Week 1: R Refresher

Import, Plot, Save and Summarize Data

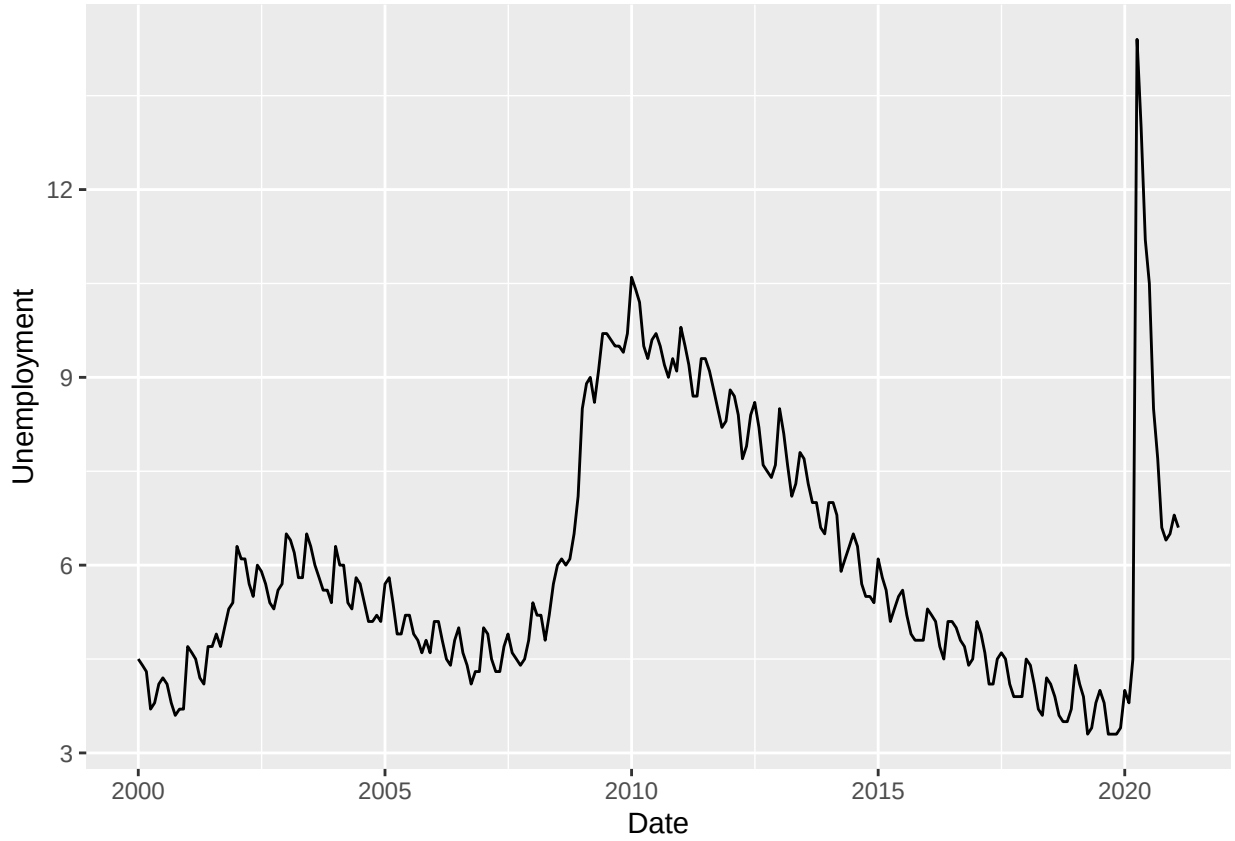
Full Dataframe:

```
##     Series.ID Year Period    Label Value
## 1 LNU04000000 1948    M01 1948 Jan   4.0
## 2 LNU04000000 1948    M02 1948 Feb   4.7
## 3 LNU04000000 1948    M03 1948 Mar   4.5
## 4 LNU04000000 1948    M04 1948 Apr   4.0
## 5 LNU04000000 1948    M05 1948 May   3.4
## 6 LNU04000000 1948    M06 1948 Jun   3.9
```

Cleaned Dataframe:

```
##           Date Unemployment
## 625 2000-01-01          4.5
## 626 2000-02-01          4.4
## 627 2000-03-01          4.3
## 628 2000-04-01          3.7
## 629 2000-05-01          3.8
## 630 2000-06-01          4.1
```

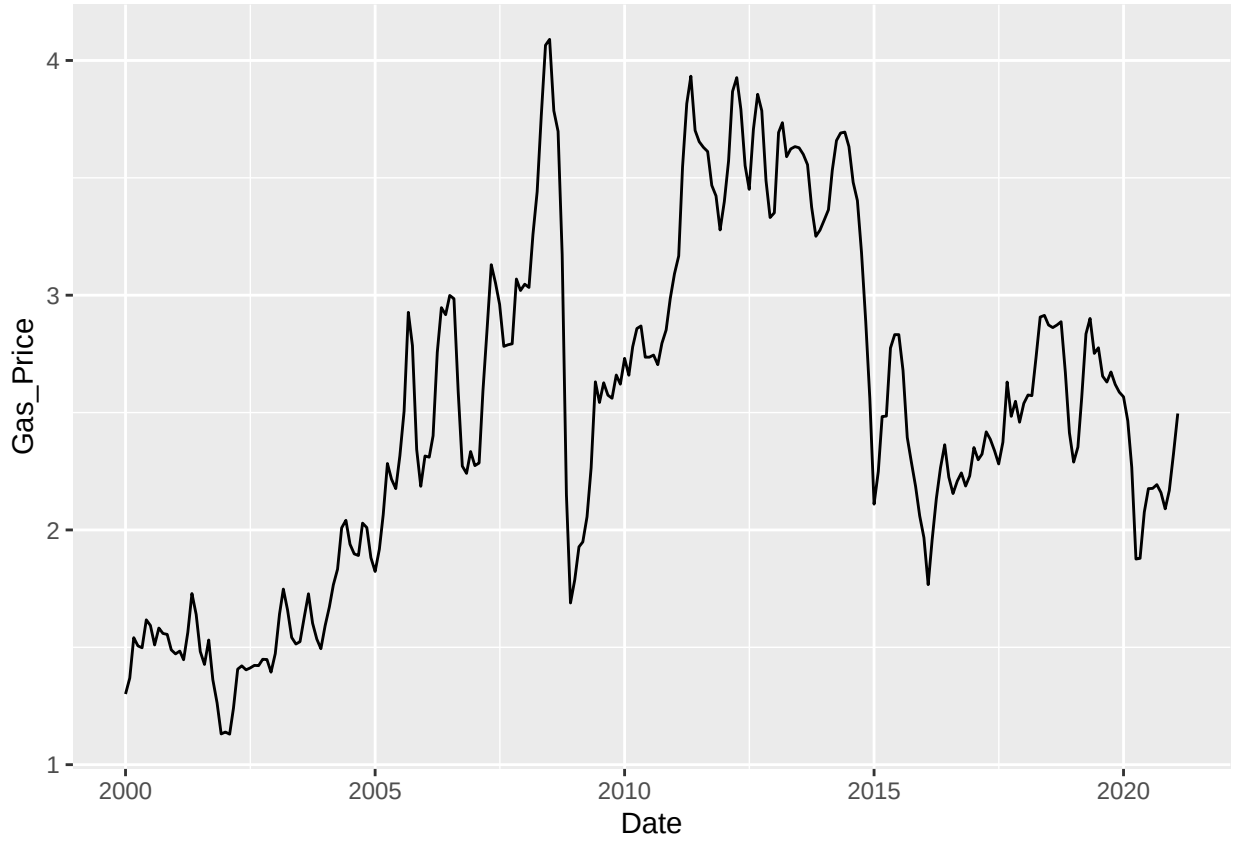
Graph of unemployment:



Cleaned Gas Price Dataframe:

```
##           Date Gas_Price
## 289 2000-01-01     1.301
## 290 2000-02-01     1.369
## 291 2000-03-01     1.541
## 292 2000-04-01     1.506
## 293 2000-05-01     1.498
## 294 2000-06-01     1.617
```

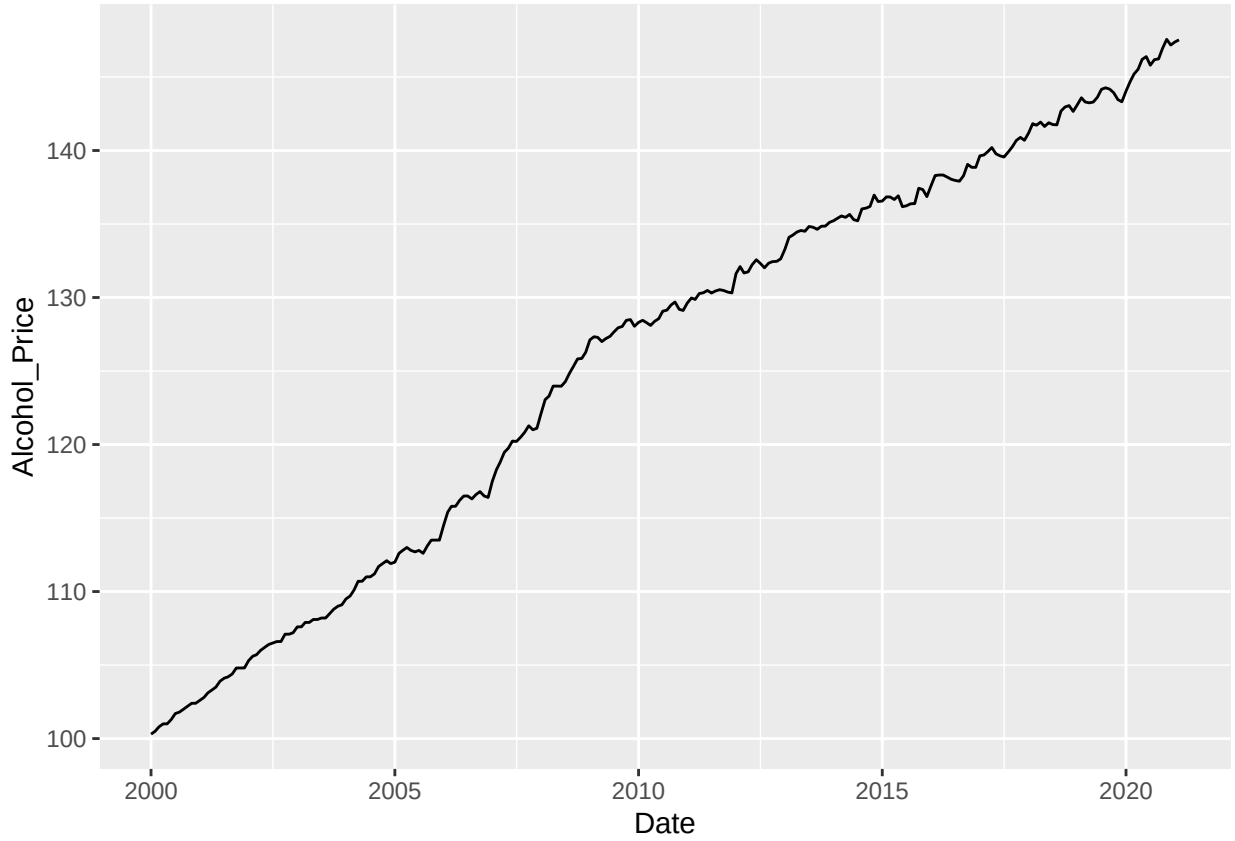
Graph of gas prices:



Cleaned Alcohol Prices Dataframe:

```
##         Date Alcohol_Price
## 2 2000-01-01         100.3
## 3 2000-02-01         100.5
## 4 2000-03-01         100.8
## 5 2000-04-01         101.0
## 6 2000-05-01         101.0
## 7 2000-06-01         101.3
```

Alcohol Price Graph:



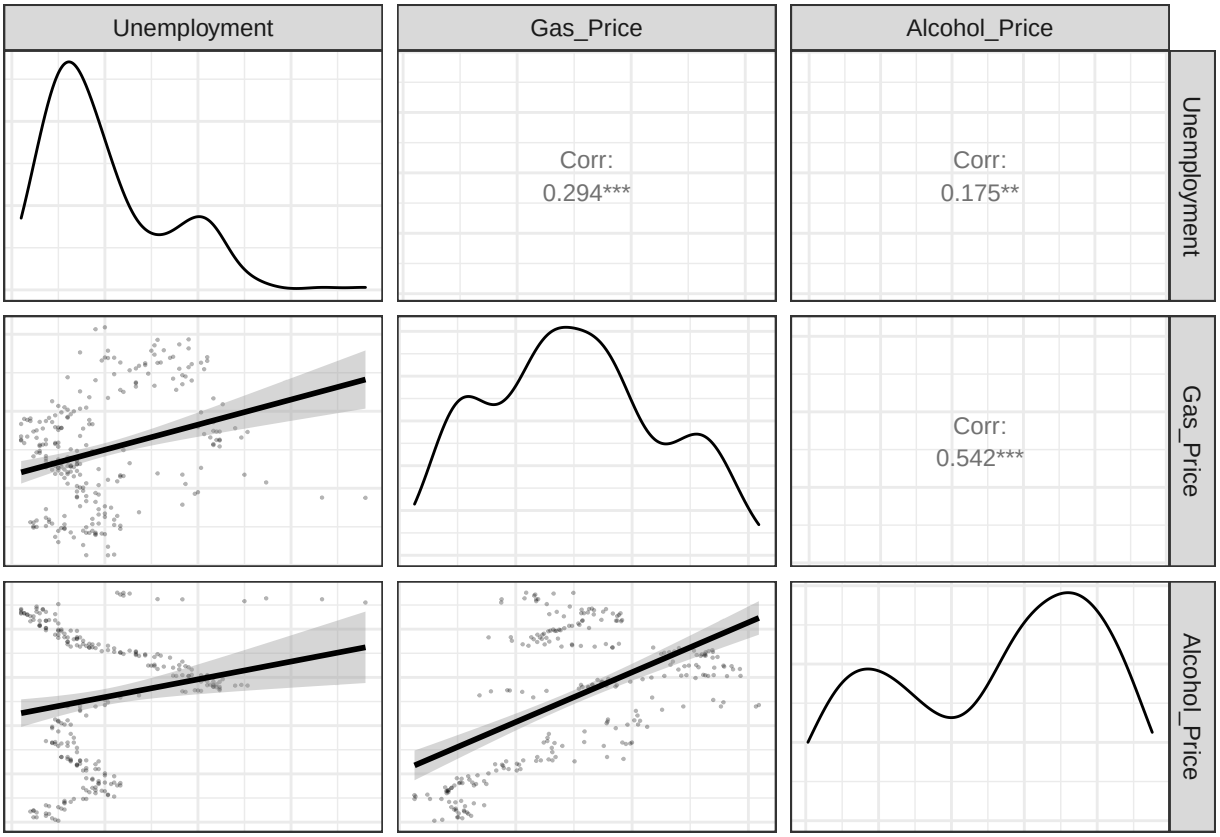
Combined Dataframe:

```
##         Date Unemployment Gas_Price Alcohol_Price
## 1 2000-01-01          4.5     1.301         100.3
## 2 2000-02-01          4.4     1.369         100.5
## 3 2000-03-01          4.3     1.541         100.8
## 4 2000-04-01          3.7     1.506         101.0
## 5 2000-05-01          3.8     1.498         101.0
## 6 2000-06-01          4.1     1.617         101.3
```

Bivariate Relations:

```
## Warning: package 'GGally' was built under R version 4.0.4
```

```
## Registered S3 method overwritten by 'GGally':
##   method from
##   +.gg   ggplot2
```



Summary of Unemployment:

```
summary(combined$Unemployment)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   3.300   4.500   5.400   5.992   7.000  14.400
```

```
str(combined$Unemployment)
```

```
##  num [1:254] 4.5 4.4 4.3 3.7 3.8 4.1 4.2 4.1 3.8 3.6 ...
```

```
class(combined$Unemployment)
```

```
## [1] "numeric"
```

Summary of Gas Prices:

```
summary(combined$Gas_Price)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   1.130   1.942   2.483   2.499   2.958   4.090
```

```r
str(combined$Gas_Price)
```

```
##  num [1:254] 1.3 1.37 1.54 1.51 1.5 ...
```

```r
class(combined$Gas_Price)
```

```
## [1] "numeric"
```

Summary of Alcohol Prices:

```r
summary(combined$Alcohol_Price)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   100.3   112.7   129.1   125.9   137.2   147.6
```

```r
str(combined$Alcohol_Price)
```

```
##  num [1:254] 100 100 101 101 101 ...
```

```r
class(combined$Alcohol_Price)
```

```
## [1] "numeric"
```

Discussion:

Unfortunately when looking into this, I wanted to find if I could look into the health of the economy along with the demand/consumption of an elastic and inelastic good. Sadly, I ended up with unemployment (as a measure of the economy), and the prices of good, not their demand. But there is still some interesting points we can draw from this:
For one, alcohol prices seem to not be affected by about anything. It may be the closest I have seen to a y=x type of graph from actual data naturally, so that is really interesting. Alcohol prices just steadily keep climbing at the same rate for the past 21 years.
Another interest aspect was how gasoline prices seemed to fluctuate with unemployment. At the big rises in unemployment, which were during the 2009 recession, and in 2020-Present during COVID lockdowns, the prices of gasoline dipped. Gasoline prices also just tend to rise over time when not in big recessions. The correlation between the two is still positive, which means that although during big points in unemployment that stand out as crises, and have a negative relationship during those times (More unemployment = Lower prices), during the rest of the 20 years in our dataframes, the prices tend to just rise in general.
In this case, I would have loved to find datasets that actually fit what I wanted to find, but it was still interesting to look into. I also would have figured that gas would have had a more steady increase in prices because that's what it seems like has been the trend, but the alcohol graph is just too perfect. (Also I wish that gas prices were what I am seeing here. It's around $3.50 out here in California)