housing madison

LingLei

2022-10-15

```
# actice list
active_list = read.csv("/Users/linglei/housing_madison/Active_Listing_Count.csv")
median_day = read.csv("/Users/linglei/housing_madison/Median_day.csv")
price_reduce = read.csv("/Users/linglei/housing_madison/Price_reduce.csv")
```

```
# clean up
active_list_2016_2022 = active_list %>%
  rename(date = DATE,
         list = ACTLISCOU31540) %>%
  mutate(month = month(date, label = TRUE),
         year = year(date)) %>%
  relocate(date, year, month)
```

```
summary_list_2016_2022 = active_list_2016_2022 %>%
  group_by(month) %>%
  summarise(mean_list = mean(list),
            max_list = max(list),
            min_list = min(list),
            sum_list = sum(list)) %>%
  mutate(season = case_when(
    month %in% c("Mar", "Apr", "May") ~ "Spring",
    month %in% c("Jun", "Jul", "Aug") ~ "Summer",
    month %in% c("Sep", "Oct", "Nov") ~ "Autumn",
    month %in% c("Dec", "Jan", "Feb") ~ "Winter"
  ))
```

```
plot_2016 = summary_list_2016_2022 %>%
  ggplot(aes(x = month, y = mean_list, fill = season)) +
  geom_col() +
  xlab("Month") +
  ylab("Mean Listing") +
  ylim(0, 1600) +
  ggtitle("Average Listing For Each Month Scaled",
          subtitle = "2016-2022") +
  theme(axis.text.x = element_text(angle = 90))
```

```
active_list_2020_2022 = active_list %>%
  rename(date = DATE,
         list = ACTLISCOU31540) %>%
  mutate(month = month(date, label = TRUE),
         year = year(date)) %>%
```

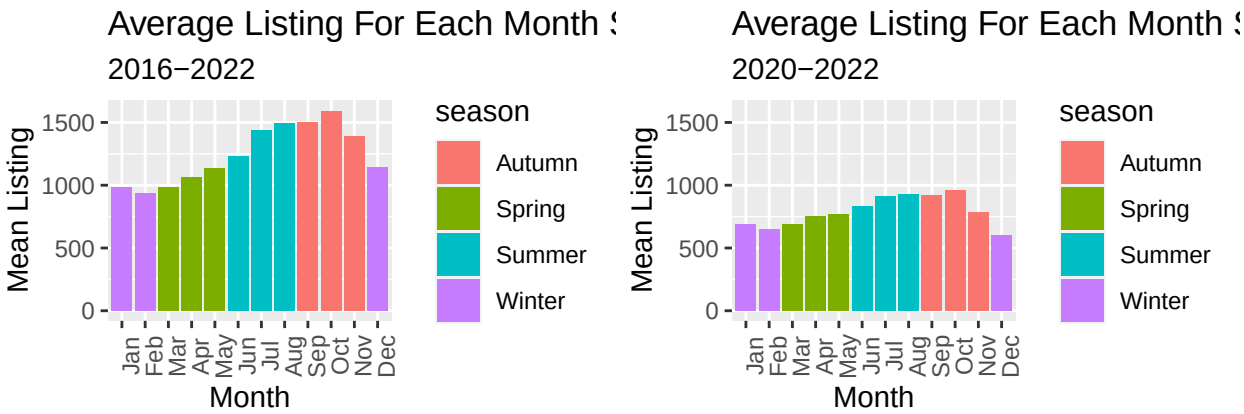
```
  filter(year >= 2020) %>%
  relocate(date, year, month)
```

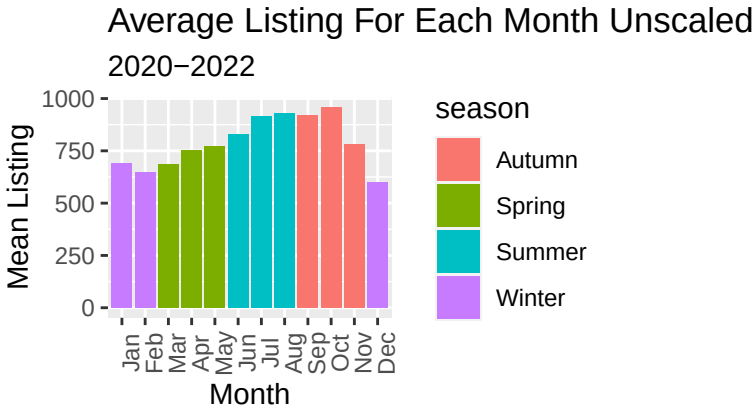
```
summary_list_2020_2022 = active_list_2020_2022 %>%
  group_by(month) %>%
  summarise(mean_list = mean(list),
            max_list = max(list),
            min_list = min(list),
            sum_list = sum(list)) %>%
  mutate(season = case_when(
    month %in% c("Mar", "Apr", "May") ~ "Spring",
    month %in% c("Jun", "Jul", "Aug") ~ "Summer",
    month %in% c("Sep", "Oct", "Nov") ~ "Autumn",
    month %in% c("Dec", "Jan", "Feb") ~ "Winter"
  ))
```

```
plot_2020 = summary_list_2020_2022  %>%
  ggplot(aes(x = month, y = mean_list, fill = season)) +
  geom_col() +
  xlab("Month") +
  ylab("Mean Listing") +
  ylim(0, 1600) +
  ggtitle("Average Listing For Each Month Scaled",
          subtitle = "2020-2022") +
  theme(axis.text.x = element_text(angle = 90))
```

```
plot_2020_unscaled = summary_list_2020_2022  %>%
  ggplot(aes(x = month, y = mean_list, fill = season)) +
  geom_col() +
  xlab("Month") +
  ylab("Mean Listing") +
  ggtitle("Average Listing For Each Month Unscaled",
          subtitle = "2020-2022") +
  theme(axis.text.x = element_text(angle = 90))
```

```
plot_grid(plot_2016, plot_2020, plot_2020_unscaled)
```



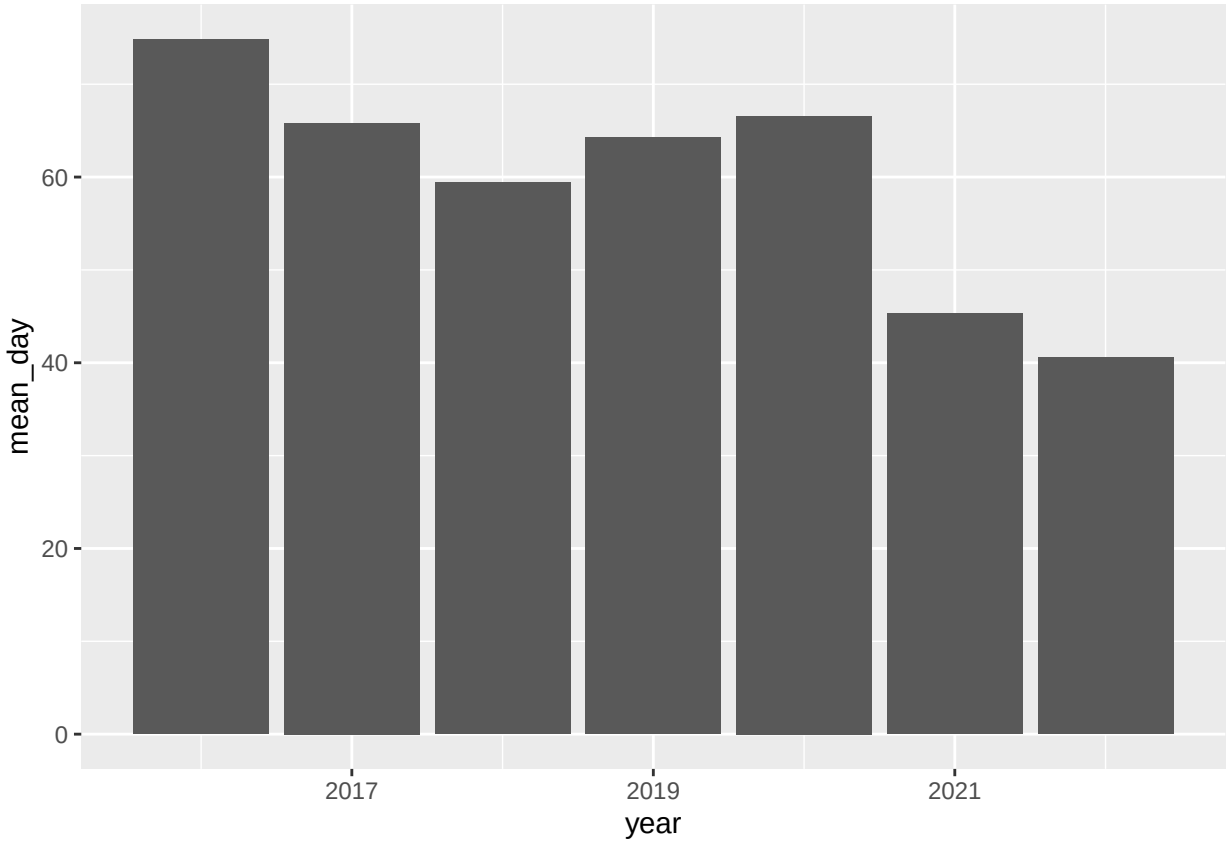


From the scaled and unscaled graphs, Madison in Wisconsin tends to keep more active listing in summer and fall while it keeps less active listing in other seasons.

```
median_day =median_day %>%
  rename(medday = MEDDAYONMAR31540,
         date = DATE) %>%
  mutate(year = year(date),
         month = month(date, label = TRUE))
```

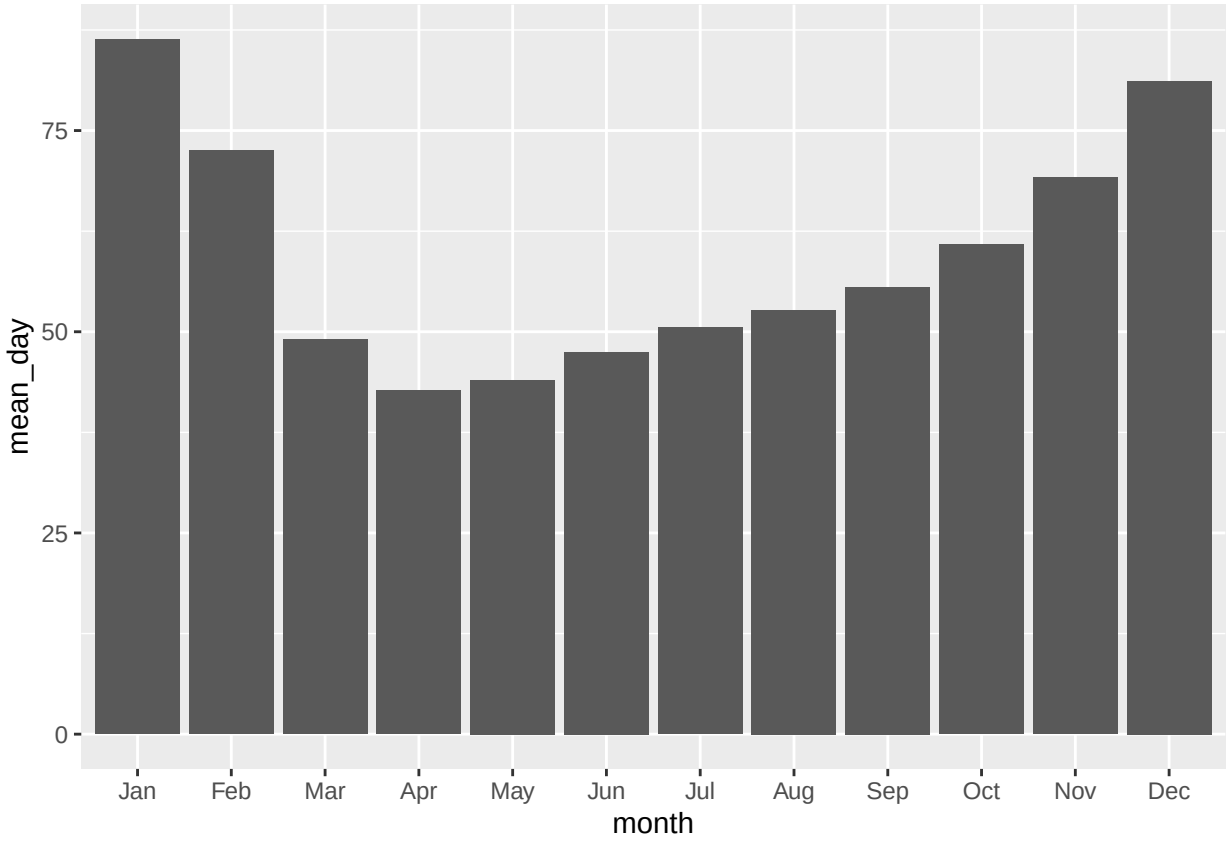
```
med_year_summary = median_day %>%
  group_by(year) %>%
  summarise(max_day = max(medday),
            min_day = min(medday),
            mean_day = mean(medday))
```

```
med_year_summary %>%
  ggplot() +
  geom_col(aes(x = year, y = mean_day))
```




```
med_month_summary = median_day %>%
  group_by(month) %>%
  summarise(max_day = max(medday),
            min_day = min(medday),
            mean_day = mean(medday))
```

```
med_month_summary %>%
  ggplot() +
  geom_col(aes(x = month, y = mean_day))
```



If the market is competitive?

$$V_{adjusted-sell} = \frac{V_{sale}}{V_{list}}$$

$$V_{sale} = \frac{1}{D_m}, D_m : median - day$$

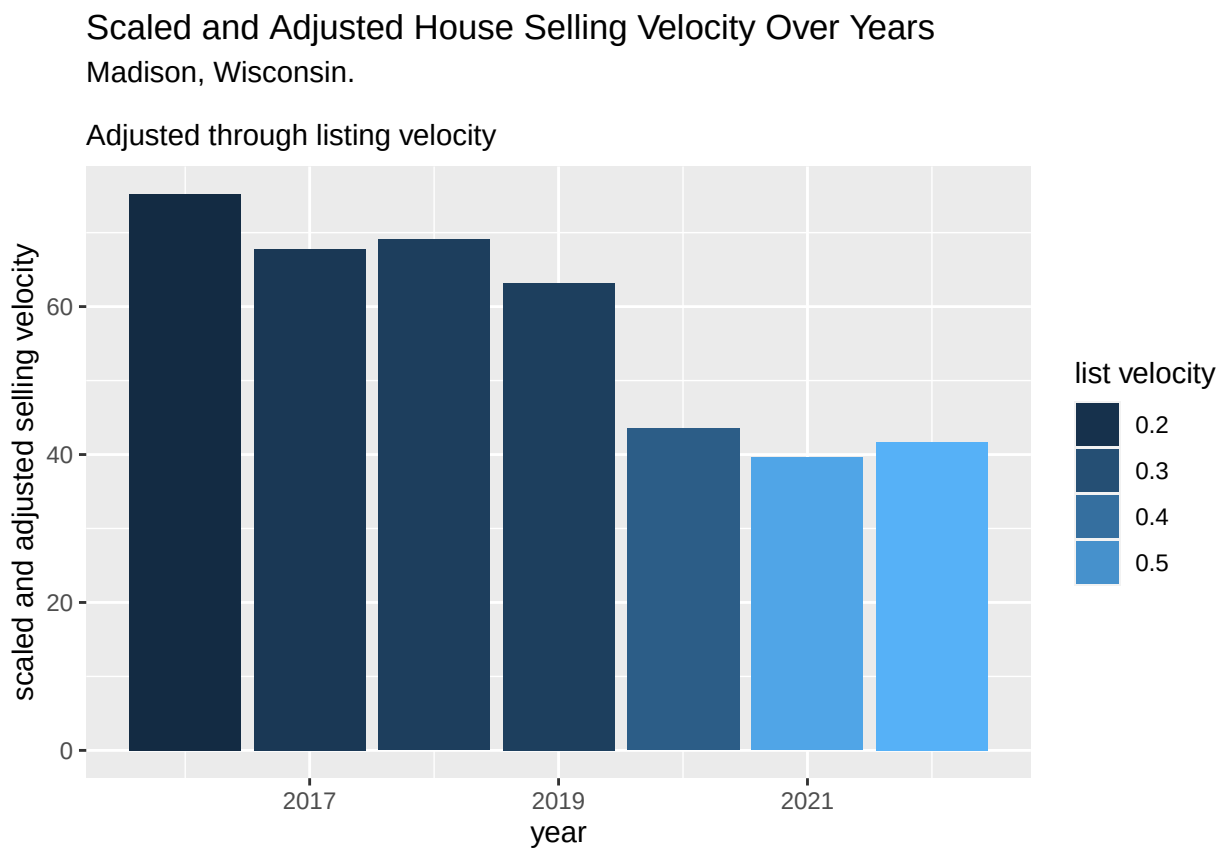
$$V_{list} = \frac{L}{365}, L : listed - houses$$

```
# combine active listing and median day tables
list_median_df = inner_join(median_day, active_list_2016_2022,
by = c("date", "year", "month"))
```

```
new_metrics_list_median = list_median_df %>%
  group_by(year) %>%
  summarise(median_days = mean(medday),
            list_year = mean(list)) %>%
  mutate(v_sale = 1/median_days,
         v_list = 365/list_year,
         v_adjust_sale = v_sale/v_list,
         scaled_v = v_adjust_sale*1000 )
```

```
new_metrics_list_median %>%
  ggplot() +
  geom_col(aes(x = year, y = scaled_v, fill = v_list)) +
```

```
  xlab("year") +
  ylab("scaled and adjusted selling velocity") +
  ggtitle("Scaled and Adjusted House Selling Velocity Over Years", subtitle = "Madison, Wisconsin. \n
Adjusted through listing velocity") +
  guides(fill = guide_legend("list velocity"))
```



Price Reduced

```
price_reduce = price_reduce %>%
  rename(reduce_price = PRIREDCOU31540,
         date = DATE) %>%
  mutate(year = year(date),
         month = month(date)) %>%
  drop_na()
```

```
price_reduce %>%
  group_by(year) %>%
  summarise(mean_reduce = mean(reduce_price))
```

```
## # A tibble: 6 x 2
##    year mean_reduce
##   <dbl>       <dbl>
```

```
## 1  2017        592.
## 2  2018        697.
## 3  2019        454.
## 4  2020        267.
## 5  2021        176.
## 6  2022        176.
```