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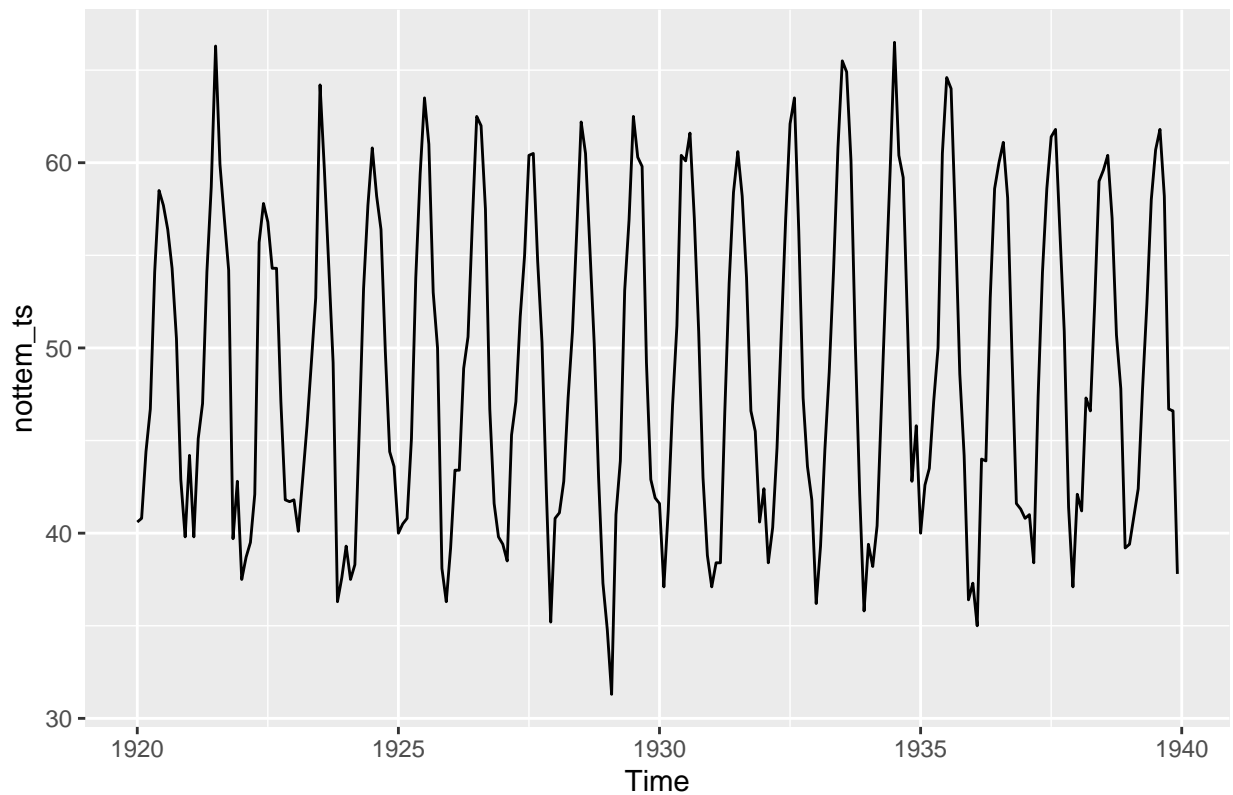
```
knitr::opts_chunk$set(echo=TRUE)

#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")
```
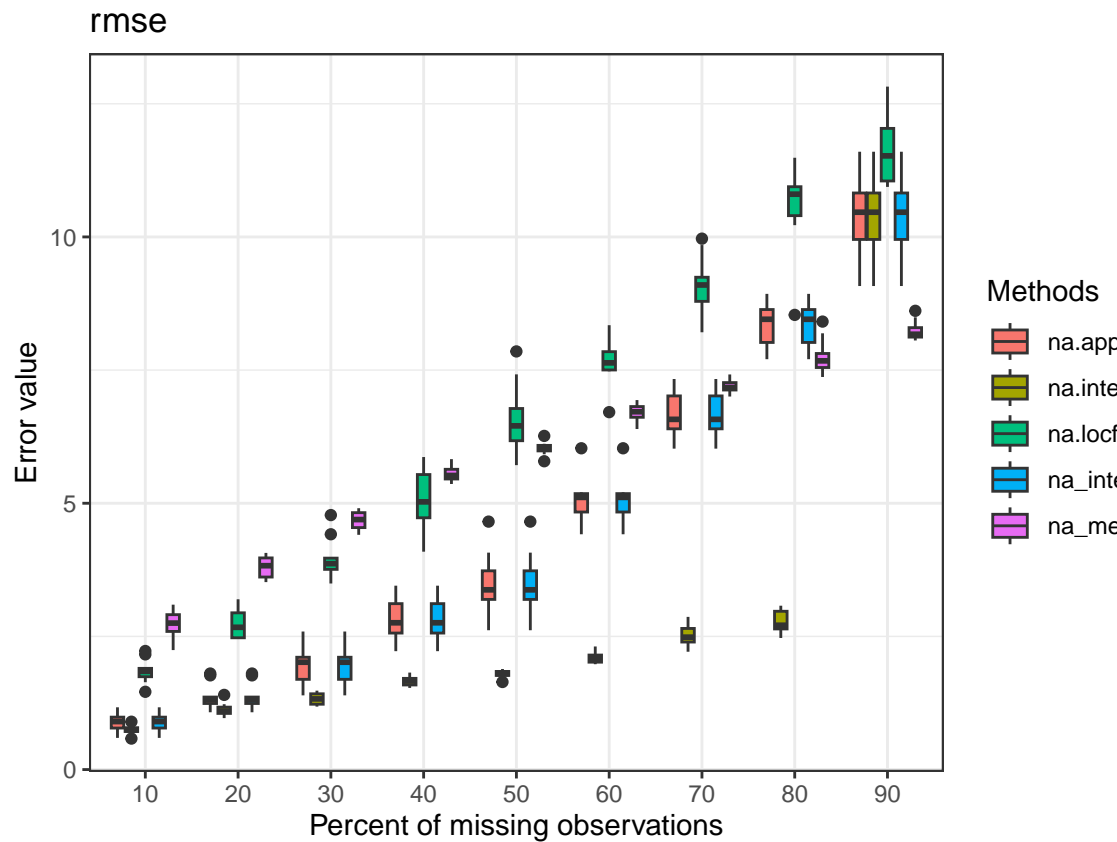
Time Series Data


```r
#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]  0.8972215  1.3738418  1.9405243  2.8146363  3.4918143  5.0628108  6.6476047
## [8]  8.3689661 10.4047680
##
## $na.interp
## [1]  0.7515512  1.1259267  1.3274837  1.6649228  1.7941232  2.0950834  2.5134635
```

```
## [8]   2.7693108 10.4047680
##
## $na_interpolation
## [1]   0.8972215  1.3738418   1.9405243   2.8146363   3.4918143   5.0628108   6.6476047
## [8]   8.3689661 10.4047680
##
## $na.locf
## [1]   1.845757   2.733951   3.945085   5.073271   6.580475   7.645756   9.099494
## [8] 10.601125 11.618421
##
## $na_mean
## [1] 2.726413 3.799683 4.672159 5.550081 6.025049 6.699951 7.204416 7.752384
## [9] 8.244470
```

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


Visualization of Errors

```
plot_impute(dataIn = nottem_ts, smps = "mcar", showmiss = TRUE)
```
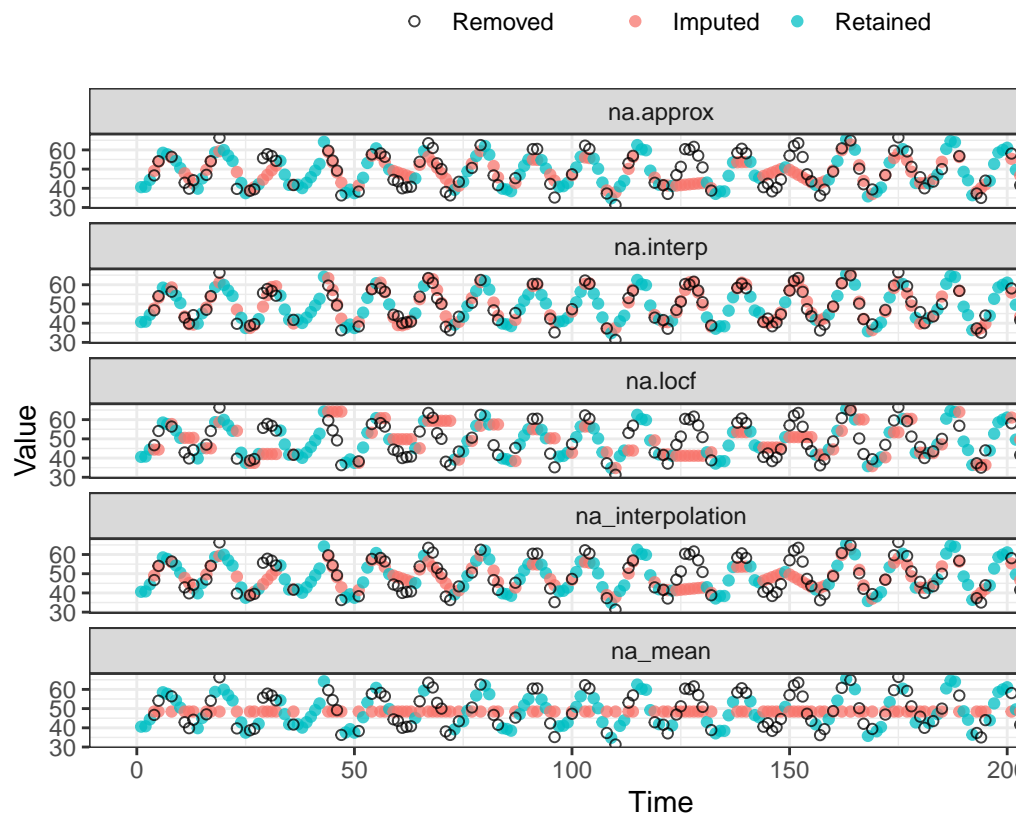
Visualization of Imputed Data

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