HW 1, CS 625, Spring 2023

Suprabha Pandey

Jan 18, 2023

## Git, GitHub

1. *What is your GitHub username?* suprabhapandey
2. *What is the URL of your remote GitHub repo (created through Mr. Kennedy’s exercises)?* <https://github.com/suprabhapandey/git-workshop>

## R

The command below will load the tidyverse package. If you have installed R, RStudio, and the tidyverse package, it should display a list of loaded packages and their versions.

library(tidyverse)

## ── Attaching packages ─────────────────────────────────────── tidyverse 1.3.2 ──  
## ✔ ggplot2 3.4.0 ✔ purrr 1.0.1   
## ✔ tibble 3.1.8 ✔ dplyr 1.0.10  
## ✔ tidyr 1.2.1 ✔ stringr 1.5.0   
## ✔ readr 2.1.3 ✔ forcats 0.5.2   
## ── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
## ✖ dplyr::filter() masks stats::filter()  
## ✖ dplyr::lag() masks stats::lag()

## R Markdown

1. *Create a bulleted list with at least 3 items*

* Bulleted list item 1
* Bulleted list item 2
* Bulleted list item 3

1. *Write a single paragraph that demonstrates the use of italics, bold, bold italics, code, and includes a link. The paragraph does not have to make sense.* **Markdown** is *designed* to be easy to **read** and easy to **write** . It is also very *easy* to learn. The guide below shows how to use Pandoc’s Markdown, a slightly extended version2 of Markdown that R Markdown2 understands
2. *Create a level 3 heading* ## 3rd Level Header

## R

#### Data Visualization Exercises

1. (Q2) *How many rows are in mpg? How many columns?*

224 rows and 11 columns

1. (Q4) *Make a scatterplot of hwy vs cyl.* ggplot(data = mpg) + geom\_point(mapping = aes(x = displ, y = hwy))

#### Workflow: basics Exercises

1. (Q2) *Tweak each of the following R commands so that they run correctly (library(tidyverse) is correct):*

library(tidyverse)  
ggplot(data = mpg) +   
 geom\_point(mapping = aes(x = displ, y = hwy))  
  
filter(mpg, cyl == 8)  
  
filter(diamond, carat > 3)

There are some spelling mistakes. “dota” should be replaced with “data”, “fliter” should be replaced with “filter”. There should be “==” instead of “=”.

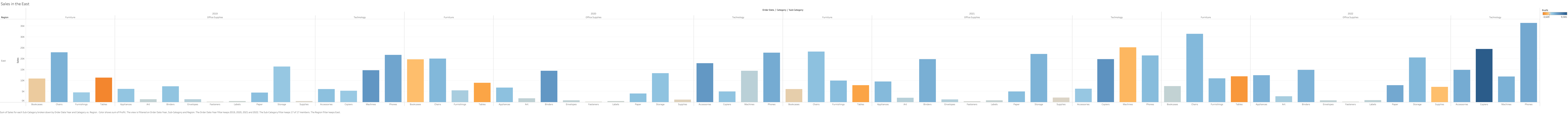
## Google Colab

1. *What are the URLs of your Google Colab notebooks (both Python and R)?* Python <https://colab.research.google.com/drive/1zJfrCMO-rlVYag6xFhUb1Deysaa04fNd?usp=sharing>

R <https://colab.research.google.com/drive/1fyn0e5aEAmUgzn8Utr_WonMWCOQHKm91?usp=sharing> ## Tableau

*Insert your the image of your final bar chart here*

knitr::include\_graphics("/Users/suprabhatripathi/Documents/GitHub/git-workshop/Sales\_in\_the\_East.png")



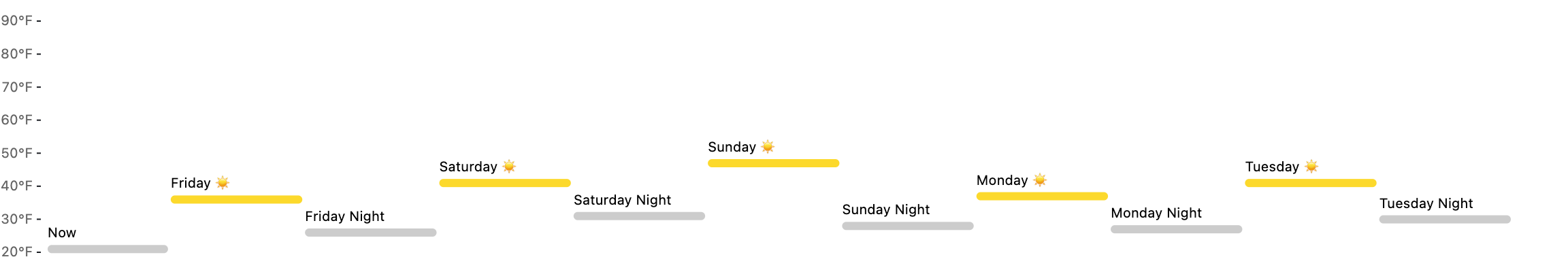
1. *What conclusions can you draw from the chart?* My conculsion as per Sales in east region for year 2019 was highest sales among furnitures was chairs, storage in Office supplies, phones in technology. On the other hand, in the next year 2020, chairs was the most sold items in the furniture cateogory, binders in the office supplies category, phones in the technology. Likewise, in the year 2021, chairs, storage, machines were in the highest sales trend. Similarly, for the year 2022, chairs, storage, phones were in the highest sale category for their respective group. Here we can observe that in technology cateogry phones were most sold item except year 2021.

## Observable and Vega-Lite

### A Taste of Observable

1. *In the “New York City weather forecast” section, try replacing Forecast: detailedForecast with Forecast: shortForecast. Then press the blue play button or use Shift-Return to run your change. What happens?* The description of Forecast changed to the short description like a one or two word.
2. *Under the scatterplot of temperature vs. name, try replacing markCircle() with markSquare(). Then press the blue play button or use Shift-Return to run your change. What happens? How about markPoint()?* markSquare() changed all the circle points with square and markPoint() changed all the Squares with the hollow circles/rings.
3. *Under “Pick a location, see the weather forecast”, pick a location on the map. Where was the point you picked near?* I picked Littlerock, CA.
4. *The last visualization on this page is a “fancy” weather chart embedded from another notebook. Click on the 3 dots next to that chart and choose ‘Download PNG’. Insert the PNG into your report.*

knitr::include\_graphics("/Users/suprabhatripathi/Documents/GitHub/git-workshop/Weather.png")



### Charting with Vega-Lite

markCircle()

1. *Pass an option of { size: 200 } to markCircle().* Size of circle will be increased
2. *Try markSquare instead of markCircle.* Circle will be chnages to square
3. *Try markPoint({ shape: 'diamond' }).* Points will be changes to hollow diamonds. vl.x().fieldQ("Horsepower"), …
4. *Change Horsepower to Acceleration* Graph changed to Miles per gallon /Accerelation. Scatter plot is mostly on the right side.
5. *Swap what fields are displayed on the x- and y-axis* Now Scatter Plot on at the center of the graph. Also, Acceleration is on Y axis and Miles per Gallon is on X Axis.

vl.tooltip().fieldN("Name")

1. *Change Name to Origin.* After Changing tooltip when we mouse hover any points it shows origin name which is country in our case. Another example, count()
2. *Remove the vl.y().fieldN("Origin") line.* After Removing, data of all the origins/country will be shown in one line.
3. *Replace count() with average("Miles\_per\_Gallon").* It is showing maximum 23 something average of all the origins. ## References

*Every report must list the references that you consulted while completing the assignment. If you consulted a webpage, you must include the URL.*

* Insert Reference 1, <https://www.example.com>
* Insert Reference 2, <https://www.example.com/reallyreallyreally-extra-long-URI/>