

Lab01MD

Xiaodong Guo

2024-03-02

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
#summary(cars)
```

Including Plots

You can also embed plots, for example:

```
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
## v dplyr      1.1.4      v readr      2.1.5
## v forcats    1.0.0      v stringr   1.5.1
## v ggplot2    3.5.0      v tibble    3.2.1
## v lubridate  1.9.3      v tidyr     1.3.1
## v purrr      1.0.2
```

```
## -- Conflicts ----- tidyverse_conflicts() --
```

```
## x dplyr::filter() masks stats::filter()
```

```
## x dplyr::lag()     masks stats::lag()
```

```
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(lubridate)
```

```
# set random seed
```

```
seed <- 230 # <Replaced by last 3 digits in Gxd's student ID here.>
```

```
set.seed(seed)
```

```
# generate alarm id, dates and region vectors
```

```
alarms_id.vt <- paste0('alarms_', sample(LETTERS, size = 10, replace = FALSE))
```

```
alarms_dates.vt <- paste0('d.', seq(Sys.Date()-60, Sys.Date(), by = '1 day')) # 60 days back.
```

```
alarms_region.vt <- c('AKL_North', 'AKL_Central', 'Waiheke', 'AKL_South', 'AKL_Others')
```

```
# generate random alarm frequency counts
```

```
alarms_count.mt <- matrix(round(runif(length(alarms_id.vt) * length(alarms_dates.vt))*seed),  
                           nrow = length(alarms_id.vt), ncol=length(alarms_dates.vt))
```

```
colnames(alarms_count.mt) <- alarms_dates.vt
```

```

# set up data frames
alarms_count.df <- data.frame(alarm_id = alarms_id.vt, alarms_count.mt)
alarms_info.df <- data.frame(alarm_ID = alarms_id.vt, alarms_region = alarms_region.vt)
#Data source alarms_count.df
#data source alarms_info.df

#End of Generate DataSource

alarms_summary.df <- alarms_count.df%>%
  pivot_longer(cols = !alarm_id, names_to = 'date'
    , values_to = 'frequency') %>%
  separate(date, c('prefix'
    , 'alarm_date'),extra = "merge") %>%
  select(-prefix) %>%
  left_join(alarms_info.df, by = c('alarm_id' = 'alarm_ID')) %>%
  group_by(alarm_date, alarms_region) %>%
  summarise(total = sum(frequency))

## `summarise()` has grouped output by 'alarm_date'. You can override using the
## `.groups` argument.

#copy from Lecture note index2-6.pdf provided by Lisa Chan
knitr::kable(alarms_summary.df)

```

alarm_date	alarms_region	total
2024.01.04	AKL_Central	253
2024.01.04	AKL_North	158
2024.01.04	AKL_Others	158
2024.01.04	AKL_South	183
2024.01.04	Waiheke	424
2024.01.05	AKL_Central	113
2024.01.05	AKL_North	202
2024.01.05	AKL_Others	364
2024.01.05	AKL_South	306
2024.01.05	Waiheke	224
2024.01.06	AKL_Central	17
2024.01.06	AKL_North	96
2024.01.06	AKL_Others	165
2024.01.06	AKL_South	153
2024.01.06	Waiheke	240
2024.01.07	AKL_Central	161
2024.01.07	AKL_North	223
2024.01.07	AKL_Others	112
2024.01.07	AKL_South	280
2024.01.07	Waiheke	206
2024.01.08	AKL_Central	371
2024.01.08	AKL_North	209
2024.01.08	AKL_Others	204
2024.01.08	AKL_South	220
2024.01.08	Waiheke	135
2024.01.09	AKL_Central	334
2024.01.09	AKL_North	312
2024.01.09	AKL_Others	236

alarm_date	alarms_region	total
2024.01.09	AKL_South	415
2024.01.09	Waiheke	149
2024.01.10	AKL_Central	322
2024.01.10	AKL_North	381
2024.01.10	AKL_Others	192
2024.01.10	AKL_South	211
2024.01.10	Waiheke	270
2024.01.11	AKL_Central	300
2024.01.11	AKL_North	235
2024.01.11	AKL_Others	305
2024.01.11	AKL_South	76
2024.01.11	Waiheke	140
2024.01.12	AKL_Central	158
2024.01.12	AKL_North	117
2024.01.12	AKL_Others	182
2024.01.12	AKL_South	123
2024.01.12	Waiheke	341
2024.01.13	AKL_Central	244
2024.01.13	AKL_North	193
2024.01.13	AKL_Others	121
2024.01.13	AKL_South	155
2024.01.13	Waiheke	75
2024.01.14	AKL_Central	193
2024.01.14	AKL_North	346
2024.01.14	AKL_Others	190
2024.01.14	AKL_South	163
2024.01.14	Waiheke	251
2024.01.15	AKL_Central	195
2024.01.15	AKL_North	148
2024.01.15	AKL_Others	226
2024.01.15	AKL_South	204
2024.01.15	Waiheke	271
2024.01.16	AKL_Central	160
2024.01.16	AKL_North	236
2024.01.16	AKL_Others	152
2024.01.16	AKL_South	221
2024.01.16	Waiheke	63
2024.01.17	AKL_Central	129
2024.01.17	AKL_North	184
2024.01.17	AKL_Others	265
2024.01.17	AKL_South	141
2024.01.17	Waiheke	284
2024.01.18	AKL_Central	386
2024.01.18	AKL_North	337
2024.01.18	AKL_Others	258
2024.01.18	AKL_South	305
2024.01.18	Waiheke	151
2024.01.19	AKL_Central	65
2024.01.19	AKL_North	326
2024.01.19	AKL_Others	202
2024.01.19	AKL_South	143
2024.01.19	Waiheke	250

alarm_date	alarms_region	total
2024.01.20	AKL_Central	185
2024.01.20	AKL_North	315
2024.01.20	AKL_Others	90
2024.01.20	AKL_South	237
2024.01.20	Waiheke	299
2024.01.21	AKL_Central	113
2024.01.21	AKL_North	164
2024.01.21	AKL_Others	188
2024.01.21	AKL_South	310
2024.01.21	Waiheke	178
2024.01.22	AKL_Central	244
2024.01.22	AKL_North	107
2024.01.22	AKL_Others	113
2024.01.22	AKL_South	93
2024.01.22	Waiheke	127
2024.01.23	AKL_Central	371
2024.01.23	AKL_North	292
2024.01.23	AKL_Others	310
2024.01.23	AKL_South	167
2024.01.23	Waiheke	164
2024.01.24	AKL_Central	380
2024.01.24	AKL_North	197
2024.01.24	AKL_Others	408
2024.01.24	AKL_South	74
2024.01.24	Waiheke	169
2024.01.25	AKL_Central	298
2024.01.25	AKL_North	277
2024.01.25	AKL_Others	232
2024.01.25	AKL_South	173
2024.01.25	Waiheke	404
2024.01.26	AKL_Central	280
2024.01.26	AKL_North	189
2024.01.26	AKL_Others	241
2024.01.26	AKL_South	375
2024.01.26	Waiheke	369
2024.01.27	AKL_Central	355
2024.01.27	AKL_North	283
2024.01.27	AKL_Others	290
2024.01.27	AKL_South	251
2024.01.27	Waiheke	250
2024.01.28	AKL_Central	166
2024.01.28	AKL_North	215
2024.01.28	AKL_Others	306
2024.01.28	AKL_South	217
2024.01.28	Waiheke	128
2024.01.29	AKL_Central	279
2024.01.29	AKL_North	144
2024.01.29	AKL_Others	265
2024.01.29	AKL_South	249
2024.01.29	Waiheke	279
2024.01.30	AKL_Central	117
2024.01.30	AKL_North	283

alarm_date	alarms_region	total
2024.01.30	AKL_Others	75
2024.01.30	AKL_South	393
2024.01.30	Waiheke	231
2024.01.31	AKL_Central	300
2024.01.31	AKL_North	407
2024.01.31	AKL_Others	137
2024.01.31	AKL_South	112
2024.01.31	Waiheke	221
2024.02.01	AKL_Central	293
2024.02.01	AKL_North	125
2024.02.01	AKL_Others	203
2024.02.01	AKL_South	145
2024.02.01	Waiheke	149
2024.02.02	AKL_Central	256
2024.02.02	AKL_North	181
2024.02.02	AKL_Others	143
2024.02.02	AKL_South	166
2024.02.02	Waiheke	352
2024.02.03	AKL_Central	311
2024.02.03	AKL_North	247
2024.02.03	AKL_Others	176
2024.02.03	AKL_South	403
2024.02.03	Waiheke	199
2024.02.04	AKL_Central	351
2024.02.04	AKL_North	93
2024.02.04	AKL_Others	118
2024.02.04	AKL_South	136
2024.02.04	Waiheke	260
2024.02.05	AKL_Central	308
2024.02.05	AKL_North	261
2024.02.05	AKL_Others	129
2024.02.05	AKL_South	390
2024.02.05	Waiheke	295
2024.02.06	AKL_Central	214
2024.02.06	AKL_North	269
2024.02.06	AKL_Others	186
2024.02.06	AKL_South	196
2024.02.06	Waiheke	235
2024.02.07	AKL_Central	119
2024.02.07	AKL_North	113
2024.02.07	AKL_Others	13
2024.02.07	AKL_South	244
2024.02.07	Waiheke	285
2024.02.08	AKL_Central	286
2024.02.08	AKL_North	301
2024.02.08	AKL_Others	239
2024.02.08	AKL_South	68
2024.02.08	Waiheke	183
2024.02.09	AKL_Central	203
2024.02.09	AKL_North	357
2024.02.09	AKL_Others	386
2024.02.09	AKL_South	265

alarm_date	alarms_region	total
2024.02.09	Waiheke	152
2024.02.10	AKL_Central	182
2024.02.10	AKL_North	333
2024.02.10	AKL_Others	267
2024.02.10	AKL_South	162
2024.02.10	Waiheke	218
2024.02.11	AKL_Central	303
2024.02.11	AKL_North	125
2024.02.11	AKL_Others	247
2024.02.11	AKL_South	133
2024.02.11	Waiheke	174
2024.02.12	AKL_Central	62
2024.02.12	AKL_North	204
2024.02.12	AKL_Others	338
2024.02.12	AKL_South	215
2024.02.12	Waiheke	327
2024.02.13	AKL_Central	103
2024.02.13	AKL_North	260
2024.02.13	AKL_Others	24
2024.02.13	AKL_South	224
2024.02.13	Waiheke	386
2024.02.14	AKL_Central	175
2024.02.14	AKL_North	225
2024.02.14	AKL_Others	264
2024.02.14	AKL_South	239
2024.02.14	Waiheke	258
2024.02.15	AKL_Central	394
2024.02.15	AKL_North	145
2024.02.15	AKL_Others	159
2024.02.15	AKL_South	325
2024.02.15	Waiheke	98
2024.02.16	AKL_Central	332
2024.02.16	AKL_North	107
2024.02.16	AKL_Others	197
2024.02.16	AKL_South	197
2024.02.16	Waiheke	146
2024.02.17	AKL_Central	393
2024.02.17	AKL_North	337
2024.02.17	AKL_Others	140
2024.02.17	AKL_South	211
2024.02.17	Waiheke	155
2024.02.18	AKL_Central	119
2024.02.18	AKL_North	444
2024.02.18	AKL_Others	120
2024.02.18	AKL_South	192
2024.02.18	Waiheke	389
2024.02.19	AKL_Central	172
2024.02.19	AKL_North	261
2024.02.19	AKL_Others	385
2024.02.19	AKL_South	73
2024.02.19	Waiheke	242
2024.02.20	AKL_Central	319

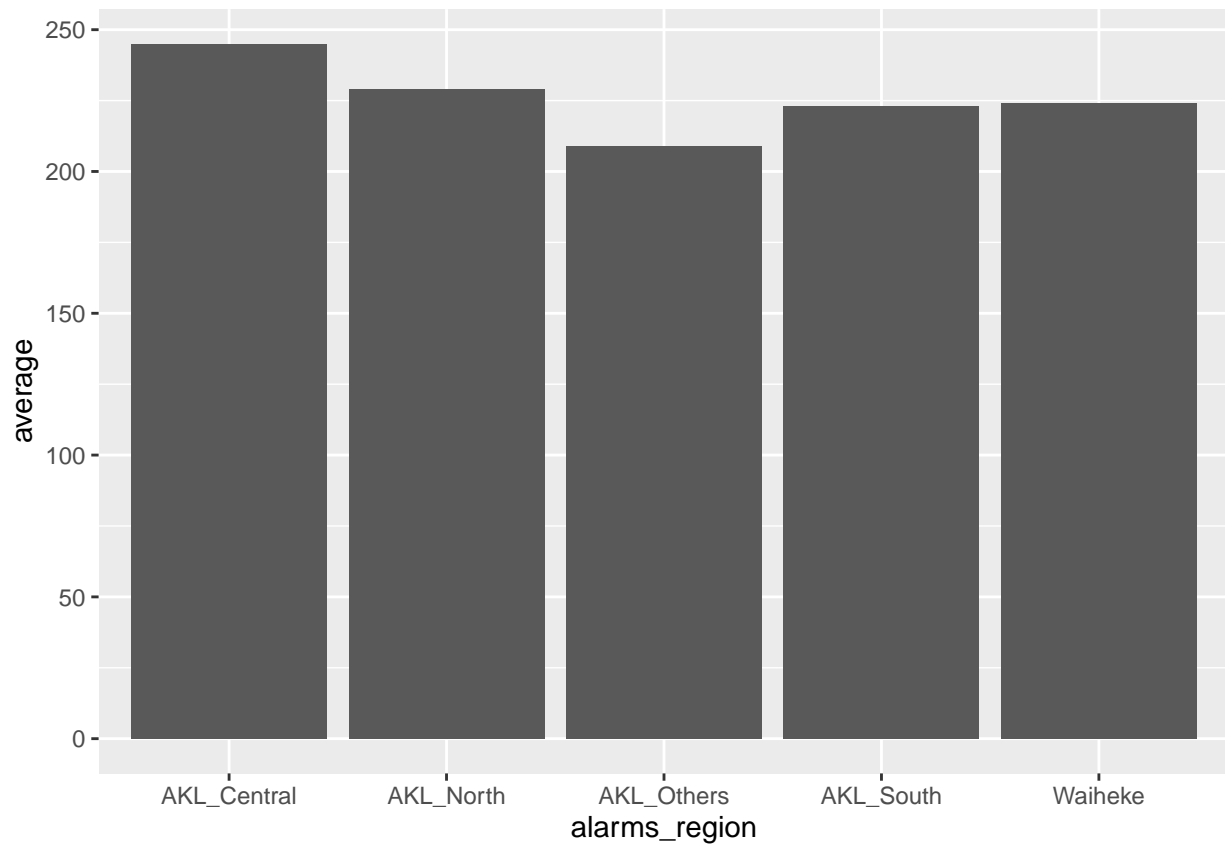
alarm_date	alarms_region	total
2024.02.20	AKL_North	78
2024.02.20	AKL_Others	186
2024.02.20	AKL_South	439
2024.02.20	Waiheke	368
2024.02.21	AKL_Central	243
2024.02.21	AKL_North	157
2024.02.21	AKL_Others	129
2024.02.21	AKL_South	106
2024.02.21	Waiheke	265
2024.02.22	AKL_Central	169
2024.02.22	AKL_North	211
2024.02.22	AKL_Others	430
2024.02.22	AKL_South	208
2024.02.22	Waiheke	210
2024.02.23	AKL_Central	350
2024.02.23	AKL_North	229
2024.02.23	AKL_Others	272
2024.02.23	AKL_South	323
2024.02.23	Waiheke	75
2024.02.24	AKL_Central	360
2024.02.24	AKL_North	129
2024.02.24	AKL_Others	196
2024.02.24	AKL_South	150
2024.02.24	Waiheke	155
2024.02.25	AKL_Central	169
2024.02.25	AKL_North	329
2024.02.25	AKL_Others	300
2024.02.25	AKL_South	307
2024.02.25	Waiheke	211
2024.02.26	AKL_Central	332
2024.02.26	AKL_North	186
2024.02.26	AKL_Others	368
2024.02.26	AKL_South	218
2024.02.26	Waiheke	191
2024.02.27	AKL_Central	251
2024.02.27	AKL_North	216
2024.02.27	AKL_Others	142
2024.02.27	AKL_South	182
2024.02.27	Waiheke	100
2024.02.28	AKL_Central	443
2024.02.28	AKL_North	399
2024.02.28	AKL_Others	155
2024.02.28	AKL_South	323
2024.02.28	Waiheke	236
2024.02.29	AKL_Central	228
2024.02.29	AKL_North	274
2024.02.29	AKL_Others	227
2024.02.29	AKL_South	379
2024.02.29	Waiheke	166
2024.03.01	AKL_Central	324
2024.03.01	AKL_North	280
2024.03.01	AKL_Others	177

alarm_date	alarms_region	total
2024.03.01	AKL_South	227
2024.03.01	Waiheke	162
2024.03.02	AKL_Central	217
2024.03.02	AKL_North	134
2024.03.02	AKL_Others	197
2024.03.02	AKL_South	137
2024.03.02	Waiheke	230
2024.03.03	AKL_Central	238
2024.03.03	AKL_North	192
2024.03.03	AKL_Others	117
2024.03.03	AKL_South	372
2024.03.03	Waiheke	226
2024.03.04	AKL_Central	260
2024.03.04	AKL_North	192
2024.03.04	AKL_Others	111
2024.03.04	AKL_South	291
2024.03.04	Waiheke	261

```
#alarms_insight.df <- alarms_count.df %>%
# pivot_longer(cols = !alarm_id, names_to = 'date'
# , values_to = 'frequency') %>%
# separate(date, c('prefix'
# ,
# 'year'
# ,
# 'month'
# ,
# 'day')) %>%
#select(-prefix) %>%
#mutate(category = if_else(frequency >= 230,#changed the shreshold because of
#the change of seed. by Gxd
# 'High'
# ,
# 'Low')) %>%
#left_join(alarms_info.df, by = c('alarm_id' = 'alarm_ID')) %>%
#group_by(year, month, day, alarms_region) %>%
#summarise(total = sum(frequency))
#by Gxd percent_high = 100*sum(category == 'High') / n())
alarms_signal_average.df<- alarms_summary.df%>%group_by(alarms_region)%>%summarise(average=round(mean(t
knitr::kable(alarms_signal_average.df)
```

alarms_region	average
AKL_Central	245
AKL_North	229
AKL_Others	209
AKL_South	223
Waiheke	224


```
ggplot(alarms_signal_average.df,aes(x=alarms_region,y=average))+geom_bar(stat="identity")
```



#end of Lab01 No.2

```
week_string<-c("Mon","Tue","Wed","Thu","Fri","Sat","Sun")
```

```
alarms_summary_by_wday.df<-
```

```
  alarms_summary.df%>%
  mutate(m_week_day=week_string[as.numeric(wday(as.Date(ymd(alarm_date),label=TRUE,week_start=1)))))%>%
  group_by(m_week_day)%>%
  summarise(week_day_total=sum(total))%>%filter((week_day_total==max(week_day_total))|(week_day_total==min(week_day_total)))
```

```
knitr::kable(alarms_summary_by_wday.df)
```

m_week_day	week_day_total
Fri	10668
Wed	9059

#end of Lab01 No.3

```
#alarms_signal_total.df<-alarms_insight.df%>%group_by(alarms_region,year)%>%summarise(totalAll=sum(total))
```

```
#ggplot(alarms_signal_insight.df,aes(x=alarms_region,y=totalAll))+geom_point(mapping = aes(color = year))
```

##Comment from Frank Guo

The rules of R language are too flexible and easy to be confused. Most functions are not easy to understand without a solid math background and, thus, not easy to be called correctly. For convenient purposes or other reasons, almost all the functions have too many vars, and they are not easy to understand deeply. It may be easy to start but hard to write well.

Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.