Assignment 1

Rex Shen

9/23/2021

```
# Load Packages
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.0 --
```

```
## v ggplot2 3.3.3     v purrr   0.3.4
## v tibble  3.1.2     v dplyr   1.0.2
## v tidyr   1.1.2     v stringr 1.4.0
## v readr   1.4.0     v forcats 0.5.0
```

```
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(ggplot2)
```

```
### Get Electricity Together

Years <- 2017:2020
Quarters <- 1:4
Type <- "Electric"

PGE_Elec <- NULL

for (Year in Years) {
    for (Quarter in Quarters) {

        Filename <- paste0("PGE/Electric/PGE_",
            Year, "_Q", Quarter, "_", Type,
            "UsageByZip.csv")

        Temp <- read_csv(Filename)
        PGE_Elec <- rbind(PGE_Elec, Temp)

    }
}
```

```
##
## -- Column specification --------------------------------------------------------
## cols(
##   ZIPCODE = col_double(),
##   MONTH = col_double(),
##   YEAR = col_double(),
##   CUSTOMERCLASS = col_character(),
##   COMBINED = col_character(),
```

```
##   TOTALCUSTOMERS = col_number(),
##   TOTALKWH = col_number(),
##   AVERAGEKWH = col_number()
## )
##
##
## -- Column specification --------------------------------------------------------
## cols(
##   ZIPCODE = col_double(),
##   MONTH = col_double(),
##   YEAR = col_double(),
##   CUSTOMERCLASS = col_character(),
##   COMBINED = col_character(),
##   TOTALCUSTOMERS = col_number(),
##   TOTALKWH = col_number(),
##   AVERAGEKWH = col_number()
## )
##
##
## -- Column specification --------------------------------------------------------
## cols(
##   ZIPCODE = col_double(),
##   MONTH = col_double(),
##   YEAR = col_double(),
##   CUSTOMERCLASS = col_character(),
##   COMBINED = col_character(),
##   TOTALCUSTOMERS = col_number(),
##   TOTALKWH = col_number(),
##   AVERAGEKWH = col_number()
## )
##
##
## -- Column specification --------------------------------------------------------
## cols(
##   ZIPCODE = col_double(),
##   MONTH = col_double(),
##   YEAR = col_double(),
##   CUSTOMERCLASS = col_character(),
##   COMBINED = col_character(),
##   TOTALCUSTOMERS = col_number(),
##   TOTALKWH = col_number(),
##   AVERAGEKWH = col_number()
## )
##
##
## -- Column specification --------------------------------------------------------
## cols(
##   ZIPCODE = col_double(),
##   MONTH = col_double(),
##   YEAR = col_double(),
##   CUSTOMERCLASS = col_character(),
##   COMBINED = col_character(),
##   TOTALCUSTOMERS = col_number(),
##   TOTALKWH = col_number(),
```

```
##   AVERAGEKWH = col_number()
## )
##
##
## -- Column specification --------------------------------------------------------
## cols(
##   ZIPCODE = col_double(),
##   MONTH = col_double(),
##   YEAR = col_double(),
##   CUSTOMERCLASS = col_character(),
##   COMBINED = col_character(),
##   TOTALCUSTOMERS = col_number(),
##   TOTALKWH = col_number(),
##   AVERAGEKWH = col_number()
## )
##
##
## -- Column specification --------------------------------------------------------
## cols(
##   ZIPCODE = col_double(),
##   MONTH = col_double(),
##   YEAR = col_double(),
##   CUSTOMERCLASS = col_character(),
##   COMBINED = col_character(),
##   TOTALCUSTOMERS = col_number(),
##   TOTALKWH = col_number(),
##   AVERAGEKWH = col_number()
## )
##
##
## -- Column specification --------------------------------------------------------
## cols(
##   ZIPCODE = col_double(),
##   MONTH = col_double(),
##   YEAR = col_double(),
##   CUSTOMERCLASS = col_character(),
##   COMBINED = col_character(),
##   TOTALCUSTOMERS = col_number(),
##   TOTALKWH = col_number(),
##   AVERAGEKWH = col_number()
## )
##
##
## -- Column specification --------------------------------------------------------
## cols(
##   ZIPCODE = col_double(),
##   MONTH = col_double(),
##   YEAR = col_double(),
##   CUSTOMERCLASS = col_character(),
##   COMBINED = col_character(),
##   TOTALCUSTOMERS = col_number(),
##   TOTALKWH = col_number(),
##   AVERAGEKWH = col_number()
## )
```

```
##
##
## -- Column specification --------------------------------------------------------
## cols(
##   ZIPCODE = col_double(),
##   MONTH = col_double(),
##   YEAR = col_double(),
##   CUSTOMERCLASS = col_character(),
##   COMBINED = col_character(),
##   TOTALCUSTOMERS = col_number(),
##   TOTALKWH = col_number(),
##   AVERAGEKWH = col_number()
## )
##
##
## -- Column specification --------------------------------------------------------
## cols(
##   ZIPCODE = col_double(),
##   MONTH = col_double(),
##   YEAR = col_double(),
##   CUSTOMERCLASS = col_character(),
##   COMBINED = col_character(),
##   TOTALCUSTOMERS = col_number(),
##   TOTALKWH = col_number(),
##   AVERAGEKWH = col_number()
## )
##
##
## -- Column specification --------------------------------------------------------
## cols(
##   ZIPCODE = col_double(),
##   MONTH = col_double(),
##   YEAR = col_double(),
##   CUSTOMERCLASS = col_character(),
##   COMBINED = col_character(),
##   TOTALCUSTOMERS = col_number(),
##   TOTALKWH = col_number(),
##   AVERAGEKWH = col_number()
## )
##
##
## -- Column specification --------------------------------------------------------
## cols(
##   ZIPCODE = col_double(),
##   MONTH = col_double(),
##   YEAR = col_double(),
##   CUSTOMERCLASS = col_character(),
##   COMBINED = col_character(),
##   TOTALCUSTOMERS = col_number(),
##   TOTALKWH = col_number(),
##   AVERAGEKWH = col_number()
## )
##
##
```

```
## -- Column specification --------------------------------------------------------
## cols(
##   ZIPCODE = col_double(),
##   MONTH = col_double(),
##   YEAR = col_double(),
##   CUSTOMERCLASS = col_character(),
##   COMBINED = col_character(),
##   TOTALCUSTOMERS = col_number(),
##   TOTALKWH = col_number(),
##   AVERAGEKWH = col_number()
## )
##
##
## -- Column specification --------------------------------------------------------
## cols(
##   ZIPCODE = col_double(),
##   MONTH = col_double(),
##   YEAR = col_double(),
##   CUSTOMERCLASS = col_character(),
##   COMBINED = col_character(),
##   TOTALCUSTOMERS = col_number(),
##   TOTALKWH = col_number(),
##   AVERAGEKWH = col_number()
## )
##
##
## -- Column specification --------------------------------------------------------
## cols(
##   ZIPCODE = col_double(),
##   MONTH = col_double(),
##   YEAR = col_double(),
##   CUSTOMERCLASS = col_character(),
##   COMBINED = col_character(),
##   TOTALCUSTOMERS = col_number(),
##   TOTALKWH = col_number(),
##   AVERAGEKWH = col_number()
## )
```

```r
Quarters = 1:2

for (Quarter in Quarters) {

    Filename <- paste0("PGE/Electric/PGE_2021_Q",
        Quarter, "_", Type, "UsageByZip.csv")

    Temp <- read_csv(Filename)
    PGE_Elec <- rbind(PGE_Elec, Temp)
}
```

```
##
## -- Column specification --------------------------------------------------------
## cols(
##   ZIPCODE = col_double(),
##   MONTH = col_double(),
##   YEAR = col_double(),
```

```
##   CUSTOMERCLASS = col_character(),
##   COMBINED = col_character(),
##   TOTALCUSTOMERS = col_number(),
##   TOTALKWH = col_number(),
##   AVERAGEKWH = col_number()
## )
##
##
## -- Column specification --------------------------------------------------------
## cols(
##   ZIPCODE = col_double(),
##   MONTH = col_double(),
##   YEAR = col_double(),
##   CUSTOMERCLASS = col_character(),
##   COMBINED = col_character(),
##   TOTALCUSTOMERS = col_number(),
##   TOTALKWH = col_number(),
##   AVERAGEKWH = col_number()
## )
```

```
PGE_Elec = data.frame(PGE_Elec)
```

```
# Get Gas Together

Years <- 2017:2020
Quarters <- 1:4
Type <- "Gas"

PGE_Gas <- NULL

for (Year in Years) {

    for (Quarter in Quarters) {

        Filename <- paste0("PGE/Gas/PGE_",
            Year, "_Q", Quarter, "_", Type,
            "UsageByZip.csv")

        Temp <- read_csv(Filename)
        PGE_Gas <- rbind(PGE_Gas, Temp)
    }
}
```

```
##
## -- Column specification --------------------------------------------------------
## cols(
##   ZIPCODE = col_double(),
##   MONTH = col_double(),
##   YEAR = col_double(),
##   CUSTOMERCLASS = col_character(),
##   COMBINED = col_character(),
##   TOTALCUSTOMERS = col_number(),
##   TOTALTHM = col_number(),
##   AVERAGETHM = col_number()
## )
```

```
##
##
## -- Column specification --------------------------------------------------------
## cols(
##   ZIPCODE = col_double(),
##   MONTH = col_double(),
##   YEAR = col_double(),
##   CUSTOMERCLASS = col_character(),
##   COMBINED = col_character(),
##   TOTALCUSTOMERS = col_number(),
##   TOTALTHM = col_number(),
##   AVERAGETHM = col_number()
## )
##
## -- Column specification --------------------------------------------------------
## cols(
##   ZIPCODE = col_double(),
##   MONTH = col_double(),
##   YEAR = col_double(),
##   CUSTOMERCLASS = col_character(),
##   COMBINED = col_character(),
##   TOTALCUSTOMERS = col_number(),
##   TOTALTHM = col_number(),
##   AVERAGETHM = col_double()
## )
##
## -- Column specification --------------------------------------------------------
## cols(
##   ZIPCODE = col_double(),
##   MONTH = col_double(),
##   YEAR = col_double(),
##   CUSTOMERCLASS = col_character(),
##   COMBINED = col_character(),
##   TOTALCUSTOMERS = col_number(),
##   TOTALTHM = col_number(),
##   AVERAGETHM = col_number()
## )
##
##
## -- Column specification --------------------------------------------------------
## cols(
##   ZIPCODE = col_double(),
##   MONTH = col_double(),
##   YEAR = col_double(),
##   CUSTOMERCLASS = col_character(),
##   COMBINED = col_character(),
##   TOTALCUSTOMERS = col_number(),
##   TOTALTHM = col_number(),
##   AVERAGETHM = col_number()
## )
##
##
## -- Column specification --------------------------------------------------------
```

```
## cols(
##   ZIPCODE = col_double(),
##   MONTH = col_double(),
##   YEAR = col_double(),
##   CUSTOMERCLASS = col_character(),
##   COMBINED = col_character(),
##   TOTALCUSTOMERS = col_number(),
##   TOTALTHM = col_number(),
##   AVERAGETHM = col_number()
## )
##
## -- Column specification --------------------------------------------------------
## cols(
##   ZIPCODE = col_double(),
##   MONTH = col_double(),
##   YEAR = col_double(),
##   CUSTOMERCLASS = col_character(),
##   COMBINED = col_character(),
##   TOTALCUSTOMERS = col_number(),
##   TOTALTHM = col_number(),
##   AVERAGETHM = col_double()
## )
##
## -- Column specification --------------------------------------------------------
## cols(
##   ZIPCODE = col_double(),
##   MONTH = col_double(),
##   YEAR = col_double(),
##   CUSTOMERCLASS = col_character(),
##   COMBINED = col_character(),
##   TOTALCUSTOMERS = col_number(),
##   TOTALTHM = col_number(),
##   AVERAGETHM = col_number()
## )
##
##
## -- Column specification --------------------------------------------------------
## cols(
##   ZIPCODE = col_double(),
##   MONTH = col_double(),
##   YEAR = col_double(),
##   CUSTOMERCLASS = col_character(),
##   COMBINED = col_character(),
##   TOTALCUSTOMERS = col_number(),
##   TOTALTHM = col_number(),
##   AVERAGETHM = col_number()
## )
##
##
## -- Column specification --------------------------------------------------------
## cols(
##   ZIPCODE = col_double(),
##   MONTH = col_double(),
```

```
##   YEAR = col_double(),
##   CUSTOMERCLASS = col_character(),
##   COMBINED = col_character(),
##   TOTALCUSTOMERS = col_number(),
##   TOTALTHM = col_number(),
##   AVERAGETHM = col_number()
## )
##
##
## -- Column specification --------------------------------------------------------
## cols(
##   ZIPCODE = col_double(),
##   MONTH = col_double(),
##   YEAR = col_double(),
##   CUSTOMERCLASS = col_character(),
##   COMBINED = col_character(),
##   TOTALCUSTOMERS = col_number(),
##   TOTALTHM = col_number(),
##   AVERAGETHM = col_number()
## )
##
##
## -- Column specification --------------------------------------------------------
## cols(
##   ZIPCODE = col_double(),
##   MONTH = col_double(),
##   YEAR = col_double(),
##   CUSTOMERCLASS = col_character(),
##   COMBINED = col_character(),
##   TOTALCUSTOMERS = col_number(),
##   TOTALTHM = col_number(),
##   AVERAGETHM = col_number()
## )
##
##
## -- Column specification --------------------------------------------------------
## cols(
##   ZIPCODE = col_double(),
##   MONTH = col_double(),
##   YEAR = col_double(),
##   CUSTOMERCLASS = col_character(),
##   COMBINED = col_character(),
##   TOTALCUSTOMERS = col_number(),
##   TOTALTHM = col_number(),
##   AVERAGETHM = col_number()
## )
##
##
## -- Column specification --------------------------------------------------------
## cols(
##   ZIPCODE = col_double(),
##   MONTH = col_double(),
##   YEAR = col_double(),
##   CUSTOMERCLASS = col_character(),
```

```
##   COMBINED = col_character(),
##   TOTALCUSTOMERS = col_number(),
##   TOTALTHM = col_number(),
##   AVERAGETHM = col_number()
## )
```

```
##
## -- Column specification --------------------------------------------------------
## cols(
##   ZIPCODE = col_double(),
##   MONTH = col_double(),
##   YEAR = col_double(),
##   CUSTOMERCLASS = col_character(),
##   COMBINED = col_character(),
##   TOTALCUSTOMERS = col_number(),
##   TOTALTHM = col_number(),
##   AVERAGETHM = col_double()
## )
```

```
##
## -- Column specification --------------------------------------------------------
## cols(
##   ZIPCODE = col_double(),
##   MONTH = col_double(),
##   YEAR = col_double(),
##   CUSTOMERCLASS = col_character(),
##   COMBINED = col_character(),
##   TOTALCUSTOMERS = col_number(),
##   TOTALTHM = col_number(),
##   AVERAGETHM = col_number()
## )
```

```
Quarters = 1:2

for (Quarter in Quarters) {

    Filename <- paste0("PGE/Gas/PGE_2021_Q",
        Quarter, "_", Type, "UsageByZip.csv")

    Temp <- read_csv(Filename)
    PGE_Gas <- rbind(PGE_Gas, Temp)
}
```

```
##
## -- Column specification --------------------------------------------------------
## cols(
##   ZIPCODE = col_double(),
##   MONTH = col_double(),
##   YEAR = col_double(),
##   CUSTOMERCLASS = col_character(),
##   COMBINED = col_character(),
##   TOTALCUSTOMERS = col_number(),
##   TOTALTHM = col_number(),
##   AVERAGETHM = col_number()
## )
##
```

```
##
## -- Column specification --------------------------------------------------------
## cols(
##   ZIPCODE = col_double(),
##   MONTH = col_double(),
##   YEAR = col_double(),
##   CUSTOMERCLASS = col_character(),
##   COMBINED = col_character(),
##   TOTALCUSTOMERS = col_number(),
##   TOTALTHM = col_number(),
##   AVERAGETHM = col_number()
## )
```

```
PGE_Gas = data.frame(PGE_Gas)
rm(Temp)
```

```
names(PGE_Elec)[7] <- "TOTAL_KWH/THM"
names(PGE_Elec)[8] <- "AVERAGE_KWH/THM"


names(PGE_Gas)[7] <- "TOTAL_KWH/THM"
names(PGE_Gas)[8] <- "AVERAGE_KWH/THM"

finalData = rbind(PGE_Elec, PGE_Gas)
```

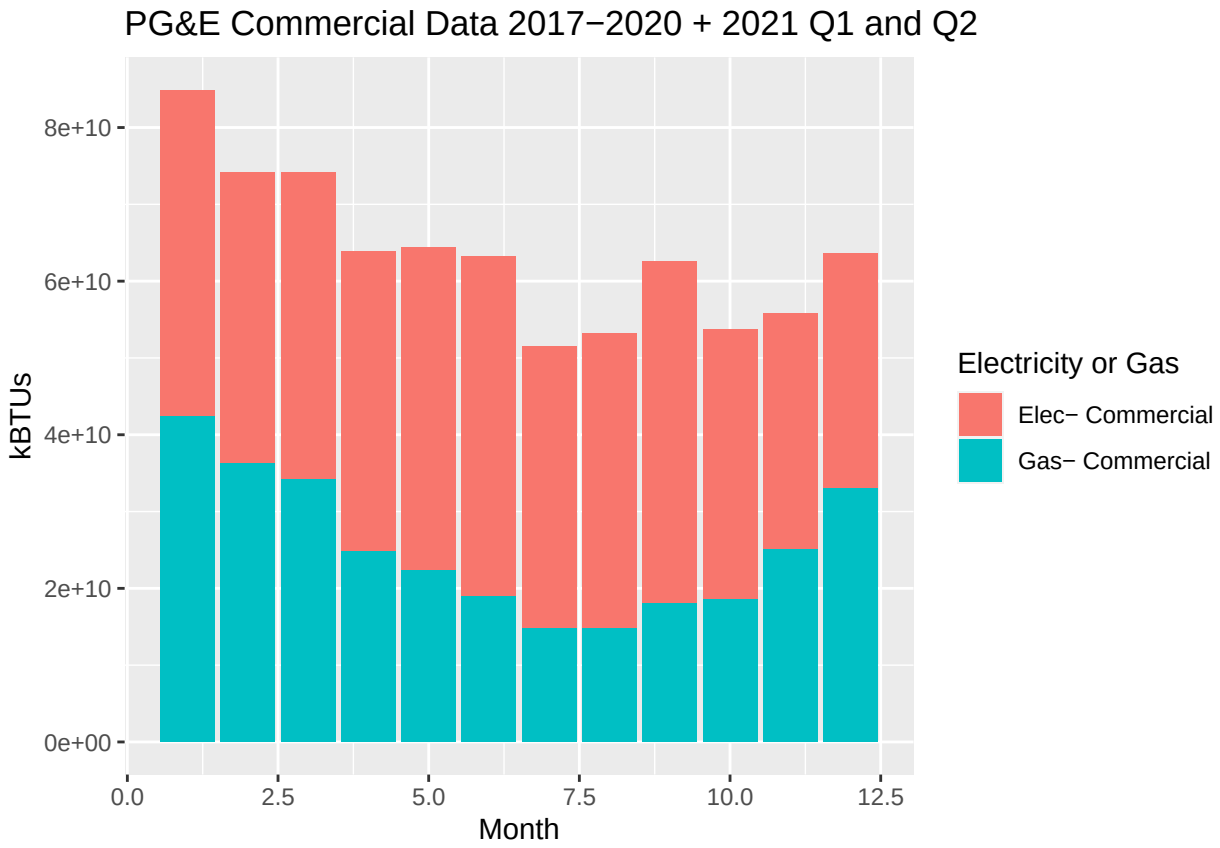
```
for (i in 1:nrow(finalData)) {

    if (startsWith(finalData$CUSTOMERCLASS[i],
        "Elec")) {
        finalData$TOTAL_kBTUs[i] = finalData$`TOTAL_KWH/THM`[i] *
            3.412

    } else if (startsWith(finalData$CUSTOMERCLASS[i],
        "Gas")) {
        finalData$TOTAL_kBTUs[i] = finalData$`TOTAL_KWH/THM`[i] *
            100
    }
}
```

```
ResidentialData = subset(finalData, finalData$CUSTOMERCLASS ==
    "Elec- Residential" | finalData$CUSTOMERCLASS ==
    "Gas- Residential")

CommercialData = subset(finalData, finalData$CUSTOMERCLASS ==
    "Elec- Commercial" | finalData$CUSTOMERCLASS ==
    "Gas- Commercial")
```

```
# Residential

newPlotOne = ggplot(data = ResidentialData) +
    geom_bar(aes(x = MONTH, y = TOTAL_kBTUs,
        fill = CUSTOMERCLASS), stat = "identity",
        position = "stack") + labs(x = "Month",
    y = "kBTUs", title = "PG&E Residential Data 2017-2020 + 2021 Q1 and Q2",
    fill = "Electricity or Gas")
```

```
newPlotOne
```


```
# Commercial

newPlotTwo = ggplot(data = CommercialData) +
    geom_bar(aes(x = MONTH, y = TOTAL_kBTUs,
        fill = CUSTOMERCLASS), stat = "identity",
        position = "stack") + labs(x = "Month",
    y = "kBTUs", title = "PG&E Commercial Data 2017-2020 + 2021 Q1 and Q2",
    fill = "Electricity or Gas")

newPlotTwo
```



COVID Pandemic 2019 - 2020 Analysis

Note, make some plots here.

Comment on any observable changes in energy consumption that may be attributable to the COVID-19 pandemic (you are encouraged to create additional plots that help emphasize the change between 2019 and 2020). Explain any key assumptions you made in the analysis, or caveats about the data sources that you think the reader should be aware of. Publish all of this work in a GitHub webpage titled "yourname_A1" (using the steps from Chapter 1.4). For this assignment, a support script is available by request from the TAs.