



# COVID-19 App

Big Data and Semantic Technologies Course  
2019/2020

Developed by:  
Antonino Durazzo  
Francesco de Pertis

## Table of Contents

1 Introduction	3
1.1 Purpose	3
1.2 Adopted Technologies	3
1.3 Overview	4
2 Implementation	5
2.1 Ontology Schema definition	5
2.2 Ontology Population	6
2.3 Ontology Interrogation and Graphical User Interface	7
2.3.1 Home Tab	7
2.3.2 Trend Tab	9
2.3.3 Comparison Tab	11
3 Conclusions	13

# 1 Introduction

In this chapter, the purpose of the project and adopted technologies will be pointed out.

## 1.1 Purpose

The purpose of the project is building an application that allows a user to display some relevant statistics on the data about the current sanitary emergency, very well known as COVID-19. The idea is to build an ontology in order to give a semantics to open raw data, daily updated and freely distributed by the Department of Civil Protection afferent to the Government of the Italian Republic.

These data represent tests, swabs and cases of COVID-19 and refer to Italian Regions and the entire Italian Country.

## 1.2 Adopted Technologies

Several software technologies have been adopted to implement this project:

- **Protégé** - a free and open-source ontology editor and framework - was used to get a graphical view of our ontology schema.
- **Eclipse IDE** - an integrated development environment - was used to write code to implement the project.
- **Java JDK** was used to develop the different component of the project and the interfaces that let these components to work together.
- **Apache Jena** - a free and open source Java framework for building Semantic Web and Linked Data applications - was used to define, populate and interrogate our ontology.
- **Window Builder** - SWT and Swing Designer plug-in for Eclipse - was used to develop the Graphical User Interface.
- **JFreeChart** - a Java library - was used to develop the charts integrated in the GUI.

## 1.3 Overview

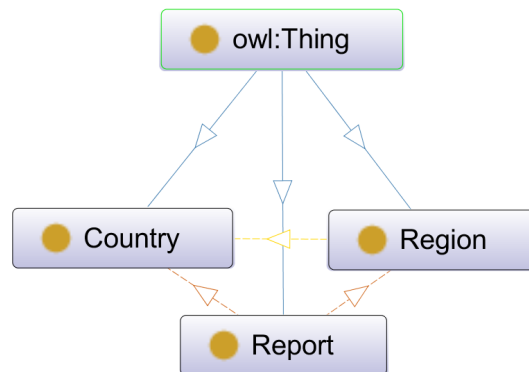
The project was developed entirely in Java language and it is divided in six classes:

- BarChart.java - defines a Bar Chart Object used in Home and Comparison Tabs of the GUI.
- CovidApp.java - defines the queries and the whole GUI.
- dataRegion.java - contains the different datatype values for every region.
- OntModelCovid.java - defines the Ontology Schema and Population.
- parseCSV.java - imports and parses data from CSV dataset
- TimeSeriesChart.java - defines a Time Series Chart Object used in Trend Tab of the GUI.

## 2 Implementation

In this chapter, the implementation phase is pointed out.

### 2.1 Ontology Schema definition



The ontology schema is defined as in the above figure. Three classes called Country, Region and Report are defined as sub-classes of owl:Thing, as requested in OWL praxis. The relations between these Classes are called Object Properties and they represent how these Classes are related to each other. In particular, Region “isPartOf” Country and a Report “belongsTo” only a Country or a Region.

Furthermore, these three sub-classes have some attribute that are called Datatype Proprieties. In particular:

- Country
  - hasName - property connected to a xsd:Name Literal
  - hasPopulation - property connected to a xsd:unsignedInt
- Region
  - hasName - property connected to a xsd:Name Literal
  - hasPopulation - property connected to a xsd:unsignedInt
  - hasCodRegion - property connected to a xsd:unsignedInt
  - hasLatitude - property connected to a xsd:double
  - hasLongitude - property connected to a xsd:double

- Report
  - hasDate - property connected to a xsd:date
  - hasDeceased - property connected to a xsd:unsignedInt
  - hasDischargedHealed - property connected to a xsd:unsignedInt
  - hasHomeIsolation - property connected to a xsd:unsignedInt
  - hasIntensiveCare - property connected to a xsd:unsignedInt
  - hasRecoveredWithSymptoms - property connected to a xsd:unsignedInt
  - hasSwabs - property connected to a xsd:unsignedInt
  - hasTestedCases - property connected to a xsd:unsignedInt
  - hasTotalCases - property connected to a xsd:unsignedInt
  - hasTotalHospitalized - property connected to a xsd:unsignedInt
  - hasTotalPositives - property connected to a xsd:unsignedInt

## 2.2 Ontology Population

Ontology Population consists of creating a sequence of individuals for every class. Country and Region population was made manually due to their small amount. In fact, in the ontology there is only one individual of the Class Country called “Italy” because the dataset under consideration regards only Italian Regions. Furthermore, there are twenty-one individuals of the Class Region that represent nineteen Italian Region and the only two Autonomous provinces of Trento and Bolzano.

Report population was made in an automatic way taking in input the dataset under consideration available to the following link:

<https://raw.githubusercontent.com/pcm-dpc/COVID-19/master/dati-regioni/dpc-covid19-ita-regioni.csv>

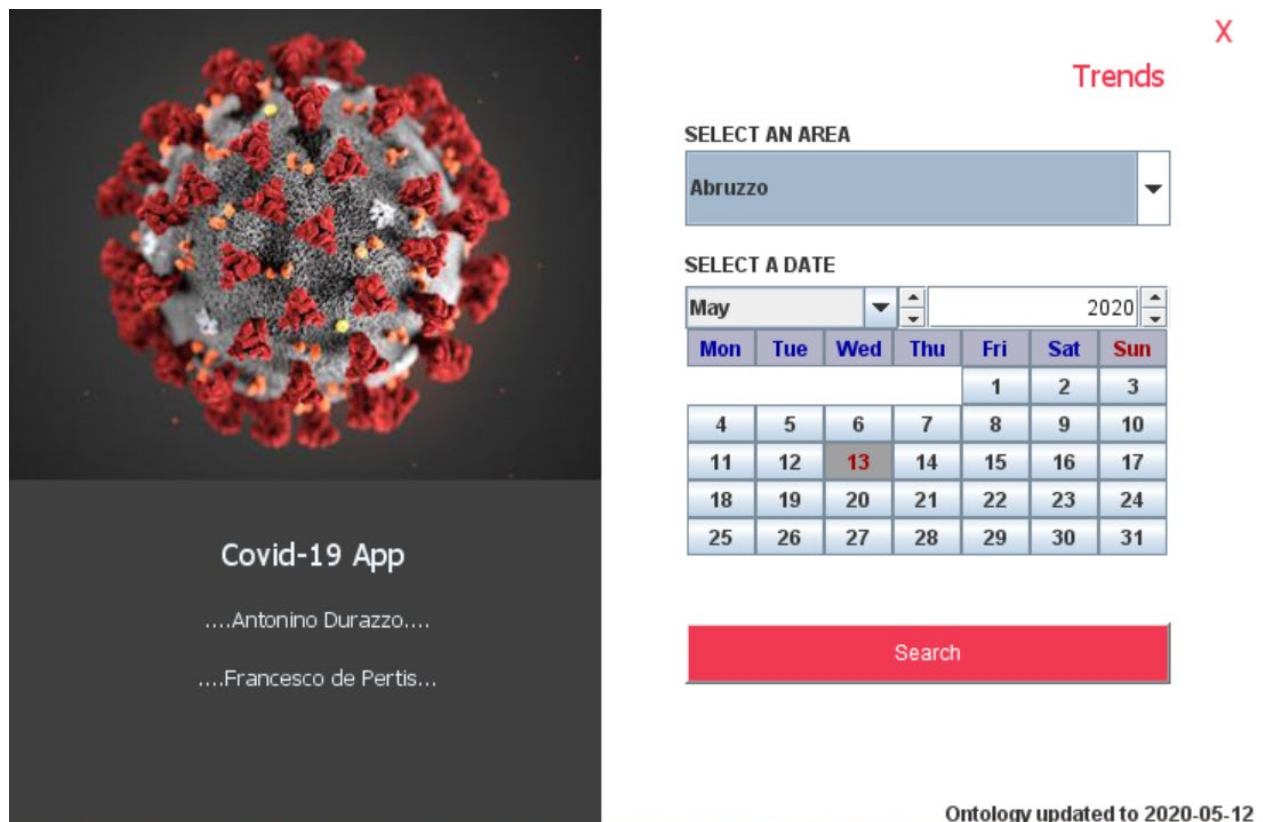
This dataset is automatically downloaded when the application starts and is parsed with the methods defined in parseCSV.java Class. In particular, the program builds a new Report Individual for every row of the dataset creating a number of individuals equal to **number\_of\_regions \* number\_of\_days**. When the whole dataset is parsed. The ontology is written on a file located in ./owl/ontologycovid.owl.

## 2.3 Ontology Interrogation and Graphical User Interface

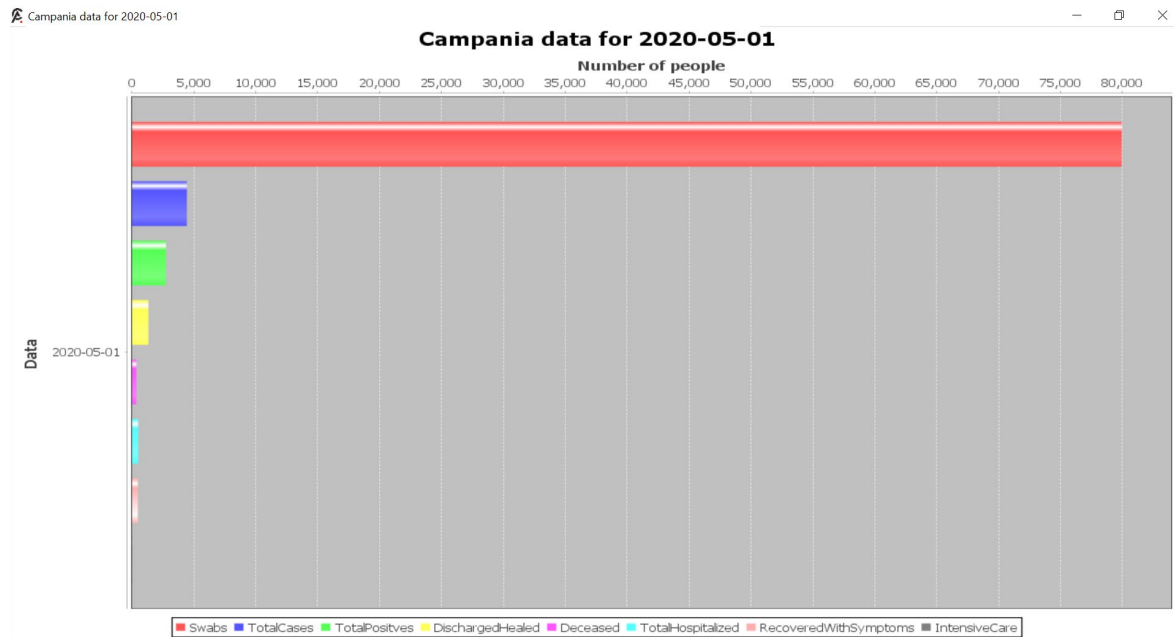
Ontology Interrogation consists of different queries that are sent to the semantic dataset in order to get a particular datum of interest. In the project, the interrogation allows the user to view some important statistics about the data collected by the ontology. Queries let the user get one or more reports and view their datatype properties.

Graphical User Interface, abbreviated GUI, represents a very simple and easy way which the user can send queries to the semantic dataset and view their results. The GUI is developed with Java Swing API and it is divided in three tabs called Home, Trend and Comparison.

### 2.3.1 Home Tab



When the application starts, Home Tab will appear. On the left side, there is a fixed image with the name of the application and the authors' names. On the right side, the interface allows to select an area (Country or Region) and a date. At bottom, there is the Search Button that allows the user to view the Datatype Properties associated to a single Report object related to the specified area and date. By clicking on the Search Button, a Bar Chart will appear in another window.



The query sent to the semantic database is:

PREFIX NS: <<http://www.covidapp.org/ontology#>>

SELECT \*

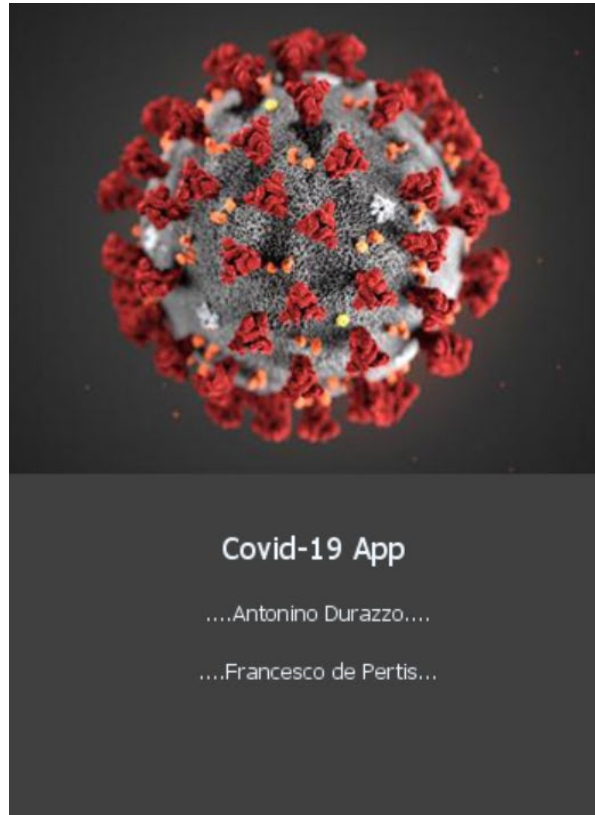
WHERE { ?Report NS:hasDate "date"^^<<http://www.w3.org/2001/XMLSchema#date>> .

?Report NS:belongsTo NS:Region .};

At bottom right, there is a label that show last data available. At top right, there is the Trends Bottom that will take the user to the Trend Tab.



### 2.3.2 Trend Tab



Home Comparison X

SELECT AN AREA

Abruzzo ▼

SELECT AN AREA

▼

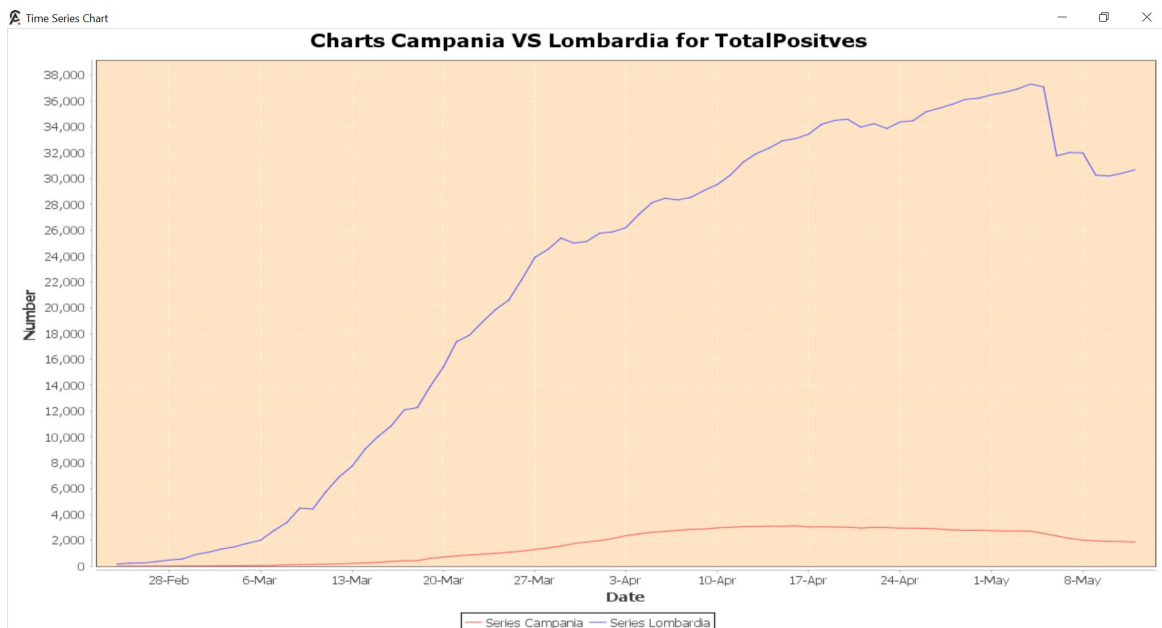
SELECT A TREND

Deceased ▼

Search

Ontology updated to 2020-05-12

Trend Tab allows the user to select one or two areas and a Datatype Property and view a Time Trend for these specified values. By clicking on the Search Button, a Time Series Chart will appear in another window.

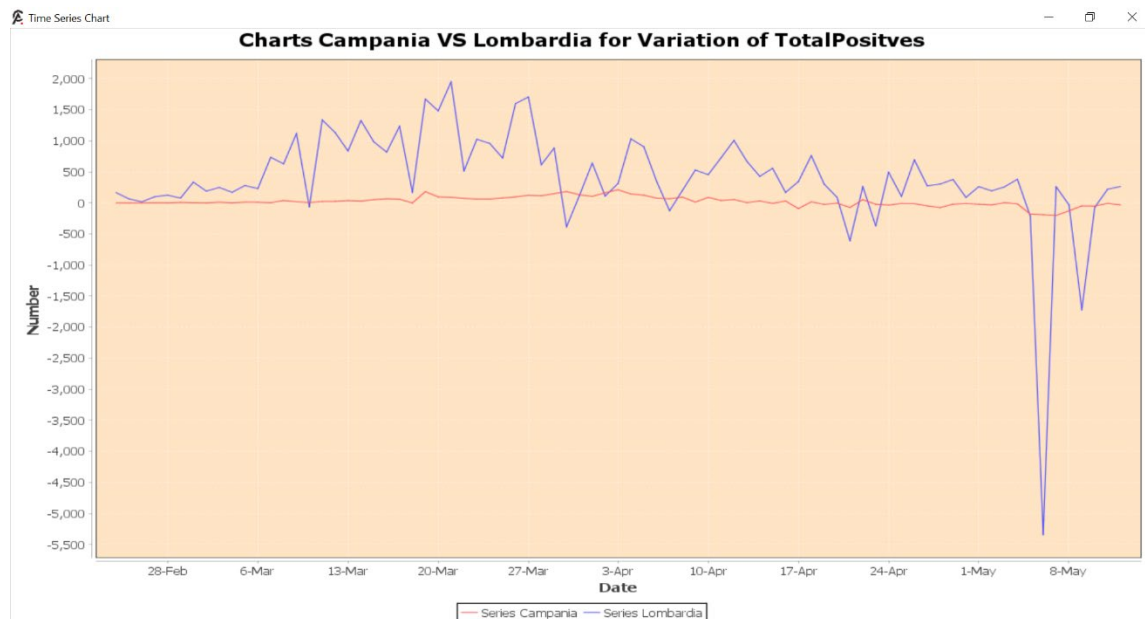


The query sent to the semantic database is:

```
PREFIX NS: <http://www.covidapp.org/ontology#>
SELECT *
WHERE { ?Report NS:belongsTo NS:Region .
        ?Report NS:hasDate ?date . }
ORDER BY ASC(?date) "
```

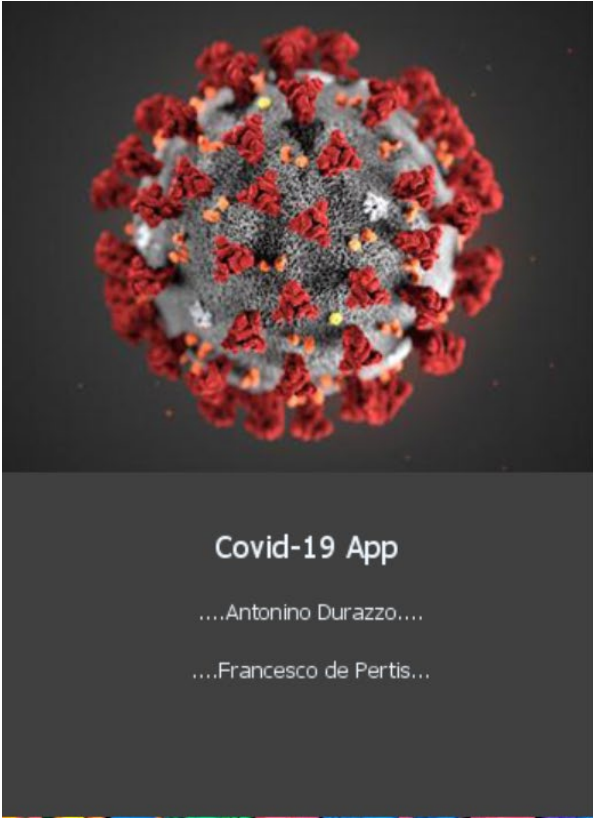
In the selection list, for every Datatype Property, there is the associated daily variation, too. It is obtained by:

**$\text{VarDatatypeProperty}(\text{date}, \text{area}) = \text{DatatypeProperty}(\text{date}, \text{area}) - \text{DatatypeProperty}(\text{date} - 1\text{d}, \text{area})$ .**



At top, on the left there is the Home Bottom to come back to the Home Tab, while, on the right there is the Comparison Button that will take the user to the Comparison Tab.

### 2.3.3 Comparison Tab



**Covid-19 App**

....Antonino Durazzo....

....Francesco de Pertis...

**Trends**

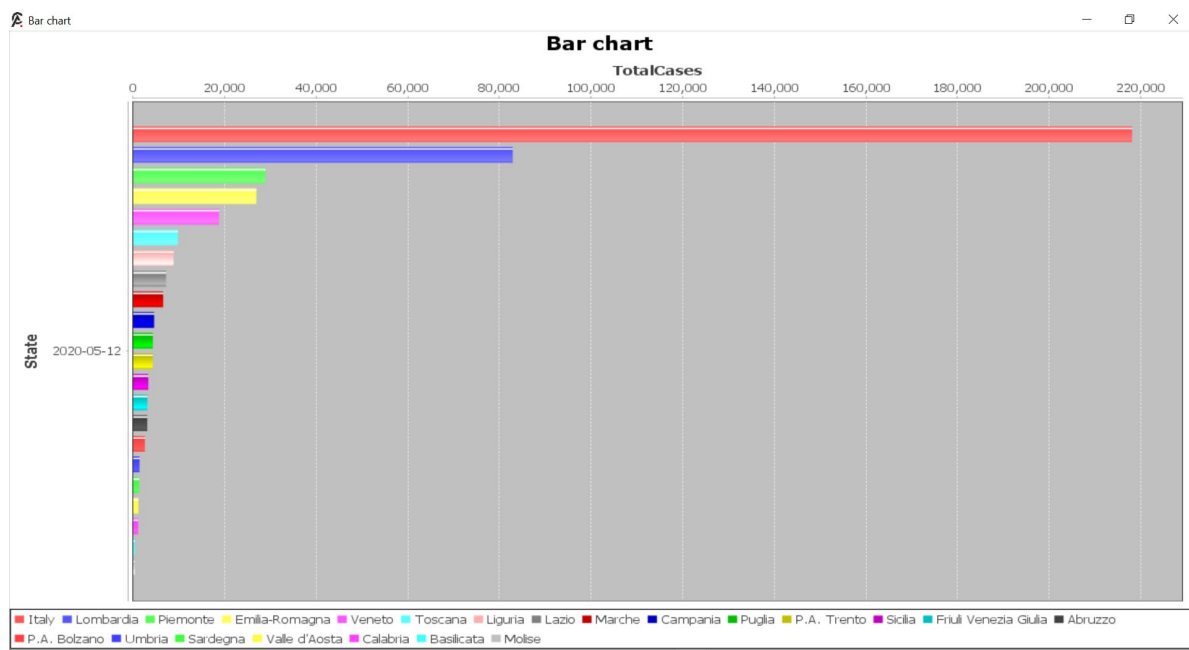
**SELECT A CATEGORY**

PercentOfCases

Search

Ontology updated to 2020-05-12

Comparison Tab allows the user to select a category and view a comparison among all present areas for the last available date. By clicking on the Search Button, a Bar Chart will appear in another window.

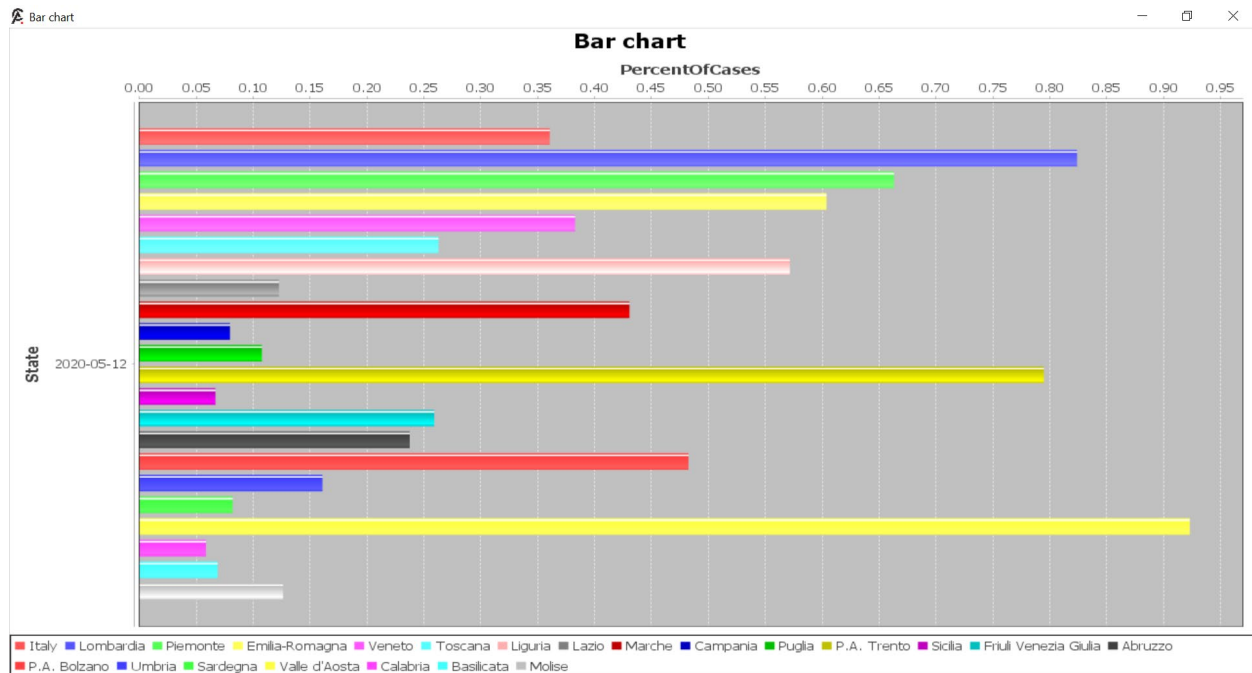


The query sent to the semantic database is:

```
PREFIX NS: <http://www.covidapp.org/ontology#>
SELECT *
WHERE { ?Report NS:hasDate "date"^^<http://www.w3.org/2001/XMLSchema#date>.
      ?Report NS:hasProperty ?value . }"
ORDER BY DESC(?value);
```

In the selection list, in addition to all the Datatype Properties, there is also the percentual values of Total cases over the population. This value is obtained by:

$$\text{PercentOfCases}(\text{area}) = \frac{\text{TotalCases}(\text{area})}{\text{Population}(\text{area})}$$



At top left, there is the Trends Bottom to come back to the Trend Tab.

### 3 Conclusions

This application let the user interact with data and get some statistics about them in a very simple way. In this way, this application can be used also by users with no coding skills. In a future development, some features could be improved such as:

- Enlarging Ontology schema in order to incorporate other Classes of data.
- Increasing the Ontology Population with other Countries and relative Regions.
- Creating a new Tab for advanced users to perform some queries according to SPARQL syntax.