Aggregated analysis

Taha Nasir

2022-10-02

```
library(readxl)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(CausalImpact)
```

```
df<-read_excel("PatientSatisfactionScores.xlsx")
time.points=seq.Date(as.Date("2020-07-01"),by=1,length.out =900 )
```

Aggregated Intervention Causal Impact Analysis

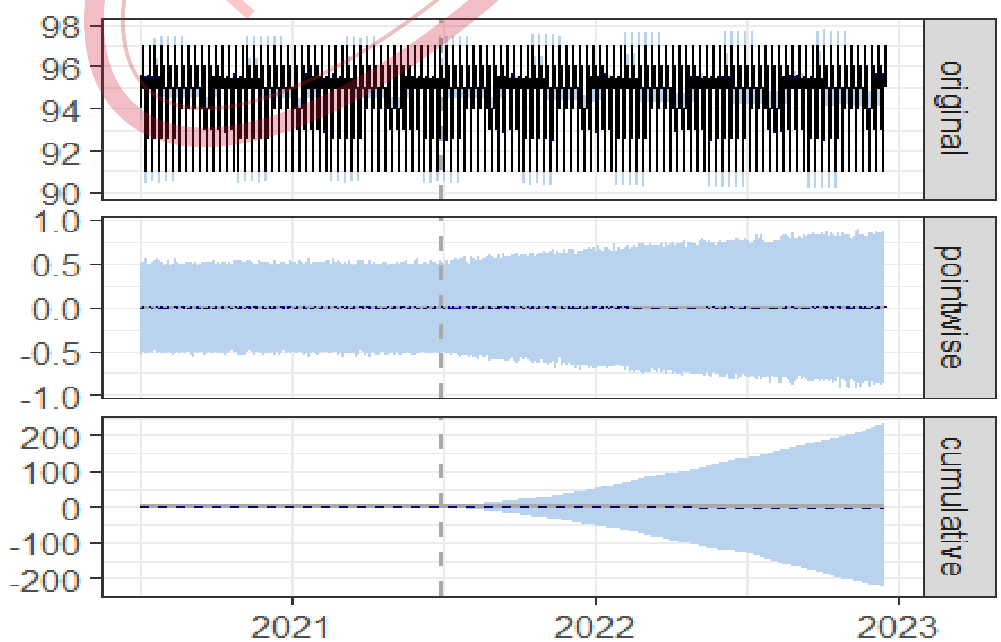
```
prov1<-ts(df$`Provider 1`)
prov2<-ts(df$`Provider 2`)
prov3<-ts(df$`Provider 3`)
prov4<-ts(df$`Provider 4`)
prov5<-ts(df$`Provider 5`)
prov6<-ts(df$`Provider 6`)
prov7<-ts(df$`Provider 7`)
prov8<-ts(df$`Provider 8`)
prov9<-ts(df$`Provider 9`)
prov10<-ts(df$`Provider 10`)
prov11<-ts(df$`Provider 11`)
prov12<-ts(df$`Provider 12`)
prov13<-ts(df$`Provider 13`)
prov14<-ts(df$`Provider 14`)
prov15<-ts(df$`Provider 15`)
prov16<-ts(df$`Provider 16`)
prov17<-ts(df$`Provider 17`)
prov18<-ts(df$`Provider 18`)
prov19<-ts(df$`Provider 19`)
prov20<-ts(df$`Provider 20`)
prov21<-ts(df$`Provider 21`)
prov22<-ts(df$`Provider 22`)
prov23<-ts(df$`Provider 23`)
```

```
prov24<-ts(df$`Provider 24`)
prov25<-ts(df$`Provider 25`)
prov26<-ts(df$`Provider 26`)
prov27<-ts(df$`Provider 27`)
prov28<-ts(df$`Provider 28`)
prov29<-ts(df$`Provider 29`)
prov30<-ts(df$`Provider 30`)
prov31<-ts(df$`Provider 31`)
prov32<-ts(df$`Provider 32`)
prov33<-ts(df$`Provider 33`)
prov34<-ts(df$`Provider 34`)
prov35<-ts(df$`Provider 35`)
data <- zoo(cbind(prov1,prov1,prov2,prov3,prov4,prov5,prov6,prov7,prov8,prov9
,prov10,prov11,              prov12,prov13,prov14,prov15,prov16,prov17,prov18
,prov19,prov20,prov21,prov22,prov23,               prov24,prov25,prov26,prov2
7,prov28,prov29,prov30,prov31,prov32,prov33,prov34,prov35), time.points)
pre.period=as.Date(c("2020-07-01","2021-06-30"))
post.period=as.Date(c("2021-07-01","2022-12-17"))
impact <- CausalImpact(data, pre.period, post.period)

## Warning in model.matrix.default(model.terms, my.model.frame, contrasts): t
he
## response appeared on the right-hand side and was dropped

## Warning in model.matrix.default(model.terms, my.model.frame, contrasts): p
roblem
## with term 1 in model.matrix: no columns are assigned

plot(impact)
```



Statistics and Report

```
summary(impact)

## Posterior inference {CausalImpact}
##
##                          Average            Cumulative
## Actual                   95                 50623
## Prediction (s.d.)        95 (0.22)          50624 (115.88)
## 95% CI                   [94, 95]           [50390, 50847]
##
## Absolute effect (s.d.)   -0.0011 (0.22)     -0.6001 (115.88)
## 95% CI                   [-0.42, 0.44]      [-224.50, 233.47]
##
## Relative effect (s.d.)   -0.0012% (0.23%)   -0.0012% (0.23%)
## 95% CI                   [-0.44%, 0.46%]    [-0.44%, 0.46%]
##
## Posterior tail-area probability p:   0.5
## Posterior prob. of a causal effect:  50%
##
## For more details, type: summary(impact, "report")

summary(impact,"report")
```

Analysis report {CausalImpact}

```
During the post-intervention period, the response variable had an average val
ue of approx. 94.62. In the absence of an intervention, we would have expecte
d an average response of 94.62. The 95% interval of this counterfactual predi
ction is [94.19, 95.04]. Subtracting this prediction from the observed respon
se yields an estimate of the causal effect the intervention had on the respon
se variable. This effect is -0.0011 with a 95% interval of [-0.42, 0.44]. For
a discussion of the significance of this effect, see below.
Summing up the individual data points during the post-intervention period (wh
ich can only sometimes be meaningfully interpreted), the response variable ha
d an overall value of 50.62K. Had the intervention not taken place, we would
have expected a sum of 50.62K. The 95% interval of this prediction is [50.39K
, 50.85K].


The above results are given in terms of absolute numbers. In relative terms,
the response variable showed a decrease of -0%. The 95% interval of this perc
entage is [-0%, +0%].

This means that, although it may look as though the intervention has exerted
a negative effect on the response variable when considering the intervention
period as a whole, this effect is not statistically significant, and so canno
t be meaningfully interpreted. The apparent effect could be the result of ran
dom fluctuations that are unrelated to the intervention. This is often the ca
se when the intervention period is very long and includes much of the time wh
en the effect has already worn off. It can also be the case when the interven
tion period is too short to distinguish the signal from the noise. Finally, f
```

```
ailing to find a significant effect can happen when there are not enough cont
rol variables or when these variables do not correlate well with the response
variable during the learning period.
The probability of obtaining this effect by chance is p = 0.5. This means the
effect may be spurious and would generally not be considered statistically si
gnificant.
```

Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.