POLITECNICO
MILANO 1863

Systems and Methods for Big and Unstructured Data

PROJECT WORK - DELIVERY 3

TEAM MEMBERS:

Christian Onorati - 10616952
Andrea Giuffrida - 10643540
Alessandro D'Onofrio - 10864445
Nicholas Nicolis - 10867841
Silvia Marino -10688672

Academic Year: 2021-22

Contents

1 | SPECIFICATION AND HYPOTHESIS

1.1. SPECIFICATION

This project is focused on the design and implementation of a database for storing vaccination data, through Elasticsearch and Kibana framework.

The data include information about how many people where vaccinated in a certain region with first, second, or additional dose. The data are also grouped by the age range, and also consider the person that were previously infected by the virus.

The use of Kibana permits to explore the data with much more details, thanks to the use of graphs and maps, in order to extract information in an easier and complete way.

As optional task, 2 additional datasets, taken from the same repository, were integrated to the database and 2 additional queries were performed over these datasets.

[Specification about 2nd NoSQL DB]
redis

1.2. HYPOTHESIS

2 | DATASET DESCRIPTION

The dataset is directly taken from a repository on Github.
Link to the repository: https://github.com/italia/covid19-opendata-vaccini

As the README files says, the repository contains data regarding delivery and administration of the vaccine, in all the Italian nation. The repository contains 10 tables, but for our purpose we used the "somministrazioni-vaccini-latest.csv",which contains data on the daily administration of vaccines, divided by region and age range.

In order to import this data in Elasticsearch, we applied slightly modification to the automatic mapping. First of all, a new field "ISO3166" was added, which had been necessary for the map in Kibana. As a matter of fact, "codice_regione_ISTAT", mapped as long or text, wasn't read correctly from Kibana because of the missing 0 in front of the numbers in the range 1-9, resulting in missing some regions from the map. It was then decided to convert "codice_regione_ISTAT" into "ISO3166" codification, which didn't create any problem for Kibana.

Furthermore, the field "nome_area" was mapped into text instead of keywords, since it represents a name and not a specific code.

```
1  {
2    "properties": {
3      "@timestamp": {
4        "type": "date"
5      },
6      "ISO3166": {
7        "type": "keyword"
8      },
9      "ISTAT": {
10       "type": "long"
11     },
12     "area": {
```

```
13        "type": "keyword"
14      },
15      "codice_NUTS1": {
16        "type": "keyword"
17      },
18      "codice_NUTS2": {
19        "type": "keyword"
20      },
21      "codice_regione_ISTAT": {
22        "type": "long"
23      },
24      "data_somministrazione": {
25        "type": "date",
26        "format": "iso8601"
27      },
28      "dose_addizionale_booster": {
29        "type": "long"
30      },
31      "fascia_anagrafica": {
32        "type": "keyword"
33      },
34      "fornitore": {
35        "type": "keyword"
36      },
37      "nome_area": {
38        "type": "text"
39      },
40      "pregressa_infezione": {
41        "type": "long"
42      },
43      "prima_dose": {
44        "type": "long"
45      },
46      "seconda_dose": {
47        "type": "long"
48      },
49      "sesso_femminile": {
```

```
50        "type": "long"
51      },
52      "sesso_maschile": {
53        "type": "long"
54      }
55    }
56 }
```

Here is a summary of the dataset fields, their respective data type in ElasticSearch and a short description of what they represent.

FIELD NAME	TYPE (ES)	DESCRIPTION
data_somministrazione	date	Vaccination date
codice_ISO3166	keyword	ISO3166 code of region
codice_ISTAT	keyword	Sequential number of the region
fornitore	keyword	Complete name of vaccine producer
area	keyword	Initials of the delivery region
fascia_anagrafica	text	Age group of people that received the vaccine
sesso_maschile	long	Number of male citizens that received the vaccine per day, region and age group
sesso_femminile	long	Number of female citizens that received the vaccine per day, region and age group
prima_dose	long	Number of first doses
seconda_dose	long	Number of second doses
pregressa_infezione	long	Number of citizens that received the vaccine following a COVID infection in the previous 3-6 months
dose_addizionale_booster	long	Number of booster doses
codice_NUTS1	keyword	Nomenclature of Territorial Units for Statistics (NUTS 1)
codice_NUTS2	keyword	Nomenclature of Territorial Units for Statistics (NUTS 2)
codice_regione_ISTAT	long	ISTAT region code
nome_area	text	Standard area denomination
ISO3166	keyword	https://en.wikipedia.org/wiki/ISO_3166

3 | QUERIES AND COMMANDS

3.1. QUERIES

3.1.1. Monthly moving average of administered first doses in Italy

Here we compute the moving average of first doses administered in Italy over a time window of 1 month.

```
GET /somministrazione_vaccini/_search
{
  "size": 0,
  "aggs": {
    "time_span": {
      "date_histogram": {
        "field": "data_somministrazione",
        "calendar_interval": "1M"
      },
      "aggs": {
        "vaccine_sum": {
          "sum": { "field": "prima_dose" }
        },
        "movavg": {
          "moving_avg": { "buckets_path": "vaccine_sum" }
        }
      }
    }
  }
}
```

Here's the result of the above query, limited to the first three months (December 2020, January 2021 and February 2021):

```
"aggregations" : {
  "time_span" : {
    "buckets" : [
      {
        "key_as_string" : "2020-12-01T00:00:00.000Z",
        "key" : 1606780800000,
        "doc_count" : 416,
        "vaccine_sum" : {
          "value" : 40670.0
        }
      },
      {
        "key_as_string" : "2021-01-01T00:00:00.000Z",
        "key" : 1609459200000,
        "doc_count" : 6025,
        "vaccine_sum" : {
          "value" : 1353293.0
        },
        "movavg" : {
          "value" : 40670.0
        }
      },
      {
        "key_as_string" : "2021-02-01T00:00:00.000Z",
        "key" : 1612137600000,
        "doc_count" : 8464,
        "vaccine_sum" : {
          "value" : 1578034.0
        },
```

Figure 3.1

3.1.2. Total number of administered vaccines by region in a given day

```
1  GET /somministrazione_vaccini/_search
2  {
3    "size": 0,
4    "query": {
5      "bool" : {
6        "filter": {
7          "term" : { "data_somministrazione": "<date>" }
8        }
9      }
```

```
10    },
11    "aggs": {
12      "by_region": {
13        "terms": {
14          "field": "area"
15        },
16        "aggs": {
17          "total_vaccines": {
18            "sum": {
19              "script": "doc['prima_dose'].value + doc['
                  seconda_dose'].value + doc['
                  dose_addizionale_booster'].value + doc['
                  pregressa_infezione'].value"
20            }
21          }
22        }
23      }
24    }
25  }
```

Here's the result of the above query using "2021-12-21" as input date (limited to 3 regions):

```
"aggregations" : {
  "by_region" : {
    "doc_count_error_upper_bound" : 0,
    "sum_other_doc_count" : 206,
    "buckets" : [
      {
        "key" : "LAZ",
        "doc_count" : 26,
        "total_vaccines" : {
          "value" : 57592.0
        }
      },
      {
        "key" : "FVG",
        "doc_count" : 25,
        "total_vaccines" : {
          "value" : 10824.0
        }
      },
      {
        "key" : "EMR",
        "doc_count" : 23,
        "total_vaccines" : {
          "value" : 41797.0
        }
      },
```

Figure 3.2

3.1.3.   Total number of administered vaccines by brand

```
1  GET /somministrazione_vaccini/_search
2  {
3    "size": 0,
4    "aggs": {
5      "fornitore_aggs": {
6        "terms": {
7          "field": "fornitore",
8          "size": 10
9        },
10       "aggs": {
11         "total_vaccines": {
12           "sum": {
13             "script": "doc['prima_dose'].value + doc['
                    seconda_dose'].value + doc['
                    dose_addizionale_booster'].value + doc['
                    pregressa_infezione'].value"
```

```
14              }
15            }
16          }
17        }
18      }
19 }
```

Here's the result of the above query, limited to 3 vaccine brands:


```
"aggregations" : {
  "fornitore_aggs" : {
    "doc_count_error_upper_bound" : 0,
    "sum_other_doc_count" : 0,
    "buckets" : [
      {
        "key" : "Pfizer/BioNTech",
        "doc_count" : 66128,
        "total_vaccines" : {
          "value" : 7.4880249E7
        }
      },
      {
        "key" : "Moderna",
        "doc_count" : 54393,
        "total_vaccines" : {
          "value" : 2.0212252E7
        }
      },
      {
        "key" : "Vaxzevria (AstraZeneca)",
        "doc_count" : 27854,
        "total_vaccines" : {
          "value" : 1.2161164E7
        }
      },
```

Figure 3.3

3.1.4. Percentage of male VS female vaccinations by region

```
1 GET /somministrazione_vaccini/_search
2 {
3   "size": 0,
4   "aggs": {
5     "region": {
6       "terms": {
7         "field": "area"
8       },
```

```
 9        "aggs": {
10      "male_vaccines": {
11        "sum": {
12          "field": "sesso_maschile"
13        }
14      },
15      "female_vaccines": {
16        "sum": {
17          "field": "sesso_femminile"
18        }
19      },
20      "total_vaccines" : {
21        "sum": {
22          "script": {
23            "source": "doc['prima_dose'].value + doc['
                 seconda_dose'].value + doc['
                 dose_addizionale_booster'].value + doc['
                 pregressa_infezione'].value"
24          }
25        }
26      },
27      "male_percentage" : {
28        "bucket_script": {
29          "buckets_path": {
30            "buck1" : "male_vaccines",
31            "buck2" : "total_vaccines"
32          },
33          "script": "params.buck1 / params.buck2 * 100",
34          "format": "long"
35        }
36      },
37      "female_percentage" : {
38        "bucket_script": {
39          "buckets_path": {
40            "buck1" : "female_vaccines",
41            "buck2" : "total_vaccines"
42          },
```

```
43          "script": "params.buck1 / params.buck2 * 100",
44          "format": "long"
45        }
46      }
47    }
48      }
49    }
50  }
```

Here's the result of the above query, limited to Lombardy region:

```
{
  "key" : "LOM",
  "doc_count" : 9461,
  "female_vaccines" : {
    "value" : 9852684.0
  },
  "total_vaccines" : {
    "value" : 1.9142667E7
  },
  "male_vaccines" : {
    "value" : 9289983.0
  },
  "male_percentage" : {
    "value" : 48.53024398324434,
    "value_as_string" : "long49"
  },
  "female_percentage" : {
    "value" : 51.46975601675566,
    "value_as_string" : "long51"
  }
},
```

Figure 3.4

3.1.5. Vaccines administered to people exposed to COVID for each year

```
1  GET /somministrazione_vaccini/_search
2  {
3    "size": 0,
4    "aggs": {
```

```
 5     "year division": {
 6       "date_histogram": {
 7         "field": "data_somministrazione",
 8         "interval": "year"
 9       },
10       "aggs": {
11         "total_vax_year": {
12           "sum": {
13             "script": {
14               "lang": "painless",
15               "source": "doc['prima_dose'].value + doc['
                   seconda_dose'].value + doc['
                   dose_addizionale_booster'].value + doc['
                   pregressa_infezione'].value"
16             }
17           }
18         },
19         "infected_vaccines" : {
20           "sum": {
21             "field": "pregressa_infezione"
22           }
23         },
24         "percent" : {
25           "bucket_script": {
26             "buckets_path": {
27               "buck1" : "infected_vaccines",
28               "buck2" : "total_vax_year"
29             },
30             "script": "params.buck1 / params.buck2 * 100"
31           }
32         }
33       }
34     }
35   }
36 }
```

Here's the result of the above query for years 2020 and 2021. "infected_vaccines" is

the total yearly number of vaccines administered to people already exposed to COVID, "total_vax_year" is the total yearly number of administered vaccines, and "percent" is the percentage of "infected_vaccines" with respect to the "total_vax_year".

```
"year division" : {
  "buckets" : [
    {
      "key_as_string" : "2020-01-01T00:00:00.000Z",
      "key" : 1577836800000,
      "doc_count" : 416,
      "infected_vaccines" : {
        "value" : 54.0
      },
      "total_vax_year" : {
        "value" : 40724.0
      },
      "percent" : {
        "value" : 0.13259994106669287
      }
    },
    {
      "key_as_string" : "2021-01-01T00:00:00.000Z",
      "key" : 1609459200000,
      "doc_count" : 166581,
      "infected_vaccines" : {
        "value" : 1641334.0
      },
      "total_vax_year" : {
        "value" : 1.08908095E8
      },
      "percent" : {
        "value" : 1.507081727946853
      }
    }
```

Figure 3.5

3.1.6.   Number of weekly vaccinations by brand

```
1  GET /somministrazione_vaccini/_search
2  {
3    "size": 0,
4    "aggs": {
5      "time_span": {
6        "date_histogram": {
7          "field": "data_somministrazione",
8          "calendar_interval": "1w"
9        },
10       "aggs": {
11         "fornitore_aggs": {
12           "terms": {
13             "field": "fornitore",
14             "size": 10
```

```
15              },
16            "aggs": {
17              "total_vaccines":{
18                "sum": {
19                  "script": "doc['prima_dose'].value+doc['
                      seconda_dose'].value+doc['
                      dose_addizionale_booster'].value+doc['
                      pregressa_infezione'].value"
20                }
21              }
22            }
23          }
24        }
25      }
26    }
27  }
```

Here's the result of the above query, limited to the week starting on January 4th 2021:

```json
{
  "key_as_string" : "2021-01-04T00:00:00.000Z",
  "key" : 1609718400000,
  "doc_count" : 1236,
  "fornitore_aggs" : {
    "doc_count_error_upper_bound" : 0,
    "sum_other_doc_count" : 0,
    "buckets" : [
      {
        "key" : "Pfizer/BioNTech",
        "doc_count" : 1182,
        "total_vaccines" : {
          "value" : 553357.0
        }
      },
      {
        "key" : "Moderna",
        "doc_count" : 42,
        "total_vaccines" : {
          "value" : 190.0
        }
      },
      {
        "key" : "Janssen",
        "doc_count" : 12,
        "total_vaccines" : {
          "value" : 16.0
        }
      }
    ]
  }
},
```

Figure 3.6

3.1.7. Number of vaccines administered in a given region by age group

```
1  GET /somministrazione_vaccini/_search
2  {
3    "size": 0,
4    "query": {
5      "match": {
6        "nome_area": "<region_name>"
7      }
8    },
9    "aggs": {
10     "by_age_group": {
11       "terms": {
12         "field": "fascia_anagrafica"
13       },
14       "aggs": {
```

```
15          "total_vaccines": {
16            "sum": {
17              "script": "doc['prima_dose'].value + doc['
                  seconda_dose'].value + doc['
                  dose_addizionale_booster'].value + doc['
                  pregressa_infezione'].value"
18            }
19          }
20        }
21      }
22    }
23 }
```

Here's the result of the above query using "Emilia Romagna" as input (limited to age groups 60-69, 70-79 and 50-59):

```
"buckets" : [
  {
    "key" : "60-69",
    "doc_count" : 1153,
    "total_vaccines" : {
      "value" : 1209702.0
    }
  },
  {
    "key" : "70-79",
    "doc_count" : 1136,
    "total_vaccines" : {
      "value" : 1085588.0
    }
  },
  {
    "key" : "50-59",
    "doc_count" : 1081,
    "total_vaccines" : {
      "value" : 1480668.0
    }
  },
```

Figure 3.7

3.1.8. Number of booster doses in the last three months

```
1  GET /somministrazione_vaccini/_search
2  {
3    "size": 0,
4    "query": {
5      "range": {
6        "data_somministrazione": {
7          "gte": "now-3M/M",
8          "lte": "now/M"
9        }
10     }
11   },
12   "aggs": {
13     "total_booster_doses": {
14       "sum": {
15         "field": "dose_addizionale_booster"
16       }
17     }
18   }
19 }
```

Here's the result of the above query:

Figure 3.8

3.1.9. Percentage of each vaccine type per region

The following query computes, for each region, the percentage of first, second, third doses, and vaccines administered to people already exposed to COVID, over the total number of vaccinations in that region.

```
1  GET /somministrazione_vaccini/_search
2  {
3    "size":0,
4    "aggs": {
5      "regions": {
6        "terms": {
7          "field": "area"
8        },
9        "aggs": {
10          "vaccines_per_region" : {
11            "sum": {
12              "script": {
```

```
13              "lang": "painless",
14              "source": "doc['prima_dose'].value + doc['
                    seconda_dose'].value + doc['
                    dose_addizionale_booster'].value + doc['
                    pregressa_infezione'].value"
15            }
16          }
17        },
18        "first_dose_count":{
19          "sum": {
20            "field": "prima_dose"
21          }
22        },
23        "second_dose_count":{
24          "sum": {
25            "field": "seconda_dose"
26          }
27        },
28        "booster_count":{
29          "sum": {
30            "field": "dose_addizionale_booster"
31          }
32        },
33        "healed_people":{
34          "sum": {
35            "field": "pregressa_infezione"
36          }
37        },
38        "first_percent" : {
39          "bucket_script": {
40            "buckets_path": {
41              "buck1" : "first_dose_count",
42              "buck2" : "vaccines_per_region"
43            },
44            "script": "params.buck1 / params.buck2 * 100"
45          }
46        },
```

```
47          "second_percent" : {
48            "bucket_script": {
49              "buckets_path": {
50                "buck1" : "second_dose_count",
51                "buck2" : "vaccines_per_region"
52              },
53              "script": "params.buck1 / params.buck2 * 100"
54            }
55          },
56          "booster_percent" : {
57            "bucket_script": {
58              "buckets_path": {
59                "buck1" : "booster_count",
60                "buck2" : "vaccines_per_region"
61              },
62              "script": "params.buck1 / params.buck2 * 100"
63            }
64          },
65          "ex_infected_percent" : {
66            "bucket_script": {
67              "buckets_path": {
68                "buck1" : "healed_people",
69                "buck2" : "vaccines_per_region"
70              },
71              "script": "params.buck1 / params.buck2 * 100"
72            }
73          }
74        }
75      }
76    }
77 }
```

Here's the result of the above query, limited to the Campania region:

```
{
  "key" : "CAM",
  "doc_count" : 9016,
  "second_dose_count" : {
    "value" : 3956198.0
  },
  "vaccines_per_region" : {
    "value" : 9928949.0
  },
  "booster_count" : {
    "value" : 1529319.0
  },
  "first_dose_count" : {
    "value" : 4281899.0
  },
  "healed_people" : {
    "value" : 161533.0
  },
  "first_percent" : {
    "value" : 43.12540028154037
  },
  "second_percent" : {
    "value" : 39.845083301364525
  },
  "booster_percent" : {
    "value" : 15.402627206565366
  },
  "ex_infected_percent" : {
    "value" : 1.6268892105297348
  }
},
```

Figure 3.9

3.2. COMMANDS

3.2.1. Insert data for a day

This command simulates the insertion of vaccination data of a given region (Lombardy) for a given date (2021-12-28).

```
1 POST _bulk
2 { "index":{ "_index": "somministrazione_vaccini" } }
3 { "data_somministrazione": "2021-12-28", "fornitore": "Pfizer
     Pediatrico", "area": "LOM", "fascia_anagrafica": "05-11",
     "sesso_maschile": 100, "sesso_femminile": 150, "
   prima_dose": 50, "seconda_dose": 50, "pregressa_infezione"
     : 0, "dose_addizionale_booster": 50, "codice_NUTS1": "ITC"
     , "codice_NUTS2": "ITC4", "codice_regione_ISTAT": 3, "
   nome_area": "Lombardia" }
```

```
4  { "index":{ "_index": "somministrazione_vaccini" } }
5  { "data_somministrazione": "2021-12-28", "fornitore": "Pfizer
      /BioNTech", "area": "LOM", "fascia_anagrafica": "12-19", "
      sesso_maschile": 100, "sesso_femminile": 150, "prima_dose"
      : 50, "seconda_dose": 50, "pregressa_infezione": 0, "
      dose_addizionale_booster": 50, "codice_NUTS1": "ITC", "
      codice_NUTS2": "ITC4", "codice_regione_ISTAT": 3, "
      nome_area": "Lombardia"  }
```

3.2.2. Delete data for a day

This command can be used to delete all data regarding a given date (in this case, 2021-12-28).

```
1  POST /somministrazione_vaccini/_delete_by_query
2  {
3    "query": {
4      "match": {
5        "data_somministrazione": "2021-12-28"
6      }
7    }
8  }
```

4 | KIBANA DASHBOARD

We expect you to deliver the dashboard, the dataset (since you could have changed it), and a chapter in the deliverable in which you describe each representation (with its picture). To export your dashboard, you can write "export dashboard" in the search box in your Elasticsearch UI (the one I showed during the lecture). The first result will be the one through which you can export the dashboard and all the individual graphs. Be sure to export ALL of them.

The dashboard graphs can be about any of the queries you built for the fourth point of the deliverable. If they are complex and are not suited for representation, you can make new and simple ones. Be sure to build a good number of them (i.e., the text asked for eight queries. Therefore, provide at least six representations). Moreover, we'd prefer the ones you built for the fourth point to the (new) simple ones.

5 | OPTIONAL TASKS

5.1. Additional dataset

In addition to the dataset used, there were added 2 datasets concerning the population of all the regions and the administration points. The 2 files are "platea.csv" and "punti-somministrazione-tipologia", which can be found in the same repository used for the precedent dataset.

Link to the repository: https://github.com/italia/covid19-opendata-vaccini

Regarding the mappings, everything was left as the automatic mapping suggest, except for "nome_area" in both dataset, which was mapped as text, as it was done with the precedent dataset.

This is the mapping applied for "platea.csv":

```
{
  "properties": {
    "area": {
      "type": "keyword"
    },
    "fascia_anagrafica": {
      "type": "keyword"
    },
    "nome_area": {
      "type": "text"
    },
    "totale_popolazione": {
      "type": "long"
    }
  }
}
```

FIELD NAME	TYPE (ES)	DESCRIPTION
area	keyword	Initials of the delivery region
fascia_anagrafica	keyword	Age group of people that received the vaccine
nome_area	text	Standard area denomination
totale_popolazione	long	Number of people living in that area and of that age range

This is the mapping applied to "punti-somministrazione-tipologia.csv":

```
{
  "properties": {
    "area": {
      "type": "keyword"
    },
    "codice_NUTS1": {
      "type": "keyword"
    },
    "codice_NUTS2": {
      "type": "keyword"
    },
    "codice_regione_ISTAT": {
      "type": "long"
    },
    "denominazione_struttura": {
      "type": "text"
    },
    "nome_area": {
      "type": "text"
    },
    "tipologia": {
      "type": "keyword"
    }
  }
}
```

FIELD NAME	TYPE (ES)	DESCRIPTION
area	keyword	Initials of the delivery region
codice_NUTS1	keyword	Nomenclature of Territorial Units for Statistics (NUTS 1)
codice_NUTS2	keyword	Nomenclature of Territorial Units for Statistics (NUTS 2)
codice_regione_ISTAT	long	ISTAT region code
denominazione_struttura	text	Name of the admistration point
nome_area	text	Standard area denomination
tipologia	keyword	Type of administration point: hospital or territorial

5.1.1. Query

We performed 2 additional queries on the new datasets. With the first one we calculated the total amount and the percentage of all the doses administrated over the overall population, thanks to data in "platea.csv".

The problem was that in the file platea.csv the age ranges 80-89 and 90+ were grouped into a single age range 80+. This doesn't permit the join on the fields with somministrazione-vaccini-latest.csv, and then it is impossible to calculate the percentage for that specific age ranges.

```
1  GET /platea,covid_vaccines/_search
2  {
3    "size": 0,
4    "aggs": {
5      "age_range":{
6        "terms": {
7          "field": "fascia_anagrafica"
8        },
9        "aggs": {
10         "first_dose_count": {
11           "sum": {
12             "field": "prima_dose"
13           }
14         },
15         "second_dose_count": {
```

```
16            "sum": {
17              "field": "seconda_dose"
18            }
19          },
20          "population_count": {
21            "sum": {
22              "field": "totale_popolazione"
23            }
24          },
25          "first_dose_percent":{
26            "bucket_script": {
27              "buckets_path": {
28                "buck1": "first_dose_count",
29                "buck2": "population_count"
30              },
31              "script": "if (params.buck2!=0) params.buck1/
                    params.buck2 * 100"
32            }
33          },
34          "second_dose_percent":{
35            "bucket_script": {
36              "buckets_path": {
37                "buck1": "second_dose_count",
38                "buck2": "population_count"
39              },
40              "script": "if (params.buck2!=0) params.buck1/
                    params.buck2 * 100"
41            }
42          }
43        }
44      }
45    }
46 }
```

Here's the result of the above query:

```
{
  "key" : "60-69",
  "doc_count" : 21581,
  "second_dose_count" : {
    "value" : 6075018.0
  },
  "population_count" : {
    "value" : 7553311.0
  },
  "first_dose_count" : {
    "value" : 6656012.0
  },
  "first_dose_percent" : {
    "value" : 88.12045472508679
  },
  "second_dose_percent" : {
    "value" : 80.42854319119125
  }
},
```

Figure 5.1

In the second query, it was calculated the number of each type of administration point for each region. Here is the query:

```
1  GET /type_administration_point/_search
2  {
3    "size": 0,
4    "aggs": {
5      "regions": {
6        "terms": {
7          "field": "area"
8        },
9        "aggs": {
10         "point_type": {
11           "terms": {
12             "field": "tipologia"
13           }
14         }
15       }
16     }
17   }
18 }
```

This is the result of the above query:

```
{
  "key" : "PUG",
  "doc_count" : 381,
  "point_type" : {
    "doc_count_error_upper_bound" : 0,
    "sum_other_doc_count" : 0,
    "buckets" : [
      {
        "key" : "STS",
        "doc_count" : 272
      },
      {
        "key" : "RIA",
        "doc_count" : 109
      }
    ]
  }
},
{
  "key" : "SIC",
  "doc_count" : 265,
  "point_type" : {
    "doc_count_error_upper_bound" : 0,
    "sum_other_doc_count" : 0,
    "buckets" : [
      {
        "key" : "STS",
        "doc_count" : 148
      },
      {
        "key" : "RIA",
        "doc_count" : 117
      }
    ]
  }
},
```

Figure 5.2

6 | REFERENCES AND SOURCES

- https://www.elastic.co/guide/index.html