

Assignment 2

Assignment 2

Task A. Plotting monthly averages of temperature, volume, price and wind production.

```
# Create a data frame of monthly averages of the variables volume, price,
# temperature and wind_production in long format.
df_avg_long <-
  df %>%

  # Select the specified variables
  select(date, volume, temperature, price, wind_production) %>%

  # Calculate monthly averages for each variable
  mutate(year_month = floor_date(date, "month")) %>%
  group_by(year_month) %>%
  summarise(
    avg_vol = mean(volume, na.rm = TRUE),
    avg_temp = mean(temperature, na.rm = TRUE),
    avg_price = mean(price, na.rm = TRUE),
    avg_wind = mean(wind_production, na.rm = TRUE)
  ) %>%

  # Scale the values of volume and wind to make the plot readable
  mutate(
    avg_vol = avg_vol / 10^5,
    avg_wind = avg_wind / 10^3
  ) %>%

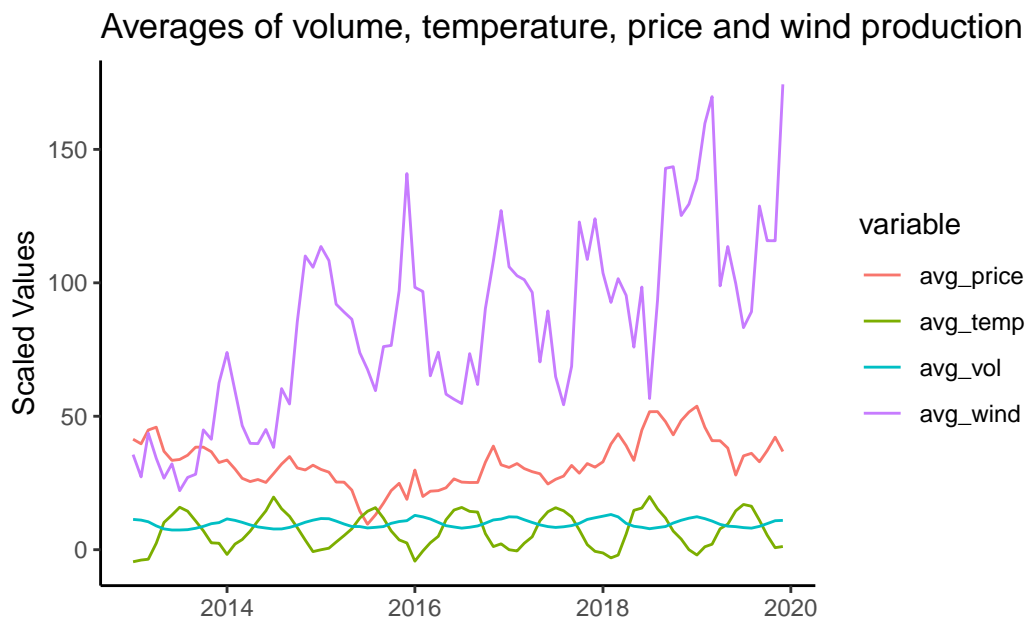
  # Pivot to long format
```

```

pivot_longer(
  cols = -year_month,
  names_to = "variable",
  values_to = "scaled values"
)

# Plot the specified variables
df_avg_long %>%
  ggplot(
    aes(
      x = year_month,
      y = `scaled values`,
      col = variable
    )
  ) +
  geom_line() +
  labs(
    x = "",
    y = "Scaled Values",
    title = "Averages of volume, temperature, price and wind production"
  ) +
  theme_classic()

```



Task B. Why will a OLS regression on quantity not provide an estimate of demand?

An OLS regression of quantity on price alone, even with controls, is not fitted to estimate demand due to many reasons:

1. Endogeneity exists as price and quantity in markets are simultaneously determined
2. Omitted variables like for example consumer income, advertising and seasonality. These can affect and lead to omitted variable bias.
3. Measurement errors in price and quantity can contribute to bias for the estimates for the parameters
4. OLS (or Gauss-Markow theorem) assumes a linear relationship. This may not hold if the demand curves are nonlinear.
5. Heteroskedasticity which is variability in the error term given any explanatory variable. In this case the variability in quantity varies across price levels which in return can affect standard errors.
6. The demand's dynamic behavior and seasonality may not be sufficiently captured. To estimate demand accurately, more advanced econometric methods are often required. These advanced methods usually contain instrumental variable regression and structural equation modelling. Controlling for relevant factors and improve our data quality can also make the estimation precision better.

Task C. Considering choice of instrument for price when estimating demand

When choosing a instrument for price when estimating demand, it's important to consider the validity of the instrument in terms of the requirements for a valid instrument. These requirements can be relevance, exogeneity, and exclusion restrictions:

1. Why is temperature not a valid instrument for price when estimating demand? The key requirements for exogeneity may not be satisfied if one were to use temperature as an instrument for price when estimating demand. (Exogeneity refers to the condition in which an independent variable is unrelated to the error term, indicating that it is not influenced by unobserved factors in the model.) Temperature is usually affected by factors that also may affect demand, like seasonality and consumer behavior. The instrument would not be valid due to endogeneity concerns if temperature is correlated with unobservable demand shocks.

2. Why can magazine levels (or deviations) and wind power production potentially be good instruments for price when estimating demand? Magazine levels and wind power production could potentially be sufficient instruments for price as they can fulfil the requirements for a valid instrument. They should be relevant in other terms they should be affecting the price. In addition, they should be exogenous, meaning that they should not be correlated with unobservable factors. But establishing their exogeneity requires thoughtful consideration and testing.
3. Why is it necessary to control for seasonality (say, calendar month) and temperature? It is necessary to control for seasonality and temperature to avoid omitted variable bias. (Omitted variable bias appear if a relevant variable that affects the dependent variable is excluded in a regression model. This leads to biased and estimates of the coefficients of the variables that are included.) Seasonality captures variation in the systematic demand throughout the year unrelated to price. Meanwhile temperature affects demand independently. When we control for these factors, we ensure precise demand estimates.
4. How could controlling for weekday and year be useful? Incorporating controls for weekday and year is beneficial to justify for demand patterns variation. Weekdays and weekends usually give different demand profiles given the consumers consumption. Meanwhile years would capture trends that are long-term and macroeconomic affects on demand. These controls strengthen the accuracy of demand estimates, while taking into account sources of variation unrelated to price.