SMBUD 2021 - Project work

NoSQL DB for supporting a data analysis scenario over data about COVID-19 vaccination statistics

Authors:

Pavithra Saravana Moorthy - 10748677
Vinaykrishna Munnaluri - 10780972
Stancioi Ioana-Ruxandra - 10855466
Roger Aldair Veliz Sedano - 10614456
Wilson Emannuelle Quispe Limay - 10182182

Contents

1 Introduction

Monitoring and analyzing data from the COVID-19 outbreak across multiple data sources such as the number of hospitalizations, deaths, vaccinations, and a variety of other types of data are critical to controlling the spread of the coronavirus.

1.1 Problem formulation & Solution

The Italian Government collects and publishes a range of data and statistics about COVID-19 vaccination.

The latest available data are those reported to or compiled from local official sources by the Italian Health Organization, the following link leads to the relevant statistics about vaccination status. https://github.com/italia/covid19-opendata-vaccini/

The locations listed in the data include all the Italian regions with their respective code and abbreviation in order to track the number of COVID-19 vaccine doses administered for each region on a daily basis.

Additional information about the gender and age of the population vaccinated is present in the dataset because they are important factors in understanding immunization, including vaccine administration and efficacy.

Studies show that after being vaccinated against COVID-19, protection against the virus and the ability to prevent infection with variants can decrease over time due to changes in variants. That's why the vaccine booster data is relevant and included in the dataset, in addition to the number of dose information.

2 Implementation

To implement the solution we use the ELK stack. It is composed of three open source projects: Elasticsearch, Logstash, and Kibana.

Elasticsearch is a NoSQL database which is based on search engine called Lucene used for search and analytics.
Logstash is a server-side pipeline which is used for managing logs and events.
Kibana acts like a visualization layer on Elasticsearch.

2.1 Data

Data file Link:
https://raw.githubusercontent.com/italia/covid19-opendata-vaccini/master/dati/somministrazioni-vaccini-latest.csv

2.2 Schema & Mapping

We imported the csv file into Elasticsearch and used the following setting as shown in fig 1.

Fig. 1. Settings

The import process is composed of different steps automatically handled by Elastichsearch. it consists in:

1. **Processing file**
2. **Creating index**
3. **Creating ingest pipeline**
4. **Uploading the data**

5. **Creating index pattern**

Within a search engine, mapping defines how a document is indexed and how it indexes and stores its fields. For analysis purposes we used the automatic mapping.

The following table reports the mapping for this project.

Mapping


```
{
  "properties": {
    "@timestamp": {
      "type": "date"
    },
    "area": {
      "type": "keyword"
    },
    "codice_NUTS1": {
      "type": "keyword"
    },
    "codice_NUTS2": {
      "type": "keyword"
    },
    "codice_regione_ISTAT": {
      "type": "long"
    },
    "data_somministrazione": {
      "type": "date",
      "format": "MM/dd/yyyy||M/d/yyyy"
    },
    "dose_addizionale_booster": {
      "type": "long"
    },
    "fascia_anagrafica": {
      "type": "keyword"
    },
    "fornitore": {
      "type": "keyword"
    },
    "nome_area": {
      "type": "keyword"
    },
    "pregressa_infezione": {
      "type": "long"
    },
```

```
    "prima_dose": {
      "type": "long"
    },
    "seconda_dose": {
      "type": "long"
    },
    "sesso_femminile": {
      "type": "long"
    },
    "sesso_maschile": {
      "type": "long"
    }
  }
}
```

Tab. 1. Mapping

It's possible to use the type "text" for the fields "nome_area", "area", "fornitore" for full text search, in that case we need to specify an analyzer for that data type. We didn't use the values of those fields for our queries as partial matching, so we decided to leave the "keyword" as a type for those fields.

Type	Fields
Date	• data_somministrazione
Long	• codice_regione_ISTAT, • prima_dose • pregressa_infezione • seconda_dose • sesso_femminile • sesso_maschile
Keyword	• area • codice_NUTS1 • codice_NUTS2 • fascia_anagrafica • fornitore • nome_area

Tab. 2. Field data type

2.3 Queries

Elasticsearch provides a powerful set of options for querying documents for various use cases.

1) To get types of vaccines taken among a range of age groups.

Query	Partial Result

GET /vaccine-covid/_search
{
"aggs": {
"0": {
"terms": {
"field": "fornitore",
"order": {
"_count": "desc"
},
"size": 5
},
"aggs": {
"1": {
"terms": {
"field": "fascia_anagrafica",
"order": {
"_count": "desc"
},
"size": 3
}
}
}
}
},
"size": 0,
"fields": [
{
"field": "@timestamp",
"format": "date_time"
},
{
"field": "data_somministrazione",
"format": "date_time"
}
],
"script_fields": {},
"stored_fields": [
"*"
],
"runtime_mappings": {},
"_source": {
"excludes": []
},
"query": {
"bool": {
"must": [],
"filter": [
{
"range": {
"@timestamp": {
	{
	"took" : 279,
	"timed_out" : false,
	"_shards" : {
	"total" : 1,
	"successful" : 1,
	"skipped" : 0,
	"failed" : 0
	},
	"hits" : {
	"total" : {
	"value" : 10000,
	"relation" : "gte"
	},
	"max_score" : null,
	"hits" : [ ]
	},
	"aggregations" : {
	"0" : {
	"doc_count_error_upper_bound" : 0,
	"sum_other_doc_count" : 0,
	"buckets" : [
	{
	"key" : "Pfizer/BioNTech",
	"doc_count" : 15086,
	"1" : {
	"doc_count_error_upper_bound" : 0,
	"sum_other_doc_count" : 10033,
	"buckets" : [
	{
	"key" : "50-59",
	"doc_count" : 1686
	},
	{
	"key" : "60-69",
	"doc_count" : 1684
	},
	{
	"key" : "40-49",
	"doc_count" : 1683
	}
	]
	}
	},
	{
	"key" : "Moderna",
	"doc_count" : 14411,
	"1" : {
	"doc_count_error_upper_bound" : 0,
	"sum_other_doc_count" : 9487,

"format": "strict_date_optional_time", "gte": "2021-10-09T22:00:00.000Z", "lte": "2022-01-08T01:59:17.576Z" } } }], "should": [], "must_not": [] } } }	"buckets" : [{ "key" : "40-49", "doc_count" : 1646 }, { "key" : "50-59", "doc_count" : 1642 }.....

2) To retrieve the count of booster dose taken along the period of 90 days

Query	Partial Result

GET /vaccine-covid/_search
{
"aggs": {
"0": {
"date_histogram": {
"field": "data_somministrazione",
"calendar_interval": "1d",
"time_zone": "Europe/Rome"
},
"aggs": {
"2": {
"percentiles": {
"field": "dose_addizionale_booster",
"percents": [
50
]
}
}
}
}
},
"size": 0,
"fields": [
{
"field": "@timestamp",
"format": "date_time"
},
{
"field": "data_somministrazione",
"format": "date_time"
}
],
"script_fields": {},
"stored_fields": [
"*"
],
"runtime_mappings": {},
"_source": {
	{
	"took" : 237,
	"timed_out" : false,
	"_shards" : {
	"total" : 1,
	"successful" : 1,
	"skipped" : 0,
	"failed" : 0
	},
	"hits" : {
	"total" : {
	"value" : 10000,
	"relation" : "gte"
	},
	"max_score" : null,
	"hits" : [ ]
	},
	"aggregations" : {
	"0" : {
	"buckets" : [
	{
	"key_as_string" : "2021-10-10T00:00:00.000+02:00",
	"key" : 1633816800000,
	"doc_count" : 357,
	"2" : {
	"values" : {
	"50.0" : 1.0
	}
	}
	},
	{
	"key_as_string" : "2021-10-11T00:00:00.000+02:00",
	"key" : 1633903200000,
	"doc_count" : 450,
	"2" : {
	"values" : {
	"50.0" : 4.0
	}

```
    "excludes": []
  },
  "query": {
    "bool": {
      "must": [],
      "filter": [
        {
          "range": {
            "data_somministrazione": {
              "format": "strict_date_optional_time",
              "gte": "2021-10-09T22:00:00.000Z",
              "lte": "2022-01-08T01:59:17.576Z"
            }
          }
        }
      ],
      "should": [],
      "must_not": []
    }
  }
}
```

```
        }
      },
      {
        "key_as_string" : "2021-10-12T00:00:00.000+02:00",
        "key" : 1633989600000,
        "doc_count" : 457,
        "2" : {
          "values" : {
            "50.0" : 4.0
          }
        }
      },
      {
        "key_as_string" : "2021-10-13T00:00:00.000+02:00",
        "key" : 1634076000000,
        "doc_count" : 484,
        "2" : {
          "values" : {
            "50.0" : 3.0
          } ……
```

3) To get the average count of first and second doses in last 3 months

Query	Partial Result

```
GET /vaccine-covid/_search
{
  "aggs": {
    "0": {
      "date_histogram": {
        "field": "data_somministrazione",
        "calendar_interval": "1d",
        "time_zone": "Europe/Rome"
      },
      "aggs": {
        "2": {
          "avg": {
            "field": "prima_dose"
          }
        },
        "3": {
          "avg": {
            "field": "seconda_dose"
          }
        }
      }
    }
  },
  "size": 0,
  "fields": [
    {
      "field": "@timestamp",
      "format": "date_time"
    },
```

```
{
  "took" : 37,
  "timed_out" : false,
  "_shards" : {
    "total" : 1,
    "successful" : 1,
    "skipped" : 0,
    "failed" : 0
  },
  "hits" : {
    "total" : {
      "value" : 10000,
      "relation" : "gte"
    },
    "max_score" : null,
    "hits" : [ ]
  },
  "aggregations" : {
    "0" : {
      "buckets" : [
        {
          "key_as_string" : "2021-10-10T00:00:00.000+02:00",
          "key" : 1633816800000,
          "doc_count" : 357,
          "2" : {
            "value" : 73.98319327731092
          },
          "3" : {
            "value" : 120.47619047619048
```

```
    {
      "field": "data_somministrazione",
      "format": "date_time"
    }
  ],
  "script_fields": {},
  "stored_fields": [
    "*"
  ],
  "runtime_mappings": {},
  "_source": {
    "excludes": []
  },
  "query": {
    "bool": {
      "must": [],
      "filter": [
        {
          "range": {
            "data_somministrazione": {
              "format": "strict_date_optional_time",
              "gte": "2021-10-09T22:00:00.000Z",
              "lte": "2022-01-08T02:20:08.825Z"
            }
          }
        }
      ],
      "should": [],
      "must_not": []
    }
  }
}
```

```
        }
      },
      {
        "key_as_string" : "2021-10-11T00:00:00.000+02:00",
        "key" : 1633903200000,
        "doc_count" : 450,
        "2" : {
          "value" : 123.29777777777778
        },
        "3" : {
          "value" : 182.45555555555555
        }
      },
      {
        "key_as_string" : "2021-10-12T00:00:00.000+02:00",
        "key" : 1633989600000,
        "doc_count" : 457,
        "2" : {
          "value" : 130.02407002188184
        },
        "3" : {
          "value" : 172.6323851203501
        }
      },
      { ..}
```

4) Retrieves the count of vaccines taken by different age groups in a period of 3 months.

Query	Result

```
GET /vaccine-covid/_search
{
  "aggs": {
    "0": {
      "terms": {
        "field": "fascia_anagrafica",
        "order": {
          "_count": "desc"
        },
        "size": 5
      }
    }
  },
  "size": 0,
  "fields": [
    {
      "field": "@timestamp",
```

```
{
  "took" : 441,
  "timed_out" : false,
  "_shards" : {
    "total" : 1,
    "successful" : 1,
    "skipped" : 0,
    "failed" : 0
  },
  "hits" : {
    "total" : {
      "value" : 10000,
      "relation" : "gte"
    },
    "max_score" : null,
    "hits" : [ ]
  },
```

```
    "format": "date_time"
   },
   {
    "field": "data_somministrazione",
    "format": "date_time"
   }
  ],
  "script_fields": {},
  "stored_fields": [
   "*"
  ],
  "runtime_mappings": {},
  "_source": {
   "excludes": []
  },
  "query": {
   "bool": {
    "must": [],
    "filter": [
     {
      "range": {
       "@timestamp": {
        "format": "strict_date_optional_time",
        "gte": "2021-10-09T22:00:00.000Z",
        "lte": "2022-01-08T01:59:17.576Z"
       }
      }
     }
    ],
    "should": [],
    "must_not": []
   }
  }
}
```

```
 "aggregations" : {
  "0" : {
   "doc_count_error_upper_bound" : 0,
   "sum_other_doc_count" : 13795,
   "buckets" : [
    {
     "key" : "60-69",
     "doc_count" : 3982
    },
    {
     "key" : "70-79",
     "doc_count" : 3741
    },
    {
     "key" : "50-59",
     "doc_count" : 3732
    },
    {
     "key" : "40-49",
     "doc_count" : 3711
    },
    {
     "key" : "30-39",
     "doc_count" : 3677
    }
   ]
  }
 }
}
```

5) Region-wise count of people vaccinated in Italy

Query	Partial Result

```
GET /vaccine-covid/_search
{
  "aggs": {
    "0": {
      "terms": {
        "field": "nome_area",
        "order": {
          "_count": "desc"
        },
        "size": 20
      }
    }
  },
  "size": 0,
  "fields": [
    {
      "field": "@timestamp",
```

```
{
  "took" : 56,
  "timed_out" : false,
  "_shards" : {
    "total" : 1,
    "successful" : 1,
    "skipped" : 0,
    "failed" : 0
  },
  "hits" : {
    "total" : {
      "value" : 10000,
      "relation" : "gte"
    },
    "max_score" : null,
    "hits" : [ ]
  },
```

```
    "format": "date_time"
   },
   {
    "field": "data_somministrazione",
    "format": "date_time"
   }
  ],
  "script_fields": {},
  "stored_fields": [
   "*"
  ],
  "runtime_mappings": {},
  "_source": {
   "excludes": []
  },
  "query": {
   "bool": {
    "must": [],
    "filter": [
     {
      "range": {
       "@timestamp": {
        "format": "strict_date_optional_time",
        "gte": "2021-10-09T22:00:00.000Z",
        "lte": "2022-01-08T02:20:08.825Z"
       }
      }
     }
    ],
    "should": [],
    "must_not": []
   }
  }
}
```

```
"aggregations" : {
  "0" : {
    "doc_count_error_upper_bound" : 0,
    "sum_other_doc_count" : 1157,
    "buckets" : [
      {
        "key" : "Lazio",
        "doc_count" : 2060
      },
      {
        "key" : "Emilia-Romagna",
        "doc_count" : 1783
      },
      {
        "key" : "Piemonte",
        "doc_count" : 1751
      },
      {
        "key" : "Lombardia",
        "doc_count" : 1713
      },
      {
        "key" : "Abruzzo",
        "doc_count" : 1620
      },
      {
        "key" : "Campania",
        "doc_count" : 1610
      },
      {...}
```

6) To get the average of male and female vaccinated in last 90 days

Query	Partial Result

GET /vaccine-covid/_search
{
"aggs": {
"0": {
"date_histogram": {
"field": "@timestamp",
"calendar_interval": "1d",
"time_zone": "Europe/Rome"
},
"aggs": {
"1": {
"avg": {
"field": "sesso_femminile"
}
},
"2": {
"avg": {
	{
	"took" : 221,
	"timed_out" : false,
	"_shards" : {
	"total" : 1,
	"successful" : 1,
	"skipped" : 0,
	"failed" : 0
	},
	"hits" : {
	"total" : {
	"value" : 10000,
	"relation" : "gte"
	},
	"max_score" : null,
	"hits" : [ ]
	},

<table>
<tr>
<td>

```
          "field": "sesso_maschile"
        }
       }
      }
    }
  }
 },
 "size": 0,
 "fields": [
  {
   "field": "@timestamp",
   "format": "date_time"
  },
  {
   "field": "data_somministrazione",
   "format": "date_time"
  }
 ],
 "script_fields": {},
 "stored_fields": [
  "*"
 ],
 "runtime_mappings": {},
 "_source": {
  "excludes": []
 },
 "query": {
  "bool": {
   "must": [],
   "filter": [
    {
     "range": {
      "@timestamp": {
       "format": "strict_date_optional_time",
       "gte": "2021-10-09T22:00:00.000Z",
       "lte": "2022-01-08T02:35:07.853Z"
      }
     }
    }
   ],
   "should": [],
   "must_not": []
  }
 }
}
```

</td>
<td>

```
"aggregations" : {
 "0" : {
  "buckets" : [
   {
    "key_as_string" : "2021-10-10T00:00:00.000+02:00",
    "key" : 1633816800000,
    "doc_count" : 357,
    "1" : {
     "value" : 108.80112044817928
    },
    "2" : {
     "value" : 129.75350140056022
    }
   },
   {
    "key_as_string" : "2021-10-11T00:00:00.000+02:00",
    "key" : 1633903200000,
    "doc_count" : 450,
    "1" : {
     "value" : 189.5822222222222
    },
    "2" : {
     "value" : 216.9311111111111
    }
   },
   {
    "key_as_string" : "2021-10-12T00:00:00.000+02:00",
    "key" : 1633989600000,
    "doc_count" : 457,
    "1" : {
     "value" : 194.26258205689277
    },
    "2" : {
     "value" : 220.20787746170677
    }
   },
   {
    "key_as_string" : "2021-10-13T00:00:00.000+02:00",
    "key" : 1634076000000,
    "doc_count" : 484,
    "1" : {..}
```

</td>
</tr>
</table>

7) People who have taken moderna without covid, be it first dose or second dose.

Query	Partial Result

GET /vaccine-covid/_search
{
"query": {
"bool": {
"must": [
{"term": {
"fornitore": {
	{
	"took" : 2044,
	"timed_out" : false,
	"_shards" : {
	"total" : 1,
	"successful" : 1,
	"skipped" : 0,

```
          "value": "Moderna"
         }
        }
       }
     ],
     "must_not": [
      {
       "term": {
        "previous_infection": {
         "value": 1
        }
       }
      }
     ],
     "should": [
      {
       "term": {
        "first_dose": {
         "value": 0
        }
       }
      }
     ]
    }
   }
}
```

```
  "failed" : 0
 },
 "hits" : {
  "total" : {
   "value" : 10000,
   "relation" : "gte"
  },
  "max_score" : 1.1187733,
  "hits" : [
   {
    "_index" : "vaccine-covid",
    "_type" : "_doc",
    "_id" : "BxGzRH4BfiVrZ9_SiQ0k",
    "_score" : 1.1187733,
    "_source" : {
     "area" : "ABR",
     "codice_regione_ISTAT" : 13,
     "nome_area" : "Abruzzo",
     "data_somministrazione" : "2021-08-03",
     "dose_addizionale_booster" : 0,
     "codice_NUTS1" : "ITF",
     "fascia_anagrafica" : "12-19",
     "prima_dose" : 390,
     "pregressa_infezione" : 6,
     ....}
```

8) Progress of infections along a span of 90 days

Query	Partial Result

```
GET /vaccine-covid/_search
{
  "aggs": {
    "0": {
      "date_histogram": {
        "field": "@timestamp",
        "calendar_interval": "1d",
        "time_zone": "Europe/Rome"
      },
      "aggs": {
        "1": {
          "percentiles": {
            "field": "pregressa_infezione",
            "percents": [
              50
            ]
          }
        }
      }
    }
  },
  "size": 0,
  "fields": [
    {
```

```
{
 "took" : 140,
 "timed_out" : false,
 "_shards" : {
  "total" : 1,
  "successful" : 1,
  "skipped" : 0,
  "failed" : 0
 },
 "hits" : {
  "total" : {
   "value" : 10000,
   "relation" : "gte"
  },
  "max_score" : null,
  "hits" : [ ]
 },
 "aggregations" : {
  "0" : {
   "buckets" : [
    {
     "key_as_string" : "2021-10-12T00:00:00.000+02:00",
     "key" : 1633989600000,
     "doc_count" : 457,
```

<table>
<tr>
<td>

<pre>
    "field": "@timestamp",
    "format": "date_time"
   },
   {
    "field": "data_somministrazione",
    "format": "date_time"
   }
  ],
  "script_fields": {},
  "stored_fields": [
   "*"
  ],
  "runtime_mappings": {},
  "_source": {
   "excludes": []
  },
  "query": {
   "bool": {
    "must": [],
    "filter": [
     {
      "range": {
       "@timestamp": {
        "format": "strict_date_optional_time",
        "gte": "2021-10-11T22:00:00.000Z",
        "lte": "2022-01-10T20:02:07.607Z"
       }
      }
     }
    ],
    "should": [],
    "must_not": []
   }
  }
}
</pre>

</td>
<td>

<pre>
      "1" : {
       "values" : {
        "50.0" : 1.2
       }
      }
     },
     {
      "key_as_string" : "2021-10-13T00:00:00.000+02:00",
      "key" : 1634076000000,
      "doc_count" : 484,
      "1" : {
       "values" : {
        "50.0" : 0.25
       }
      }
     },
     {
      "key_as_string" : "2021-10-14T00:00:00.000+02:00",
      "key" : 1634162400000,
      "doc_count" : 477,
      "1" : {
       "values" : {
        "50.0" : 1.0
       }
      }
     },
     {
      "key_as_string" : "2021-10-15T00:00:00.000+02:00",
      "key" : 1634248800000,
      "doc_count" : 453,
      "1" : {..}
</pre>

</td>
</tr>
</table>

2.4 Commands

1) Update mapping: Add a new field to the existing index. For example to add the "terza_dose" field

<table>
<tr>
<th>Command</th>
<th>Result using:

GET /vaccine-covid/_mapping</th>
</tr>
<tr>
<td>

<pre>
PUT /vaccine-covid/_mapping
{
  "properties": {
    "terza_dose": {
      "type": "long"
    }
  }
</pre>

</td>
<td>

<pre>
{
  "vaccine-covid" : {
    "mappings" : {
      "_meta" : {
        "created_by" : "file-data-visualizer"
      },
      "properties" : {
</pre>

</td>
</tr>
</table>

```
}
```

```
      "@timestamp" : {
        "type" : "date"
      },
      "area" : {
        "type" : "keyword"
      },
      "codice_NUTS1" : {
        "type" : "keyword"
      },
      "codice_NUTS2" : {
        "type" : "keyword"
      },
      "codice_regione_ISTAT" : {
        "type" : "long"
      },
      "data_somministrazione" : {
        "type" : "date",
        "format" : "iso8601"
      },
      "dose_addizionale_booster" : {
        "type" : "long"
      },
      "fascia_anagrafica" : {
        "type" : "keyword"
      },
      "fornitore" : {
        "type" : "keyword"
      },
      "nome_area" : {
        "type" : "keyword"
      },
      "pregressa_infezione" : {
        "type" : "long"
      },
      "prima_dose" : {
        "type" : "long"
      },
      "seconda_dose" : {
        "type" : "long"
      },
      "sesso_femminile" : {
        "type" : "long"
      },
      "sesso_maschile" : {
        "type" : "long"
      },
      "terza_dose" : {
        "type" : "long"
      }
    }
  }
}
```

2) Creating a new entry document (with auto-generation of id)

Command	Result

POST /vaccine-covid/_doc
{
"data_somministrazione": "2022-01-10",
"fornitore": "Moderna",
"area": "SIC",
"fascia_anagrafica": "90+",
"sesso_maschile": 1,
"sesso_femminile": 0,
"prima_dose": 1,
"seconda_dose": 0,
"pregressa_infezione": 0,
"dose_addizionale_booster": 0,
"codice_NUTS1": "ITG",
"codice_NUTS2":"ITG1",
"codice_regione_ISTAT": 19,
"nome_area": "Sicilia"
}
	{
	"_index" : "vaccine-covid",
	"_type" : "_doc",
	"_id" : "gBI5RX4BfiVrZ9_S3Bbq",
	"_version" : 1,
	"result" : "created",
	"_shards" : {
	"total" : 2,
	"successful" : 1,
	"failed" : 0
	},
	"_seq_no" : 168266,
	"_primary_term" : 1
	}

3. Kibana Visualization

For data visualization we used Kibana. Kibana visualizations are based on aggregations performed by Elasticsearch.

Dashboard_1. pdf:
https://drive.google.com/file/d/11_FE9u2SLMqMsgwtNlVjJQGbM-5gx-F0/view?usp=sharing

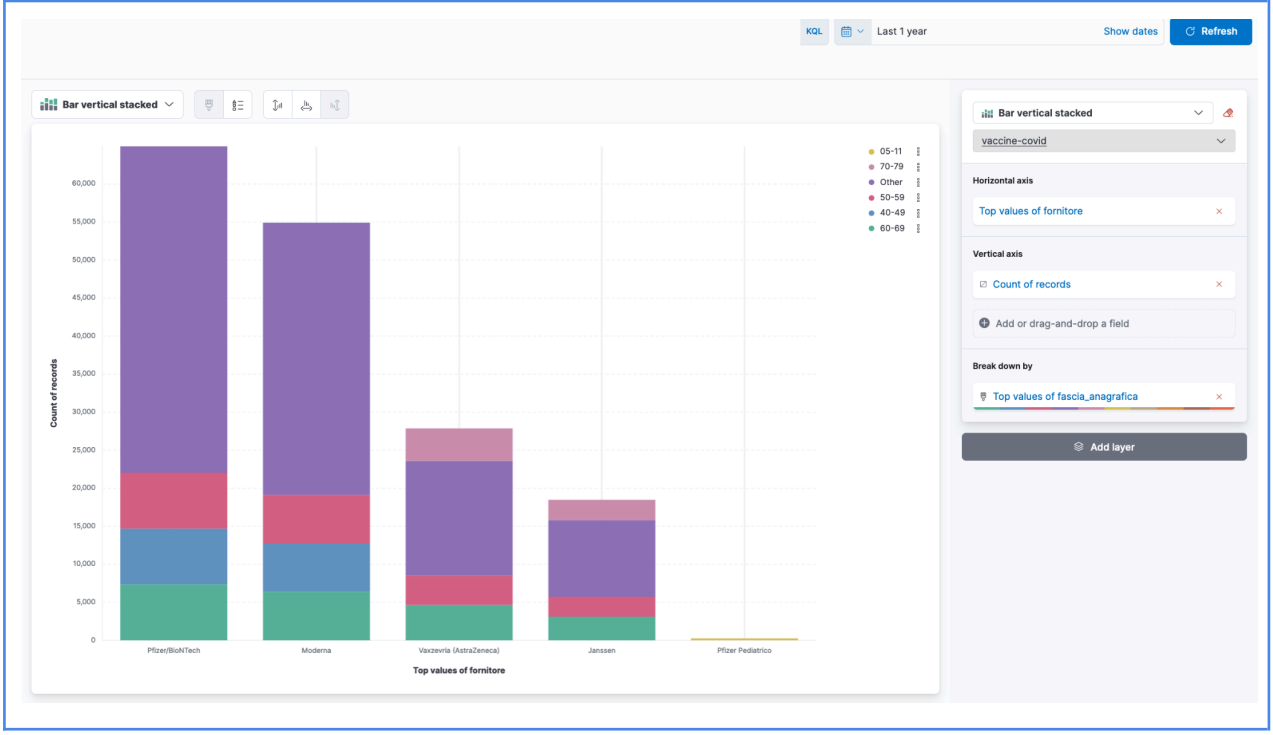
Dashboard_2.pdf:
https://drive.google.com/file/d/10XYhLd8RlzTZ7jlYpmkupt19h4_GWSRg/view?usp=sharing

Data csv:
https://drive.google.com/file/d/1lQ_nvtxXegGTJnxKubfpXOpHLIHLOEt3/view?usp=sharing

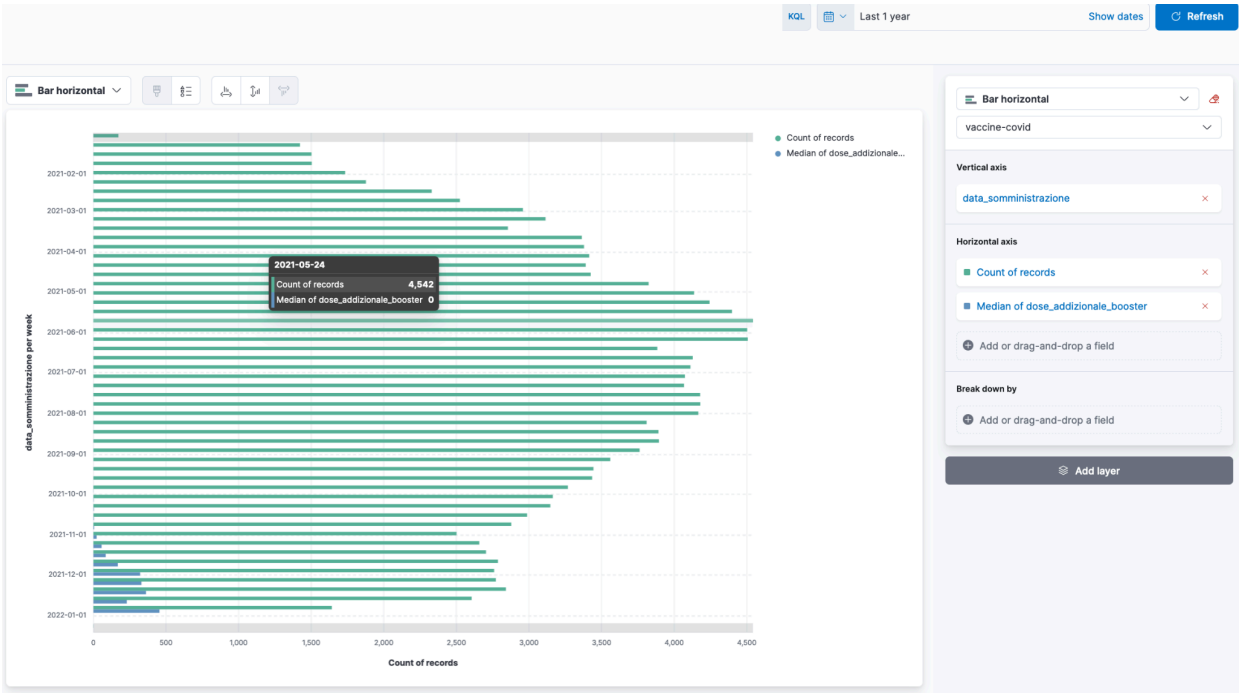
export.ndjson:
https://drive.google.com/file/d/1oKdJnv1mzTfG0xYfEgk4y3wDmQfEh9Mi/view?usp=sharing

List of visualization and their description

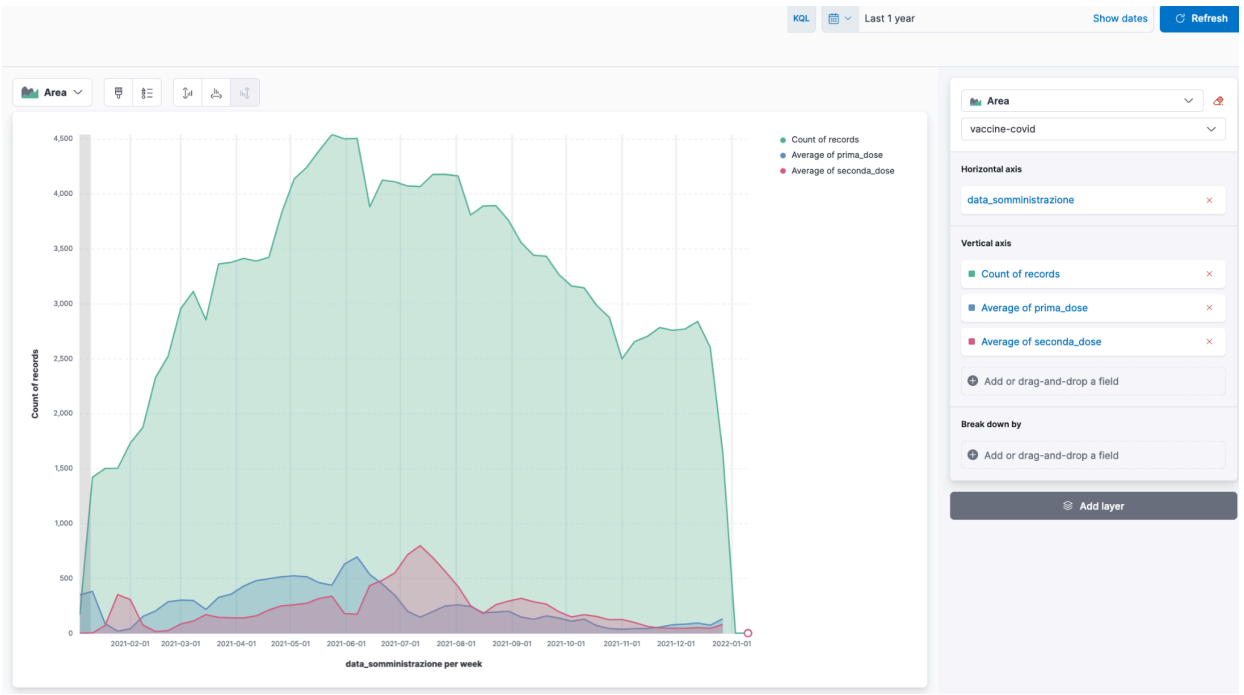
1) To get types of vaccines taken among a range of age groups.



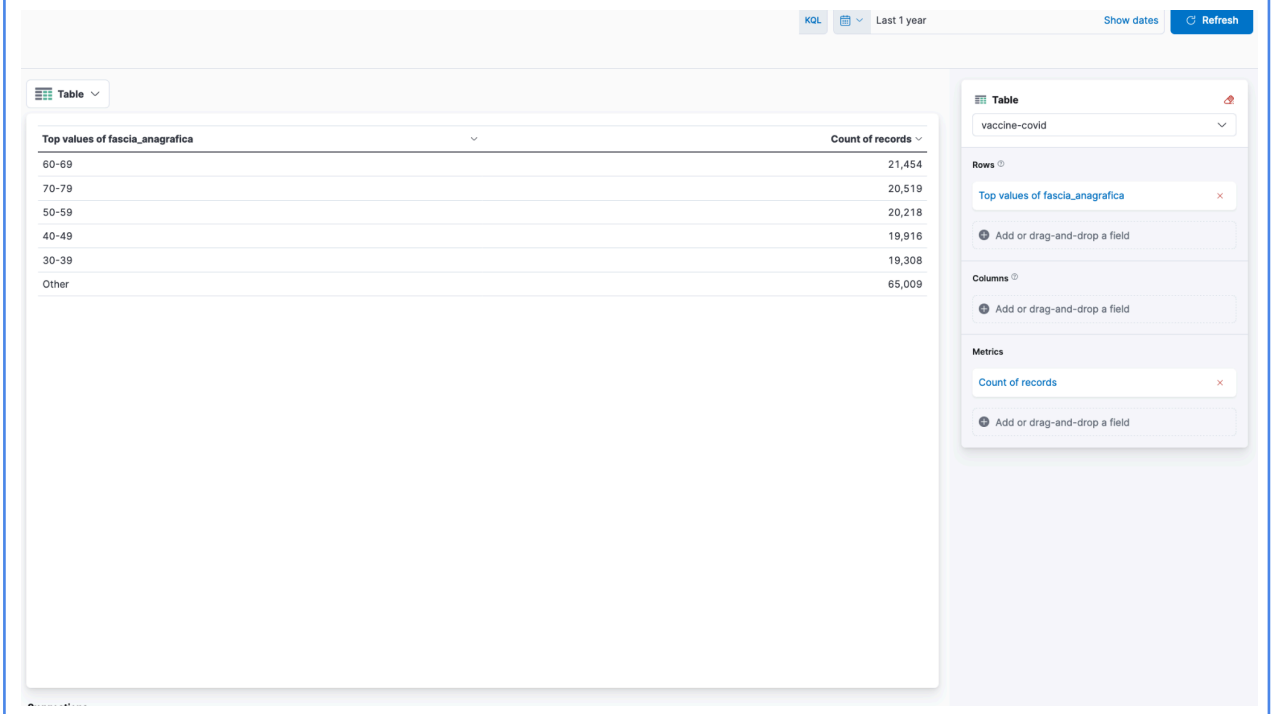
2) Booster dose taken along the period



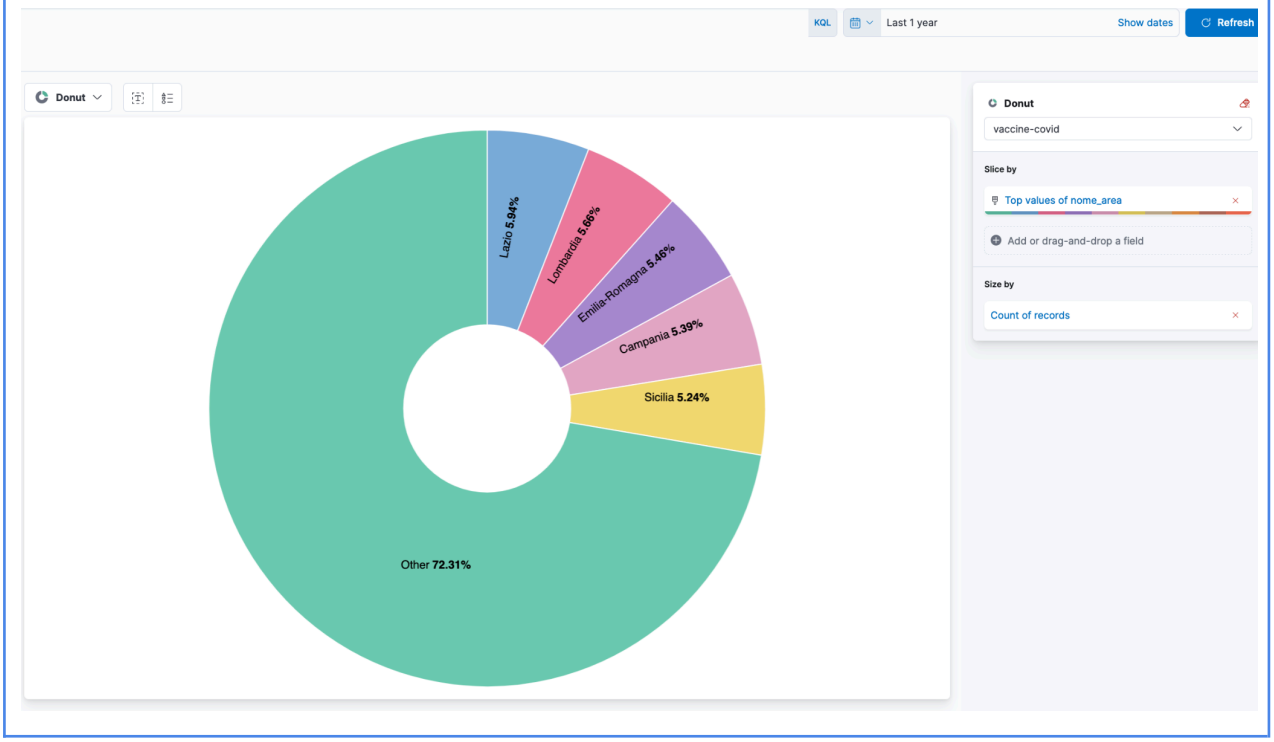
3) Average count of 1st and 2nd vaccine doses along the period



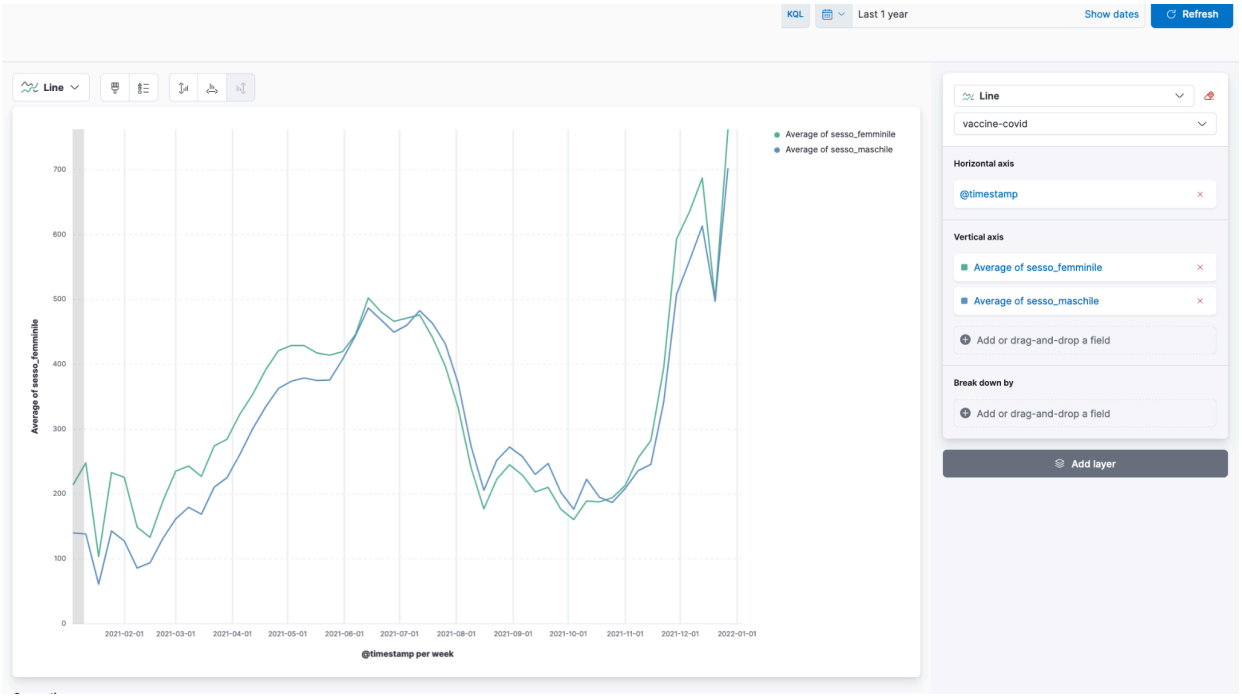
4) Count of vaccines along different age groups



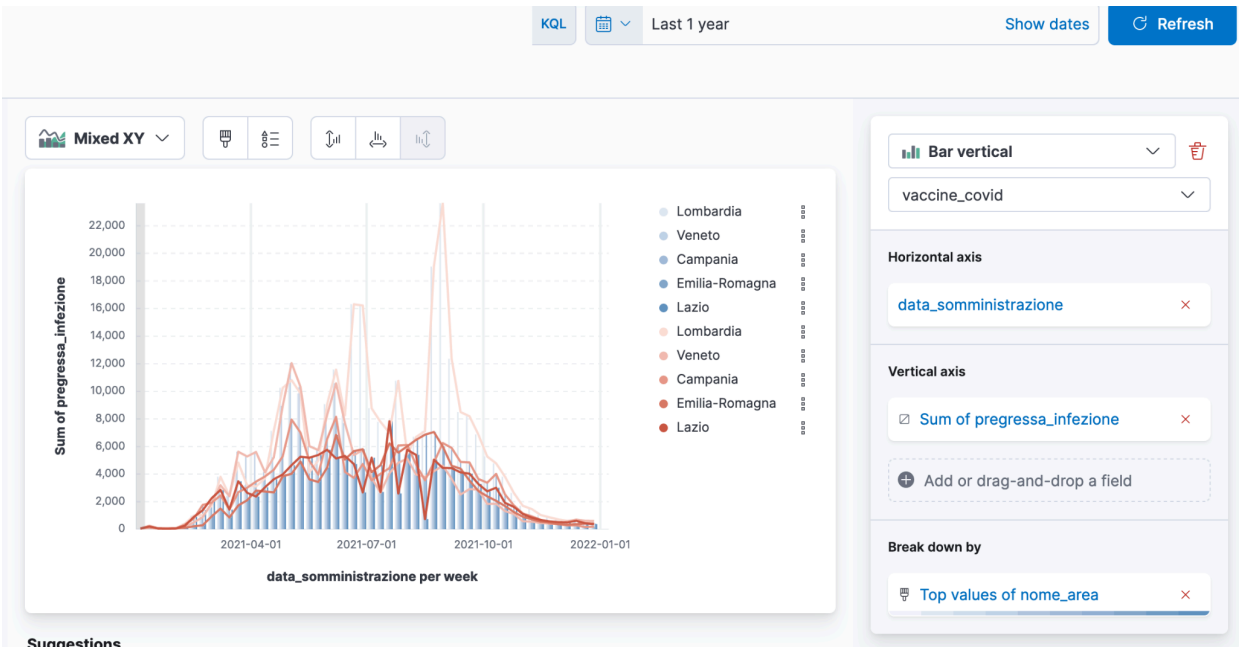
5) Region Wise (top 5) split of vaccine count



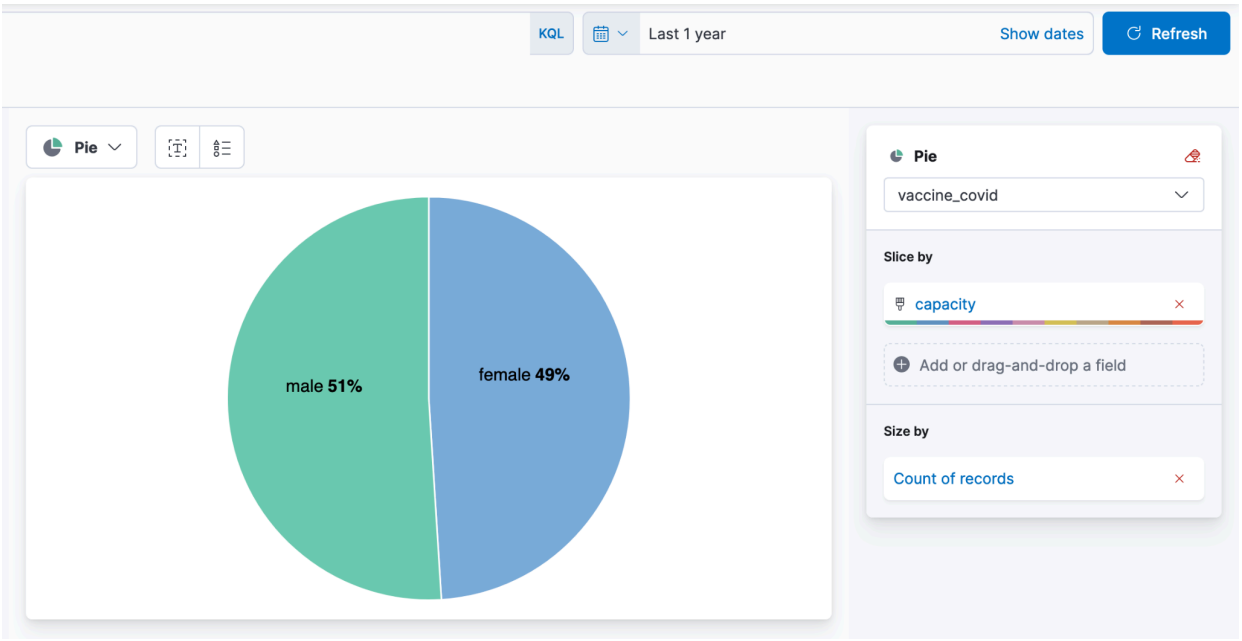
6) Average male and female taking doses



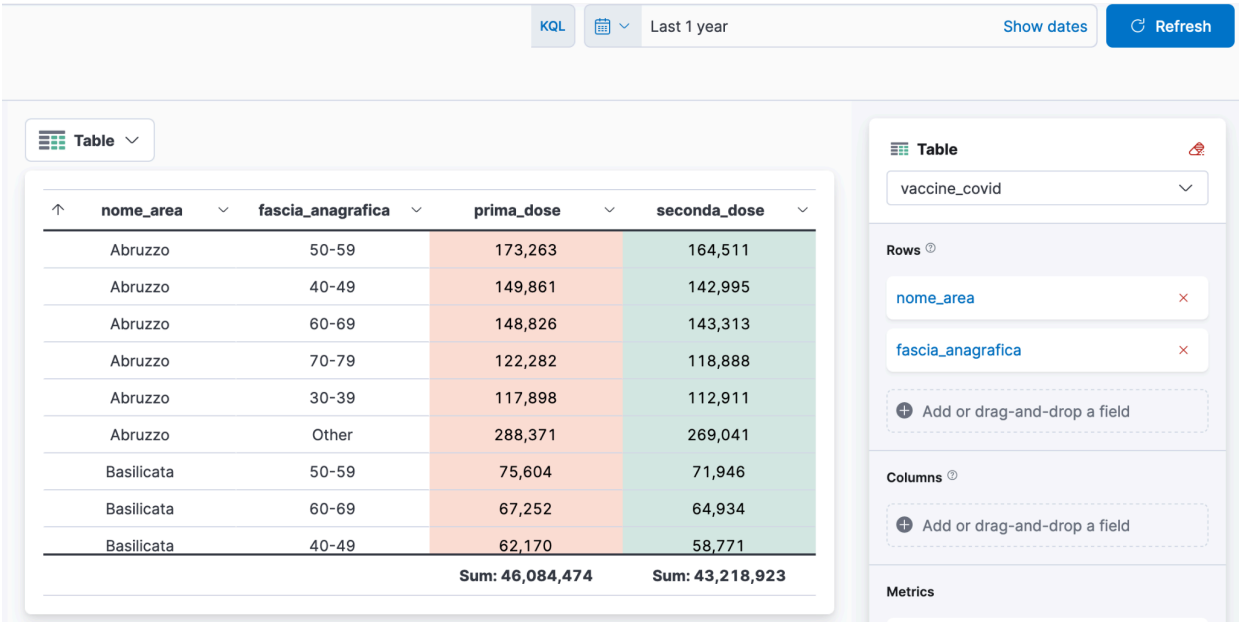
7) previous infection count by region top 5 regions and their trends over last one year



8) *vaccination count percentage categorized by gender*



9) *summary of doses by region and age for last one year*



nome_area	fascia_anagrafica	prima_dose	seconda_dose
Abruzzo	50-59	173,263	164,511
Abruzzo	40-49	149,861	142,995
Abruzzo	60-69	148,826	143,313
Abruzzo	70-79	122,282	118,888
Abruzzo	30-39	117,898	112,911
Abruzzo	Other	288,371	269,041
Basilicata	50-59	75,604	71,946
Basilicata	60-69	67,252	64,934
Basilicata	40-49	62,170	58,771
		Sum: 46,084,474	**Sum: 43,218,923**

10) number of vaccine doses given in each region every day since there are so many regions we just took top 5 for last one year