Posted on GitHub at https://github.com/1fastgranada/CSC302_HW5

1. Load house_prices.rda in R and perform the followings:
    a. Use facet_wrap to draw line plots to represent the trend for house_price_index over years for each state. Please have three ticks on the x axis which are for years 1980, 2000, and 2020. Since you have limited space your labels can be ’80, ’00, and ’20 for these ticks.
    b. Use gather() function from tidyr to reshape your data in a way that it will have measure and value columns for house_price_index and unemploy_perc, respectively, in two columns. You should leave out remaining attributes by using ”-” in front them inside the function. Then, direct the output to a new data frame called house_reshaped Please refer to examples we did in class.
    c. Use house_reshaped to replicate the graph in 1a with two lines where each represents house_price_index and unemploy_perc over the years for each state. Once you have the chart, please comment on it as to whether it’s a good graph to present those two pieces of information.
    d. When you achieve the goals above, please create a R markdown file and create an HTML page which has your codes and charts including your comments.

Please see R Markdown file “HW5_ARoss.Rmd” and resultant “HW5_ARoss.html” on GitHub linked in the header.

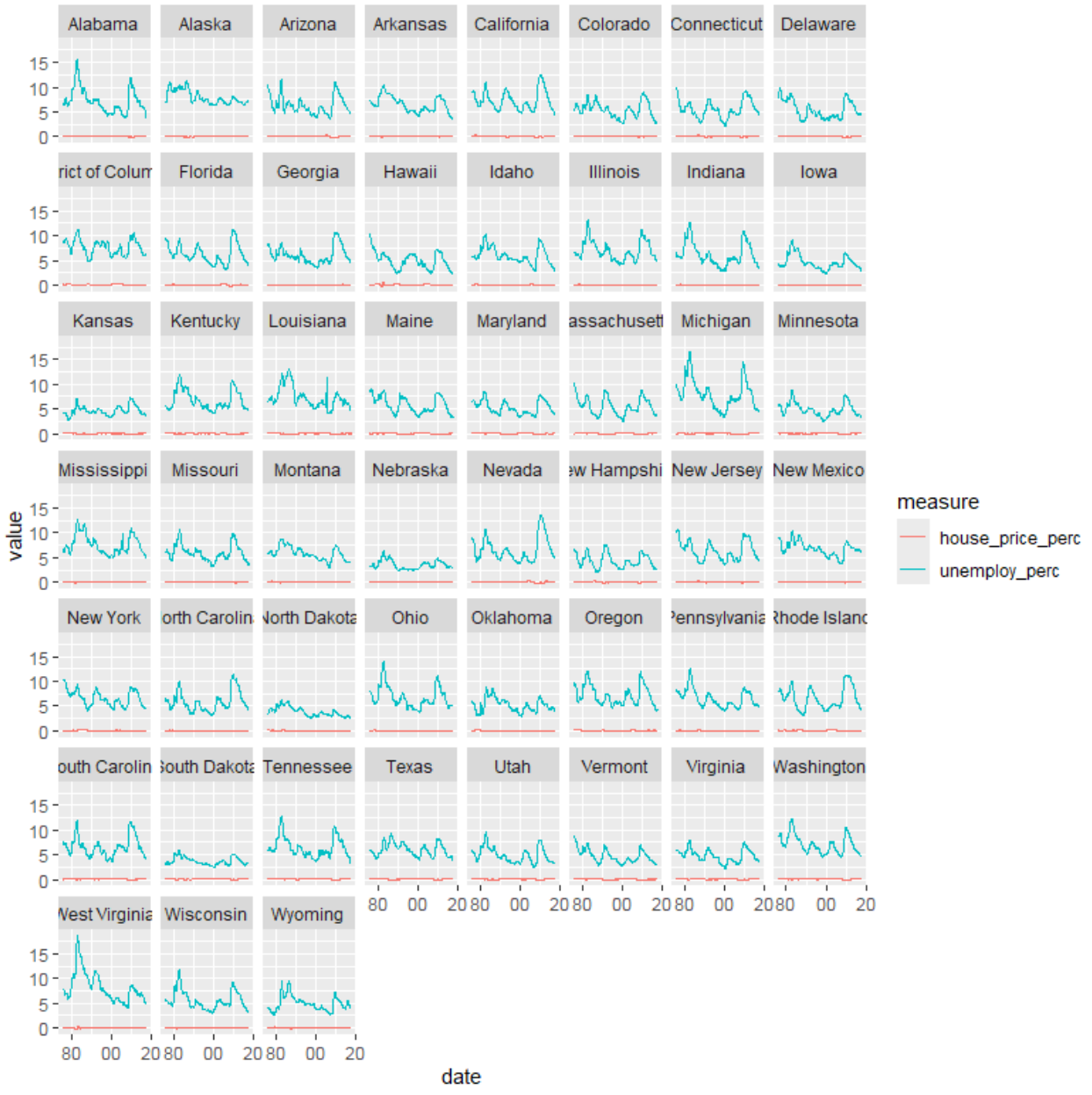
2. Load house_prices.rda in R, perform the following operation, and explain how it changes the original data set. Then, create facet_wrap plot which can visualize your values for each state and colorize the lines based on the measurement, i.e., unemploy_perc and house_perc. Please consider using scale_x_continuous(breaks = ymd(c("1980-01-01", "2000-01-01", "2020-01-01")), labels = c("80", "00", "20")) to have a few breaks on your x-axis for a more clear view. Don’t forget to include library(lubridate) in your script.

```
house_prices%>%gather(key=measure, value=value, -c(house_price_index, date, state))->house_reshaped
```

This provides a very similar result as in question 1b above, replacing “house_price_index” with “house_price_perc”. The chart now allows us to visualize “unemploy_perc” better, but “house_price_percent” is so small, it’s in the noise at the bottom.

Posted on GitHub at https://github.com/1fastgranada/CSC302_HW5

```
# (2) Load house_prices.rda in R, perform the following operation, and
#     explain how it changes the original data set. Then, create facet_wrap plot
#     which can visualize your values for each state and colorize the lines
#     based on the measurement, i.e., unemploy_perc and house_perc.

house_prices%>%gather(key=measure, value=value, -c(house_price_index, date, state))->house_reshaped

head(house_reshaped)

# This provides a very similar result as in question 1b above, replacing house_price_index with house_price_perc

p3 <- ggplot(house_reshaped, aes(x = date, y = value, color=measure)) +
  geom_line() +
  facet_wrap(~state) +
  scale_x_continuous(breaks = ymd(c("1980-01-01", "2000-01-01", "2020-01-01")), labels = c("80", "00", "20"))
labs(title = "House Price Percent Trend by State",
     x = "Year",
     y = "House Price Percent")

print(p3)
```



Posted on GitHub at https://github.com/1fastgranada/CSC302_HW5

3. Please see the Python notebook here and make a copy to yourself. Then follow the instructions in the notebook to perform your visualization. Please submit your notebook as a response to this question.

Please see Python notebook "HW5_ARoss.ipynb" linked here or on GitHub linked in the header.