

Homework 2

Ethan Chang - ehc586

This homework is due on Jan. 24, 2023 at 11:00pm. Please submit as a pdf file on Canvas.

Problem 1: (3 pts) We will work with the dataset `txhouse` that has been derived from the `txhousing` dataset provided by **ggplot2**. See here for details of the original dataset:

<https://ggplot2.tidyverse.org/reference/txhousing.html> (<https://ggplot2.tidyverse.org/reference/txhousing.html>).

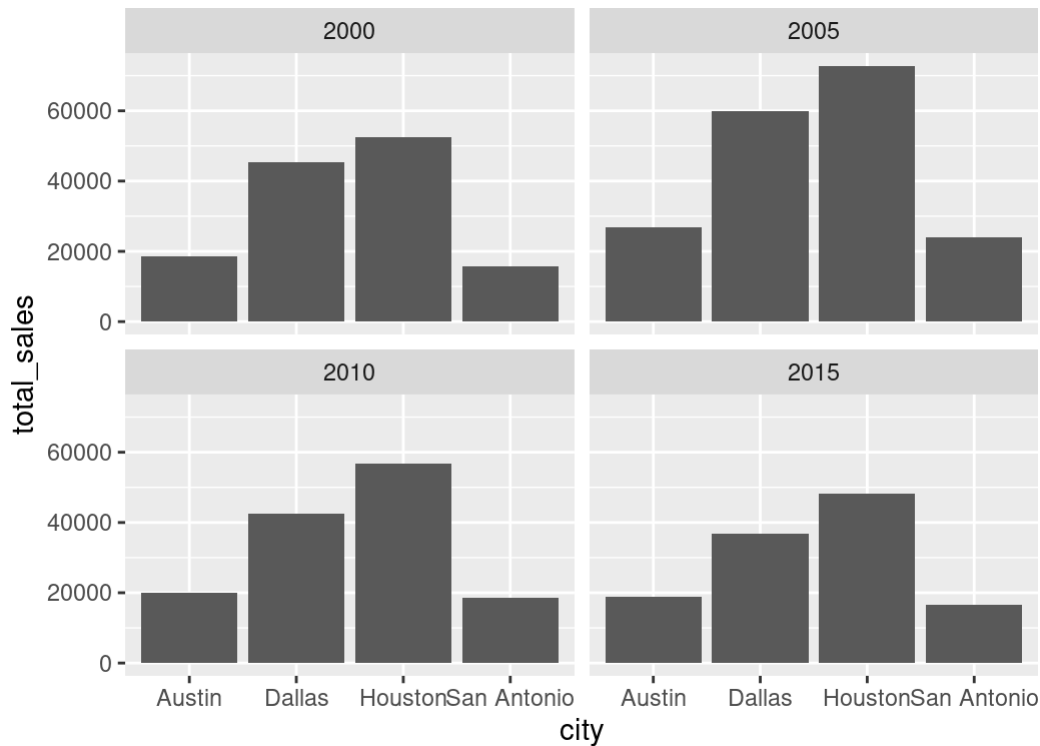
`txhouse` contains three columns: `city` (containing four Texas cities), `year` (containing four years between 2000 and 2015) and `total_sales` indicating the total number of sales for the specified year and city.

`txhouse`

```
## # A tibble: 16 × 3
## # Groups:   city [4]
##   city      year total_sales
##   <chr>    <int>      <dbl>
## 1 Austin    2000      18621
## 2 Austin    2005      26905
## 3 Austin    2010      19872
## 4 Austin    2015      18878
## 5 Dallas    2000      45446
## 6 Dallas    2005      59980
## 7 Dallas    2010      42383
## 8 Dallas    2015      36735
## 9 Houston   2000      52459
## 10 Houston  2005      72800
## 11 Houston  2010      56807
## 12 Houston  2015      48109
## 13 San Antonio 2000      15590
## 14 San Antonio 2005      24034
## 15 San Antonio 2010      18449
## 16 San Antonio 2015      16455
```

Use `ggplot` to make a bar plot of the total housing sales (column `total_sales`) for each `city` and show one panel per `year`. You do not have to worry about the order of the bars. Hint: Use `facet_wrap()`. See slides from Class 2.

```
ggplot(txhouse, aes(city, total_sales)) +
  facet_wrap(vars(year)) +
  geom_col()
```



Problem 2: (3 pts) Use ggplot to make a bar plot of the total housing sales (column `total_sales`) for each year. Color the bar borders with color `"gray20"` and assign a fill color based on the `city` column.

```
ggplot(txhouse, aes(year, total_sales, fill = city)) +  
  geom_col(color = "gray20")
```



Problem 3: (4 pts) Modify the plot from Problem 2 by placing the bars for each city side-by-side rather than stacked. Next, reorder the bars for each year by `total_sales` in descending order. See slides from Class 4.

```
ggplot(txhouse, aes(fct_reorder(as.factor(year), -total_sales), total_sales, fill = city)) +  
  geom_col(position = "dodge", color = "gray20") +  
  ylab("Total Sales") +  
  xlab("Year")
```

