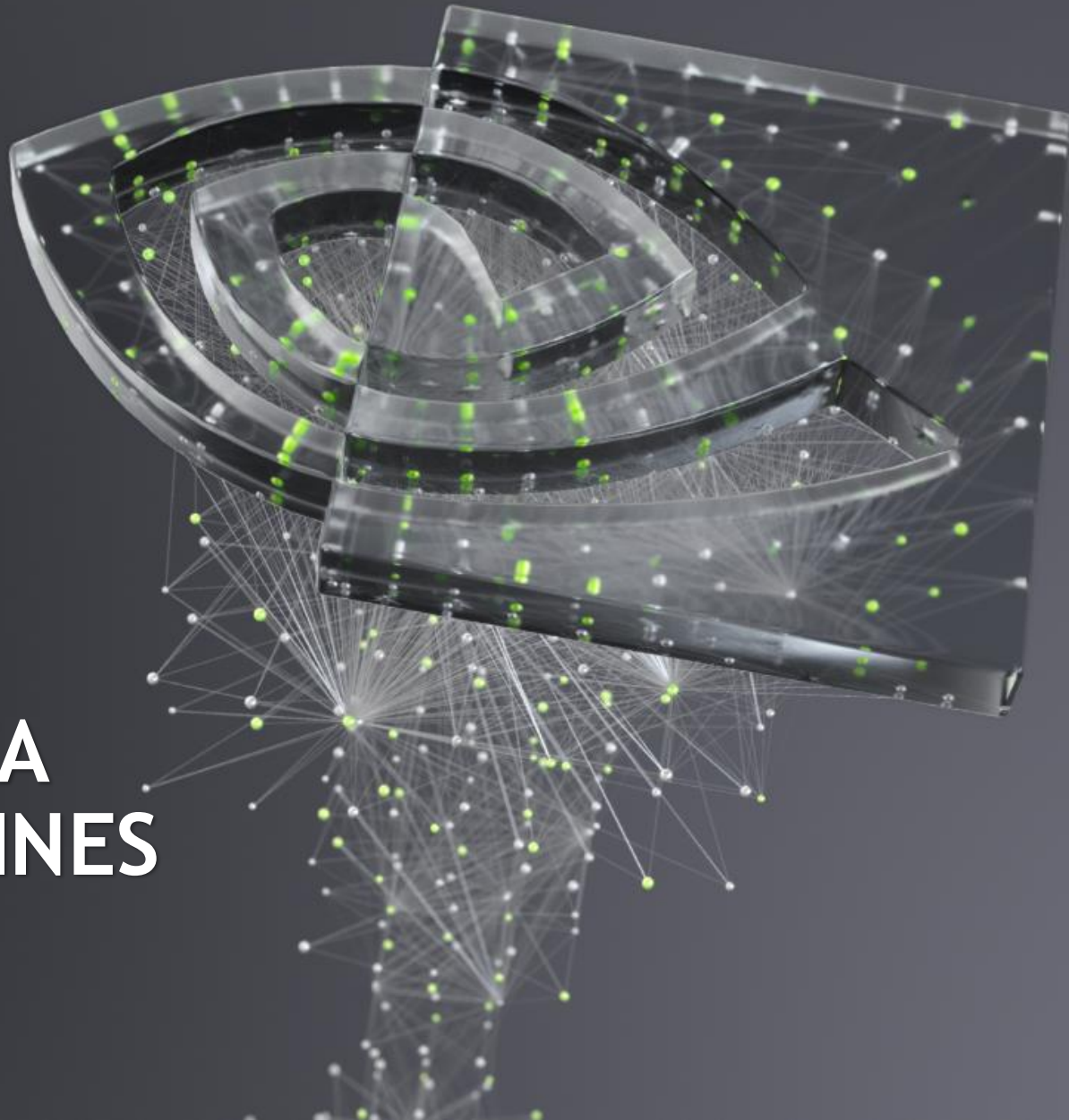




DEEP  
LEARNING  
INSTITUTE

# ACCELERATING DATA ENGINEERING PIPELINES

Part 3: Data Visualization



---

# AGENDA

Part 1: Data Formats

Part 2: ETL with NVTabular

Part 3: Data Visualization

# AGENDA – PART 3

- Data Quality
- Interactive Dashboards
- Plotly



DATA QUALITY

# BECOMING A DATA DETECTIVE

What is wrong with the following data?



# WHAT IS WRONG WITH THIS DATA?

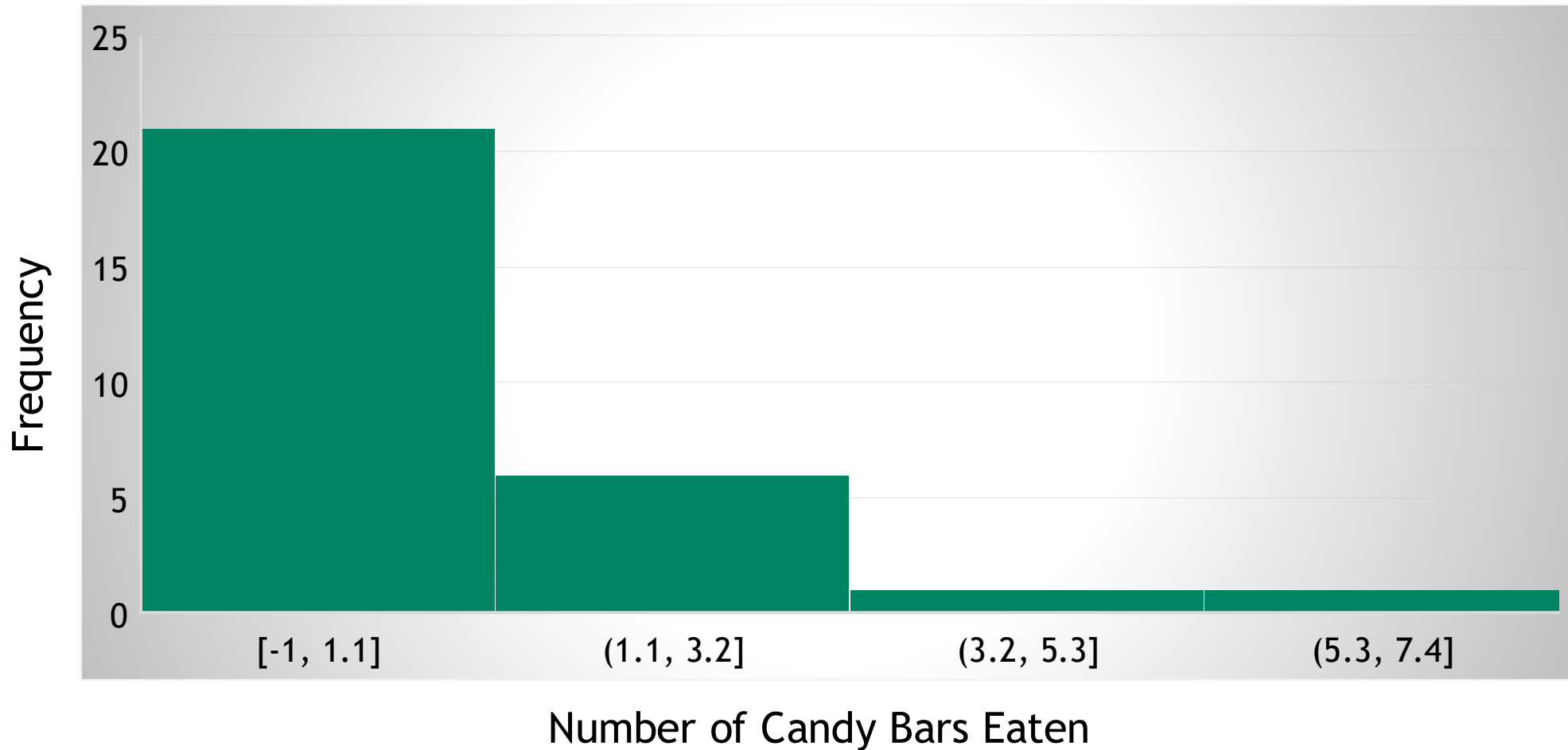
## North America Temperature

Temperature	Mexico City, Mexico	Washington, DC, USA	Ottawa, Canada
Average	14	36	-7
Median	12	37	-8
Mode	12	36	-8
Min	5	25	-14
Max	22	43	-5



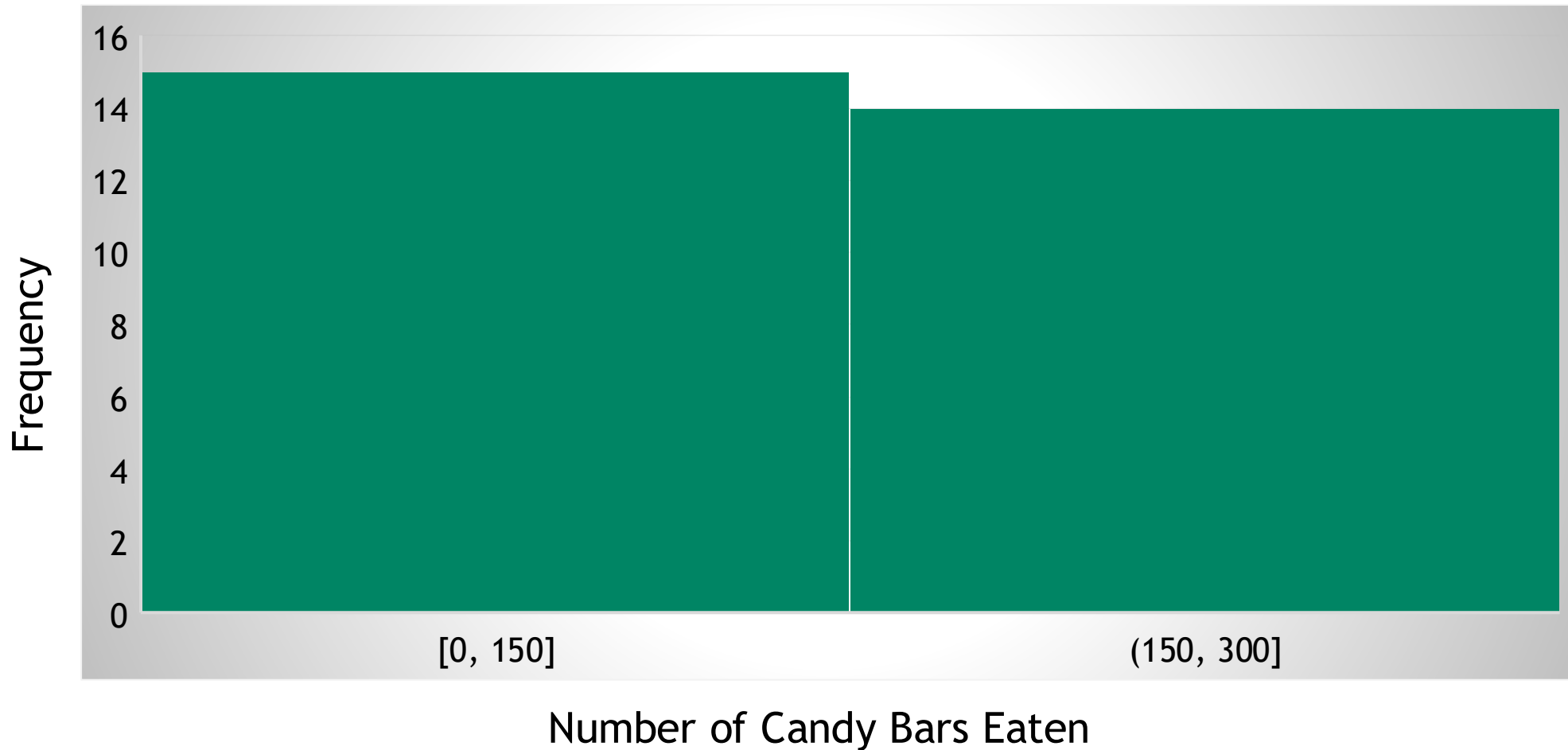
# WHAT IS WRONG WITH THIS DATA?

Candy Bars Eaten in a Day



# WHAT IS WRONG WITH THIS DATA?

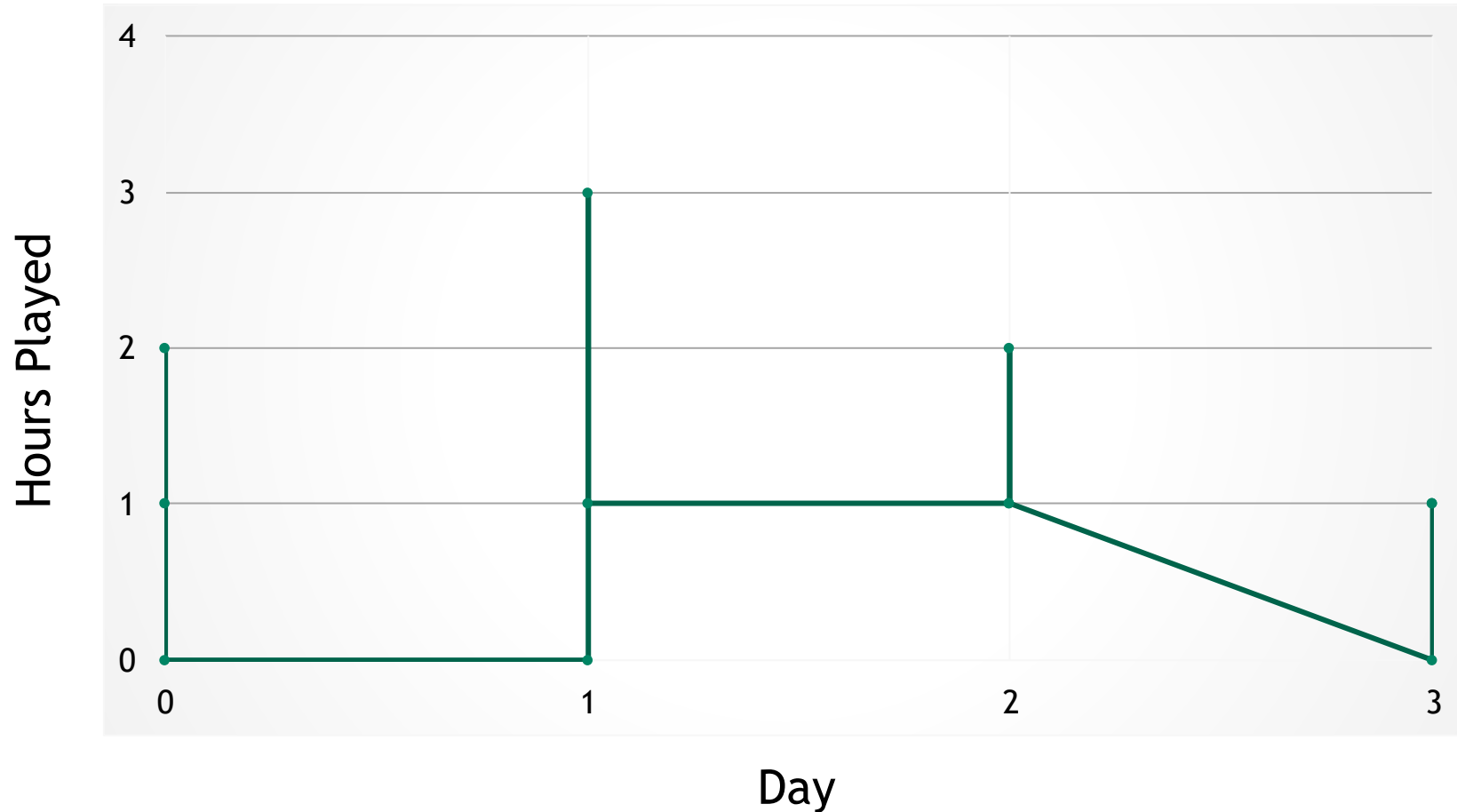
Candy Bars Eaten in a Day





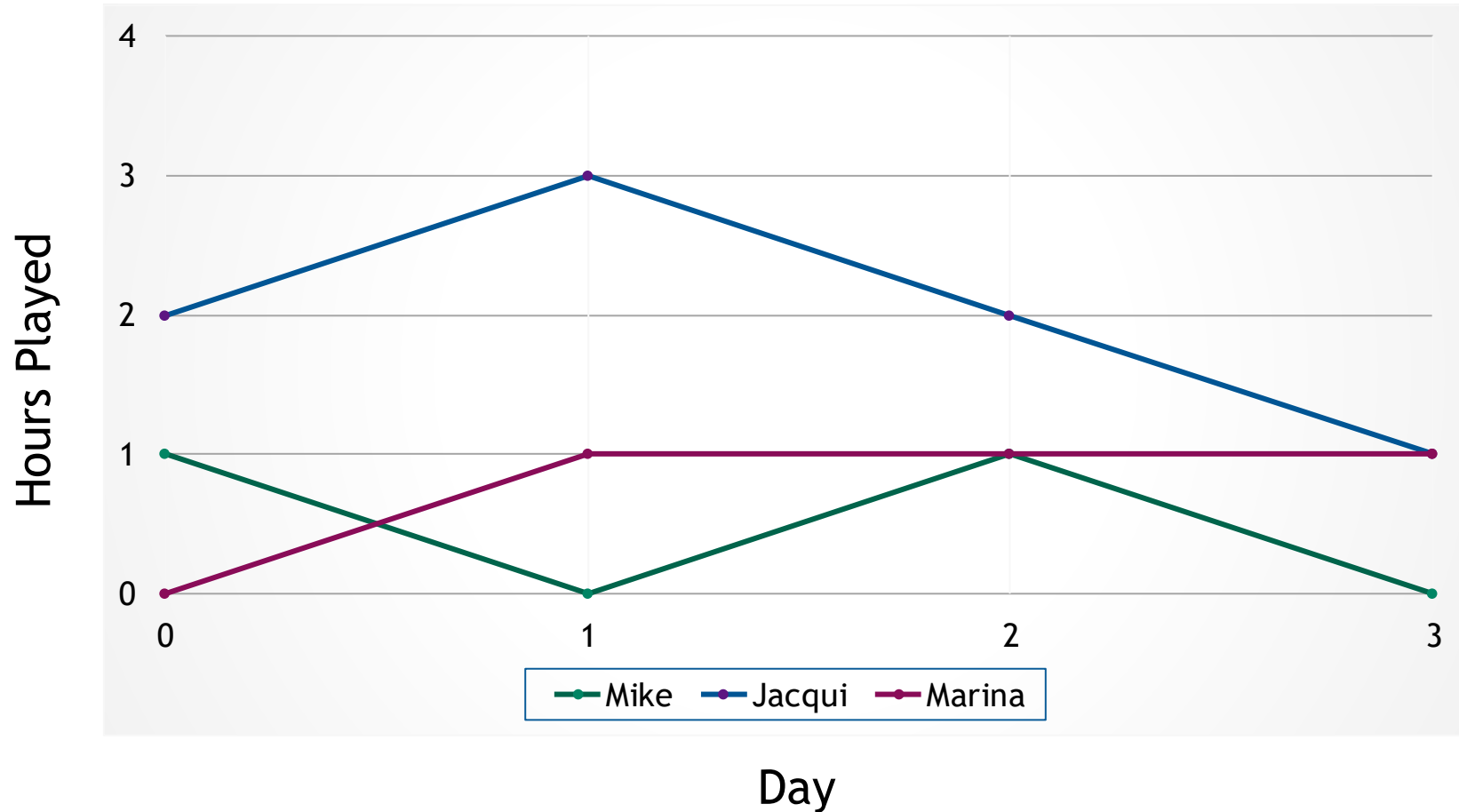
# WHAT IS WRONG WITH THIS DATA?

Hours Playing Video Games



# WHAT IS WRONG WITH THIS DATA?

Hours Playing Video Games



# WHAT IS WRONG WITH THIS DATA?

United States Precipitation

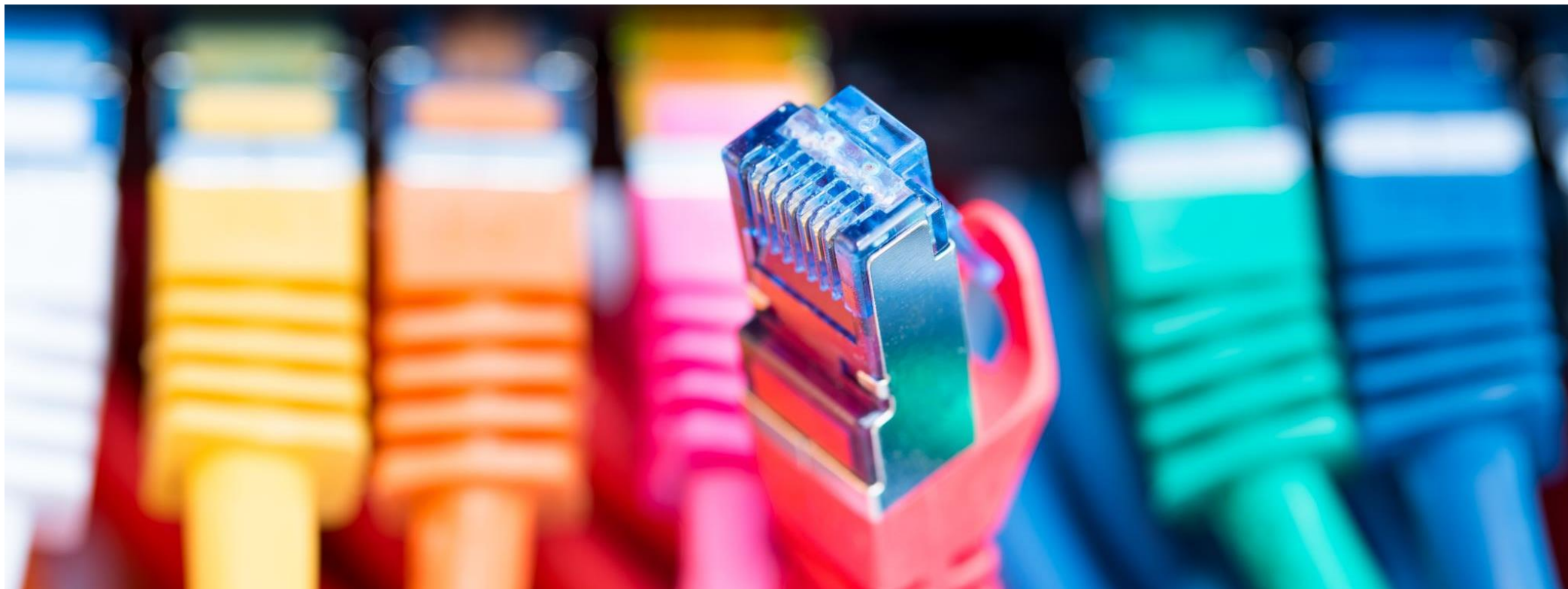




# INTERACTIVE DASHBOARDS

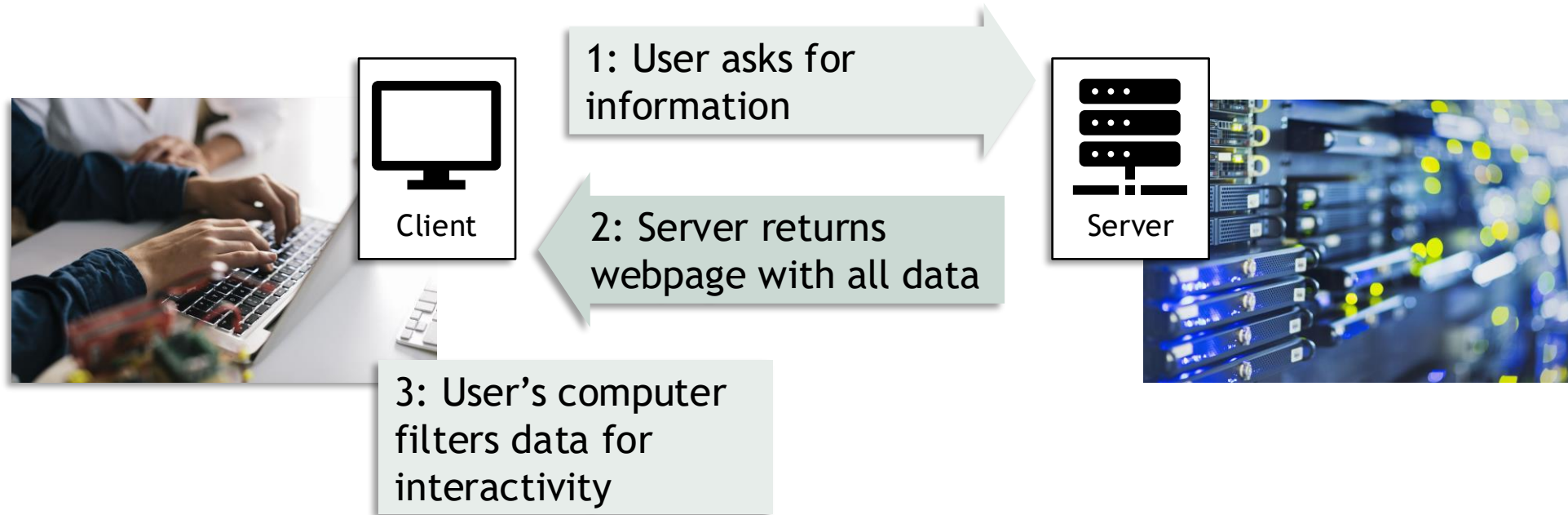
# THE INTERNET

The world's largest system of digital data?



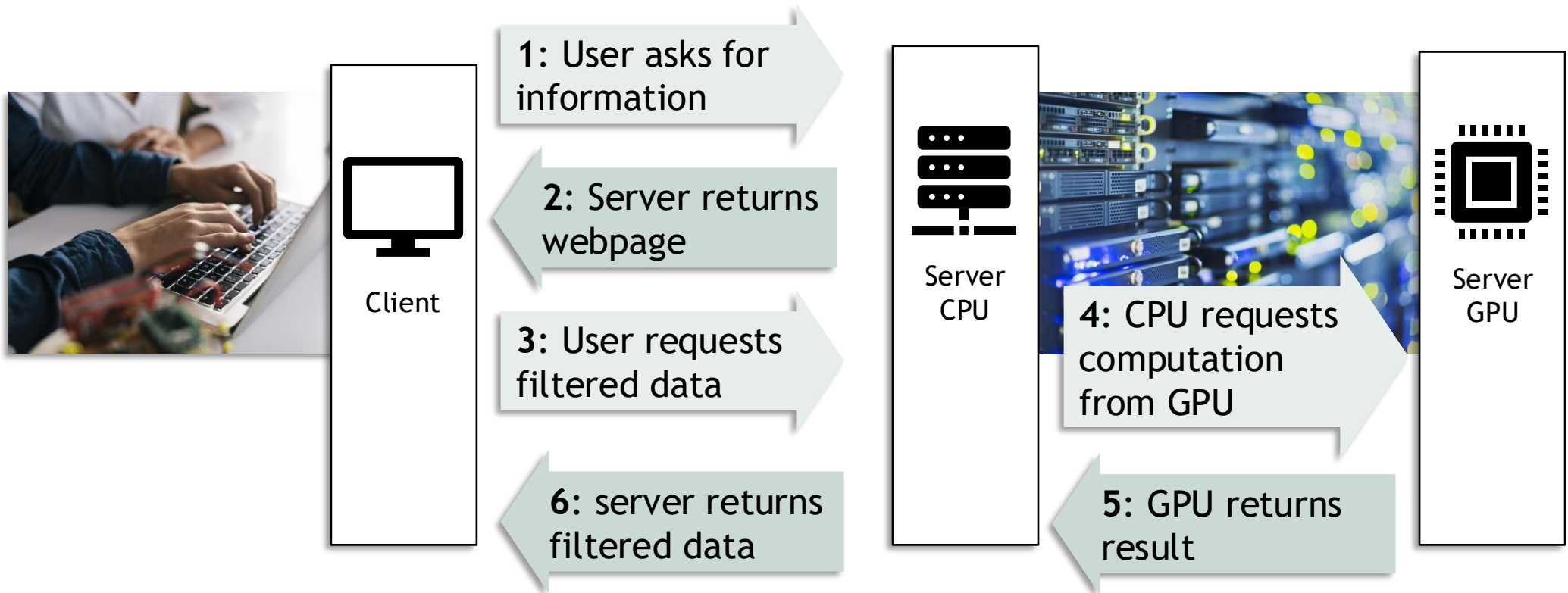
# A DASHBOARD DATA PIPELINE

## A Simple Server



# DASHBOARDS FOR BIG DATA

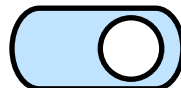
## A More Complex Server





# MAKING A GOOD DASHBOARD

Date: 1969-07-20



Show Zeros



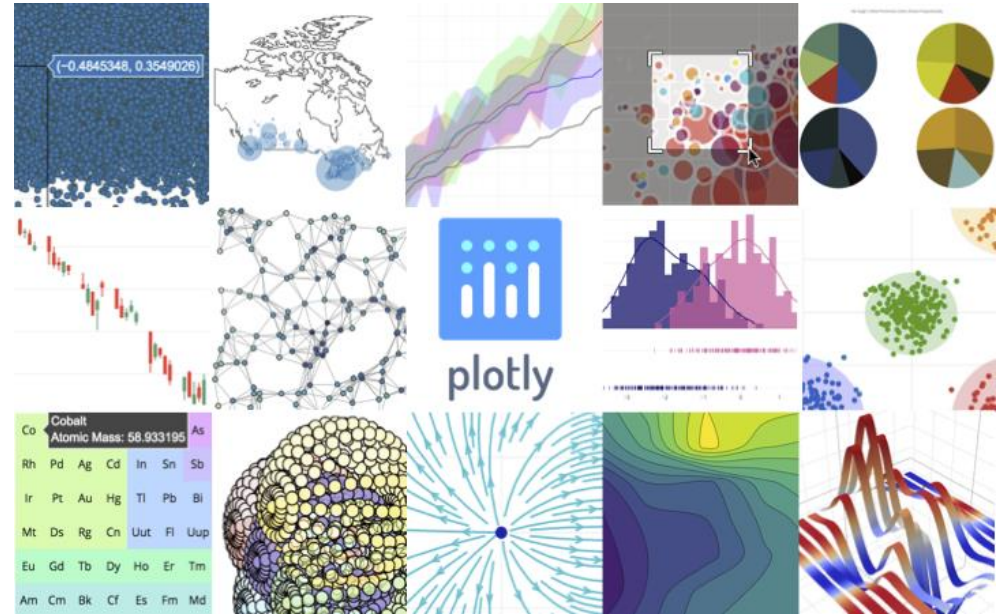


PLOTLY DASH

# PLOTLY DASH

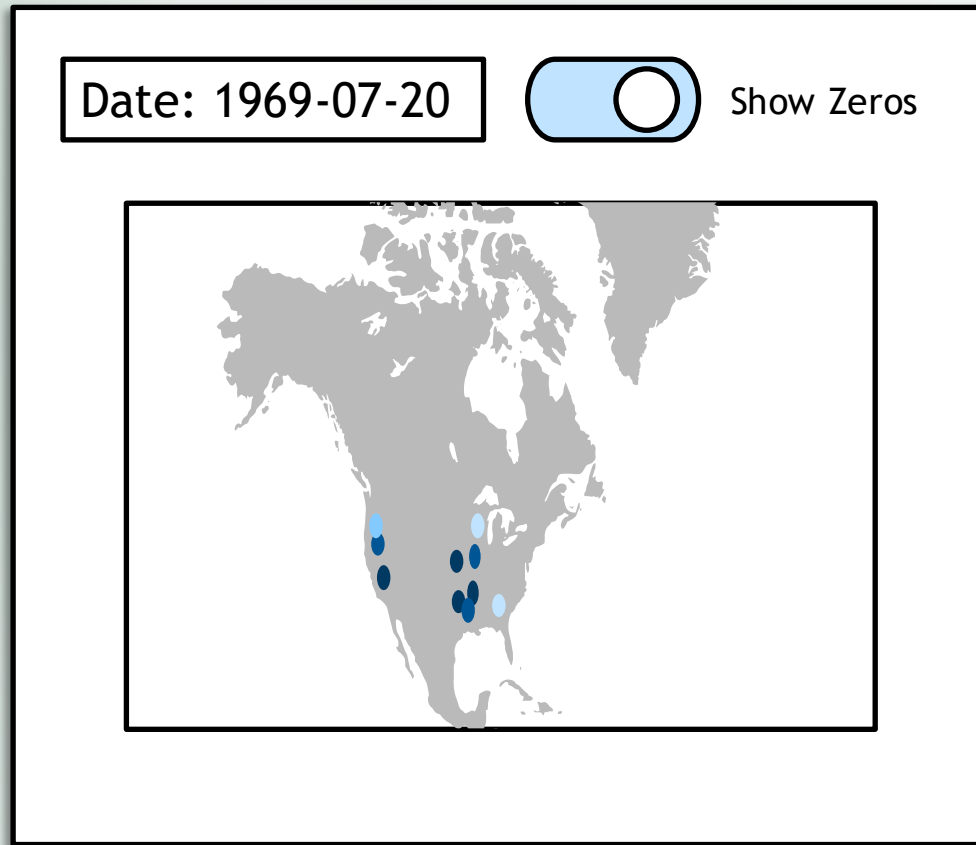
## The front end for ML and data science models

- ▶ **plotly.py** is an interactive, open-source, and browser-based graphing library for Python
- ▶ Supports line plots, scatter plots, area charts, bar charts, error bars, box plots, histograms, heatmaps, subplots, multiple-axes, polar charts, and bubble charts.



<https://github.com/plotly/plotly.py>

# MAKING A GOOD DASHBOARD



Div

Div

DatePickerSingle  
id="date-picker"

BooleanSwitch  
id="show-zeros"

Graph  
id="precipitation-map"

# MAKING A GOOD DASHBOARD

```
app.layout = html.Div([
    html.Div([
        dcc.DatePickerSingle(
            id='date-picker',
            min_date_allowed=date_min,
            max_date_allowed=date_max,
            initial_visible_month=initial_date,
            date=initial_date
        ),
        daq.BooleanSwitch(
            id='show-zeros',
            on=True,
            label="Show Zeros",
            style={'display': 'inline-block'}
        )
    ]),
    dcc.Graph(id='precipitation-map')
])
```

Div

Div

DatePickerSingle  
id="date-picker"

BooleanSwitch  
id="show-zeros"

Graph

id="precipitation-map"

# MAKING A GOOD DASHBOARD

```
app.layout = html.Div([
    html.Div([
        dcc.DatePickerSingle(
            id='date-picker',
            min_date_allowed=date_min,
            max_date_allowed=date_max,
            initial_visible_month=initial_date,
            date=initial_date
        ),
        daq.BooleanSwitch(
            id='show-zeros',
            on=True,
            label="Show Zeros",
            style={'display': 'inline-block'}
        )
    ]),
    dcc.Graph(id='precipitation-map')
])
```

```
@app.callback(
    Output('precipitation-map', `figure`),
    [Input('date-picker', 'date'), # 1st input
     Input('show-zeros', 'on')] # 2nd input
    def make_graph(first_input, second_input):
        ...

        regular python function stuff

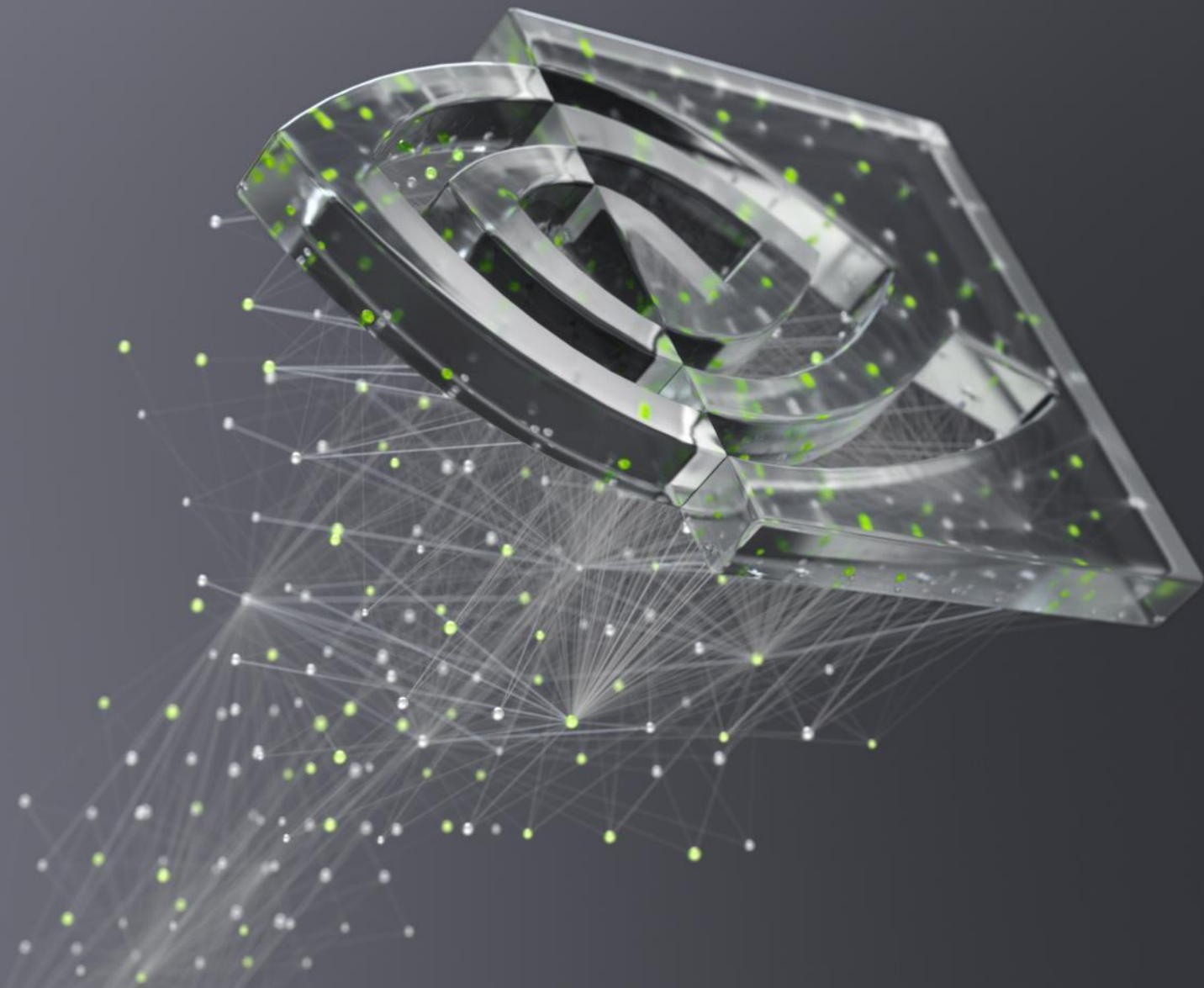
        ...

    return fig # Must match Output data type
```



LET'S GO!





DEEP  
LEARNING  
INSTITUTE