FIT3179 Data Visualisation

Week 9 Studio: Interactive Charts with Vega-Lite

Part 1. Interactions in Vega-Lite

In this part, we will create a bubble plot in Vega-Lite using a COVID-19 dataset, and then add some interactions to the visualisation.

The dataset contains the COVID-19 statistical data of all the countries in the world on 10 Oct 2020. The data is available here:
https://raw.githubusercontent.com/KaneSec/vega_lite/main/4_interactive_scatter_plot/data/COVID_19_10_Oct_2020.csv

The final visualisation will look like this:
https://kanesec.github.io/vega_lite/4_interactive_scatter_plot/

The example GitHub repository is available here:
https://github.com/KaneSec/vega_lite/tree/main/4_interactive_scatter_plot

1.1 Building a bubble plot

We will first build a basic bubble plot by defining the mark as "circle", i.e., each circle represents a country. The following visual variables are used to encode four data attributes:

- x-axis (position on a common scale): Confirmed Cases
- y-axis (position on a common scale): Deaths
- Colour hue: Continent
- Size (2D Size, Area): Population

The bubble plot and the corresponding Vega-Lite JSON code is shown below.

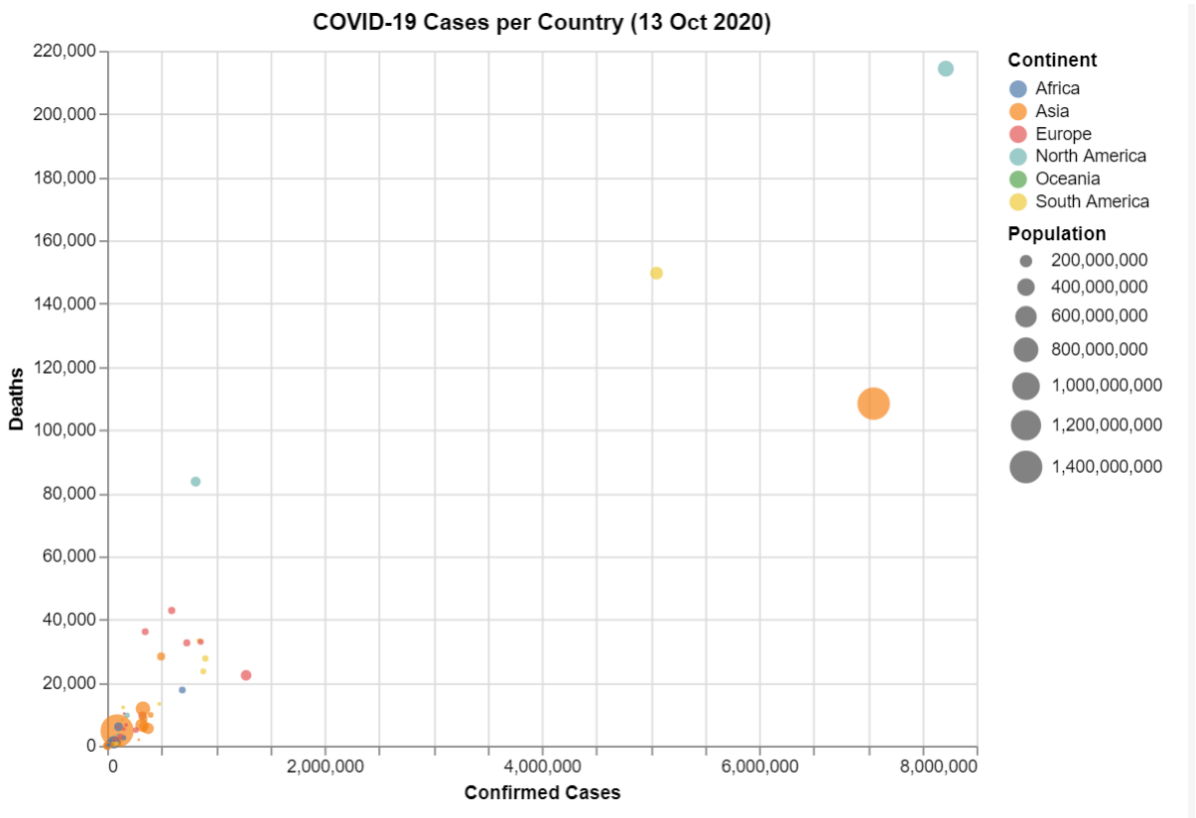

Figure 1. A basic bubble plot.

```
{
  "$schema": "https://vega.github.io/schema/vega-lite/v5.json",
  "width": 500,
  "height": 400,
  "title": "COVID-19 Cases per Country (13 Oct 2020)",
  "data": {
    "url": "https://raw.githubusercontent.com/KaneSec/vega_lite/main/4_interac
tive_scatter_plot/data/COVID_19_10_Oct_2020.csv"
  },
  "mark": "circle",
  "encoding": {
    "x": {
      "field": "Confirmed",
      "type": "quantitative",
      "title": "Confirmed Cases"
    },
    "y": {
      "field": "Deaths",
      "type": "quantitative"
    },
    "color": {
      "field": "Continent",
      "type": "nominal"
    },
    "size": {
      "field": "Population",
       "type": "quantitative"
    }
  }
}
```

1.2 Customising the bubble plot

Next, we will change some parameter settings (e.g., colour scales, axis scales, etc.) to make our bubble plot more informative.

1.2.1 x-axis and y-axis

Both the Confirmed Cases and the Deaths data distribution are right-skewed (since some countries such as the US, Brazil, India have significantly larger numbers of cases compared to the other countries). The default axis uses a linear scale, which results in that most countries are clustered in the bottom-left corner (Figure 1). We will use a log scale for both our x-axis and y-axis. The result and the corresponding code are shown below.

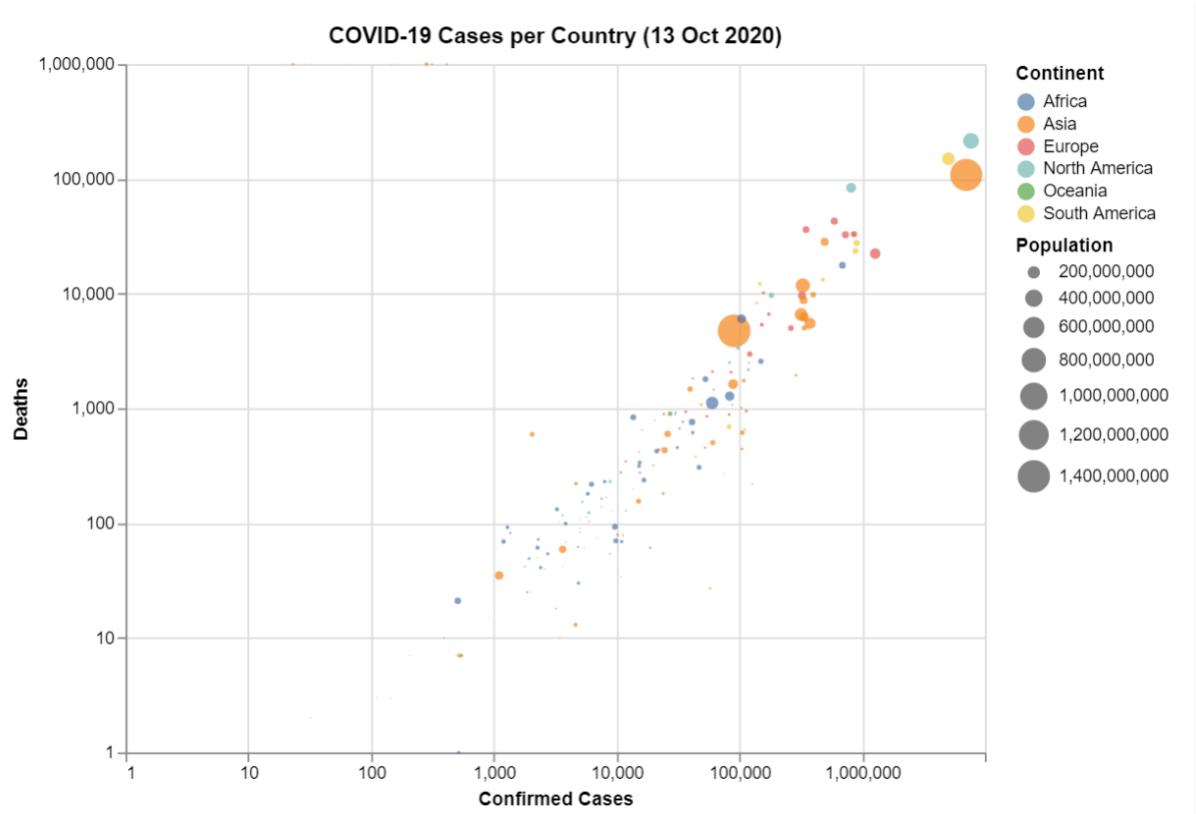

Figure 2. log scales of the x-axis and y-axis.

We need to define a scale with a type of "log". To remove the complex background grid, we will also change the axis ticks with "tickCount" (which is the total number of ticks of x-axis and y-axis).

```
  "x": {
     "field": "Confirmed",
     "type": "quantitative",
     "title": "Confirmed Cases",
     "axis": {"tickCount": 7},
     "scale": {"type": "log", "domain": [1, 10000000]}
   },
   "y": {
     "field": "Deaths",
     "type": "quantitative",
     "axis": {"tickCount": 6},
     "scale": {"type": "log", "domain": [1, 1000000]}
   }
```

Since log transformation does not work for zero or negative numbers, we define a "filter" to filter out such values. To do so, we just need to define a "filter" inside the "transform" list, which can be added after the data loading part. See the code below:

```
  "data": {
    "url": "https://raw.githubusercontent.com/KaneSec/vega_lite/main/4_interac
tive_scatter_plot/data/COVID_19_10_Oct_2020.csv"
  },
  "transform": [
```

```
    {"filter": "datum.Active > 0"},
    {"filter": "datum.Deaths > 0"}
  ]
```

Note: Vega-Lite is capable of some basic data cleaning and aggregations. For more complicated data cleaning and integration, we can use JavaScript to pre-process the data and then send them for visualisation. If you are not familiar with JavaScript, it is recommended that you clean or integrate your data first before uploading them for visualisation in Vega-Lite.

1.2.2 Classification: changing the scale for "size"

Another parameter we would like to tune is the size of the points, which encodes the population of each country. Here we define a classification scale, similar to what we had in our week-8 Studio. The result and the code are shown below:

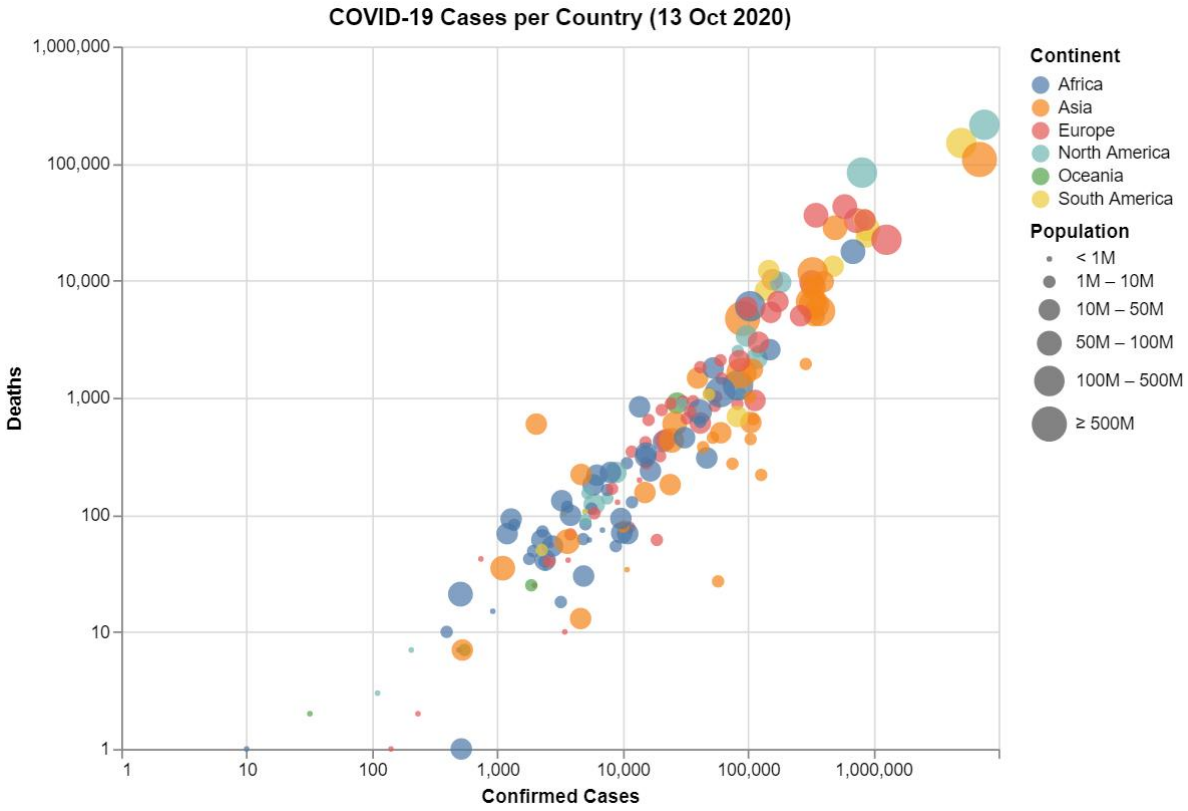

Figure 3. Classification scale for size.

Here the scale type is defined as "threshold". Then we use five thresholds to divide the domain into six classes. The correspondences are:

- Population below 1M -> size of the bubble is 10
- Population between 1M and 10M -> size of the bubble is 50
- …
- Population above 500M -> the bubble size is 400

We also define the number format to make the legend easier to read (Figure 3). ".1s" means SI-prefix with two significant digits: e.g., 1000000 will be represented as 1M. If you change the format to ".2s", then 1000000 will be represented as 1.0M. A format of "," will give you

“1,000,000”. For more details of format in Vega-lite, please check it here: https://vega.github.io/vega-lite/docs/format.html.

```
"size": {
    "field": "Population",
    "type": "quantitative",
    "scale": {
      "type": "threshold",
      "domain": [1000000, 10000000, 50000000, 100000000, 500000000],
      "range": [10, 50, 150, 200, 300, 400]
    },
    "legend": {"format": ".1s"}
  }
```

1.2.3 Colour scale

The default colour is already effective but could be further improved. Here we change the colour scale of each continent and try to differentiate those continents that are doing well in handling COVID-19 (e.g., Asia, Oceania) and those that are struggling (e.g., North America, Europe, etc.).

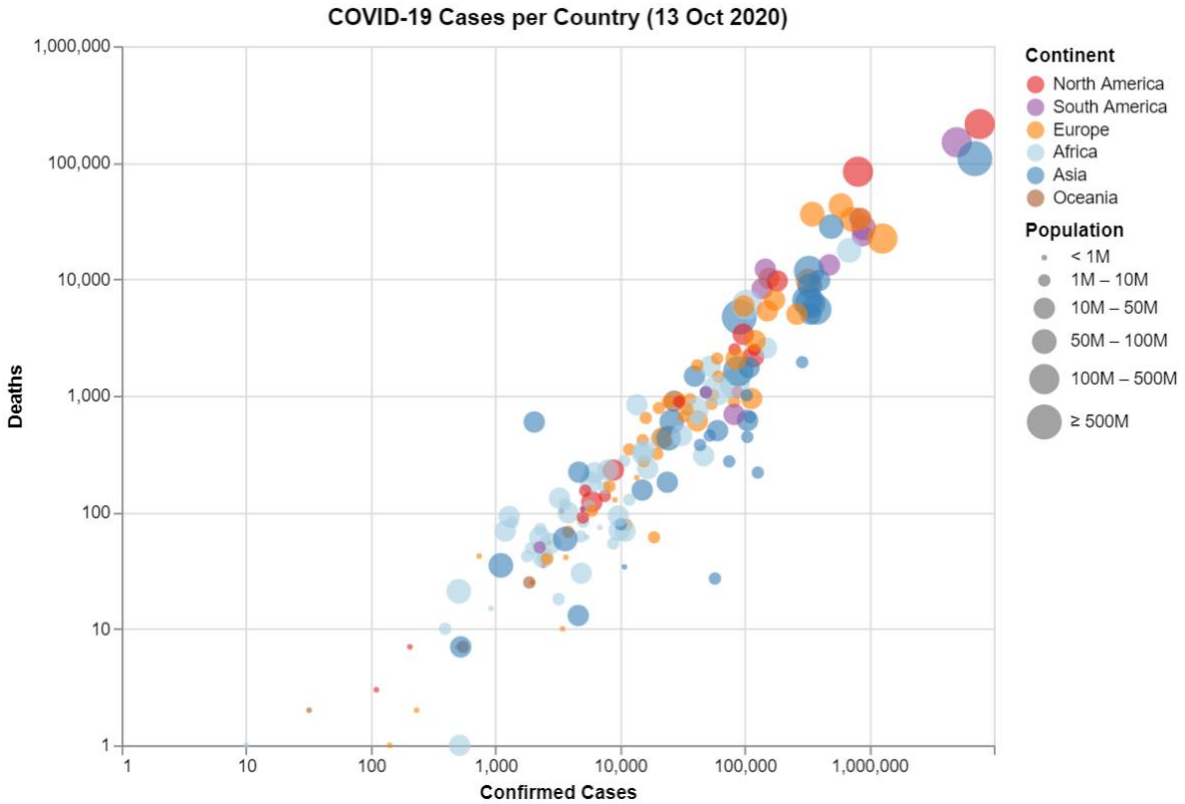

Figure 4. Changing the colour scale.

The code is shown below. We customise the colour scale with colours picked from https://colorbrewer2.org/. We can also define the transparency of the bubbles (“opacity”) to reduce visual clutter.

```
"color": {
```

```
        "field": "Continent",
        "type": "nominal",
        "scale": {
          "domain": [
            "North America",
            "South America",
            "Europe",
            "Africa",
            "Asia",
            "Oceania"
          ],
          "range": [
            "#e41a1c",
            "#984ea3",
            "#ff7f00",
            "#a6cee3",
            "#377eb8",
            "#a65628"
          ]
        }
      },
    "opacity": {
          "value": 0.6
        }
```

End of 1.2: Let's wrap up the code that we have so far. You can copy the following JSON code to the Vega Editor and view the result.

```
{
  "$schema": "https://vega.github.io/schema/vega-lite/v5.json",
  "width": 500,
  "height": 400,
  "title": "COVID-19 Cases per Country (13 Oct 2020)",
  "data": {
    "url": "https://raw.githubusercontent.com/KaneSec/vega_lite/main/4_interac
tive_scatter_plot/data/COVID_19_10_Oct_2020.csv"
  },
  "transform": [
    {"filter": "datum.Active > 0"},
    {"filter": "datum.Deaths > 0"}],
  "mark": "circle",
  "encoding": {
    "x": {
      "field": "Confirmed",
      "type": "quantitative",
      "title": "Confirmed Cases",
      "axis": {"tickCount": 7},
      "scale": {"type": "log", "domain": [1, 10000000]}
    },
    "y": {
      "field": "Deaths",
      "type": "quantitative",
      "axis": {"tickCount": 6},
      "scale": {"type": "log", "domain": [1, 1000000]}
    },
```

```
    "color": {
          "field": "Continent",
          "type": "nominal",
          "scale": {
            "domain": [
              "North America",
              "South America",
              "Europe",
              "Africa",
              "Asia",
              "Oceania"
            ],
            "range": [
              "#e41a1c",
              "#984ea3",
              "#ff7f00",
              "#a6cee3",
              "#377eb8",
              "#a65628"
            ]
          }
        },
    "opacity": {
      "value": 0.6
    },
    "size": {
          "field": "Population",
          "type": "quantitative",
          "scale": {
            "type": "threshold",
            "domain": [1000000, 10000000, 50000000, 100000000, 500000000],
            "range": [10, 50, 150, 200, 300, 400]
          },
          "legend": {"format": ".1s"}
        }
  }
}
```

1.3 Adding a tooltip

In this section, we will add a tooltip to our bubble plot. Tooltips can provide some detailed information when users hover over the data points (countries) in our bubble plot. The information that we would like to show in our tooltip includes:

- Country name
- Confirmed cases, Active case, Deaths, Recovered number, and
- Cases per 10,000 Population

We have no information related to “Cases per 10,000 Population” yet, so we will need to calculate this first. To do so, we can add the following code to the “transform” part:

```
    {
      "calculate": "datum.Confirmed/datum.Population * 10000",
      "as": "Cases per 10,000 Population"
    }
```

Then, we can use the following code to define our tooltip. They should be part of the "encoding".

```
"tooltip": [
    {"field": "Country", "type": "nominal"},
    {"field": "Confirmed", "type": "quantitative", "format": ","},
    {"field": "Active", "type": "quantitative", "format": ","},
    {"field": "Deaths", "type": "quantitative", "format": ","},
    {"field": "Recovered", "type": "quantitative", "format": ","},
    {
      "field": "Cases per 10,000 Population",
      "type": "quantitative",
      "format": ".2f"
    }
  ]
```

The result is shown below. We can see a nicely formatted tooltip when we hover over a point in our bubble plot.

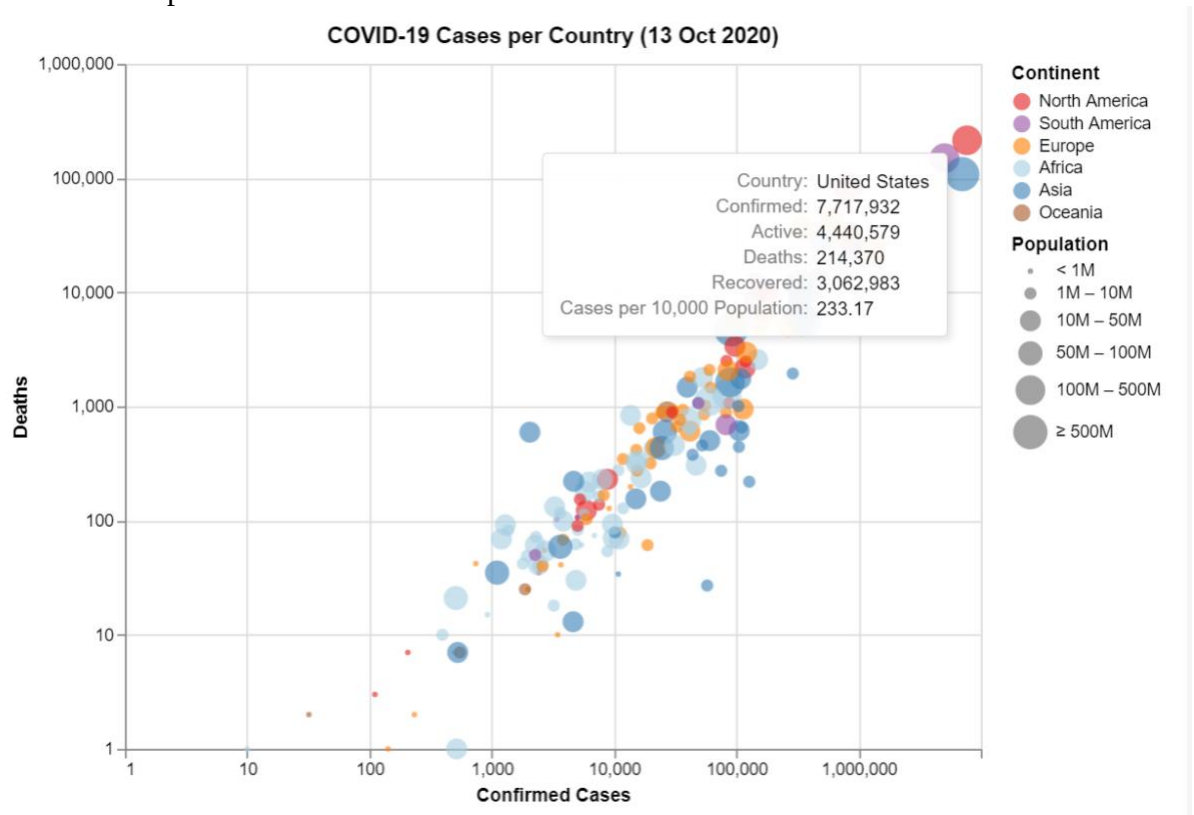

Figure 5. Adding a tooltip

The complete code is presented below:

```
{
  "$schema": "https://vega.github.io/schema/vega-lite/v5.json",
  "width": 500,
  "height": 400,
  "title": "COVID-19 Cases per Country (13 Oct 2020)",
  "data": {
```

```
    "url": "https://raw.githubusercontent.com/KaneSec/vega_lite/main/4_interac
tive_scatter_plot/data/COVID_19_10_Oct_2020.csv"
  },
  "transform": [
    {"filter": "datum.Active > 0"},
    {"filter": "datum.Deaths > 0"},
    {
      "calculate": "datum.Confirmed/datum.Population * 10000",
      "as": "Cases per 10,000 Population"
    }],
  "mark": "circle",
  "encoding": {
    "x": {
      "field": "Confirmed",
      "type": "quantitative",
      "title": "Confirmed Cases",
      "axis": {"tickCount": 7},
      "scale": {"type": "log", "domain": [1, 10000000]}
    },
    "y": {
      "field": "Deaths",
      "type": "quantitative",
      "axis": {"tickCount": 6},
      "scale": {"type": "log", "domain": [1, 1000000]}
    },
    "color": {
          "field": "Continent",
          "type": "nominal",
          "scale": {
            "domain": [
              "North America",
              "South America",
              "Europe",
              "Africa",
              "Asia",
              "Oceania"
            ],
            "range": [
              "#e41a1c",
              "#984ea3",
              "#ff7f00",
              "#a6cee3",
              "#377eb8",
              "#a65628"
            ]
          }
        },
    "opacity": {
      "value": 0.6
    },
    "size": {
          "field": "Population",
          "type": "quantitative",
          "scale": {
            "type": "threshold",
            "domain": [1000000, 10000000, 50000000, 100000000, 500000000],
```

```
            "range": [10, 50, 150, 200, 300, 400]
          },
          "legend": {"format": ".1s"}
    },
    "tooltip": [
          {"field": "Country", "type": "nominal"},
          {"field": "Confirmed", "type": "quantitative", "format": ","},
          {"field": "Active", "type": "quantitative", "format": ","},
          {"field": "Deaths", "type": "quantitative", "format": ","},
          {"field": "Recovered", "type": "quantitative", "format": ","},
          {
            "field": "Cases per 10,000 Population",
            "type": "quantitative",
            "format": ".2f"
          }
        ]
  }
}
```

1.4 Filtering/selections

Overview first, zoom and filter, then details-on-demand.
[Visual information seeking mantra by Shneiderman]

Another powerful interaction technique is filtering/selection. In this section, we will demonstrate three different types of filtering/selection implementation supported in Vega-Lite, including selection with a legend, a selection menu, and a slider. For other types of selections, please check the corresponding Vega-Lite documentation: https://vega.github.io/vega-lite/docs/selection.html.

1.4.1 Direct selection on the legend

The first selection that we are going to implement is based on the legend. When the user clicks on a continent (in the legend), we will highlight all the countries belonging to the selected continent. To achieve this, we first define a "selection" before the mark ("points") is described.

```
 "selection": {
    "continent_highlight": {
      "type": "multi",
      "fields": ["Continent"],
      "bind": "legend"
    }
  },
```

Then, we change the opacity of the mark based on the legend selection. We define a condition to do so: if the countries satisfy the condition, the opacity value is 0.6; otherwise, the opacity value will be 0.2.

```
    "opacity": {
          "condition": {"selection": "continent_highlight", "value": 0.6},
          "value": 0.2
        }
```

The visualisation result is shown below:

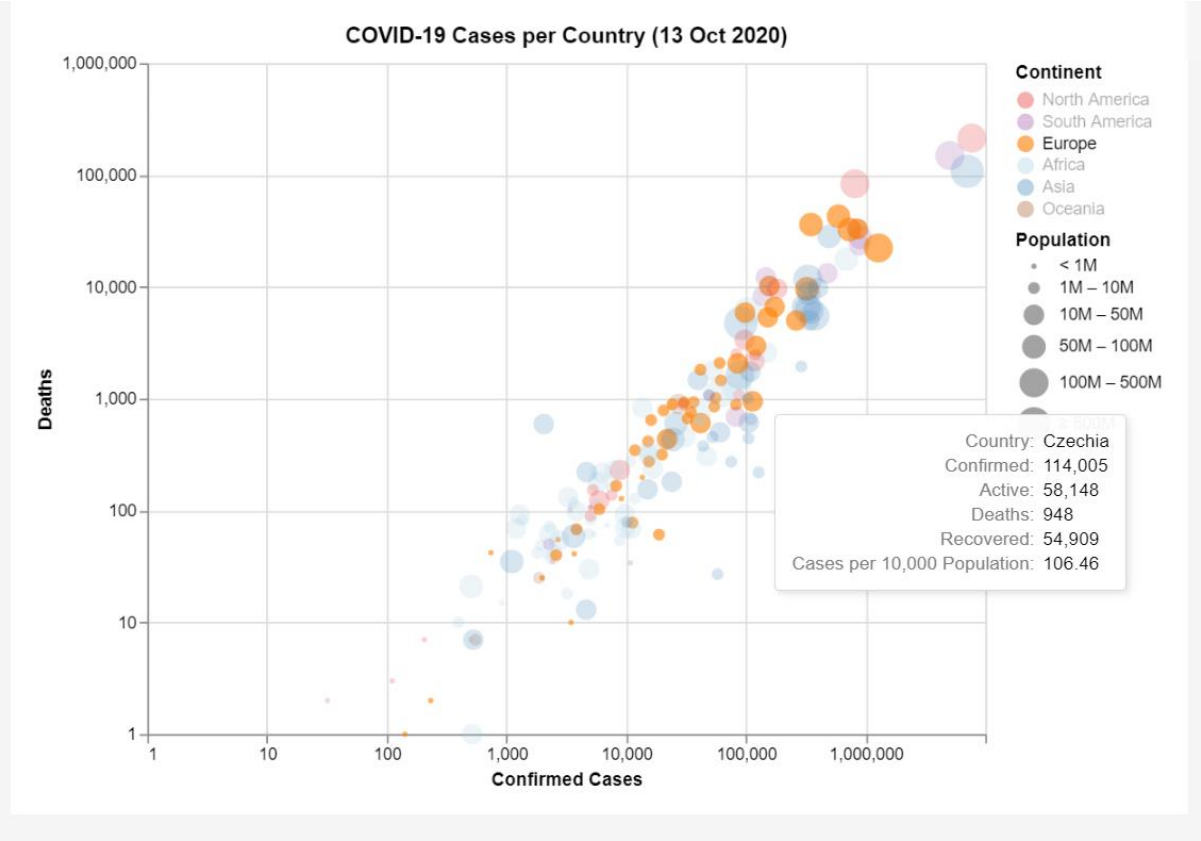

Figure 6. Legend Selection.

The full code is shown below:

```
{
  "$schema": "https://vega.github.io/schema/vega-lite/v5.json",
  "width": 500,
  "height": 400,
  "title": "COVID-19 Cases per Country (13 Oct 2020)",
  "data": {
    "url": "https://raw.githubusercontent.com/KaneSec/vega_lite/main/4_interac
tive_scatter_plot/data/COVID_19_10_Oct_2020.csv"
  },
  "transform": [
    {"filter": "datum.Active > 0"},
    {"filter": "datum.Deaths > 0"},
    {
      "calculate": "datum.Confirmed/datum.Population * 10000",
      "as": "Cases per 10,000 Population"
    }],
  "selection": {
        "continent_highlight": {
          "type": "multi",
          "fields": ["Continent"],
          "bind": "legend"
        }
      },
```

```
"mark": "circle",
"encoding": {
  "x": {
    "field": "Confirmed",
    "type": "quantitative",
    "title": "Confirmed Cases",
    "axis": {"tickCount": 7},
    "scale": {"type": "log", "domain": [1, 10000000]}
  },
  "y": {
    "field": "Deaths",
    "type": "quantitative",
    "axis": {"tickCount": 6},
    "scale": {"type": "log", "domain": [1, 1000000]}
  },
  "color": {
        "field": "Continent",
        "type": "nominal",
        "scale": {
          "domain": [
            "North America",
            "South America",
            "Europe",
            "Africa",
            "Asia",
            "Oceania"
          ],
          "range": [
            "#e41a1c",
            "#984ea3",
            "#ff7f00",
            "#a6cee3",
            "#377eb8",
            "#a65628"
          ]
        }
      },
  "opacity": {
        "condition": {"selection": "continent_highlight", "value": 0.6},
        "value": 0.2
  },
  "size": {
        "field": "Population",
        "type": "quantitative",
        "scale": {
          "type": "threshold",
          "domain": [1000000, 10000000, 50000000, 100000000, 500000000],
          "range": [10, 50, 150, 200, 300, 400]
        },
        "legend": {"format": ".1s"}
  },
  "tooltip": [
        {"field": "Country", "type": "nominal"},
        {"field": "Confirmed", "type": "quantitative", "format": ","},
        {"field": "Active", "type": "quantitative", "format": ","},
        {"field": "Deaths", "type": "quantitative", "format": ","},
```

```
            {"field": "Recovered", "type": "quantitative", "format": ","},
            {
              "field": "Cases per 10,000 Population",
              "type": "quantitative",
              "format": ".2f"
            }
          ]
    }
}
```

Please note that, from Vega-lite v5, there is an additional way to implement the legend selection, i.e., Legend binding. The code below will generate the same result in Figure 6. The main differences are highlighted in yellow. For more details about legend binding, please check it here: https://vega.github.io/vega-lite/docs/bind.html

```
{
  "$schema": "https://vega.github.io/schema/vega-lite/v5.json",
  "width": 500,
  "height": 400,
  "title": "COVID-19 Cases per Country (13 Oct 2020)",
  "data": {
    "url":
"https://raw.githubusercontent.com/KaneSec/vega_lite/main/4_interactive_scatter_plot/data/COVID_19_10_Oct_2
020.csv"
  },
  "transform": [
    {"filter": "datum.Active > 0"},
    {"filter": "datum.Deaths > 0"},
    {
      "calculate": "datum.Confirmed/datum.Population * 10000",
      "as": "Cases per 10,000 Population"
    }],
  "params":[
    {
      "name": "continent_highlight",
      "select": {"type": "point", "fields": ["Continent"]},
      "bind": "legend"
    }
  ],
  "mark": "circle",
  "encoding": {
    "x": {
      "field": "Confirmed",
      "type": "quantitative",
```

```
    "title": "Confirmed Cases",
    "axis": {"tickCount": 7},
    "scale": {"type": "log", "domain": [1, 10000000]}
  },
  "y": {
    "field": "Deaths",
    "type": "quantitative",
    "axis": {"tickCount": 6},
    "scale": {"type": "log", "domain": [1, 1000000]}
  },
  "color": {
      "field": "Continent",
      "type": "nominal",
      "scale": {
        "domain": [
          "North America",
          "South America",
          "Europe",
          "Africa",
          "Asia",
          "Oceania"
        ],
        "range": [
          "#e41a1c",
          "#984ea3",
          "#ff7f00",
          "#a6cee3",
          "#377eb8",
          "#a65628"
        ]
      }
    },
  "opacity": {
      "condition": {"param": "continent_highlight", "value": 0.6},
      "value": 0.2
  },
  "size": {
      "field": "Population",
      "type": "quantitative",
      "scale": {
```

```
        "type": "threshold",
        "domain": [1000000, 10000000, 50000000, 100000000, 500000000],
        "range": [10, 50, 150, 200, 300, 400]
      },
      "legend": {"format": ".1s"}
  },
  "tooltip": [
      {"field": "Country", "type": "nominal"},
      {"field": "Confirmed", "type": "quantitative", "format": ","},
      {"field": "Active", "type": "quantitative", "format": ","},
      {"field": "Deaths", "type": "quantitative", "format": ","},
      {"field": "Recovered", "type": "quantitative", "format": ","},
      {
        "field": "Cases per 10,000 Population",
        "type": "quantitative",
        "format": ".2f"
      }
    ]
 }
}
```

1.4.2 Filtering with a selection menu

We can also define a selection menu outside the legend to filter out some countries. Here we will still use "Continent" as an example. We will define a list of "params", and then add a param which is our continent selection. We add a "select" bind, and define the options as different continents. We also add a *null* option. *null* is a reserved keyword which means no-selection. To reduce the confusion of *null*, we change the label of this selection to "Show All" in our "labels", which controls the options that we show on the screen. For more details regarding binding in vague-lite, please check it here: https://vega.github.io/vega-lite/docs/bind.html.

Please note that, there are two names here. The first "name" is the selection name that we are going to use to reference this selection result in our later code; the second "name" is the value that is displayed on our page (in front of the selection menu), as shown in Figure 7.

```
  "params": [
    {
      "name": "Continent_selection",
      "bind": {
        "input": "select",
        "options": [
          null,
          "North America",
```

```
          "South America",
          "Europe",
          "Africa",
          "Asia",
          "Oceania"
        ],
        "labels":[
          "Show All",
          "North America",
          "South America",
          "Europe",
          "Africa",
          "Asia",
          "Oceania"
        ],

        "name": "Continent Selection: "
      }
    }
  ]
```

Then, just inside the “transform”, we add another filter option to show only the continent that is selected, or we will show all the continents if the selection is *null*. (Please note that: here we should use the value in the “Options” not the labels.)

```
    {"filter": "Continent_selection == null || datum.Continent == Continent_se
lection"}
```

The visualisation is shown in Figure 7. Only European countries are displayed after the selection.

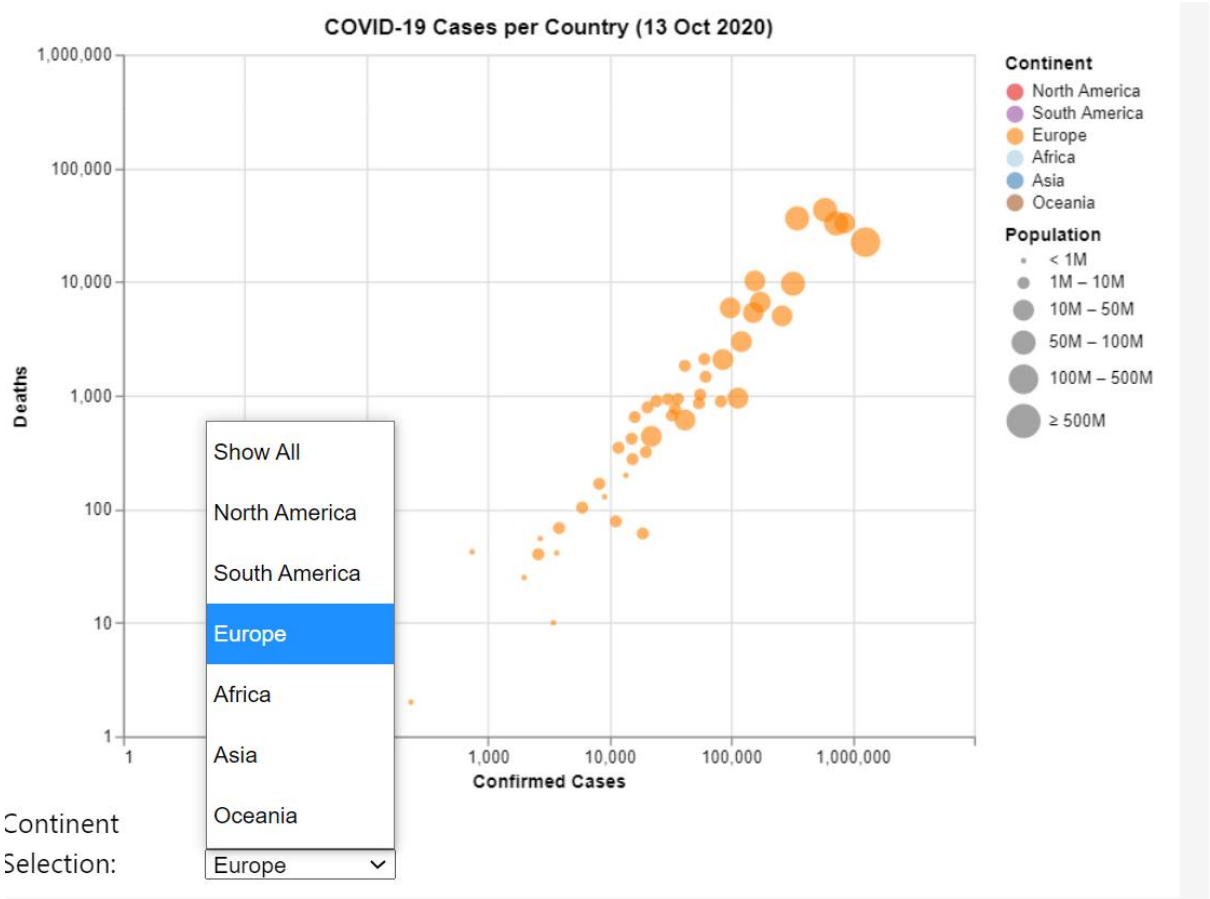

Figure 7. Selection menu.

The full code so far is shown below.

```
{
  "$schema": "https://vega.github.io/schema/vega-lite/v5.json",
  "width": 500,
  "height": 400,
  "title": "COVID-19 Cases per Country (13 Oct 2020)",
  "data": {
    "url": "https://raw.githubusercontent.com/KaneSec/vega_lite/main/4_interac
tive_scatter_plot/data/COVID_19_10_Oct_2020.csv"
  },
  "params": [
    {
      "name": "Continent_selection",
      "bind": {
        "input": "select",
        "options": [
          null,
          "North America",
          "South America",
          "Europe",
          "Africa",
          "Asia",
          "Oceania"
        ],
        "labels":[
```

```
          "Show All",
          "North America",
          "South America",
          "Europe",
          "Africa",
          "Asia",
          "Oceania"
        ],
        "name": "Continent Selection: "
      }
    }
  ],
  "transform": [
    {"filter": "datum.Active > 0"},
    {"filter": "datum.Deaths > 0"},
    {"filter": "Continent_selection == null || datum.Continent == Continent_se
lection"},
    {
      "calculate": "datum.Confirmed/datum.Population * 10000",
      "as": "Cases per 10,000 Population"
    }],
  "selection": {
        "continent_highlight": {
          "type": "multi",
          "fields": ["Continent"],
          "bind": "legend"
        }
      },
  "mark": "circle",
  "encoding": {
    "x": {
      "field": "Confirmed",
      "type": "quantitative",
      "title": "Confirmed Cases",
      "axis": {"tickCount": 7},
      "scale": {"type": "log", "domain": [1, 10000000]}
    },
    "y": {
      "field": "Deaths",
      "type": "quantitative",
      "axis": {"tickCount": 6},
      "scale": {"type": "log", "domain": [1, 1000000]}
    },
    "color": {
          "field": "Continent",
          "type": "nominal",
          "scale": {
            "domain": [
              "North America",
              "South America",
              "Europe",
              "Africa",
              "Asia",
```

```
            "Oceania"
          ],
          "range": [
            "#e41a1c",
            "#984ea3",
            "#ff7f00",
            "#a6cee3",
            "#377eb8",
            "#a65628"
          ]
        }
      },
    "opacity": {
          "condition": {"selection": "continent_highlight", "value": 0.6},
          "value": 0.2
    },
    "size": {
          "field": "Population",
          "type": "quantitative",
          "scale": {
            "type": "threshold",
            "domain": [1000000, 10000000, 50000000, 100000000, 500000000],
            "range": [10, 50, 150, 200, 300, 400]
          },
          "legend": {"format": ".1s"}
    },
    "tooltip": [
          {"field": "Country", "type": "nominal"},
          {"field": "Confirmed", "type": "quantitative", "format": ","},
          {"field": "Active", "type": "quantitative", "format": ","},
          {"field": "Deaths", "type": "quantitative", "format": ","},
          {"field": "Recovered", "type": "quantitative", "format": ","},
          {
            "field": "Cases per 10,000 Population",
            "type": "quantitative",
            "format": ".2f"
          }
        ]
  }
}
```

1.4.3 Filtering with a slider

Another filtering option is to use a slider – this is more suitable for quantitative attributes (e.g., the population). Here we define a range bind to filter out the countries with a population value above a given threshold, which is defined by the slider. The code and the result are shown below.

```
  {
    "name": "Population_Above",
    "value": 0,
    "bind": {
      "input": "range",
      "min": 0,
      "max": 100000000,
      "step": 1000000,
```

```
      "name": "Minimum Population: "
    }
  }

  {"filter": "datum.Population > Population_Above"}
```

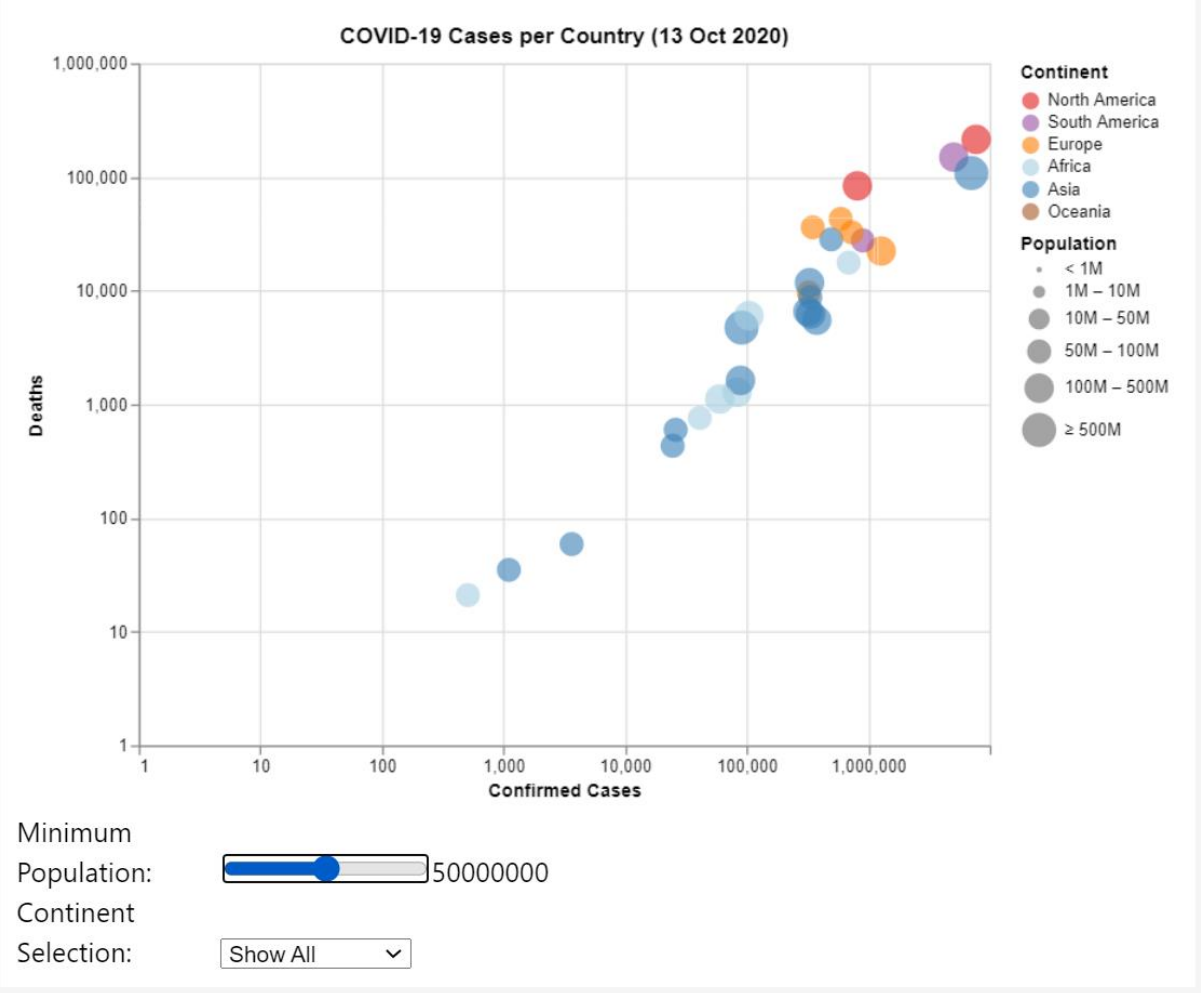

Figure 8. Filtering with a Slider.

The code that we have so far is shown below.

```
{
  "$schema": "https://vega.github.io/schema/vega-lite/v5.json",
  "width": 500,
  "height": 400,
  "title": "COVID-19 Cases per Country (13 Oct 2020)",
  "data": {
    "url": "https://raw.githubusercontent.com/KaneSec/vega_lite/main/4_interac
tive_scatter_plot/data/COVID_19_10_Oct_2020.csv"
  },
  "params": [
    {
      "name": "Population_Above",
      "value": 0,
      "bind": {
        "input": "range",
```

```
        "min": 0,
        "max": 100000000,
        "step": 1000000,
        "name": "Minimum Population: "
      }
    },
    {
      "name": "Continent_selection",
      "bind": {
        "input": "select",
        "options": [
          null,
          "North America",
          "South America",
          "Europe",
          "Africa",
          "Asia",
          "Oceania"
        ],
        "labels":[
          "Show All",
          "North America",
          "South America",
          "Europe",
          "Africa",
          "Asia",
          "Oceania"
        ],

        "name": "Continent Selection: "
      }
    }
  ],
  "transform": [
    {"filter": "datum.Active > 0"},
    {"filter": "datum.Deaths > 0"},
    {"filter": "datum.Population > Population_Above"},
    {"filter": "Continent_selection == null || datum.Continent == Continent_se
lection"},
    {
      "calculate": "datum.Confirmed/datum.Population * 10000",
      "as": "Cases per 10,000 Population"
    }],
  "selection": {
        "continent_highlight": {
          "type": "multi",
          "fields": ["Continent"],
          "bind": "legend"
        }
      },
  "mark": "circle",
  "encoding": {
    "x": {
```

```
  "field": "Confirmed",
  "type": "quantitative",
  "title": "Confirmed Cases",
  "axis": {"tickCount": 7},
  "scale": {"type": "log", "domain": [1, 10000000]}
},
"y": {
  "field": "Deaths",
  "type": "quantitative",
  "axis": {"tickCount": 6},
  "scale": {"type": "log", "domain": [1, 1000000]}
},
"color": {
      "field": "Continent",
      "type": "nominal",
      "scale": {
        "domain": [
          "North America",
          "South America",
          "Europe",
          "Africa",
          "Asia",
          "Oceania"
        ],
        "range": [
          "#e41a1c",
          "#984ea3",
          "#ff7f00",
          "#a6cee3",
          "#377eb8",
          "#a65628"
        ]
      }
    },
"opacity": {
      "condition": {"selection": "continent_highlight", "value": 0.6},
      "value": 0.2
},
"size": {
      "field": "Population",
      "type": "quantitative",
      "scale": {
        "type": "threshold",
        "domain": [1000000, 10000000, 50000000, 100000000, 500000000],
        "range": [10, 50, 150, 200, 300, 400]
      },
      "legend": {"format": ".1s"}
},
"tooltip": [
      {"field": "Country", "type": "nominal"},
      {"field": "Confirmed", "type": "quantitative", "format": ","},
      {"field": "Active", "type": "quantitative", "format": ","},
      {"field": "Deaths", "type": "quantitative", "format": ","},
      {"field": "Recovered", "type": "quantitative", "format": ","},
      {
        "field": "Cases per 10,000 Population",
```

```
            "type": "quantitative",
            "format": ".2f"
          }
        ]
  }
}
```

1.5 Text Annotations

Annotations can normally be used to provide some key information on top of the existing graph. Vega-Lite provides both text and line annotations. In this section, we will add text annotations on top of our bubble plot. For more examples related to other types of annotations, please check the Vega-Lite documentation here: https://vega.github.io/vega-lite/examples/.

To add an annotation, we need to define layers first: our original bubble plot will be our bottom layer, and the text annotations will be our top layer. We add a selection of country names on top of the bubble plot. As the points in the bubble plot and the text annotation share the same location, we can have the x and y encoding outside the layers. The code structure is shown below:

```
  "encoding": {
    "x": …
    "y":
  },
  "layer": [
      {
            Layer 1: bubble plot
      },
      {
            Layer 2: text annotations
      }
   ]
```

The following code creates the text annotation layer. We define the type of the mark as "text" and then adjust the position, alignment, and the font. Under "encoding", we will use the attribute "Country" (Country name) as our annotated text. We also define a test condition (https://vega.github.io/vega-lite/docs/condition.html) to only show those countries that we selected (other countries have an opacity of 0). We also add a tooltip to our text. Some additional information can be presented when users hover over either the points or the texts on the bubble plot.

```
   {
      "mark": {
        "type": "text",
        "align": "right",
        "baseline": "middle",
        "dx": -12,
        "fontSize": 11.5,
        "fontStyle": "italic"
      },
      "encoding": {
```

```
        "text": {"field": "Country", "type": "nominal"},
        "color": {"value": "black"},
        "opacity": {
          "condition": {
            "test": "datum['Country'] == 'China' || datum['Country'] == 'Singa
pore' || datum['Country'] == 'Australia' || datum['Country'] == 'New Zealand'
|| datum['Country'] == 'Italy' || datum['Country'] == 'Yemen' ||  datum['Count
ry'] == 'United States'",
            "value": 1
          },
          "value": 0
        },
        "tooltip": [
          {"field": "Country", "type": "nominal"},
          {"field": "Confirmed", "type": "quantitative", "format": ","},
          {"field": "Active", "type": "quantitative", "format": ","},
          {"field": "Deaths", "type": "quantitative", "format": ","},
          {"field": "Recovered", "type": "quantitative", "format": ","},
          {
            "field": "Cases per 10,000 Population",
            "type": "quantitative",
            "format": ".2f"
          }
        ]
      }
    }
```

The resulting visualisation is shown below. An interactive version is available at https://kanesec.github.io/vega_lite/4_interactive_scatter_plot/.

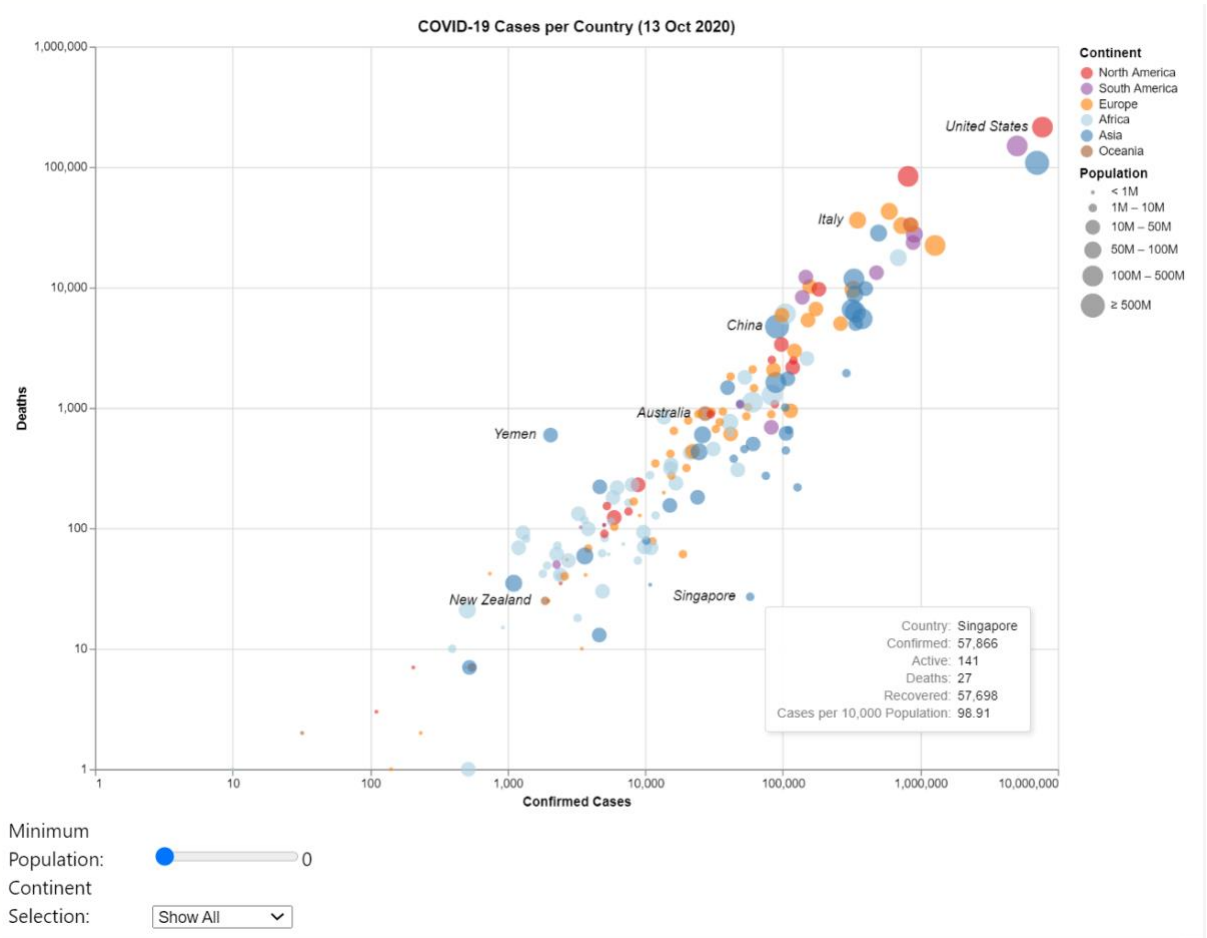

Figure 9. The final interactive visualisation.

The full code is shown below.

```
{
  "$schema": "https://vega.github.io/schema/vega-lite/v5.json",
  "width": 800,
  "height": 600,
  "title": "COVID-19 Cases per Country (13 Oct 2020)",
  "data": {
    "url":
"https://raw.githubusercontent.com/KaneSec/vega_lite/main/4_interactive_scatter_plot/data/COVID_19_10_Oct_2020.csv"
  },
  "params": [
    {
      "name": "Population_Above",
      "value": 0,
      "bind": {
        "input": "range",
```

```
        "min": 0,
        "max": 100000000,
        "step": 1000000,
        "name": "Minimum Population: "
      }
    },
    {
      "name": "Continent_selection",
      "bind": {
        "input": "select",
        "options": [
          null,
          "North America",
          "South America",
          "Europe",
          "Africa",
          "Asia",
          "Oceania"
        ],
        "labels":[
          "Show All",
          "North America",
          "South America",
          "Europe",
          "Africa",
          "Asia",
          "Oceania"
        ],
        "name": "Continent Selection: "
      }
    }
  ],
  "transform": [
    {"filter": "datum.Active > 0"},
    {"filter": "datum.Deaths > 0"},
    {"filter": "datum.Population > Population_Above"},
    {"filter": "Continent_selection == null || datum.Continent ==
Continent_selection"},
    {
      "calculate": "datum.Confirmed/datum.Population * 10000",
      "as": "Cases per 10,000 Population"
    }
```

```
],
"encoding": {
  "x": {
    "field": "Confirmed",
    "type": "quantitative",
    "title": "Confirmed Cases",
    "axis": {"tickCount": 7},
    "scale": {"type": "log", "domain": [1, 10000000]}
  },
  "y": {
    "field": "Deaths",
    "type": "quantitative",
    "axis": {"tickCount": 6},
    "scale": {"type": "log", "domain": [1, 1000000]}
  }
},
"layer": [
  {
    "selection": {
      "continent_highlight": {
        "type": "multi",
        "fields": ["Continent"],
        "bind": "legend"
      }
    },
    "mark": "circle",
    "encoding": {
      "size": {
        "field": "Population",
        "type": "quantitative",
        "scale": {
          "type": "threshold",
          "domain": [1000000, 10000000, 50000000, 100000000, 500000000],
          "range": [10, 50, 150, 200, 300, 400]
        },
        "legend": {"format": ".1s"}
      },
      "color": {
        "field": "Continent",
        "type": "nominal",
        "scale": {
          "domain": [
```

```
            "North America",
            "South America",
            "Europe",
            "Africa",
            "Asia",
            "Oceania"
          ],
          "range": [
            "#e41a1c",
            "#984ea3",
            "#ff7f00",
            "#a6cee3",
            "#377eb8",
            "#a65628"
          ]
        }
      },
      "opacity": {
        "condition": {"selection": "continent_highlight", "value": 0.6},
        "value": 0.2
      },
      "tooltip": [
        {"field": "Country", "type": "nominal"},
        {"field": "Confirmed", "type": "quantitative", "format": ","},
        {"field": "Active", "type": "quantitative", "format": ","},
        {"field": "Deaths", "type": "quantitative", "format": ","},
        {"field": "Recovered", "type": "quantitative", "format": ","},
        {
          "field": "Cases per 10,000 Population",
          "type": "quantitative",
          "format": ".2f"
        }
      ]
    }
  },
  {
    "mark": {
      "type": "text",
      "align": "right",
      "baseline": "middle",
      "dx": -12,
      "fontSize": 11.5,
```

```
        "fontStyle": "italic"
      },
      "encoding": {
        "text": {"field": "Country", "type": "nominal"},
        "color": {"value": "black"},
        "opacity": {
          "condition": {
            "test": "datum['Country'] == 'China' || datum['Country'] ==
'Singapore' || datum['Country'] == 'Australia' || datum['Country'] == 'New
Zealand' || datum['Country'] == 'Italy' || datum['Country'] == 'Yemen' ||
datum['Country'] == 'United States'",
            "value": 1
          },
          "value": 0
        },
        "tooltip": [
          {"field": "Country", "type": "nominal"},
          {"field": "Confirmed", "type": "quantitative", "format": ","},
          {"field": "Active", "type": "quantitative", "format": ","},
          {"field": "Deaths", "type": "quantitative", "format": ","},
          {"field": "Recovered", "type": "quantitative", "format": ","},
          {
            "field": "Cases per 10,000 Population",
            "type": "quantitative",
            "format": ".2f"
          }
        ]
      }
    }
  ]
}
```

Part 2. Multiple Visualisations on a Dashboard

Similar to a Tableau dashboard, we can include multiple visualisations in the same webpage with Vega-Lite. This is similar to the dashboard in Tableau.

There are a few options for creating multiple visualisations. One of the simplest options is to use the HTML+CSS. The following example uses Pure.css (that we learn in the week-5 studio) to integrate a few visualisations on the webpage. (Note: the visualisations and text on this page are for demonstration purpose only)

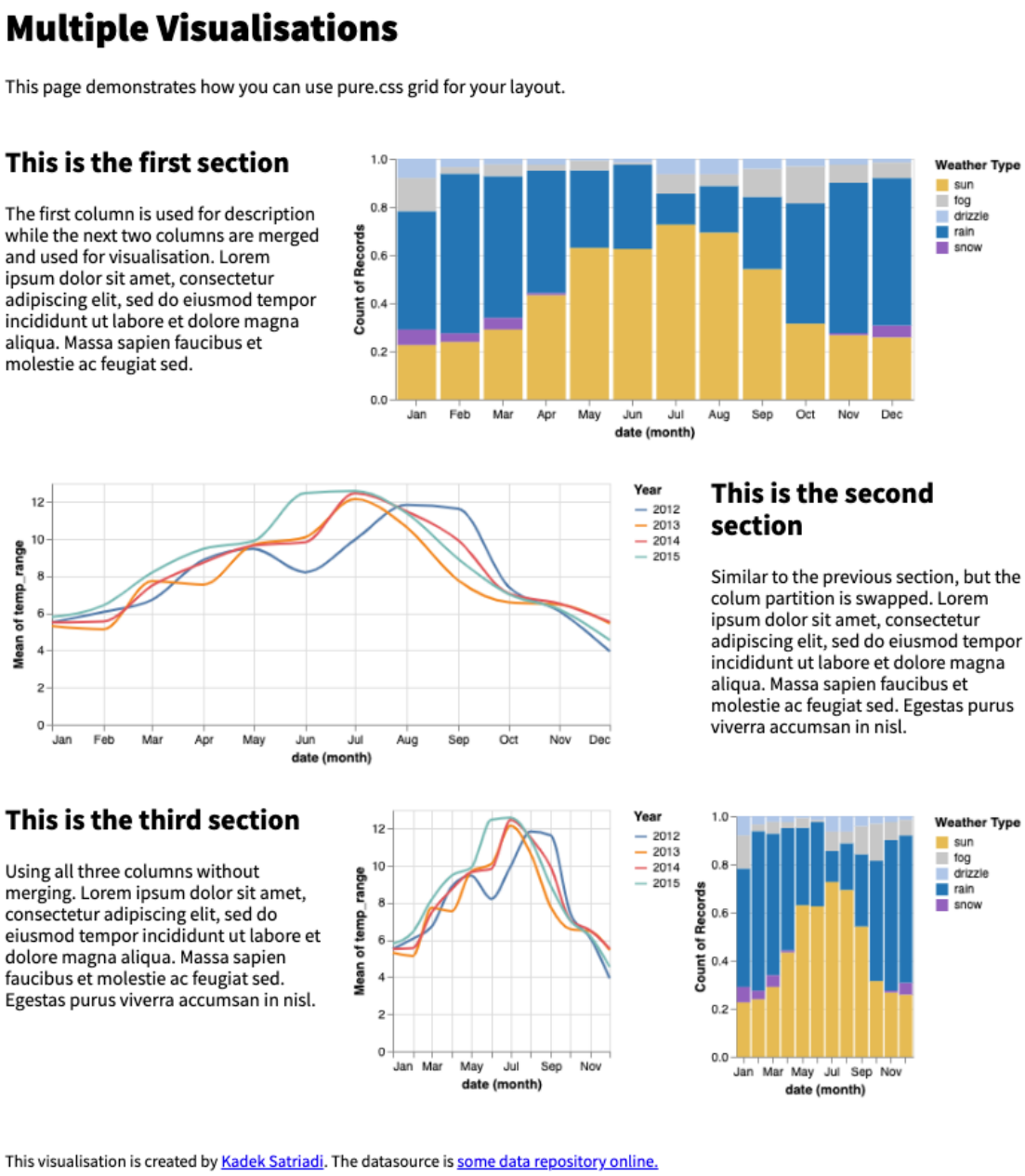

Figure 10. Multiple visualisations on a page.

- Example Page: https://kanesec.github.io/vega_lite/5_multiple_charts_html/
- Github: https://github.com/KaneSec/vega_lite/tree/main/5_multiple_charts_html

Pay attention to the following aspects while you are checking this example:

- How to control the layout of multiple visualisations with CSS (pure.CSS and self-defined CSS style).
- How to embed multiple vega-lite visualisation on the page with Javascript.
- How to hide the “Open in Vega Editor” function.

Vega-Lite itself also provides some options for multi-view displays (faceting) and interactive multi-view displays. We will discuss this in our next week’s studio. For preview, please check the examples on this page: https://vega.github.io/vega-lite/examples/.