

Hand picking similar behaving group of customers to check  
clustering results

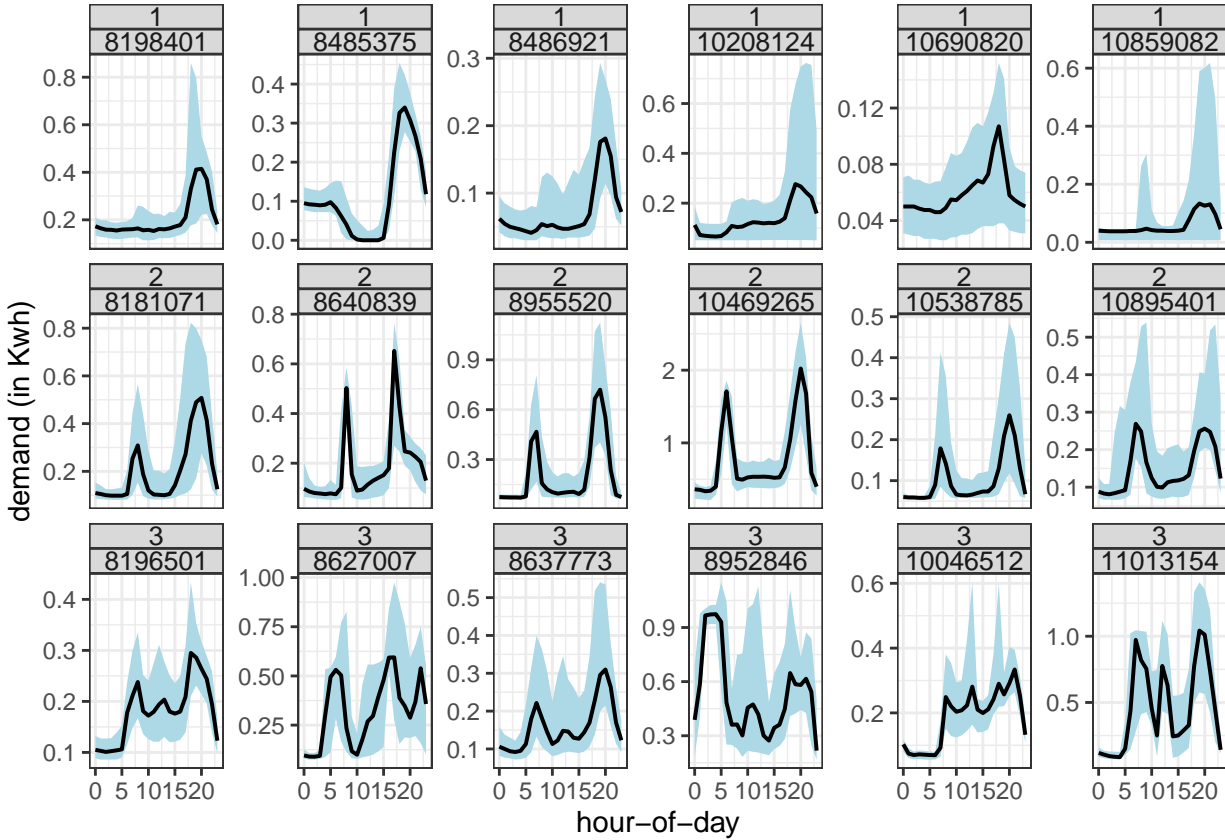


Figure 1: Median (black) and quartile deviation (blue region) of hourly demand drawn for few customers showing similar behaviors. Roughly speaking, Design 1 has one evening peak, Design 2 has two peaks and Design 3 has three peaks in a day. Each of design 1, 2 and 3 have six similar behaving customers resulting to 18 time series. We want our clustering results to group each of the designs together.

```
#data_pick <- data_pick %>% filter(!(customer_id %in% c(8485375, 8952846)))
library(gracsr)
v2 <- suppressWarnings(
  scaled_dist_gran(data_pick, "hour_day",
    response = "general_supply_kwh",
    quantile_prob_val = c(0.5, 0.75, 0.9))) %>% rename("dist_hod" = "dist")
v3 <- suppressWarnings(
  scaled_dist_gran(data_pick, "day_month",
    response = "general_supply_kwh",
    quantile_prob_val = c(0.5, 0.75, 0.9))) %>% rename("dist_dom" = "dist")

data_dist <- v3 %>%
```

```

left_join(v2) %>%
mutate(dist = dist_dom) %>%
  pivot_wider(-c(3, 4),
              names_from = customer_to,
              values_from = dist) %>%
rename("customer_id" = "customer_from")

```

```
## # A tibble: 3 x 2
```

```
##   group      n
##   <int> <int>
## 1     1     7
## 2     2     8
## 3     3     3
```

```
## # A tibble: 7 x 2
```

```
##   group customer_id
##   <int>         <int>
## 1     1      8181071
## 2     1      8196501
## 3     1      8198401
## 4     1      8485375
## 5     1     10469265
## 6     1     10859082
## 7     1     11013154
```

```
## # A tibble: 8 x 2
```

```
##   group customer_id
##   <int>         <int>
## 1     2      8486921
## 2     2      8637773
## 3     2      8640839
## 4     2      8952846
## 5     2     10046512
## 6     2     10538785
## 7     2     10690820
## 8     2     10895401
```

```
## # A tibble: 3 x 2
```

```
##   group customer_id
##   <int>         <int>
## 1     3      8627007
## 2     3      8955520
## 3     3     10208124
```