5 designs and 5 iterations

Sayani Gupta

15/09/2021

1 *Simulation design*

```r
t = 300
t <- seq(0, t, 1)
g1 <- t %%2
g2 <- t %%3
g3 <- t %%5
mu1= c(0, 2)
mu2 = c(7, 5, 1)
mu3 = c(1, 2, 5, 3, 1)
```

Three circular granularities g1, g2 and g3 are considered with levels 2, 3 and 5 respectively. Many time series with 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300 observations are created using the five designs below, each of which is iterated five times. We anticipate to have 3 clusters, each with five time series conforming to the same design, once we execute the clustering.

2 *Distance computation*

1) wpd for three granularities computed and then distance between designs are computed by computing the euclidean distances between them - gran as variables, designs across rows and cell values as wpd.

2) Robust scaling done for each customer and then JSD is computed between same categories for a granularity.

3   Granularities design

design	g1	g2	g3
design-1	no	no	no
design-2	yes	no	no
design-3	no	yes	no
design4	no	no	yes
design5	yes	yes	yes

4   Simulate data
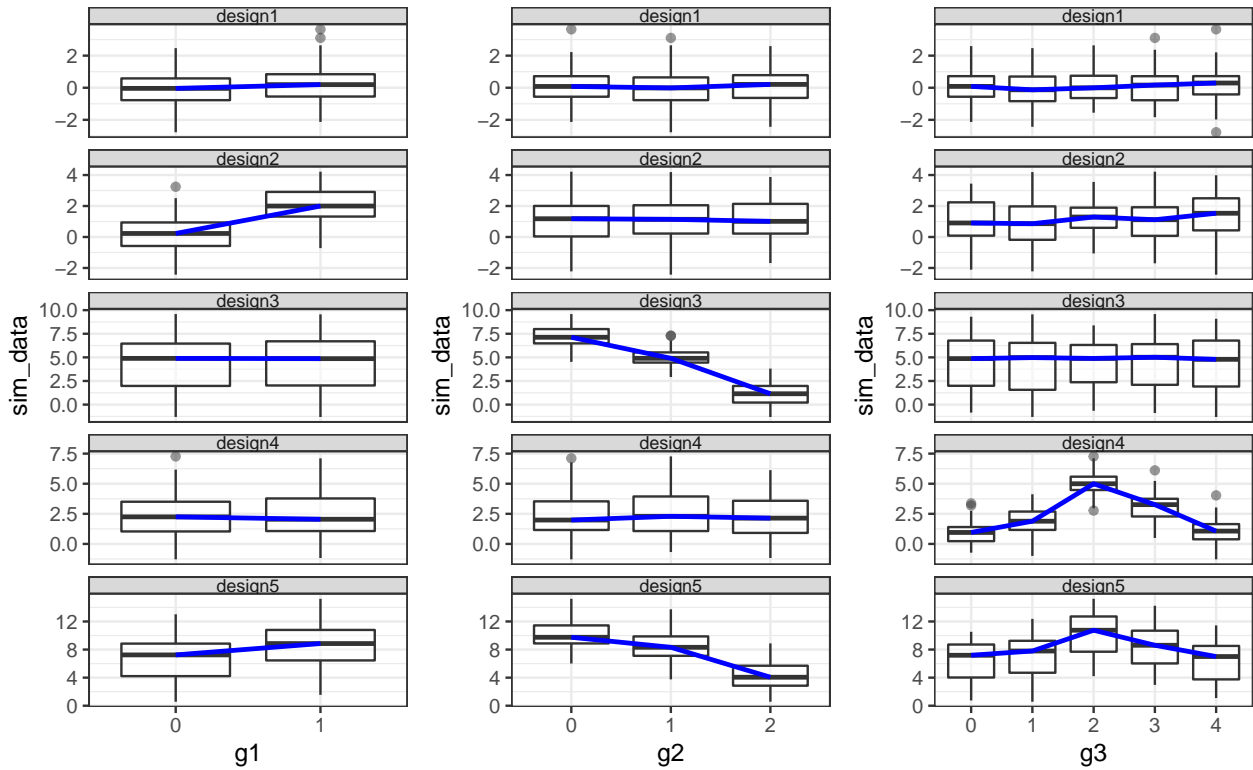
```
## # A tibble: 1,505 x 6
##     index    g1    g2    g3 design   sim_data
##     <dbl> <dbl> <dbl> <dbl> <chr>       <dbl>
## 1       0     0     0     0 design1     0.863
## 2       0     0     0     0 design2     1.04
## 3       0     0     0     0 design3     7.16
## 4       0     0     0     0 design4    -0.400
## 5       0     0     0     0 design5     7.19
## 6       1     1     1     1 design1    -0.828
## 7       1     1     1     1 design2    -0.0651
## 8       1     1     1     1 design3     5.17
## 9       1     1     1     1 design4     2.68
## 10      1     1     1     1 design5     8.21
## # ... with 1,495 more rows
```

5   Plot raw data
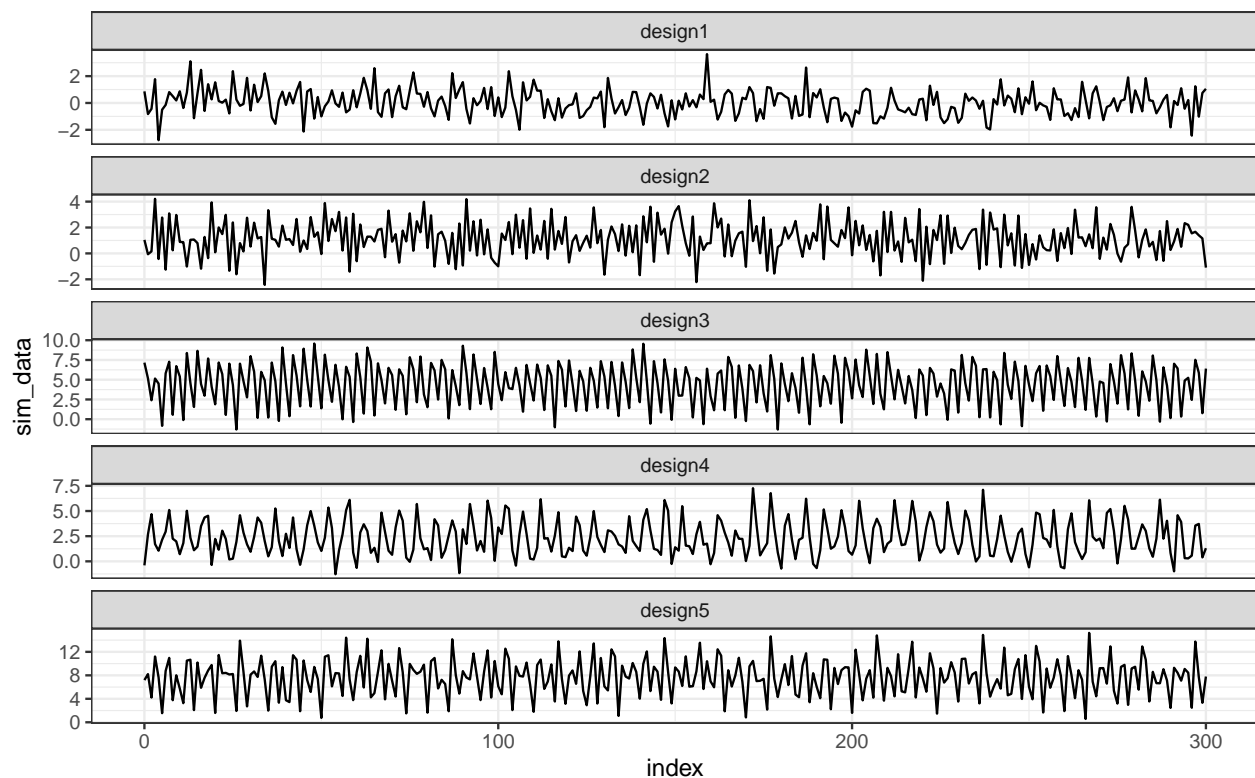
6 Plot distribution across granularities

7 Iterate designs

```
## # A tsibble: 7,525 x 8 [1]
## # Key:        customer_id [25]
##     seed_id index      g1      g2      g3 design  sim_data customer_id
##     <chr>   <dbl>   <dbl>   <dbl>   <dbl> <chr>      <dbl> <chr>
##  1 1           0       0       0       0 design1    0.586 design1-1
##  2 1           1       1       1       1 design1    0.709 design1-1
##  3 1           2       0       2       2 design1   -0.109 design1-1
##  4 1           3       1       0       3 design1   -0.453 design1-1
##  5 1           4       0       1       4 design1    0.606 design1-1
##  6 1           5       1       2       0 design1   -1.82  design1-1
##  7 1           6       0       0       1 design1    0.630 design1-1
##  8 1           7       1       1       2 design1   -0.276 design1-1
##  9 1           8       0       2       3 design1   -0.284 design1-1
## 10 1           9       1       0       4 design1   -0.919 design1-1
## # ... with 7,515 more rows
```

8 Cluster designs: distance using wpd

```
##              Reference
## Prediction design1 design2 design3 design4 design5
##     design1       5       0       0       0       0
##     design2       0       5       0       0       0
##     design3       0       0       5       0       0
##     design4       0       0       0       5       0
```

```
##     design5        0        0        0        0        5
```

9   Cluster designs: distance using js distance between categories (robust scaling)

```
##              Reference
## Prediction design1 design2 design3 design4 design5
##     design1        5        0        0        0        0
##     design2        0        5        0        0        0
##     design3        0        0        5        0        0
##     design4        0        0        0        5        0
##     design5        0        0        0        0        5
```

10   Cluster designs: distance using js distance between categories (nqt)

```
##              Reference
## Prediction design1 design2 design3 design4 design5
##     design1        5        0        0        0        0
##     design2        0        5        0        0        0
##     design3        0        0        5        0        0
##     design4        0        0        0        5        0
##     design5        0        0        0        0        5
```