BSTS Synthetic Control Generated Revenue Data

Jonathan Hershaff

2025-06-29

```
# Define the file path
file_path <- "/Users/jonathan/Desktop/projects/udacity_course/lessons/lesson4_synthetic_
control/data/synthetic_control_revenue_data.csv"

# Read the CSV file
revenue_data <- read_csv(file_path)
```

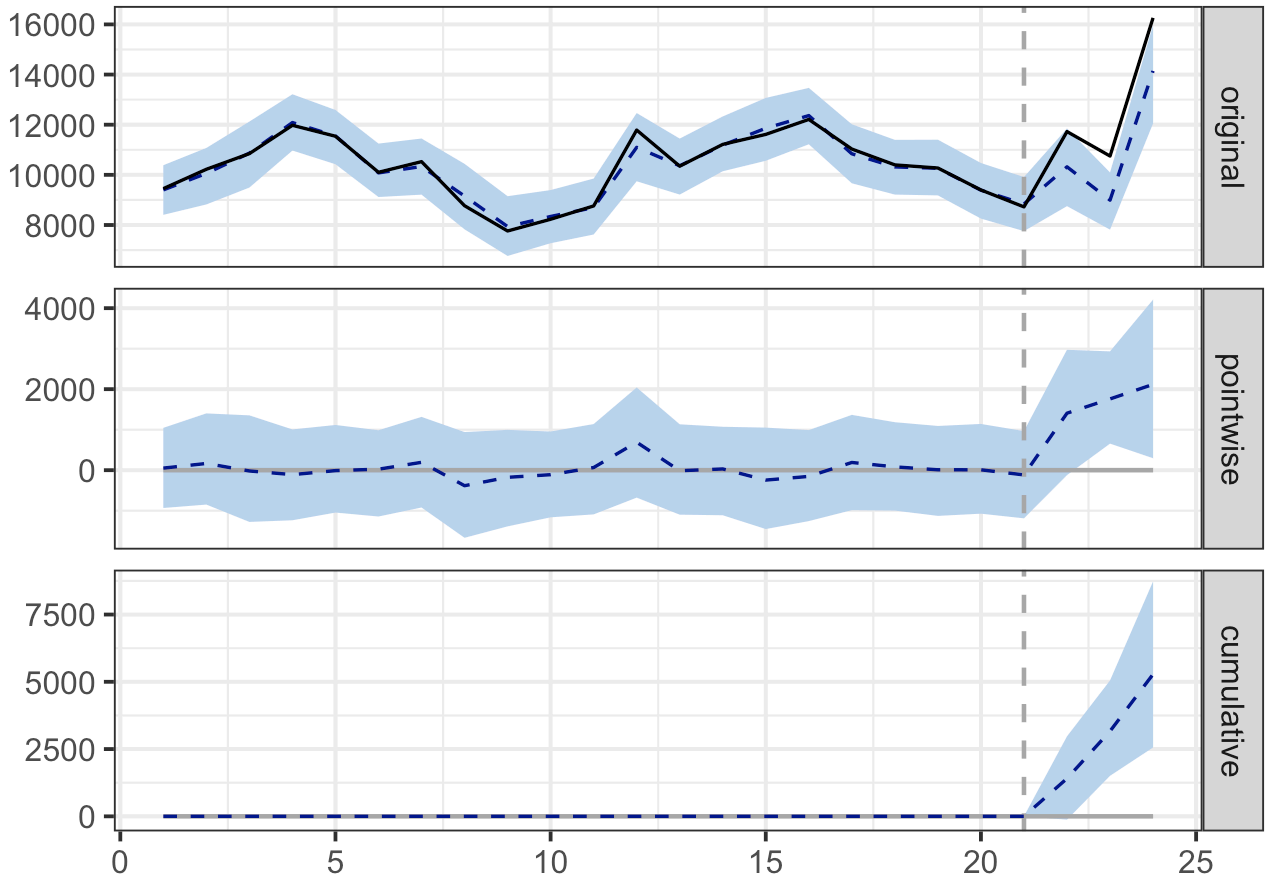
```
## Rows: 24 Columns: 11
## ── Column specification ────────────────────────────────────────────────────────
────────────────────────────────────────────────────────
## Delimiter: ","
## dbl  (10): Region 1, Region 2, Region 3, Region 4, Region 5, Region 6, Region 7, Regi
on 8, Region 9, Region 10
## date  (1): Month
##
## ℹ Use `spec()` to retrieve the full column specification for this data.
## ℹ Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
# Replace spaces with underscores in column names
colnames(revenue_data) <- gsub(" ", "_", colnames(revenue_data))

# View the first few rows of the data
print(head(revenue_data))
```

```
# Prepare the data for CausalImpact
# Extract the treated region (Region 5) and the control regions (Region 1 to Region 4, R
egion 6 to Region 10)
treated_region <- revenue_data$Region_5
control_regions <- revenue_data[, c("Region_1", "Region_2", "Region_3", "Region_4",
                                    "Region_6", "Region_7", "Region_8", "Regio
n_9", "Region_10")]

# Combine treated and control regions into a matrix
impact_data <- cbind(treated_region, control_regions)

# Define the pre-treatment and post-treatment periods
pre_period <- c(1, which(revenue_data$Month == "2024-09-01"))
post_period <- c(which(revenue_data$Month == "2024-10-01"), nrow(revenue_data))

# Run the CausalImpact analysis
impact <- CausalImpact(impact_data, pre_period, post_period)

# Print the summary of the impact analysis
summary(impact)
```

```
## Posterior inference {CausalImpact}
##
##                         Average          Cumulative
## Actual                  12913            38738
## Prediction (s.d.)       11152 (533)      33455 (1599)
## 95% CI                  [10003, 12059]   [30009, 36178]
##
## Absolute effect (s.d.)  1761 (533)       5282 (1599)
## 95% CI                  [853, 2910]      [2560, 8729]
##
## Relative effect (s.d.)  16% (5.7%)       16% (5.7%)
## 95% CI                  [7.1%, 29%]      [7.1%, 29%]
##
## Posterior tail-area probability p:   0.00102
## Posterior prob. of a causal effect:  99.89848%
##
## For more details, type: summary(impact, "report")
```

```
# Plot the results
plot(impact)
```




```
summary(impact, "report")
```