

# sprint1eda

*me*

*28 September 2018*

## R Markdown

This is an R Markdown document that contains graphs, tables and comments about the data provided by honeywell's chemical manufacturing process. The sample of the data used here is the PI tags data generated from machines in Line 2. The file is named "L2 Oct to Dec 2017" and available on shared drive. The structure of the data shows there 142 columns with sensor readings and 142 columns with respective timestamps. The file is ~276MB in size in XLS converted to .CSV format. Using data.table's fread the time taken to load the data into the machine is ~ 25s.

```
## [1] 8033
```

```
## Time difference of -31.68954 secs
```

Each sensor value is a concatenation of Sensor line number eg. "SC2" and sensor function eg. "CB.Sol" and Sensor variable eg. ".SP"

Some sensor names are in the lower case and therefore advisable to convert all the names of the sensors to uppercase for correct tabulation. There are also some sensor values without "\_".

## Including Plots

You can also embed plots, for example:

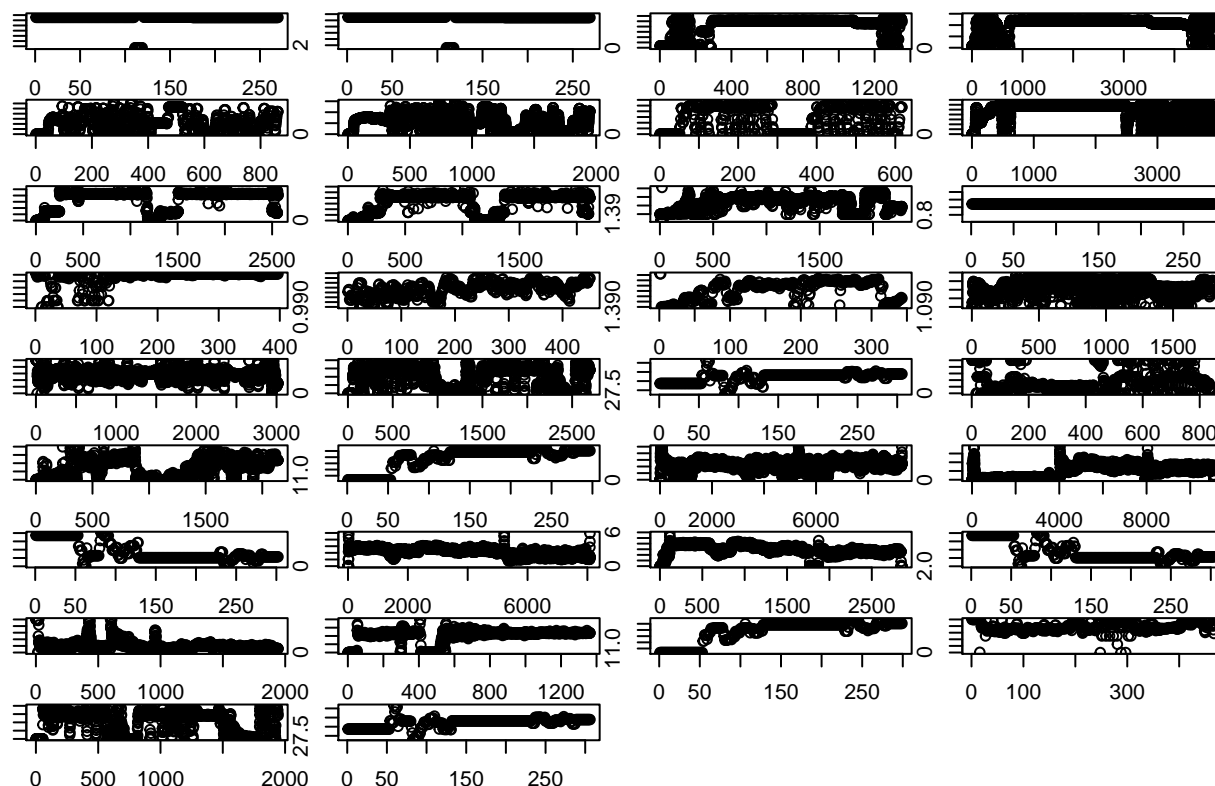
```
## block1
## SC2  sc2_ SC2_
##    2    1  139

## block3
## FIC TIC CVO  FI  TI  AI LIC PIC  CV  IT
##  30  20  18  11   7   6   6   6   4   4

## block4
## .PV .SP PEN .OP GAL LBS TUS LOW .CL 495
##  60  25  18  16   3   3   3   2   1   1
```

Thus the above table shows the frequency of each sensor type

```
## [1] 8033
```



The above time series plots shows the sensors 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38 & 40 without na values

```
## [1] 8033
```

```
## [1] . .E- Bad Bad Input
## [5] Calc Failed Closed Down Error
## [9] I/O Timeout No more values: Open Over Range
## [13] Process Recycle Running Shutdown
## [17] Under Range
## 23 Levels: - -. . . . .E- Bad Bad Input Calc Failed ... Under Range
```