

# Demonstrating the imputeTestBench package

Mayank

**Introduction** I will demonstrate the use of `imputeTestbench` on dataset `nottem` which contains monthly average air temperatures measured at Nottingham Castle from 1920 to 1939. This demonstration tries to use maximum functions available in the package.

```
knitr::opts_chunk$set(echo=TRUE)
```

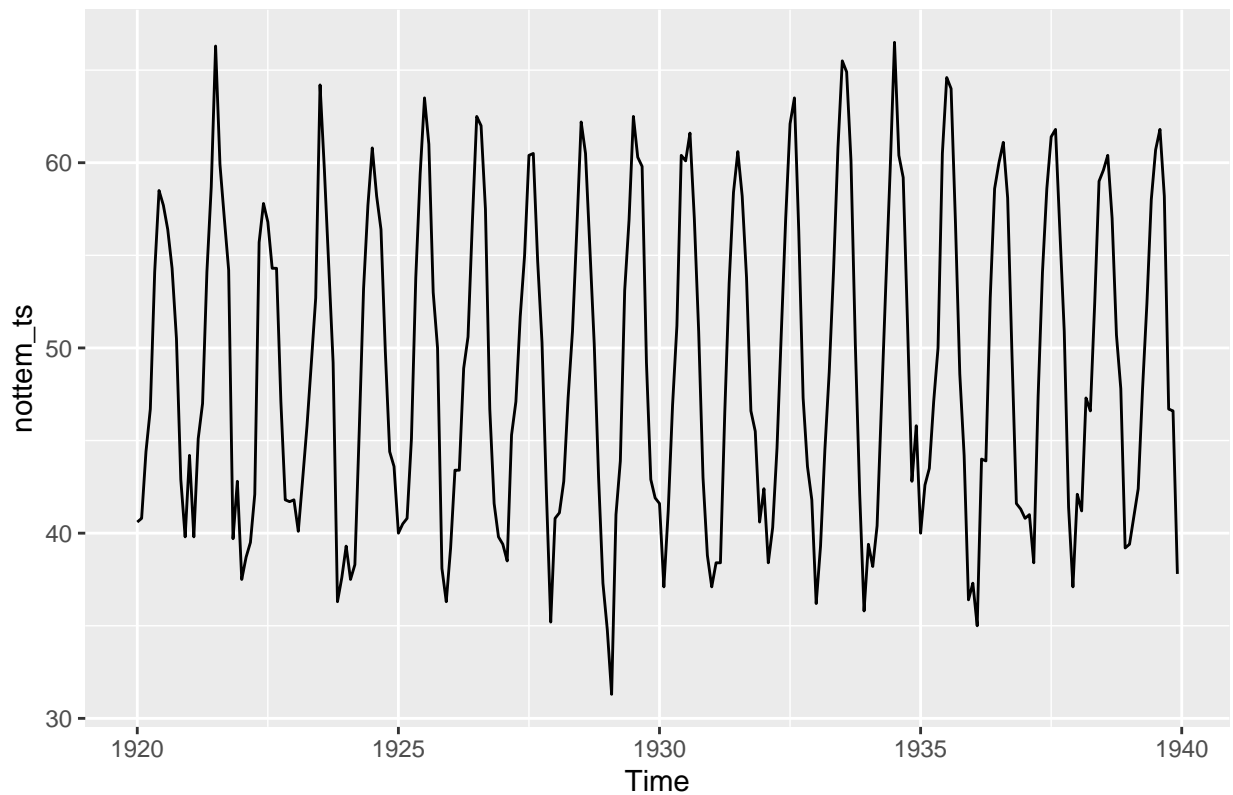
```
#load required packages  
library(imputeTestbench)
```

```
## Registered S3 method overwritten by 'quantmod':  
##   method      from  
##   as.zoo.data.frame zoo
```

```
library(ggplot2)
```

```
#loading Data  
data("nottem")  
nottem_ts <- as.ts(nottem)  
  
autoplot(nottem_ts) + ggtitle("Time Series Data")
```

## Time Series Data



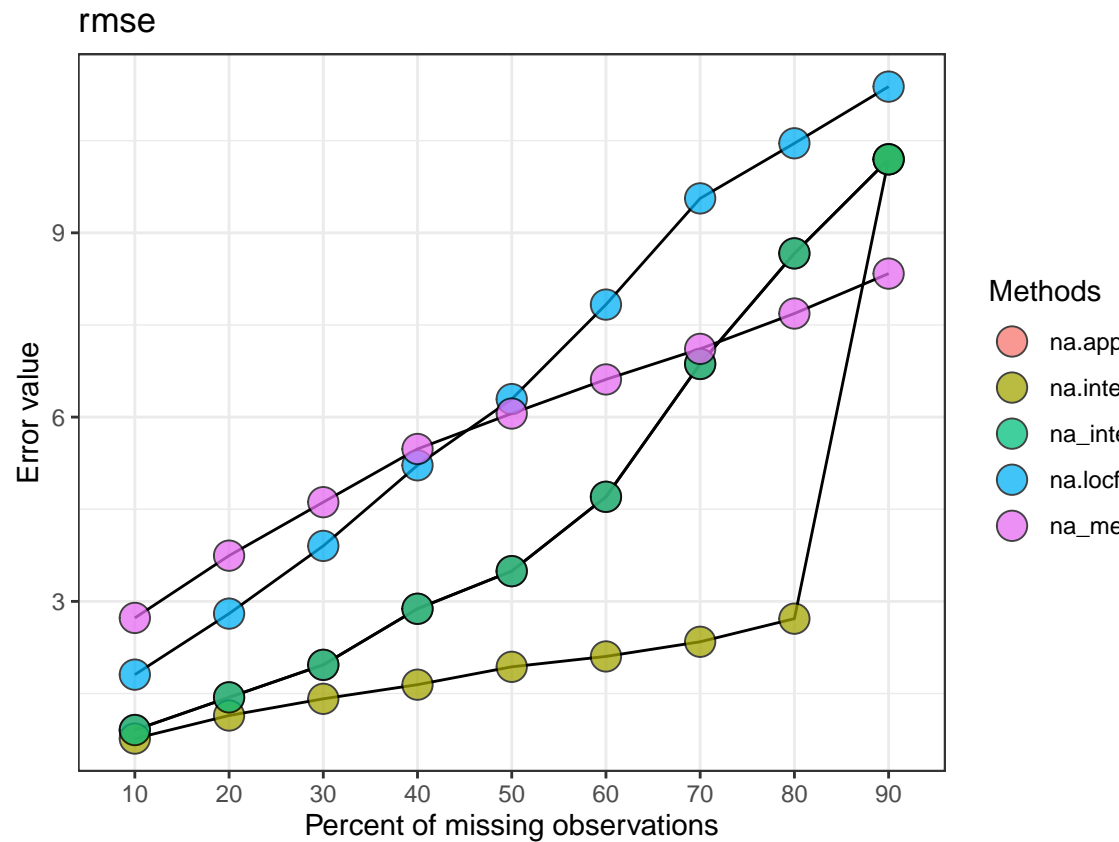
```
#Applying imputation using
nottem_ts_imputed <- impute_errors(dataIn = nottem_ts,
  smps = "mcar",
  methods = c("na.approx",
    "na.interp", "na_interpolation", "na.locf", "na_mean"),
  errorParameter = "rmse",
  repetition = 10)
print(nottem_ts_imputed)
```

Using imputeTestBench for simulating missing value and applying impute methods

```
## $Parameter
## [1] "rmse"
##
## $MissingPercent
## [1] 10 20 30 40 50 60 70 80 90
##
## $na.approx
## [1] 0.9059266 1.4375152 1.9669107 2.8797430 3.4921005 4.7026933 6.8651897
## [8] 8.6651016 10.1967712
##
## $na.interp
## [1] 0.7666661 1.1421000 1.4139511 1.6442980 1.9346878 2.1037422 2.3424783
```

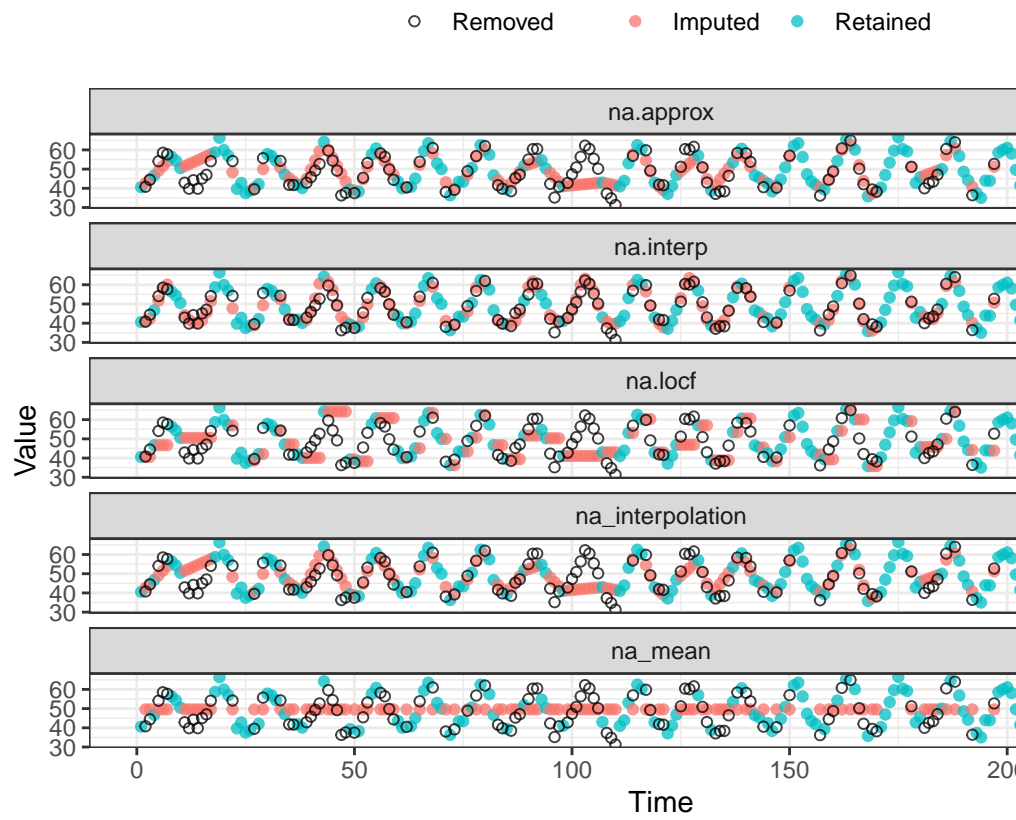
```
## [8] 2.7185591 10.1967712
##
## $na_interpolation
## [1] 0.9059266 1.4375152 1.9669107 2.8797430 3.4921005 4.7026933 6.8651897
## [8] 8.6651016 10.1967712
##
## $na.locf
## [1] 1.807615 2.800882 3.903354 5.215311 6.293113 7.831606 9.560068
## [8] 10.457303 11.380141
##
## $na_mean
## [1] 2.730757 3.744600 4.614833 5.480829 6.058097 6.613037 7.110680 7.686004
## [9] 8.336021
```

```
plot_errors(nottem_ts_imputed, plotType = "line")
```



Visualization of Errors

```
plot_impute(dataIn = nottem_ts, smps = "mcar", showmiss = TRUE)
```



### Visualization of Imputed Data

**Conclusion** This analysis demonstrates the effectiveness of different imputation methods on missing time series data