

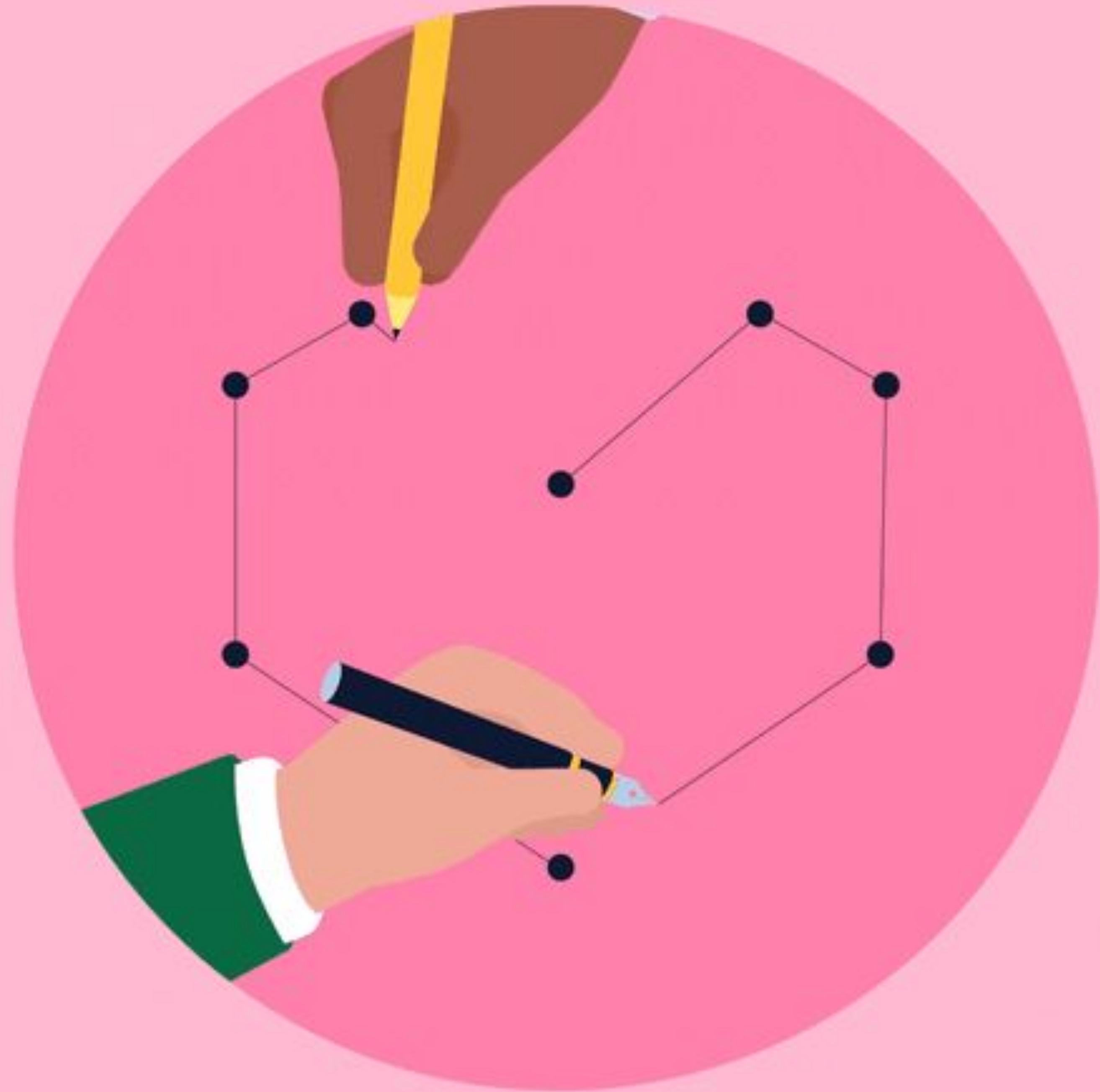
Connecting the Dots

Vishwajeetsingh Desurkar
@vishwadesurkar



SimplySmart

Lets connect the dots



Time series data



Challenges



Proactive Maintenance



Abnormal Usage



SO, WHAT'S THE SOLUTION?

Machine Learning



Yes. Yes!

Machine Learning with Ruby





Andrew Kane
ankane

Unfollow

5.1k followers · 133 following

San Francisco, CA
andrew@ankane.org
https://ankane.org

Organizations



Pinned

pgvector/pgvector Public

Open-source vector similarity search for Postgres

C 5.3k 207

searchkick Public

Intelligent search made easy

Ruby 6.2k 759

ahoy Public

Simple, powerful, first-party analytics for Rails

Ruby 3.9k 373

pghero Public

A performance dashboard for Postgres

Ruby 7.1k 414

chartkick Public

Create beautiful JavaScript charts with one line of Ruby

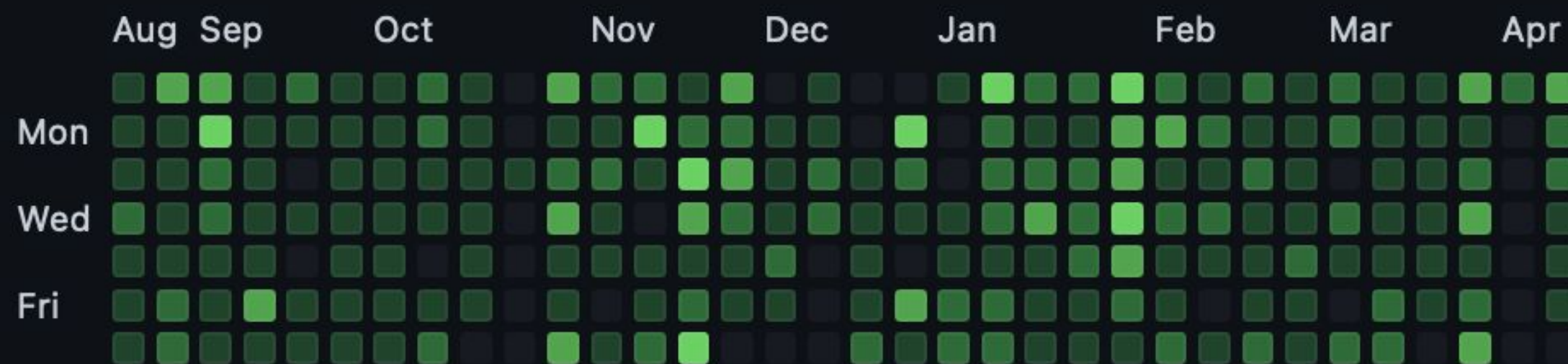
Ruby 6.1k 556

blazer Public

Business intelligence made simple

Ruby 3.7k 453

5,317 contributions in the last year



Learn how we count contributions



peacock

```
[1]: require "numo/narray"  
a = Numo::DFloat.new(3,5).seq
```

```
[1]: Numo::DFloat#shape=[3,5]  
[[0, 1, 2, 3, 4],  
 [5, 6, 7, 8, 9],  
 [10, 11, 12, 13, 14]]
```

```
[2]: a.shape
```

```
[2]: [3, 5]
```

Numo Array

```
require "rover-df"
data = Rover.read_csv("./Downloads/data.csv")
```

| time_stamp | thing_key | counts | pulsel_avg | pulsel_max | pulsel_min | pulseh_avg | ... | mainpower_min | syncrtc_avg | syncrtc_max | syncrtc_min |
|-------------------------------------|------------|--------|----------------------|------------|------------|------------|-----|---------------|-------------|-------------|-------------|
| 2023-06-21 19:30:00 +0000 UTC | 14bc91tb76 | 55 | 60507.0 | 60507 | 60507 | 0 | | 0 | 0 | 0 | 0 |
| 2023-06-21 20:30:00 +0000 UTC | 14bc91tb76 | 3 | 60507.0 | 60507 | 60507 | 0 | ... | 0 | 0 | 0 | 0 |
| 2023-06-21 21:30:00 +0000 UTC | 14bc91tb76 | 63 | 60507.0 | 60507 | 60507 | 0 | | 0 | 0 | 0 | 0 |
| ⋮ | | | | | | | | | | | |
| 2023-06-22 16:30:00 +0000 UTC | 14bc91tb76 | 12 | 60698.83333333333336 | 60699 | 60697 | 0 | | 0 | 0 | 0 | 0 |
| 2023-06-22 17:30:00 +0000 UTC | 14bc91tb76 | 140 | 60699.0 | 60699 | 60699 | 0 | ... | 0 | 0 | 0 | 0 |
| 2023-06-22 18:30:00 +0000 UTC | 14bc91tb76 | 88 | 60699.0 | 60699 | 60699 | 0 | | 0 | 0 | 0 | 0 |

Rover

• [6]: `require 'rumale'`

```
samples, labels = Rumale::Dataset.load_libsvm_file('./Downloads/pendigits')
```

```
transformer = Rumale::KernelApproximation::RBF.new(gamma: 0.0001, n_components: 1024, random_seed: 1)
```

```
transformed = transformer.fit_transform(samples)
```

```
classifier = Rumale::LinearModel::SVC.new(reg_param: 0.0001)
```

```
classifier.fit(transformed, labels)
```

```
File.open('transformer.dat', 'wb') { |f| f.write(Marshal.dump(transformer)) }
```

```
File.open('classifier.dat', 'wb') { |f| f.write(Marshal.dump(classifier)) }
```

[6]: 82502

• [7]: `samples, labels = Rumale::Dataset.load_libsvm_file('./Downloads/pendigits.t')`

```
transformer = Marshal.load(File.binread('transformer.dat'))
```

```
classifier = Marshal.load(File.binread('classifier.dat'))
```

```
transformed = transformer.transform(samples)
```

```
puts("Accuracy: %.1f%%" % (100.0 * classifier.score(transformed, labels)))
```

Accuracy: 98.5%

Rumale

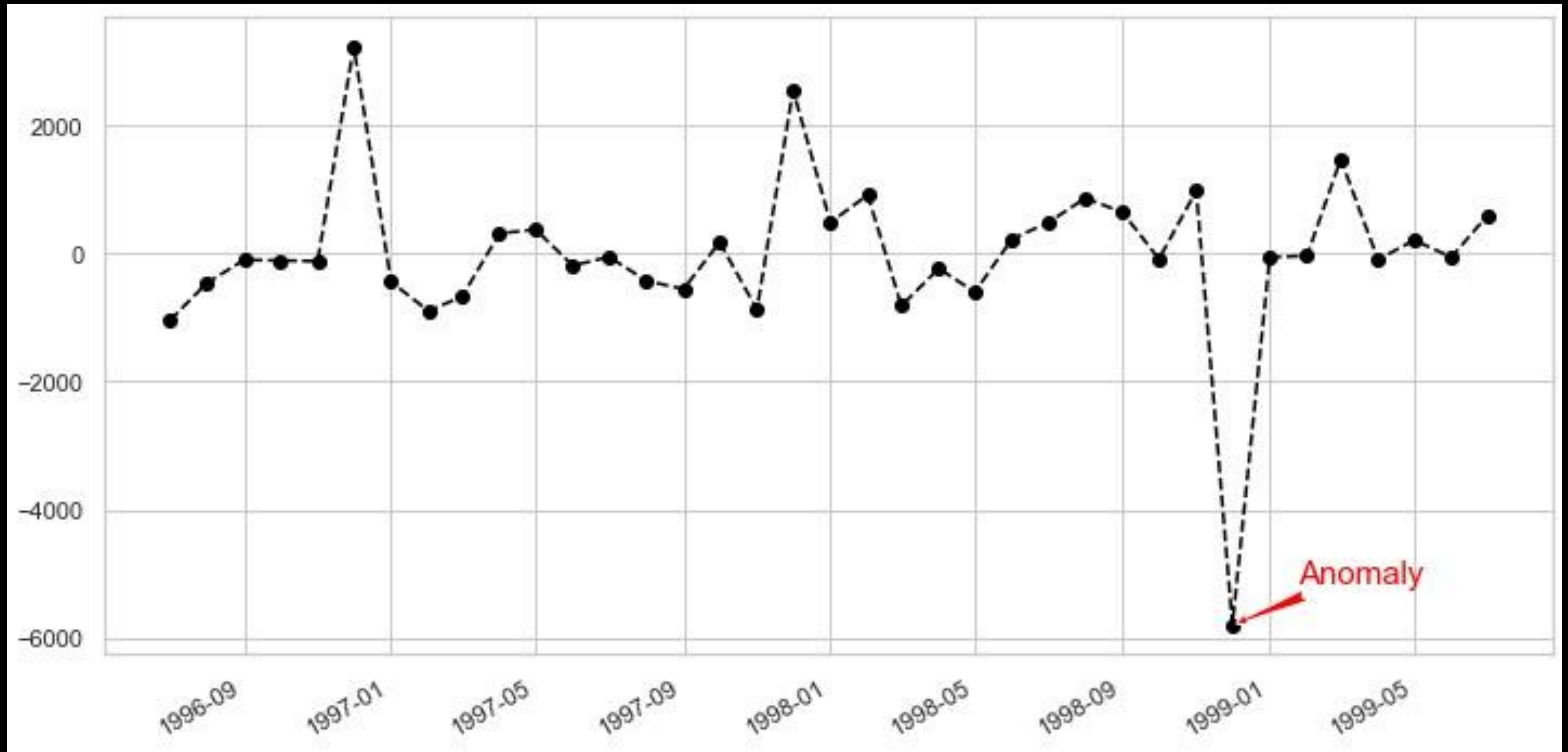


DARU

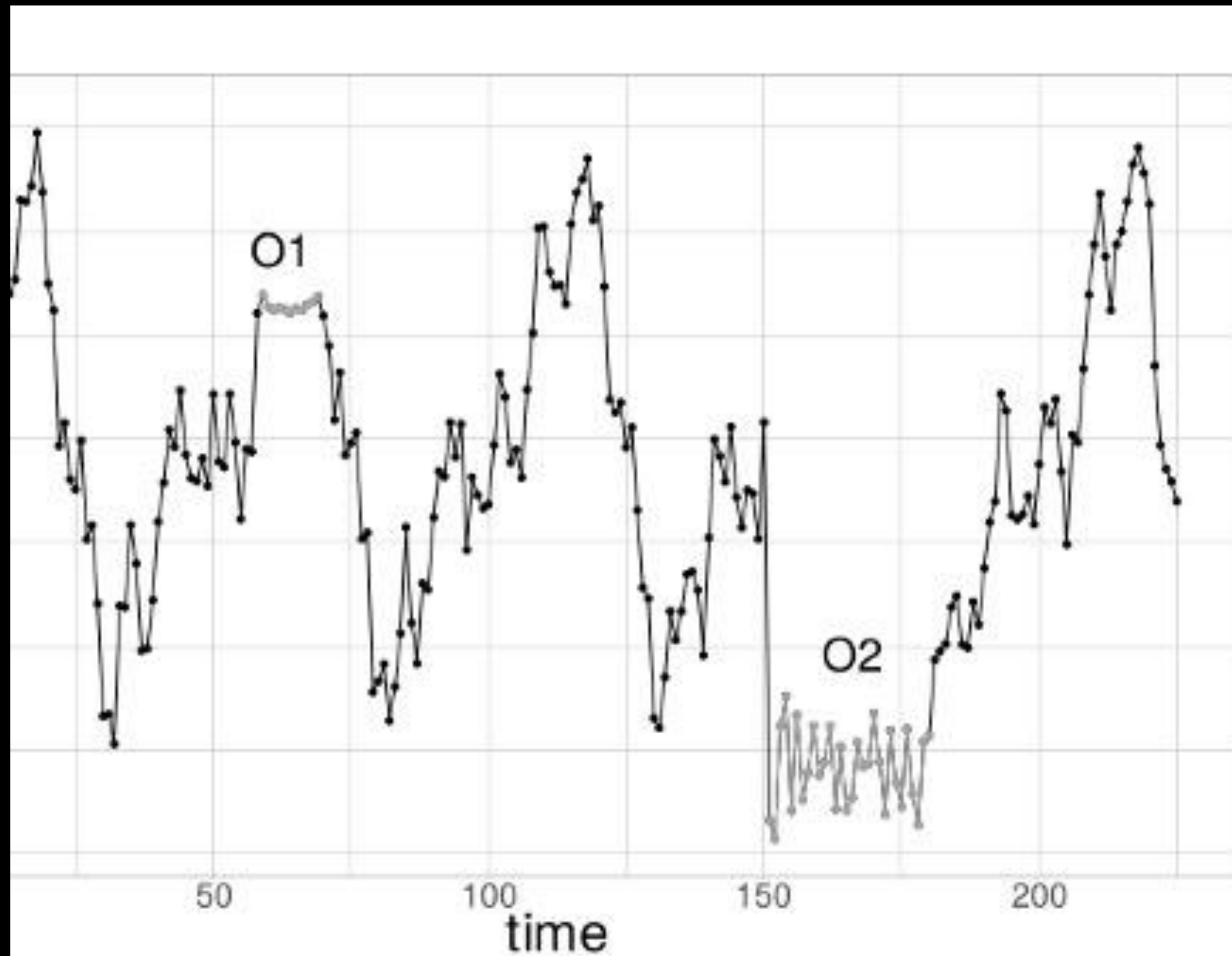
fastText
Nyaplot
Vega

...

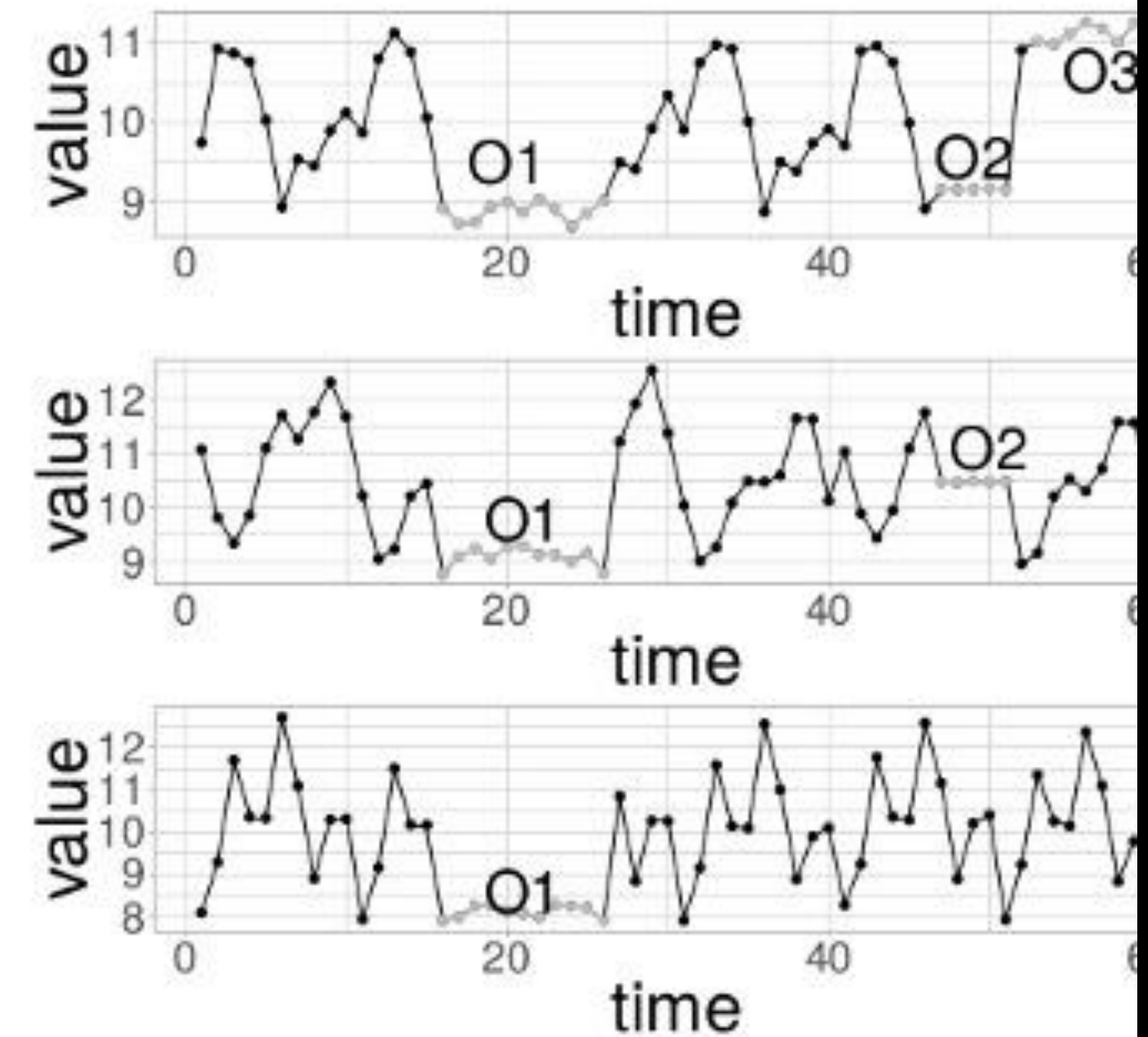
Anomaly Detection



Point Outliers



(a) Univariate time series.



(b) Multivariate time series.

Subsequence Anomaly

```
require "prophet-rb"  
require "csv"
```

```
series = {}
```

```
CSV.foreach('./Downloads/timeseries_sample_data.csv', headers: true) do |row|  
  date = Date.parse(row['ds'])  
  value = row['y'].to_f  
  series[date] = value  
end
```

```
Prophet.anomalies(series)
```

```
[#<Date: 2023-01-04 ((2459949j,0s,0n),+0s,2299161j)>, #<Date: 2023-01-02 ((2459947j,0s,0n),+0s,2299161j)>, #<Date: 2023-01-03 ((2459948j,0s,0n),+0s,2299161j)>]
```

Detect the Outliers



Detection by Forecasting

```
require "rover-df"|  
data = Rover.read_csv("./Downloads/timeseries_sample_data - no anomalies.csv")
```

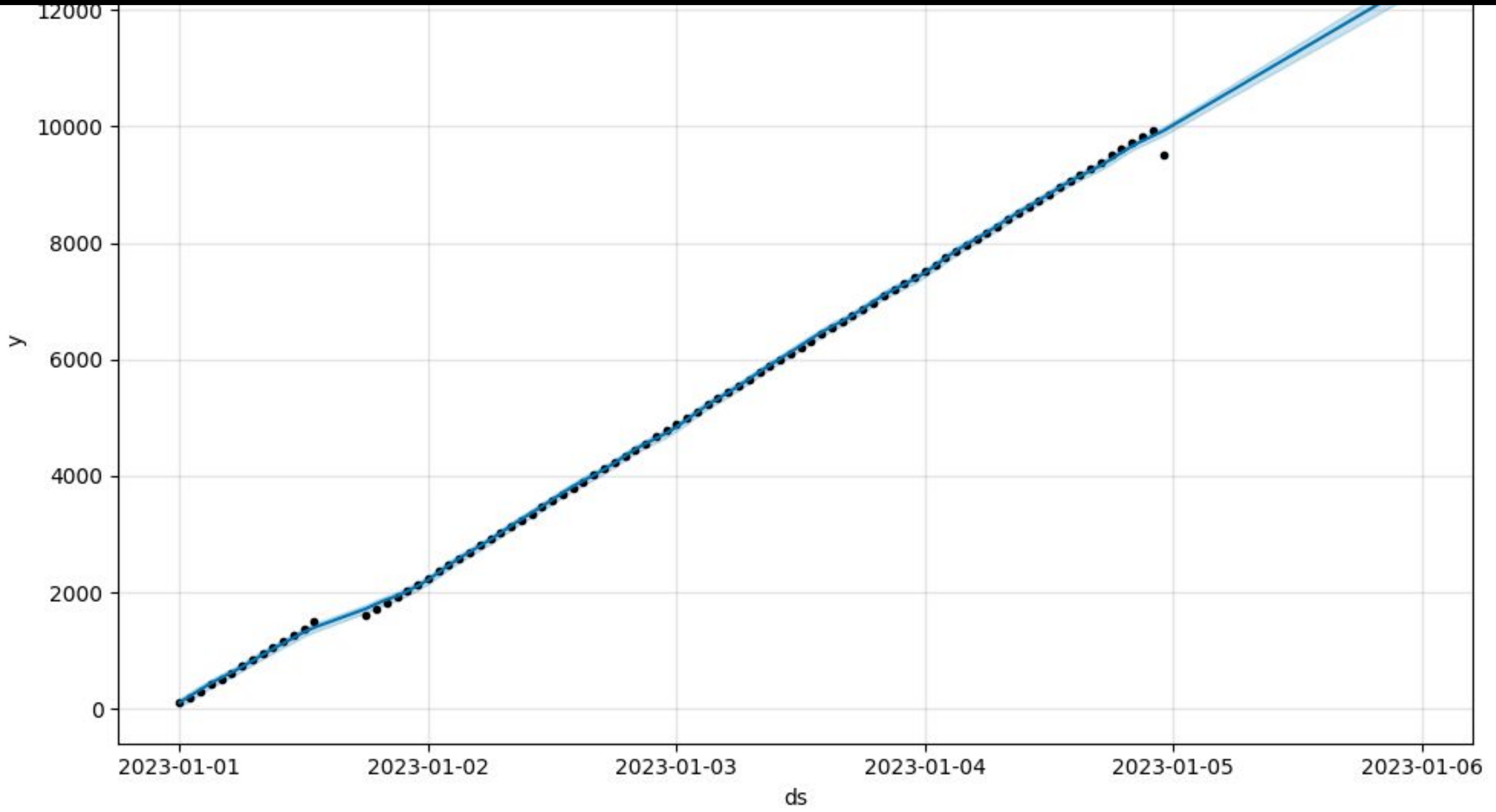
| ds | y |
|---------------------|------|
| 2023-01-01 0:00:00 | 100 |
| 2023-01-01 1:00:00 | 201 |
| 2023-01-01 2:00:00 | 305 |
| | ⋮ |
| 2023-01-04 21:00:00 | 9829 |
| 2023-01-04 22:00:00 | 9939 |
| 2023-01-04 23:00:00 | 9500 |

Load the Ideal Data

```
model = Prophet.new
model.fit(data)
future = model.make_future_dataframe(periods: 1)
future.tail
forecast = model.predict(future)
future_data = forecast[["ds", "yhat"]]
model.plot(forecast).savefig("./forecast-reading.png")
```

```
[prophet] Disabling yearly seasonality. Run prophet with yearly_seasonality: true to override this.
[prophet] Disabling weekly seasonality. Run prophet with weekly_seasonality: true to override this.
```

Forecast



Sumarize

- Time Series Data
- Challenges
 - IOT Maintenance
 - Abnormal Usage
- Machine Learning with Ruby
- Anomaly Detection



**YOU LISTENED TO ME —
THAT REALLY MEANS A LOT.**



@vishwadesurkar

simplysmart.tech