



Statistical Methods for Database Integration

Examination

DATABASES

The exam consists in two parts:

- 1) PART A: **The exam is closed-book, closed-notes;**
- 2) PART B: **You are allowed to use lecture and labs notes.**

Each questions is assigned points expressed in cents.

PART A

Ex. 1 “Bologna tourism office offers visitors to explore <http://bolognawelcome.com/en/experiences> web site, to search for experiences to live in”. Consider cards in the picture and report displayed data in a semi-structured data model (XML Language). **Describe at least two cards!**

(a) (10 points)

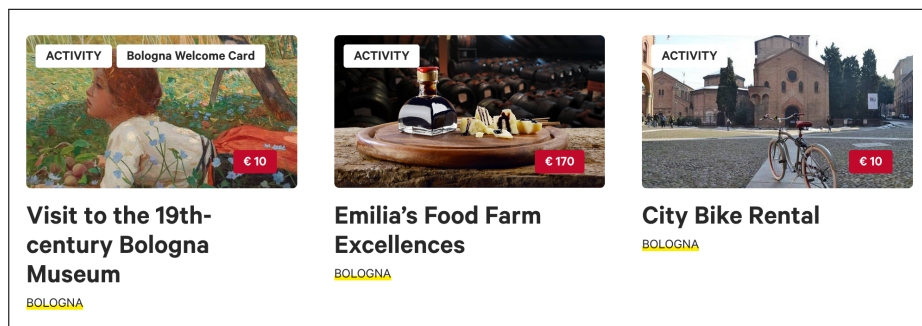


Figure 1: source: **The experience of Bologna - Bologna Welcome**



Sol.:

```
<tourism-experiences>
  <year> "2023" </year>
  <cards-list>
    <card>
      <title> Emilia's Food Farm Excellences </title>
      <location> Bologna </location>
      <type> ACTIVITY </type>
      <price currency="Euro"> 170 </price>
    </card>
    <card>
      <title> City Bike Rental </title>
      <location> Bologna </location>
      <type> ACTIVITY </type>
      <price currency="Euro"> 10 </price>
    </card>
  </cards-list>
</tourism-experiences>
```

- (c) (10 points) We are able to gather data from a web page by scraping process. By inspection of part of the web site at the point (a) we have the text document below. Basically which language describes these data? How we can locate and gather data described at the point (a) (provide examples)?

```
<div class="card event-card">
  <app-card>
    <a title="City Bike Rental" href="/en/experiences/315118/City-Bike-Rental">
      <span class="img is-wide">
        
      <span class="badges"> <flex>
        <span class="badge">ACTIVITY</span>
      </span>
      <span class="badges-price"> <flex>
        <span>
          <span class="badge">
            <span>€ 10</span>
          </span>
        </span>
      </span>
      <span class="title">City Bike Rental</span>
    <span class="highlighted-badges">
      <label>
        <span>
          ::before
            Bologna
          </span>
        </label>
        ::after
      </span>
    </a>
  </app-card>
</div>
```



Sol.:

- The web page is written by HTML language
- Data could be gathered by **class** name of sections:
 - “.title”, City Bike Rental
 - “.badge”, ACTIVITY
 - “.badges-price”, € 10
 - “.highlighted-badges”, Bologna

(c) **(optional 10 points)** The table tells the performances of **Bologna** and **Milano** about culture and creativity.

Index Performance	2019	
	Bologna	Milano
Cultural vibrancy	38.3	17.2
Creative economy	34.8	37.2
Enabling environment	25	28.9

Table 1: Cultural and creatives cities

Create by means of a SQL statement a table holding the given data.

Sol.:

```
CREATE TABLE cccity(
  year CHAR(4),
  city VARCHAR(50),
  index_per VARCHAR(100),
  value NUMERIC(4,1),
  PRIMARY KEY(year, city, index_per)
);
```

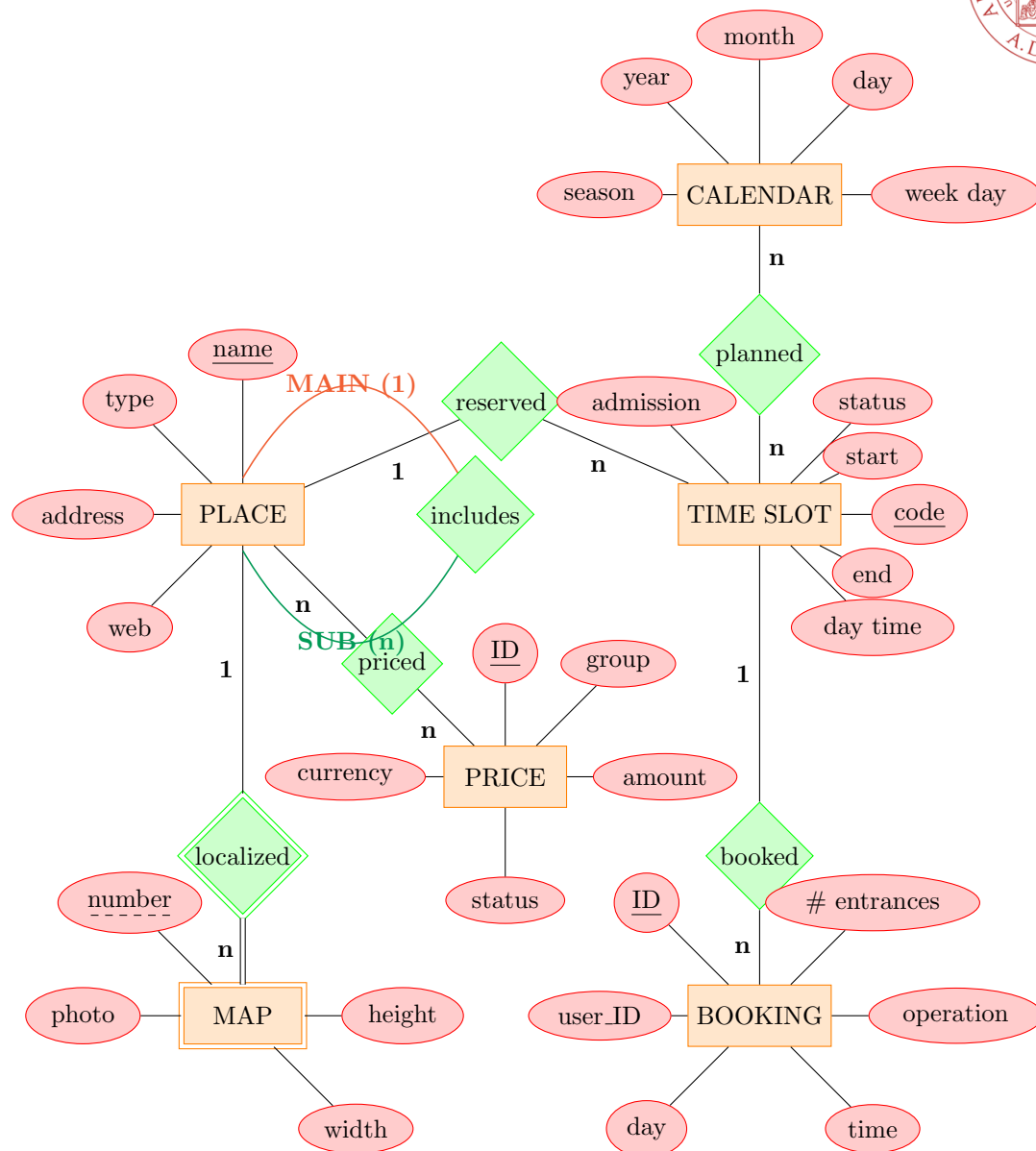
**Es. 2 - Data Modeling**

- (1) **(35 points)** “Bologna is a relevant city for its historical heritage”. An amazing experience should be easily planned, best fitting custom interests and available time, thanks to the aim of the municipality to realize the app “**experience@Bologna**”.

Draw the E/R diagram that capture the requirements stated below aiming to develop the database supporting the application.

Use “ID” as key only if strictly necessary.

- (a) **Place** identifies an artistic/historical attraction, with at least information related to the name (assuming uniqueness), the type (i.e, museum, archeological site, historical shop, ...), the address in Bologna, the url to access to the web site (if it exists).
- (b) In order to localize the place, the app should offers one or more **maps**, accessible by the user. Each map is labeled by a number, which is insufficient to uniquely identify a map (it looks that the entity is weak).
- (c) Consider that one place, like an historical building, includes more specific places (artistic attractions), like a lodge, a fountain, a collection of statues, which could be singly visited, therefore this relationship must be tracked.
- (d) Generally entrance is not free, therefore all applicable **prices** must be codified and registered: code, amount, currency and people category price is addressed to (for example, students, senior, architect,), type of services (i.e. basic entrance, basic entrance+expositions).
- (e) The app helps visitors to book in advance their visit, reserving time slots for each attraction. **Slot** [time slot] is codified, with at least information related to day time (morning, afternoon, evening), beginning hour, ending hour, number of entrances admitted, status (available / not available). Slot defines entrances during opening hours.
- (f) Entrances depend not only on slot time by also on opening days, registered in a **calendar** which specifies season (winter, summer), week day, day, month and year.
- (g) Visitors books their visit by the app, selecting and **booking** one slot. This tasks seems a transaction, that is a task that register slot reservation, and specifically user identifier, date and time of the transaction, number of entrances, and the transaction type that could be not only *reserve*, but also *modify*, or *cancel*.



- (2) (Optional: 5 points). Write “relations” of the **Relational Data Model** which describe **Place** and **Map**.

Place(name, type, address, web, main_name)

Map(number, photo, width, eight, Place.name)



PARTE B

Es. 3 - SQL (45 points) Let assume the database “online-market”.

Region(name, description)

Producer(name, description)

Sheet(ID, description, Region.name, Producer.name)

Ingredient(name, description)

Made(Ingredient.name, Sheet.ID)

Menu(name, description, main)

Food(name, weight, unit, label, price, status, Menu.menu_name, Sheet.ID)

Saving(code, type, start, end)

Discounted(Food.name, Food.weight, Food.unit, Saving.code, perc)

GiftBasket(name, description)

BasketCombines(GiftBasket.name, Food.name, Food.weight, Food.unit)

Session(ID, date, hour, network)

Choices(Session.ID, Menu.name, time)

Selection(Session.ID, Food.name, Food.weight, Food.unit, quantity, timestamp)

Questions

- 1) We aim to prepare a special shop window! We would like to exhibit pasta and condiments. Analyze the food sheet description and detect products including the word **pasta**. Identify condiments considering the **main** menu **condiments**. Reports in alphabetical order food name, weight, unit and price.

```
(SELECT Food.name, weight, unit, price
  FROM Food, Sheet
 WHERE Food.sheet_id = Sheet.ID
    AND description LIKE '%pasta%')
UNION
(SELECT Food.name, weight, unit, price
  FROM Food, Menu
 WHERE Food.menu_name = Menu.name
    AND main = 'Condiments')
ORDER BY name;
```



- 2) Marketing aims to promote some products applying a special discount. Products must be a selection of those never discounted (not registered as discounted) and not to be cheap (that is a price smaller than 5 €). Reports in alphabetical order food name, weight, unit and a new column reporting 10% discount.

```
SELECT F.name, weight, unit, price, '10%' AS save
FROM Food F
WHERE NOT EXISTS (SELECT *
                  FROM Discounted D
                  WHERE D.name = F.name
                  AND D.weight = F.weight
                  AND D.unit = F.unit)
AND F.price >= 5
ORDER BY F.name;
```

- 3) We aim to stock some food into the pantry. The goal is to have something of **Condiments** and **Stuffed pasta**. Products weights are expressed in grams or milliliters. Report for each food category (menu), total number of products, the corresponding total weight (now expressed in kilos or litres).

```
SELECT Food.menu_name, COUNT(*) AS nr, SUM(Food.weight/100) AS total
FROM Food JOIN Menu ON Food.menu_name = Menu.name
WHERE Menu.main IN ('Condiments', 'Stuffed Pasta')
AND Food.unit IN ('g', 'ml')
GROUP BY Food.menu_name;
```