creator = Nutritionist.idNutritionist
WHERE NutritionalDB.Diet.idDiet IN (
    SELECT NutritionalDB.Diet.idDiet
    FROM NutritionalDB.ListOfFood
    GROUP BY idDiet HAVING COUNT(*) > 10
);
```

· Get the names and protein content of foods that have more protein than the average protein of all foods (Fig. 6):



Diets with more than ten foods	Creator
► Dieta rica en proteínas para deportistas	Pablo Moreno Garcia-Espina
Dieta para reducir el colesterol	Pablo Moreno Garcia-Espina
Dieta para aumentar la masa muscular	Pablo Moreno Garcia-Espina
Dieta para reducir la inflamación	Pablo Moreno Garcia-Espina
Dieta para mejorar la concentración	Pablo Moreno Garcia-Espina
Dieta para aliviar la artritis	Pablo Moreno Garcia-Espina
Dieta para mejorar la digestión	Pablo Moreno Garcia-Espina
Dieta para reducir la hinchazón abdominal	Pablo Moreno Garcia-Espina
Dieta para reducir la retención de líquidos	Pablo Moreno Garcia-Espina
Dieta sin lactosa para veganos	Susana Rocio Fernandez Giacomassi
Dieta rica en fibra para mejorar la digestión	Susana Rocio Fernandez Giacomassi
Dieta sin sal para mejorar la circulación	Susana Rocio Fernandez Giacomassi
Dieta para mejorar la salud intestinal	Susana Rocio Fernandez Giacomassi
Dieta para reducir el estrés	Susana Rocio Fernandez Giacomassi
Dieta para mejorar la salud ocular	Susana Rocio Fernandez Giacomassi
Dieta para mejorar la circulación sanguínea	Susana Rocio Fernandez Giacomassi
Dieta para mejorar la salud del hígado	Susana Rocio Fernandez Giacomassi
Dieta para mejorar la salud de las articulaciones	Susana Rocio Fernandez Giacomassi
Dieta baja en grasas	Ismael Navas Delgado
Dieta equilibrada para mantener el peso	Ismael Navas Delgado
Dieta sin cafeína para mejorar el sueño	Ismael Navas Delgado

Figure 5: Fourth query output

```
SELECT NutritionalDB.Food.Name, NutritionalDB.Food.
    Protein
FROM NutritionalDB.Food
WHERE NutritionalDB.Food.Protein > (
SELECT AVG(NutritionalDB.Food.Protein)
FROM NutritionalDB.Food
);
```

## 0.4 RDBMS Optimization

Firstly, we have changed the first query in order to make a deeper search. We can also search in the database according to the type of food, as we made a join between the table of Food and the table of Type of Food.

In order to make the search easier when it comes to diets related with sporty people, we have decided to design two indexes: one with Name and Kcal and another with Protein and Name. This index becomes useful with our first query, as we are looking firstly at the food with protein's value of more than 20 grams.

For the previous index, when we try to execute the first query apparently the

Name	Protein
► Bagel	11
Buckwheat	13.3
Einkorn wheat	16.5
Kamut	14.3
Millet	11
Quinoa	14.1
Soy flour	39.6
Soybeans cooked	16.6
Spelt raw	14.6
Wheat bran	15.6
Wheat germ	27.2
Wheat whole grain	12.6
Whole-grain bread	10.4
Anchovy	20.3
Barracuda	25.4
Bluefish	20.4
Catfish	18

Figure 6: Fifth query output

index is not used. That could be because the improvement when the index is used is not considerably high when you compare it with the performance without the index.

This picture (Fig. 7) shows the performance of the first query without the index.

This picture (Fig. 8) shows the performance of the first query with the index.

Furthermore, as we will frequently look for the different diets that one nutritionist has created, we will create an index for the column “Full Name” in the table Nutritionist. We already have an index for the id of a nutritionist, but we will typically get a diet of a nutritionist by his concrete name. With this index we will get an advantage in the second and third query, as they use the full name of the nutritionist in order to make the query.

This picture (Fig. 9) shows the time which lasts the second query without the index.

This picture (Fig. 10) shows the time which lasts the second query with the index.

This picture (Fig. 11) shows the performance of the first query without the index.

This picture (Fig. 12) shows the performance of the third query with the index.

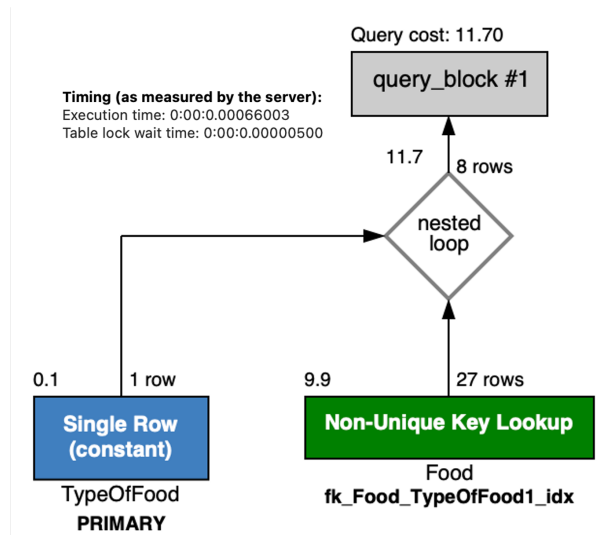


Figure 7: First query without index

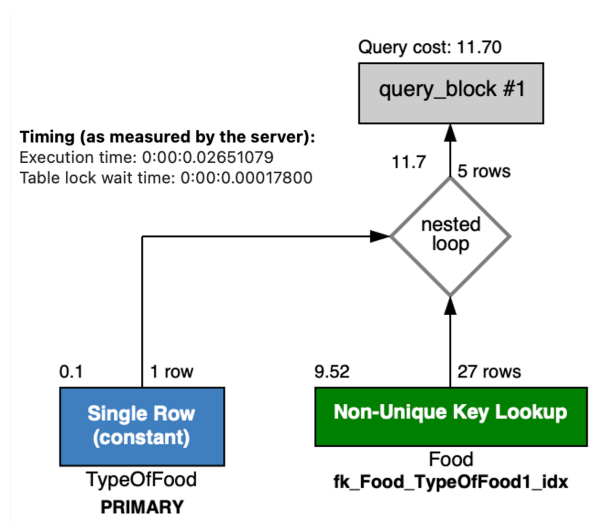


Figure 8: First query with index

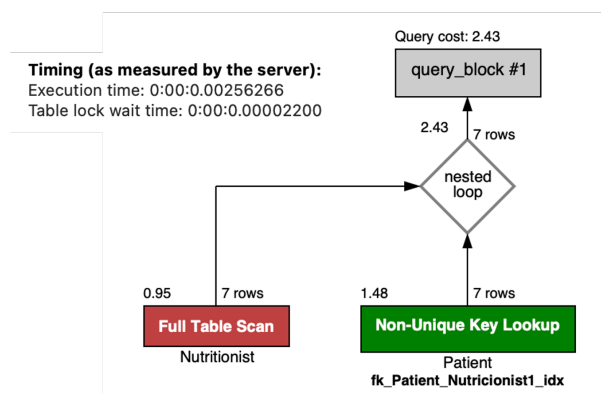


Figure 9: Second query without index

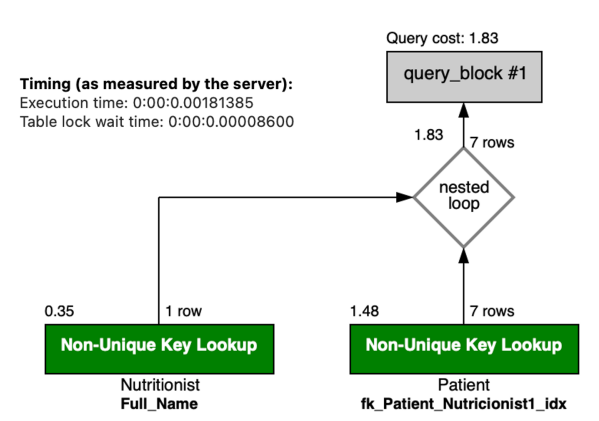


Figure 10: Second query with index

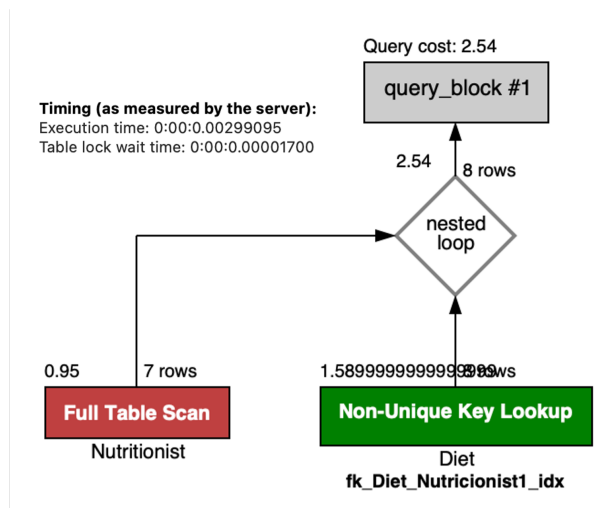


Figure 11: Third query without index

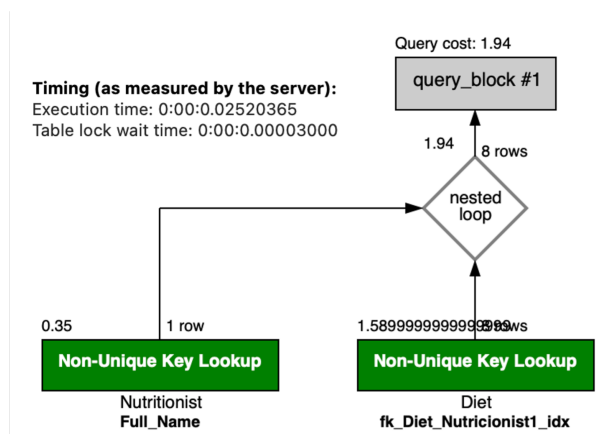


Figure 12: Third query with index

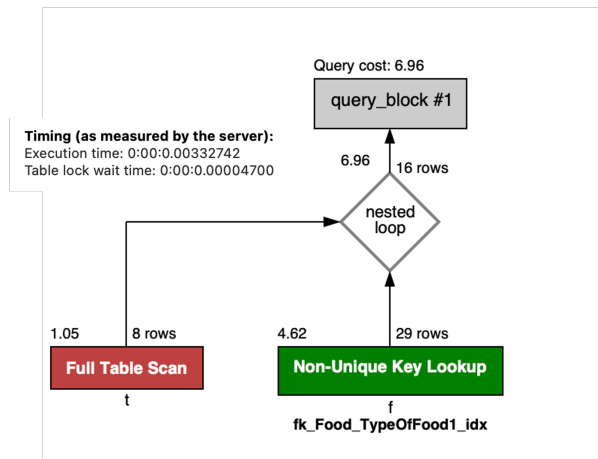


Figure 13: Extra query with index

We see how the time elapsed in the second time is lower than in the first time. The existence of the index makes the performance of the query higher.

We will also do queries which involve the name of the different patients we have in our database. Therefore, it would be suitable too to make an index with the name of the patients. Every time that we do a query in which the name of the patient is involved, we will get an advantage in terms of time, as we can see before.

We think that nutritionists may want to filter the aliments by the nutritional information and origin, so the query would be (Fig.9):

```
SELECT f.Name
FROM Food f
JOIN TypeOfFood t ON f.TypeOfFood = t.Type
WHERE t.Origin = 'animal' AND (f.Protein > 15 OR f.
  Iron > 5);
```

In this case, we have created an index for the table Food including the columns Protein and Iron, but it did not use it, neither the protein index created before.

## 0.5 Python

In order to develop the Python application to test the queries from an user interface, we have created a new project in PyCharm named NutritionalDB.

Specifically in the class NutritionalDB.py we have created three definitions (Fig. 9): one to initiate the connection to the database in SQL, one close the connection and another one to print the result of the queries we will implement next.

```
import pymysql
import pymysql.cursors
from pymysql import DatabaseError
class NutritionalQuerys(object):

    def __init__(self, host, user, password, database):
        self.connection = pymysql.connect(host=host,
                                           user=user,
                                           password=password,
                                           db=database,
                                           charset='utf8mb4',
                                           cursorclass=
pymysql.cursors.DictCursor)
        print("Successfully Connected!")

    def close(self):
        self.connection.close()
        print("Connection Closed!")

    def print_query(self, query):
        cursor = self.exec_query(query)
        print()
        for row in cursor:
            print(row)
```

However, if we analyze the code for the "print query" definition we see that

another definition called "exec query" is called. This new definition is used in order to send the query to the cursor created with the connection (Fig. 10)

```
def exec_query(self, query):
    try:
        with self.connection.cursor() as cursor:
            cursor.execute(query)
            return cursor
    except DatabaseError as e:
        if e.args[1] == '#42000Unknown Database':
            print(e)
        else:
            print(e)
```

After these first definitions, we have the definitions for each of the queries. When we made the queries in SQL we asked for concrete values of the columns (for example 20 grams of protein), however here we have parameterized the functions in order to make the query according to the user's goal.

When the user in Python calls the definition in order to ask for a food whose protein content is greater than X (Fig. 11), he has to specify this X value in the call to the method.

```
def gtProtein(self, protein):
    sql = 'SELECT NutritionalDB.Food.Name FROM Food
    WHERE Food.Protein > ' + str(protein) + ';'
    print()
    print('Food with more than ' + str(protein) + '
    grams of protein:')
    self.print_query(sql)
```

As we can see (Fig. 12), all of the definitions we have implemented in Python have the same structure. Firstly we create a string called sql in which we write the proper query. We have to insert here the different parameters we have in the header of the definition. Secondly, we print an empty line in order to make a separation in the terminal between queries. Then we print a text in order to make

the user understand about the meaning of the query, and finally we execute the method "print query" in order to show in the terminal the results of the query.

We have created a method that receives a doctor as input and returns the patients treated by the input doctor.

```
def treatedBy(self, doctor):
    sql = "SELECT NutritionalDB.Patient.Name FROM
    NutritionalDB.Patient JOIN Nutritionist " \
        "ON Patient.Doctor = Nutritionist.
    idNutritionist " \
        "WHERE Nutritionist.FullName = '" + doctor + "'
    ';"
    print()
    print('Patients treated by ' + str(doctor) + ':')
    self.print_query(sql)
```

The following definition receives an integer value and returns the diets with more than that number of food.

```
def dietsWithMoreFood(self, quantity):
    sql = "SELECT NutritionalDB.Diet.Description as " \
        "'Diets with more than " + str(quantity) + "
    foods ', " \
        "NutritionalDB.Nutritionist.FullName as '
    Creator' FROM " \
        "NutritionalDB.Diet JOIN NutritionalDB.
    Nutritionist ON " \
        "Diet.Creator = Nutritionist.idNutritionist
    " \
        "WHERE NutritionalDB.Diet.idDiet IN (SELECT
    " \
        "NutritionalDB.Diet.idDiet FROM
    NutritionalDB.ListOfFood " \
        "GROUP BY idDiet HAVING COUNT(*) > " \
```



```

        """ + str(quantity) + ");"

    print()
    print('Diets with more than ' + str(quantity) + '
foods:')
    self.print_query(sql)

```

We can also get the food with more than the average of an input column. And the last definition receives two columns and two values, so we can filter by columns and grams.

```

def moreThanAvg(self, column):
    sql = "SELECT NutritionalDB.Food.Name, NutritionalDB
.Food." + column + \
        " FROM NutritionalDB.Food" \
        " WHERE NutritionalDB.Food." + column + " > (
SELECT AVG(NutritionalDB.Food." \
        + column + ") FROM NutritionalDB.Food);"
    print()
    print('Food with more ' + str(column) + ' than the
average:')
    self.print_query(sql)

def filterByInfoAndOrigin(self, info1, info2, value1,
value2, origin):
    sql = "SELECT NutritionalDB.Food.Name FROM
NutritionalDB.Food" \
        " JOIN NutritionalDB.TypeOfFood ON Food.
TypeOfFood = TypeOfFood.Type" \
        " WHERE TypeOfFood.Origin = '" + origin + \
        "' AND (Food." + info1 + " > '" + str(value1) +
" OR Food." + info2 + \
        " > '" + str(value2) + ");"
    print()

```

```

        print('Food with more than ' + str(value1) + ' grams
of ' + str(info1) + ' or more than '
            + str(value2) + ' grams of ' + str(info2) + '
from ' + str(origin) + ' origin :')
        self.print_query(sql)

```

Lastly, we have created a main class to test these definitions:

```

if __name__ == '__main__':
    password = 'Your_Password'
    ndb = NutritionalQuerys('localhost', 'root', password, '
NutritionalDB')
    ndb.gtProtein(20)
    ndb.treatedBy('Pablo Moreno Garcia-Espina')
    ndb.dietsWithMoreFood(10)
    ndb.moreThanAvg('Protein')
    ndb.filterByInfoAndOrigin('Protein', 'Iron', 15, 5, '
animal')

```

A preview of the outputs can be seen below in figures 15-18:

## 0.6 NoSQL Database Design

In this section we are going to create an alternative to our SQL database. This is going to be done in MongoDB, using JSON documents to create the collections. The structure that our NoSQL database is going to have is simple, having one collection for each table on the relational SQL database.

To export our tables in the JSON format, we used a powerful tool of MySQL Workbench. We selected each table and chose the option "export wizard table". Then, we selected the csv format, and changed it for JSON file when saving it into the repository folder.

In order to create the different collections in MongoDB, we opened a new connection in MongoDB Compass and started to create new collections in a folder. We created the collections with the same name as the tables in SQL and import the data from the JSON files we have just generated.

```

Successfully Connected!

Food with more than 20 grams of protein:

{'Name': 'Pollock'}
{'Name': 'Sole'}
{'Name': 'Anchovy'}
{'Name': 'Trout'}
{'Name': 'Bluefish'}
{'Name': 'Salmon'}
{'Name': 'Beef chuck arm pot roast separable lean and fat trimmed to 0" fat all grades raw'}
{'Name': 'Rainbow trout'}
{'Name': 'Shoulder roast'}
{'Name': 'Swordfish'}
{'Name': 'Shark'}
{'Name': 'Beef round bottom round roast separable lean and fat trimmed to 0" fat all grades raw'}
{'Name': 'Blue cheese'}
{'Name': 'Goat cheese'}
{'Name': 'Yellowtail'}
{'Name': 'Pork ground 96% lean 4% fat raw'}
{'Name': 'Mung beans'}
{'Name': 'Wild boar chop'}
{'Name': 'Wild boar ribs'}
{'Name': 'Wild boar roast'}
{'Name': 'Beef shoulder pot roast or steak boneless separable lean only trimmed to 0" fat all grades raw'}
{'Name': 'Mozzarella cheese'}
{'Name': 'Sardines'}
{'Name': 'Veal chop'}
{'Name': 'Veal cutlet'}
{'Name': 'Veal roast'}
{'Name': 'Veal shank'}
{'Name': 'Cheese cheddar'}
{'Name': 'Barracuda'}
{'Name': 'Strip steak'}
{'Name': 'T-bone steak'}
{'Name': 'Top loin steak'}
{'Name': 'Tri-tip roast'}
{'Name': 'Wheat germ'}
{'Name': 'Venison steak'}

```

Figure 14: gtProtein() output

```

Patients treated by Pablo Moreno Garcia-Espina:

{'Name': 'Juan'}
{'Name': 'Ines'}
{'Name': 'Pedro'}
{'Name': 'Ines'}
{'Name': 'Isabel'}
{'Name': 'Antonio'}
{'Name': 'Sofia'}
{'Name': 'Pablo'}

```

Figure 15: treatedBy() output

```
Diets with more than 10 foods:

{'Diets with more than 10 foods': 'Dieta vegetariana equilibrada', 'Creator': 'Antonio Rodriguez Morcas'}
{'Diets with more than 10 foods': 'Dieta detox para eliminar toxinas', 'Creator': 'Antonio Rodriguez Morcas'}
{'Diets with more than 10 foods': 'Dieta sin carbohidratos para adelgazar rapido', 'Creator': 'Antonio Rodriguez Morcas'}
{'Diets with more than 10 foods': 'Dieta para aumentar la energia', 'Creator': 'Antonio Rodriguez Morcas'}
{'Diets with more than 10 foods': 'Dieta para aliviar la migraña', 'Creator': 'Antonio Rodriguez Morcas'}
{'Diets with more than 10 foods': 'Dieta sin gluten para celíacos', 'Creator': 'Antonio Rodriguez Morcas'}
{'Diets with more than 10 foods': 'Dieta para reducir la hipertension', 'Creator': 'Antonio Rodriguez Morcas'}
{'Diets with more than 10 foods': 'Dieta sin sulfitos para aliviar alergias', 'Creator': 'Antonio Rodriguez Morcas'}
{'Diets with more than 10 foods': 'Dieta para mejorar la salud ósea', 'Creator': 'Antonio Rodriguez Morcas'}
{'Diets with more than 10 foods': 'Dieta cetogenica para adultos', 'Creator': 'Cecilia Gonzalez Perez'}
{'Diets with more than 10 foods': 'Dieta sin azucar para cuidar la piel', 'Creator': 'Cecilia Gonzalez Perez'}
{'Diets with more than 10 foods': 'Dieta rica en vitaminas para fortalecer el sistema inmunologico', 'Creator': 'Cecilia Gonzalez Perez'}
{'Diets with more than 10 foods': 'Dieta para fortalecer los huesos', 'Creator': 'Cecilia Gonzalez Perez'}
{'Diets with more than 10 foods': 'Dieta sin soja para aliviar alergias', 'Creator': 'Cecilia Gonzalez Perez'}
{'Diets with more than 10 foods': 'Dieta para mejorar la salud del cerebro', 'Creator': 'Cecilia Gonzalez Perez'}
{'Diets with more than 10 foods': 'Dieta para mejorar la salud respiratoria', 'Creator': 'Cecilia Gonzalez Perez'}
{'Diets with more than 10 foods': 'Dieta para mejorar la salud bucal', 'Creator': 'Cecilia Gonzalez Perez'}
{'Diets with more than 10 foods': 'Dieta para aliviar la depresion', 'Creator': 'Cecilia Gonzalez Perez'}
{'Diets with more than 10 foods': 'Dieta hipercalórica', 'Creator': 'Cecilia Gonzalez Perez'}
{'Diets with more than 10 foods': 'Dieta baja en grasas', 'Creator': 'Ismael Navas Delgado'}
{'Diets with more than 10 foods': 'Dieta equilibrada para mantener el peso', 'Creator': 'Ismael Navas Delgado'}
{'Diets with more than 10 foods': 'Dieta sin cafeina para mejorar el sueño', 'Creator': 'Ismael Navas Delgado'}
{'Diets with more than 10 foods': 'Dieta para mejorar la salud mental', 'Creator': 'Ismael Navas Delgado'}
{'Diets with more than 10 foods': 'Dieta para mejorar la memoria', 'Creator': 'Ismael Navas Delgado'}
{'Diets with more than 10 foods': 'Dieta para mejorar la salud del sistema nervioso', 'Creator': 'Ismael Navas Delgado'}
{'Diets with more than 10 foods': 'Dieta para diabeticos tipo 2', 'Creator': 'Maria Ines Paz Bueno'}
{'Diets with more than 10 foods': 'Dieta para reducir la ansiedad', 'Creator': 'Maria Ines Paz Bueno'}
{'Diets with more than 10 foods': 'Dieta para mejorar la salud del corazón', 'Creator': 'Maria Ines Paz Bueno'}
{'Diets with more than 10 foods': 'Dieta sin frutos secos para aliviar alergias', 'Creator': 'Maria Ines Paz Bueno'}
{'Diets with more than 10 foods': 'Dieta para aliviar el dolor menstrual', 'Creator': 'Maria Ines Paz Bueno'}
{'Diets with more than 10 foods': 'Dieta sin histamina para aliviar alergias', 'Creator': 'Maria Ines Paz Bueno'}
{'Diets with more than 10 foods': 'Dieta para mejorar la salud renal', 'Creator': 'Maria Ines Paz Bueno'}
{'Diets with more than 10 foods': 'Dieta sin maiz para aliviar alergias', 'Creator': 'Maria Ines Paz Bueno'}
{'Diets with more than 10 foods': 'Dieta sin gluten para niños', 'Creator': 'Pablo Daniel Bouzon'}
{'Diets with more than 10 foods': 'Dieta antiinflamatoria para mejorar la salud', 'Creator': 'Pablo Daniel Bouzon'}
```

Figure 16: dietsWithMoreFood() output

```
Food with more Protein than the average:

{'Name': 'Whole-grain bread', 'Protein': 10.4}
{'Name': 'Reese's Peanut Butter Cups', 'Protein': 10.5}
{'Name': 'Bagel', 'Protein': 11.0}
{'Name': 'Millet', 'Protein': 11.0}
{'Name': 'Cottage cheese', 'Protein': 11.1}
{'Name': 'Ricotta cheese', 'Protein': 11.3}
{'Name': 'Soybeans', 'Protein': 12.0}
{'Name': 'Wheat whole grain', 'Protein': 12.6}
{'Name': 'Buckwheat', 'Protein': 13.3}
{'Name': 'Quinoa', 'Protein': 14.1}
{'Name': 'Feta cheese', 'Protein': 14.2}
{'Name': 'Kamut', 'Protein': 14.3}
{'Name': 'Spelt raw', 'Protein': 14.6}
{'Name': 'Squid', 'Protein': 15.6}
{'Name': 'Wheat bran', 'Protein': 15.6}
{'Name': 'Sausage', 'Protein': 16.1}
{'Name': 'Wild boar sausage', 'Protein': 16.1}
{'Name': 'Einkorn wheat', 'Protein': 16.5}
{'Name': 'Soybeans cooked', 'Protein': 16.6}
{'Name': 'Haddock', 'Protein': 17.1}
{'Name': 'Beef ground 85% lean meat 15% fat cooked pan-broiled', 'Protein': 17.34}
{'Name': 'Beef ground 85% lean meat 15% fat raw', 'Protein': 17.34}
{'Name': 'Zebra fish', 'Protein': 17.4}
{'Name': 'Carp', 'Protein': 17.5}
{'Name': 'Monkfish', 'Protein': 17.5}
{'Name': 'Catfish', 'Protein': 18.0}
{'Name': 'Herring', 'Protein': 18.0}
{'Name': 'Shad', 'Protein': 18.0}
{'Name': 'Whitefish', 'Protein': 18.0}
{'Name': 'Flounder', 'Protein': 18.1}
{'Name': 'Halibut', 'Protein': 18.2}
{'Name': 'Eel', 'Protein': 18.4}
{'Name': 'Cod', 'Protein': 18.5}
{'Name': 'Tilefish', 'Protein': 18.5}
{'Name': 'Mackerel', 'Protein': 18.6}
{'Name': 'Zander', 'Protein': 18.6}
```

Figure 17: moreThanAvg() output

```
Food with more Protein than the average:

{'Name': 'Whole-grain bread', 'Protein': 10.4}
{'Name': 'Reese's Peanut Butter Cups', 'Protein': 10.5}
{'Name': 'Bagel', 'Protein': 11.0}
{'Name': 'Millet', 'Protein': 11.0}
{'Name': 'Cottage cheese', 'Protein': 11.1}
{'Name': 'Ricotta cheese', 'Protein': 11.3}
{'Name': 'Soybeans', 'Protein': 12.0}
{'Name': 'Wheat whole grain', 'Protein': 12.6}
{'Name': 'Buckwheat', 'Protein': 13.3}
{'Name': 'Quinoa', 'Protein': 14.1}
{'Name': 'Feta cheese', 'Protein': 14.2}
{'Name': 'Kamut', 'Protein': 14.3}
{'Name': 'Spelt raw', 'Protein': 14.6}
{'Name': 'Squid', 'Protein': 15.6}
{'Name': 'Wheat bran', 'Protein': 15.6}
{'Name': 'Sausage', 'Protein': 16.1}
{'Name': 'Wild boar sausage', 'Protein': 16.1}
{'Name': 'Einkorn wheat', 'Protein': 16.5}
{'Name': 'Soybeans cooked', 'Protein': 16.6}
{'Name': 'Haddock', 'Protein': 17.1}
{'Name': 'Beef ground 85% lean meat 15% fat cooked pan-broiled', 'Protein': 17.34}
{'Name': 'Beef ground 85% lean meat 15% fat raw', 'Protein': 17.34}
{'Name': 'Zebra fish', 'Protein': 17.4}
{'Name': 'Carp', 'Protein': 17.5}
{'Name': 'Monkfish', 'Protein': 17.5}
{'Name': 'Catfish', 'Protein': 18.0}
{'Name': 'Herring', 'Protein': 18.0}
{'Name': 'Shad', 'Protein': 18.0}
{'Name': 'Whitefish', 'Protein': 18.0}
{'Name': 'Flounder', 'Protein': 18.1}
{'Name': 'Halibut', 'Protein': 18.2}
{'Name': 'Eel', 'Protein': 18.4}
{'Name': 'Cod', 'Protein': 18.5}
{'Name': 'Tilefish', 'Protein': 18.5}
{'Name': 'Mackerel', 'Protein': 18.6}
{'Name': 'Zander', 'Protein': 18.6}
```

Figure 18: `filterByInfoAndOrigin()` output

<code>_id: 1</code>
<code>Description: "Dieta baja en grasas"</code>
<code>Creator: 3</code>
<code>_id: 2</code>
<code>Description: "Dieta sin lactosa para veganos"</code>
<code>Creator: 2</code>
<code>_id: 3</code>
<code>Description: "Dieta rica en proteínas para deportistas"</code>
<code>Creator: 1</code>
<code>_id: 4</code>
<code>Description: "Dieta vegetariana equilibrada"</code>
<code>Creator: 4</code>

Figure 19: Diet Collection in MongoDB

Before we add the data to the collections we have to do a change in the JSON files. As MongoDB always renames the primary key with ”\_id”, in order to relate the primary key in SQL with the primary key in MongoDB we have renamed the primary key in the JSON files as ”\_id”.

These is the appearance of the different collections in MongoDB: (Fig. 19-24)

## 0.7 NoSQL Database Queries and Pipelines

In this part of the project we are going to translate the different SQL queries we designed for our database into MongoDB language.

In the first MongoDB query we see the following:

```
[
  {
    $match: {
      Protein: {
        $gt: 20,
      },
    },
  },
  {
    $match: {
      TypeOfFood: "meat",
    },
  }
]
```

<pre> _id: 1 Name: "Asparagus" Kcal: 20 KJ: 84 Fat: 0.2 Protein: 2.2 Carbohydrates: 3.9 Fiber: 2.1 Sodium: 2 Potassium: 202 Calcium: 24 Iron: 1.1 TypeOfFood: "vegetables" </pre>
<pre> _id: 2 Name: "Bell Pepper" Kcal: 20 KJ: 84 Fat: 0.2 Protein: 0.9 Carbohydrates: 4.6 Fiber: 1.7 Sodium: 3 Potassium: 211 </pre>

Figure 20: Food Collection in MongoDB

<pre> _id: ObjectId('6438338635aa65b68feebd89') idDiet: 26 idFood: 1 </pre>
<pre> _id: ObjectId('6438338635aa65b68feebd8a') idDiet: 46 idFood: 1 </pre>
<pre> _id: ObjectId('6438338635aa65b68feebd8b') idDiet: 30 idFood: 4 </pre>
<pre> _id: ObjectId('6438338635aa65b68feebd8c') idDiet: 1 idFood: 5 </pre>

Figure 21: ListOfFood Collection in MongoDB

<b>_id:</b> 1
<b>FullName:</b> "Pablo Moreno Garcia-Espina"
<b>Speciality:</b> "Deporte"

---

<b>_id:</b> 2
<b>FullName:</b> "Susana Rocio Fernandez Giacomassi"
<b>Speciality:</b> "Descenso"

---

<b>_id:</b> 3
<b>FullName:</b> "Ismael Navas Delgado"
<b>Speciality:</b> "Deporte"

---

<b>_id:</b> 4
<b>FullName:</b> "Antonio Rodriguez Horcas"
<b>Speciality:</b> "Ascenso"

Figure 22: Nutritionist Collection in MongoDB

<b>_id:</b> "02358124Z"
<b>Name:</b> "Carmen"
<b>LastNames:</b> "Ruiz Soto"
<b>BirthDate:</b> "1952-08-08"
<b>State:</b> "descenso"
<b>WeightKG:</b> 89.71
<b>Height:</b> 1.86
<b>Telephone:</b> 608765432
<b>Intolerances:</b> 0
<b>Doctor:</b> 3
<b>CurrentDiet:</b> 44

---

<b>_id:</b> "05622134W"
<b>Name:</b> "Pilar"
<b>LastNames:</b> "Ortega Jimenez"
<b>BirthDate:</b> "1959-12-21"
<b>State:</b> "ascenso"
<b>WeightKG:</b> 161.1
<b>Height:</b> 1.91
<b>Telephone:</b> 650123456
<b>Intolerances:</b> 0
<b>Doctor:</b> 7
<b>CurrentDiet:</b> 27

Figure 23: Patient Collection in MongoDB



<code>_id: "cereals and bread"</code> <code>Origin: "processed"</code>
<code>_id: "chocolates and sweets"</code> <code>Origin: "processed"</code>
<code>_id: "dairy"</code> <code>Origin: "animal"</code>
<code>_id: "fish"</code> <code>Origin: "animal"</code>
<code>_id: "fruits"</code> <code>Origin: "vegetal"</code>
<code>_id: "legumes"</code> <code>Origin: "vegetal"</code>

Figure 24: TypeOfFood Collection in MongoDB

```

    },
  },
  {
    $project: {
      _id: 0,
      Name: 1,
      Protein: 1,
    },
  },
]

```

Firstly we have extracted all of the food with protein levels higher than twenty grams, and then we have extracted out of them the "meat" food. In order to obtain the same query as in SQL, we have made a project in order to stay just with the fields name and protein.

In the second query we have:

```

[
  {
    $lookup: {
      from: "Patient",
      localField: "_id",
      foreignField: "Doctor",
      as: "patients",
    },
  },
  {
    $match: {
      FullName: "Pablo Moreno Garcia-Espina",
    },
  },
  {
    $project: {
      _id: 0,

```

```

        "patients.Name": 1,
    },
},
]

```

We make a join from the table Nutritionist to the table Patient relating the identifier field and the Doctor one. Then we filter the FullName avoiding all excepting "Pablo Moreno Garcia-Espina" and finally we project the patient's name.

The third query in MongoDB is:

```

[
  {
    $lookup: {
      from: "Nutritionist",
      localField: "Creator",
      foreignField: "_id",
      as: "nutritionist",
    },
  },
  {
    $match: {
      "nutritionist.FullName":
        "Susana Rocio Fernandez Giacomassi",
    },
  },
  {
    $project: {
      _id: 0,
      Description: 1,
    },
  },
]

```

We join the table Diet with Nutritionist by relating the fields Creator in Diet and identifier in Nutritionist. Then we remain only with the diets made by the

nutritionist "Susana Rocio Fernandez Giacomassi" and we project the diet's description.

The fourth query is:

```
client = MongoClient('mongodb+srv://susanfg:bS23eImV0TRer7Xe
@cluster0.f34dy4x.mongodb.net/test')
result = client['NutritionalDB']['Diet'].aggregate([
    {
        '$lookup': {
            'from': 'Nutritionist',
            'localField': 'Creator',
            'foreignField': '_id',
            'as': 'nutritionistDoctor'
        }
    }, {
        '$lookup': {
            'from': 'ListOfFood',
            'localField': '_id',
            'foreignField': 'idDiet',
            'as': 'foods'
        }
    }, {
        '$match': {
            'foods': {
                '$gt': {
                    '$size': 10
                }
            }
        }
    }, {
        '$project': {
            'Diets with more than ten foods': '$Description',
            'Creator': '$nutritionistDoctor.FullName'
```

```
        }  
    }  
    ])
```

Here we are doing an aggregation from the Diet collection, first we use a lookup to create an array with the nutritionists of each diet. Then we do another lookup with ListOfFood to have all the foods of a diet in the foods array. Then we use match to filter the arrays food with size greater than 10. Lastly, we project the description of the diets and their creator. Basically, we are searching the diets with more than ten foods.

# Contents

0.1	Initial project description . . . . .	3
0.2	RDBMS Deployment . . . . .	3
0.3	Query Design . . . . .	5
0.4	RDBMS Optimization . . . . .	8
0.5	Python . . . . .	13
0.6	NoSQL Database Design . . . . .	17
0.7	NoSQL Database Queries and Pipelines . . . . .	21
<b>1</b>	<b>Introduction</b>	<b>31</b>
1.1	Motivation . . . . .	31
1.2	Objectives . . . . .	32
1.3	Structure of the document . . . . .	33
1.4	Technologies used . . . . .	33
<b>2</b>	<b>Conclusions and Futures Lines of Research</b>	<b>35</b>
2.1	Conclusions . . . . .	35
2.2	Future lines of Research . . . . .	35
<b>3</b>	<b>Conclusiones y Líneas Futuras</b>	<b>37</b>
3.1	Conclusiones . . . . .	37
3.2	Líneas Futuras . . . . .	37
<b>Appendix A Installation</b>		
	<b>Manual</b>	<b>39</b>



# 1

# Introduction

## 1.1 Motivation

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien.



Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

## 1.2 Objectives

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit

amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

### **1.3 Structure of the document**

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

### **1.4 Technologies used**

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.



# Conclusions and Futures Lines of Research

## 2.1 Conclusions

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

## 2.2 Future lines of Research

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna

dictum turpis accumsan semper.

# Conclusiones y Líneas Futuras

## 3.1 Conclusiones

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

## 3.2 Líneas Futuras

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.



# Appendix A

# Installation

# Manual

**Requirements:**





UNIVERSIDAD  
DE MÁLAGA

| **uma.es**

E.T.S. DE INGENIERÍA INFORMÁTICA

E.T.S de Ingeniería Informática  
Bulevar Louis Pasteur, 35  
Campus de Teatinos  
29071 Málaga