

```

# Load packages
library(plyr)
library(fpp3)
library(tsibble)
library(forecast)
library(zoo)
library(plyr)
library(dplyr)
library(tidyverse)
library(ggplot2)
library(quantmod)
library(janitor)
library(lubridate)

# Import phoenix->chicago reefer lane with yuma weather, volume, and lags
shipping <- readr::read_csv(file = 'data/data_phoenix_with_yuma_weather_and_volume_and_lags.csv') %>%
  dplyr::select(-X1) %>%
  relocate(yw, Mode, ORegionDAT, DRegionDAT, approx_cost)

# Import lettuce price spreadsheets for iceberg, green leaf, red leaf, and romaine
# iceberg <- readr::read_csv(file = 'data/lettuce_wholesale/iceberg_prices.csv') %>%
#   clean_names() %>%
#   select(record_date, farm_price, chicago_retail) %>%
#   dplyr::rename(yw = record_date,
#                 iceberg_farm = farm_price,
#                 iceberg_retail = chicago_retail) %>%
#   mutate(yw = yearweek(yw))
#
# green_leaf <- readr::read_csv(file = 'data/lettuce_wholesale/green_leaf_prices.csv') %>%
#   clean_names() %>%
#   select(record_date, farm_price, chicago_retail) %>%
#   dplyr::rename(yw = record_date,
#                 green_leaf_farm = farm_price,
#                 green_leaf_retail = chicago_retail) %>%
#   mutate(yw = yearweek(yw))
#
# red_leaf <- readr::read_csv(file = 'data/lettuce_wholesale/red_leaf_prices.csv') %>%
#   clean_names() %>%
#   select(record_date, farm_price, chicago_retail) %>%
#   dplyr::rename(yw = record_date,
#                 red_leaf_farm = farm_price,
#                 red_leaf_retail = chicago_retail) %>%
#   mutate(yw = yearweek(yw))
#
# romaine <- readr::read_csv(file = 'data/lettuce_wholesale/romaine_prices.csv') %>%
#   clean_names() %>%
#   select(record_date, farm_price, chicago_retail) %>%
#   dplyr::rename(yw = record_date,
#                 romaine_farm = farm_price,
#                 romaine_retail = chicago_retail) %>%
#   mutate(yw = yearweek(yw))
#
# iceberg

```

```
# green_leaf
# red_leaf
# romaine
```

```
#define function for converting format of price columns
#the input chr should be in a form like "$1.23" The output would be "1.23"
```

```
convert_price <- function(chr){
  output = sub('.', '', chr) #remove first character
  return(output)
}
```

```
#test case
test_price = convert_price("$1.23")
test_price
```

```
## [1] "1.23"
```

```
#Import lettuce price spreadsheets for iceberg, green leaf, red leaf, and romaine
iceberg <- readr::read_csv(file = 'data/lettuce_wholesale/iceberg_prices.csv') %>%
  clean_names() %>%
```

```
  select(commodity, record_date, farm_price, chicago_retail) %>%
  mutate(yw = yearweek(record_date)) %>%
  mutate(commodity = "IC")%>%
  select(-record_date) %>%
  mutate(farm_price = as.numeric(modify(farm_price, convert_price)), chicago_retail = as.numeric(modify
```

```
##
## -- Column specification -----
## cols(
##   Commodity = col_character(),
##   'Record Date' = col_date(format = ""),
##   'Farm Price' = col_character(),
##   'Atlanta Retail' = col_character(),
##   'Chicago Retail' = col_character(),
##   'Los Angeles Retail' = col_character(),
##   'New York Retail' = col_character(),
##   'Average Spread' = col_character()
## )
```

```
green_leaf <- readr::read_csv(file = 'data/lettuce_wholesale/green_leaf_prices.csv') %>%
  clean_names() %>%
```

```
  select(commodity, record_date, farm_price, chicago_retail) %>%
  mutate(yw = yearweek(record_date)) %>%
  mutate(commodity = "GL")%>%
  select(-record_date)%>%
  mutate(farm_price = as.numeric(modify(farm_price, convert_price)), chicago_retail = as.numeric(modify
```

```
##
## -- Column specification -----
## cols(
##   Commodity = col_character(),
```

```
## 'Record Date' = col_date(format = ""),
## 'Farm Price' = col_character(),
## 'Atlanta Retail' = col_character(),
## 'Chicago Retail' = col_character(),
## 'Los Angeles Retail' = col_character(),
## 'New York Retail' = col_character(),
## 'Average Spread' = col_character()
## )
```

```
red_leaf <- readr::read_csv(file = 'data/lettuce_wholesale/red_leaf_prices.csv') %>%
  clean_names() %>%
  select(commodity, record_date, farm_price, chicago_retail) %>%
  mutate(yw = yearweek(record_date)) %>%
  mutate(commodity = "RL") %>%
  select(-record_date) %>%
  mutate(farm_price = as.numeric(modify(farm_price, convert_price)), chicago_retail = as.numeric(modify
```

```
##
## -- Column specification -----
## cols(
##   Commodity = col_character(),
##   'Record Date' = col_date(format = ""),
##   'Farm Price' = col_character(),
##   'Atlanta Retail' = col_character(),
##   'Chicago Retail' = col_character(),
##   'Los Angeles Retail' = col_character(),
##   'New York Retail' = col_character(),
##   'Average Spread' = col_character()
## )
```

```
romaine <- readr::read_csv(file = 'data/lettuce_wholesale/romaine_prices.csv') %>%
  clean_names() %>%
  select(commodity, record_date, farm_price, chicago_retail) %>%
  mutate(yw = yearweek(record_date)) %>%
  mutate(commodity = "RO") %>%
  select(-record_date) %>%
  mutate(farm_price = as.numeric(modify(farm_price, convert_price)), chicago_retail = as.numeric(modify
```

```
##
## -- Column specification -----
## cols(
##   Commodity = col_character(),
##   'Record Date' = col_date(format = ""),
##   'Farm Price' = col_character(),
##   'Atlanta Retail' = col_character(),
##   'Chicago Retail' = col_character(),
##   'Los Angeles Retail' = col_character(),
##   'New York Retail' = col_character(),
##   'Average Spread' = col_character()
## )
```

iceberg

```
## # A tibble: 123 x 4
##   commodity farm_price chicago_retail    yw
##   <chr>         <dbl>         <dbl>  <week>
## 1 IC           0.35           1.46 2019 W20
## 2 IC           0.36           1.46 2019 W19
## 3 IC           0.38           1.46 2019 W18
## 4 IC           0.35           1.56 2019 W17
## 5 IC           0.42           1.56 2019 W16
## 6 IC           0.57           1.61 2019 W15
## 7 IC           0.5           1.61 2019 W14
## 8 IC           0.51           1.56 2019 W13
## 9 IC           0.56           1.61 2019 W12
## 10 IC          0.92           1.56 2019 W11
## # ... with 113 more rows
```

green_leaf

```
## # A tibble: 109 x 4
##   commodity farm_price chicago_retail    yw
##   <chr>         <dbl>         <dbl>  <week>
## 1 GL           0.31           1.7 2019 W20
## 2 GL           0.35           1.7 2019 W19
## 3 GL           0.34           1.71 2019 W18
## 4 GL           0.35           1.7 2019 W17
## 5 GL           0.43           1.77 2019 W16
## 6 GL           0.43           1.82 2019 W15
## 7 GL           0.39           1.93 2019 W14
## 8 GL           0.34           1.8 2019 W13
## 9 GL           0.34           1.93 2019 W12
## 10 GL          0.38           1.78 2019 W11
## # ... with 99 more rows
```

red_leaf

```
## # A tibble: 123 x 4
##   commodity farm_price chicago_retail    yw
##   <chr>         <dbl>         <dbl>  <week>
## 1 RL           0.32           1.84 2019 W20
## 2 RL           0.34           1.75 2019 W19
## 3 RL           0.34           1.73 2019 W18
## 4 RL           0.38           1.75 2019 W17
## 5 RL           0.45           1.98 2019 W16
## 6 RL           0.44           2.03 2019 W15
## 7 RL           0.43           1.86 2019 W14
## 8 RL           0.34           2    2019 W13
## 9 RL           0.35           1.85 2019 W12
## 10 RL          0.6           1.73 2019 W11
## # ... with 113 more rows
```

```
romaine
```

```
## # A tibble: 122 x 4
##   commodity farm_price chicago_retail    yw
##   <chr>         <dbl>         <dbl>  <week>
## 1 R0           0.35           2    2019 W20
## 2 R0           0.36           2    2019 W19
## 3 R0           0.36           2    2019 W18
## 4 R0           0.39           2    2019 W17
## 5 R0           0.51           2    2019 W16
## 6 R0           0.56          2.12  2019 W15
## 7 R0           0.46          2.28  2019 W14
## 8 R0           0.39          2.12  2019 W13
## 9 R0           0.34          2.12  2019 W12
## 10 R0          0.43          2.12  2019 W11
## # ... with 112 more rows
```

```
#check for duplicates
duplicates(iceberg, key = commodity, index = yw)
```

```
## # A tibble: 0 x 4
## # ... with 4 variables: commodity <chr>, farm_price <dbl>,
## #   chicago_retail <dbl>, yw <week>
```

```
duplicates(red_leaf, key = commodity, index = yw)
```

```
## # A tibble: 0 x 4
## # ... with 4 variables: commodity <chr>, farm_price <dbl>,
## #   chicago_retail <dbl>, yw <week>
```

```
duplicates(green_leaf, key = commodity, index = yw)
```

```
## # A tibble: 2 x 4
##   commodity farm_price chicago_retail    yw
##   <chr>         <dbl>         <dbl>  <week>
## 1 GL           2.04           1.72  2018 W48
## 2 GL           0.68           1.68  2018 W48
```

```
duplicates(romaine, key = commodity, index = yw)
```

```
## # A tibble: 0 x 4
## # ... with 4 variables: commodity <chr>, farm_price <dbl>,
## #   chicago_retail <dbl>, yw <week>
```

```
#aggregate the duplicates by averaging
green_leaf_duplicates_removed = plyr::ddply(green_leaf, .(commodity, yw), numcolwise(mean))

green_leaf
```

```
## # A tibble: 109 x 4
##   commodity farm_price chicago_retail    yw
##   <chr>         <dbl>         <dbl>  <week>
## 1 GL           0.31           1.7  2019 W20
## 2 GL           0.35           1.7  2019 W19
## 3 GL           0.34           1.71 2019 W18
## 4 GL           0.35           1.7  2019 W17
## 5 GL           0.43           1.77 2019 W16
## 6 GL           0.43           1.82 2019 W15
## 7 GL           0.39           1.93 2019 W14
## 8 GL           0.34           1.8  2019 W13
## 9 GL           0.34           1.93 2019 W12
## 10 GL          0.38           1.78 2019 W11
## # ... with 99 more rows
```

green_leaf_duplicates_removed

```
##   commodity    yw farm_price chicago_retail
## 1      GL 2016 W52      0.33          1.43
## 2      GL 2017 W01      0.30          1.38
## 3      GL 2017 W02      0.30          1.48
## 4      GL 2017 W03      0.31          1.23
## 5      GL 2017 W04      0.34          1.30
## 6      GL 2017 W05      0.41          1.23
## 7      GL 2017 W06      0.62          1.30
## 8      GL 2017 W07      0.75          1.25
## 9      GL 2017 W08      0.64          1.83
## 10     GL 2017 W09      0.60          1.50
## 11     GL 2017 W10      0.70          1.45
## 12     GL 2017 W11      1.04          1.50
## 13     GL 2017 W12      1.73          1.50
## 14     GL 2017 W13      2.14          1.79
## 15     GL 2017 W14      2.14          1.83
## 16     GL 2017 W15      2.08          2.26
## 17     GL 2017 W16      1.14          1.83
## 18     GL 2017 W17      0.52          1.91
## 19     GL 2017 W18      0.45          1.83
## 20     GL 2017 W19      0.39          1.91
## 21     GL 2017 W20      0.33          1.83
## 22     GL 2017 W21      0.29          1.71
## 23     GL 2017 W22      0.29          1.91
## 24     GL 2017 W23      0.29          1.83
## 25     GL 2017 W24      0.31          1.91
## 26     GL 2017 W25      0.31          1.83
## 27     GL 2017 W26      0.32          1.91
## 28     GL 2017 W27      0.32          1.42
## 29     GL 2017 W28      0.33          1.38
## 30     GL 2017 W29      0.29          1.33
## 31     GL 2017 W30      0.29          1.67
## 32     GL 2017 W31      0.29          1.58
## 33     GL 2017 W32      0.29          1.67
## 34     GL 2017 W33      0.29          1.58
## 35     GL 2017 W34      0.31          1.23
## 36     GL 2017 W35      0.41          1.67
```

## 37	GL 2017 W36	0.38	1.58
## 38	GL 2017 W37	0.32	1.62
## 39	GL 2017 W38	0.36	1.75
## 40	GL 2017 W39	0.40	1.69
## 41	GL 2017 W40	0.43	1.61
## 42	GL 2017 W41	0.41	1.75
## 43	GL 2017 W42	0.36	1.75
## 44	GL 2017 W43	0.37	1.69
## 45	GL 2017 W44	0.36	1.61
## 46	GL 2017 W45	0.35	1.69
## 47	GL 2017 W46	0.31	1.75
## 48	GL 2017 W47	0.30	1.68
## 49	GL 2017 W48	0.32	1.87
## 50	GL 2017 W49	0.34	1.92
## 51	GL 2017 W50	1.26	2.41
## 52	GL 2017 W51	0.33	1.92
## 53	GL 2017 W52	0.34	1.88
## 54	GL 2018 W01	0.35	1.73
## 55	GL 2018 W02	0.30	1.65
## 56	GL 2018 W03	0.29	1.73
## 57	GL 2018 W04	0.29	1.65
## 58	GL 2018 W05	0.30	1.61
## 59	GL 2018 W06	0.30	1.68
## 60	GL 2018 W23	0.30	1.71
## 61	GL 2018 W24	0.30	1.63
## 62	GL 2018 W25	0.29	1.71
## 63	GL 2018 W26	0.30	1.63
## 64	GL 2018 W27	0.33	1.71
## 65	GL 2018 W28	0.34	1.59
## 66	GL 2018 W29	0.32	1.71
## 67	GL 2018 W30	0.30	1.59
## 68	GL 2018 W31	0.30	1.48
## 69	GL 2018 W32	0.34	1.63
## 70	GL 2018 W33	0.35	1.78
## 71	GL 2018 W34	0.34	1.57
## 72	GL 2018 W35	0.35	1.63
## 73	GL 2018 W36	0.35	1.71
## 74	GL 2018 W37	0.37	1.63
## 75	GL 2018 W38	0.56	1.71
## 76	GL 2018 W39	0.67	1.63
## 77	GL 2018 W40	0.58	1.71
## 78	GL 2018 W41	0.43	1.63
## 79	GL 2018 W42	0.45	1.71
## 80	GL 2018 W43	0.55	1.57
## 81	GL 2018 W44	0.66	1.63
## 82	GL 2018 W45	0.69	1.76
## 83	GL 2018 W47	1.84	1.75
## 84	GL 2018 W48	1.36	1.70
## 85	GL 2018 W49	2.03	1.72
## 86	GL 2018 W50	0.90	1.92
## 87	GL 2018 W51	0.64	1.92
## 88	GL 2018 W52	0.61	1.84
## 89	GL 2019 W01	0.64	1.84
## 90	GL 2019 W02	0.66	1.85

```
## 91      GL 2019 W03      0.51      1.75
## 92      GL 2019 W04      0.41      1.75
## 93      GL 2019 W05      0.38      1.90
## 94      GL 2019 W06      0.42      1.90
## 95      GL 2019 W07      0.65      1.93
## 96      GL 2019 W08      0.82      1.93
## 97      GL 2019 W09      0.65      1.78
## 98      GL 2019 W10      0.48      1.93
## 99      GL 2019 W11      0.38      1.78
## 100     GL 2019 W12      0.34      1.93
## 101     GL 2019 W13      0.34      1.80
## 102     GL 2019 W14      0.39      1.93
## 103     GL 2019 W15      0.43      1.82
## 104     GL 2019 W16      0.43      1.77
## 105     GL 2019 W17      0.35      1.70
## 106     GL 2019 W18      0.34      1.71
## 107     GL 2019 W19      0.35      1.70
## 108     GL 2019 W20      0.31      1.70
```

```
duplicates(green_leaf_duplicates_removed, key = commodity, index = yw)
```

```
## # A tibble: 0 x 4
## # ... with 4 variables: commodity <chr>, yw <week>, farm_price <dbl>,
## #   chicago_retail <dbl>
```

```
#join lettuce data together
```

```
two_lettuce <- join(iceberg, green_leaf_duplicates_removed, type = "full", match = "all")
```

```
## Joining by: commodity, farm_price, chicago_retail, yw
```

```
three_lettuce <- join(two_lettuce, red_leaf, type = "full", match = "all")
```

```
## Joining by: commodity, farm_price, chicago_retail, yw
```

```
all_lettuce <- join(three_lettuce, romaine, type = "full", match = "all")
```

```
## Joining by: commodity, farm_price, chicago_retail, yw
```

```
all_lettuce <- all_lettuce %>%
  as_tsibble(key = commodity,
             index = yw) %>%
  fill_gaps(.full = TRUE) %>%
  filter_index("2017 W01" ~ "2019 W20") #remove anything before 2017

all_lettuce
```

```
## # A tsibble: 496 x 4 [1W]
## # Key:      commodity [4]
##   commodity farm_price chicago_retail      yw
##   <chr>      <dbl>      <dbl>    <week>
```

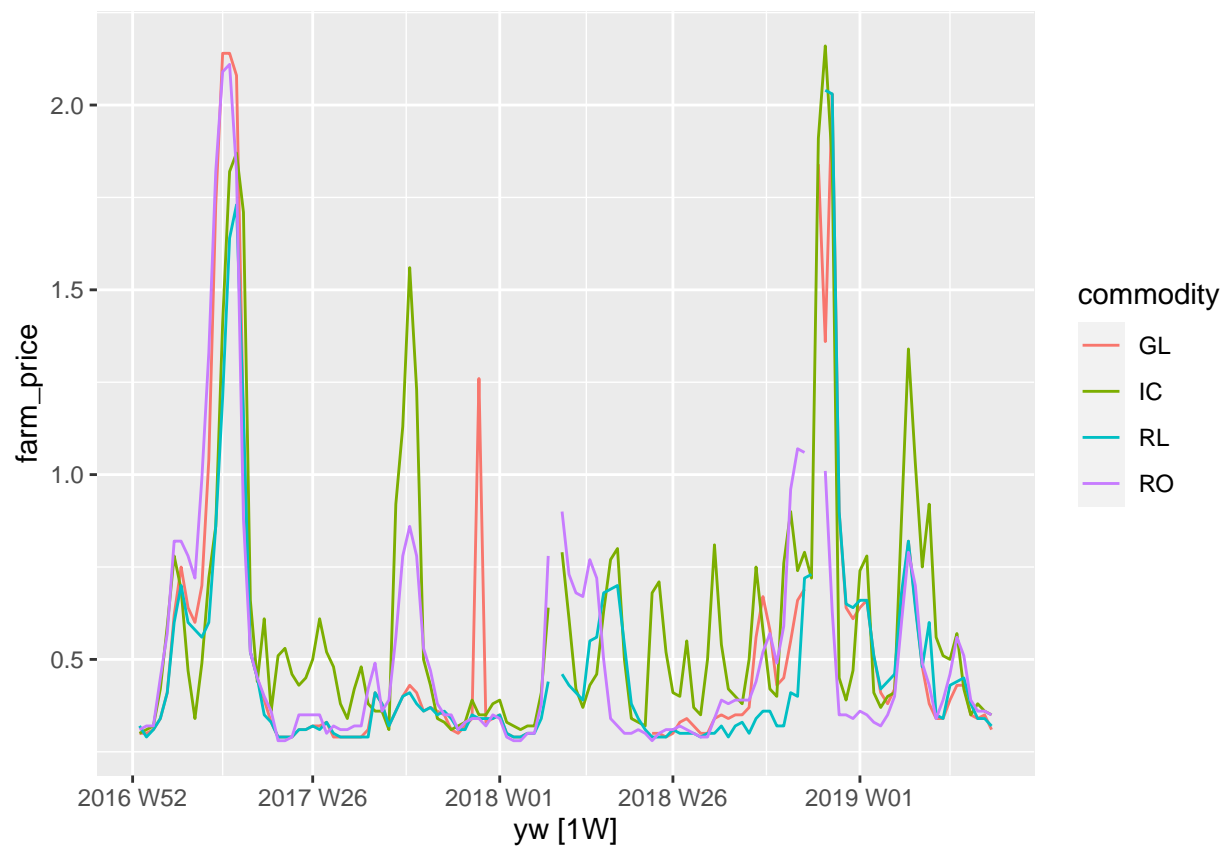


```
## 1 GL          0.3          1.38 2017 W01
## 2 GL          0.3          1.48 2017 W02
## 3 GL          0.31         1.23 2017 W03
## 4 GL          0.34         1.3   2017 W04
## 5 GL          0.41         1.23 2017 W05
## 6 GL          0.62         1.3   2017 W06
## 7 GL          0.75         1.25 2017 W07
## 8 GL          0.64         1.83 2017 W08
## 9 GL          0.6          1.5   2017 W09
## 10 GL         0.7          1.45 2017 W10
## # ... with 486 more rows
```

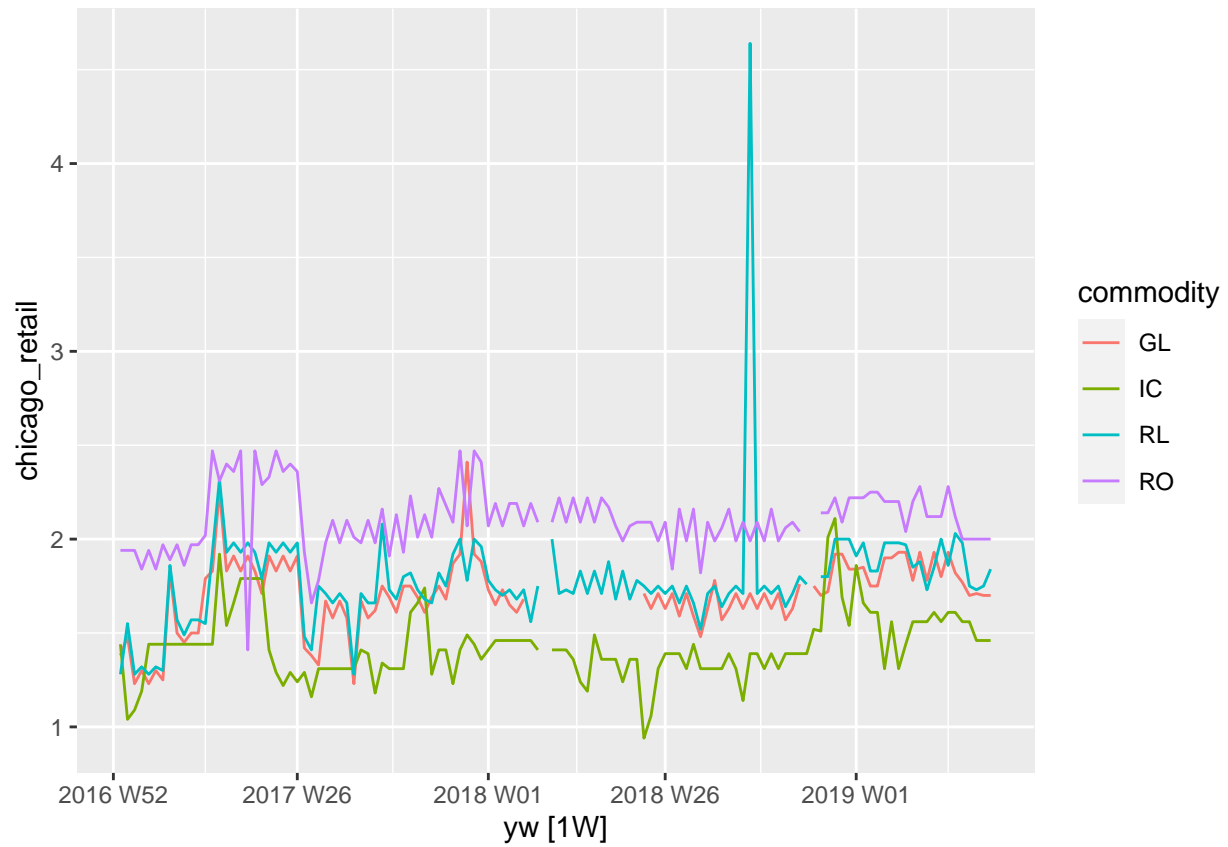
```
unique(all_lettuce$commodity)
```

```
## [1] "GL" "IC" "RL" "RO"
```

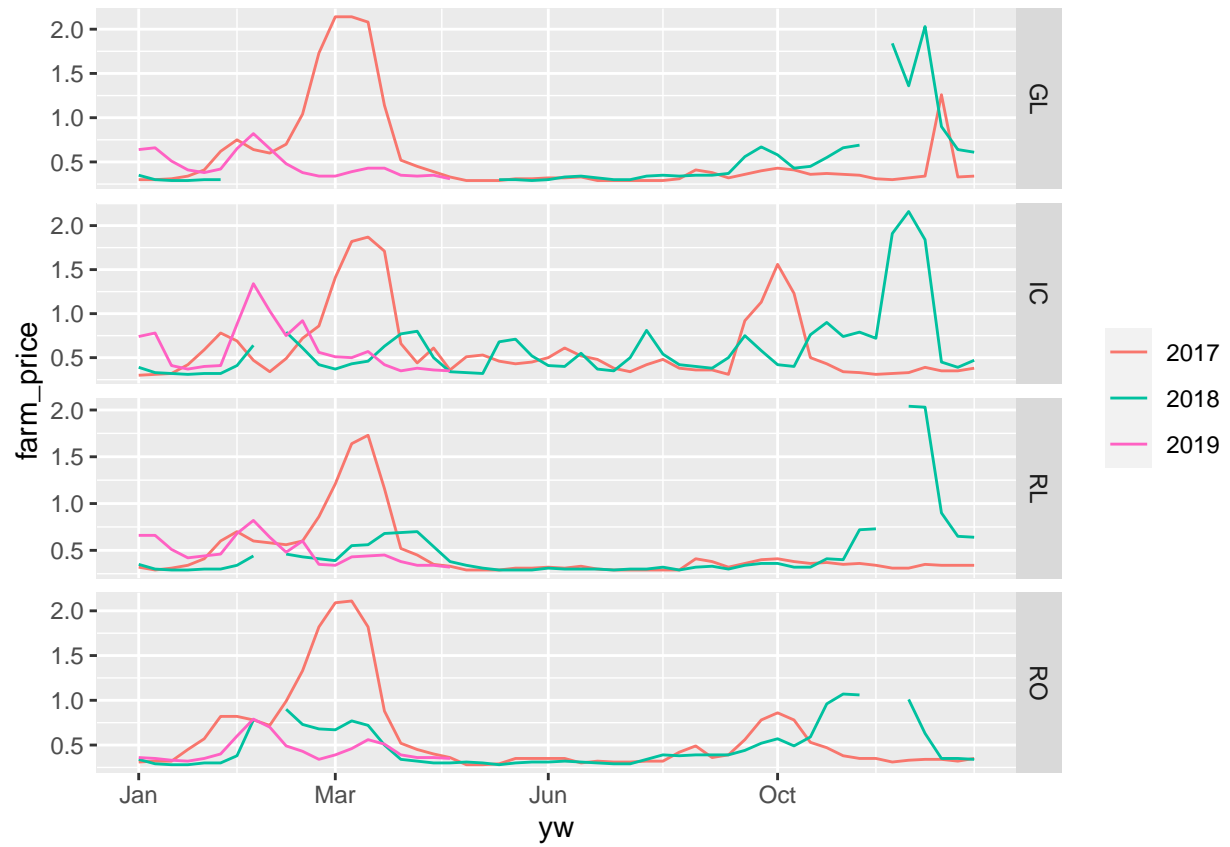
```
#visualization
autoplot(all_lettuce, farm_price)
```



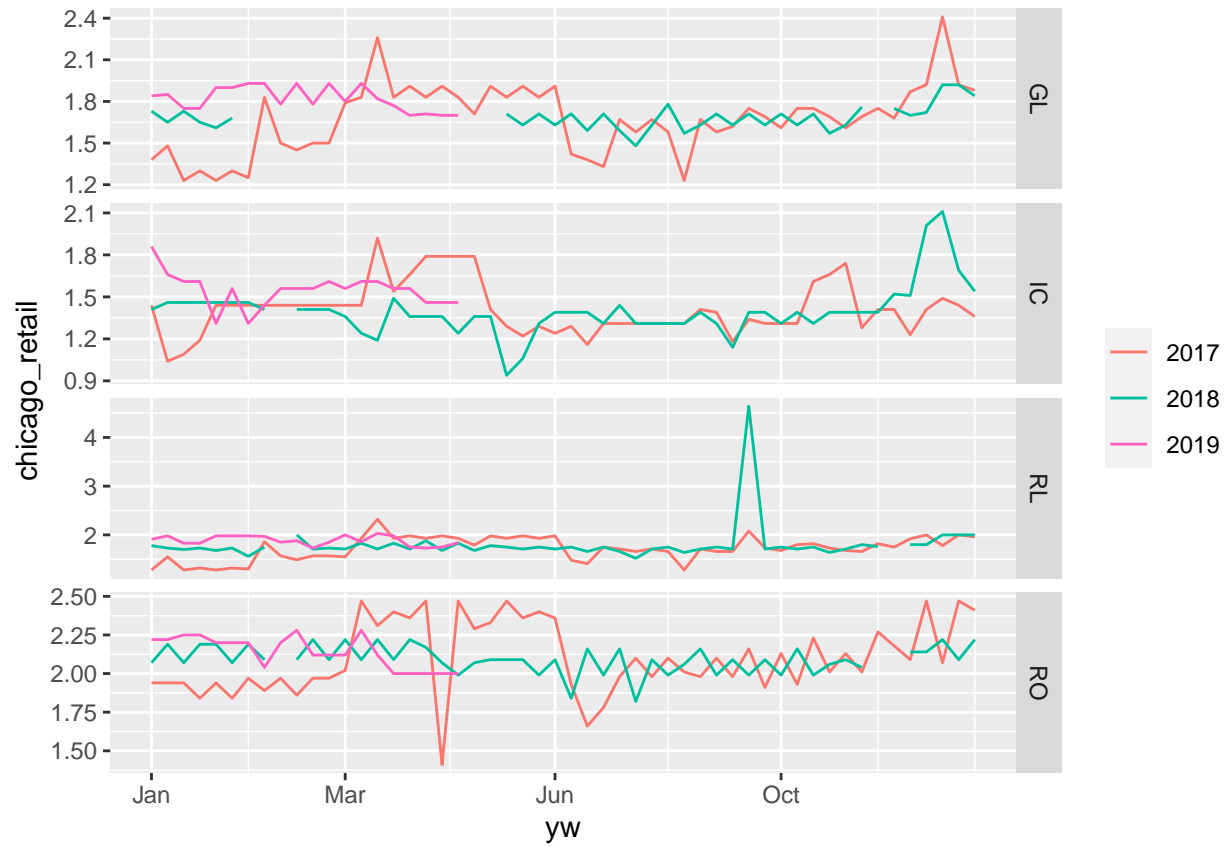
```
autoplot(all_lettuce, chicago_retail)
```



```
gg_season(all_lettuce, farm_price)
```



```
gg_season(all_lettuce, chicago_retail)
```



```
#write to csv
write.csv(all_lettuce, "data/lettuce_wholesale/all_lettuce.csv")
```

```
fill_gaps(.full= TRUE)
```