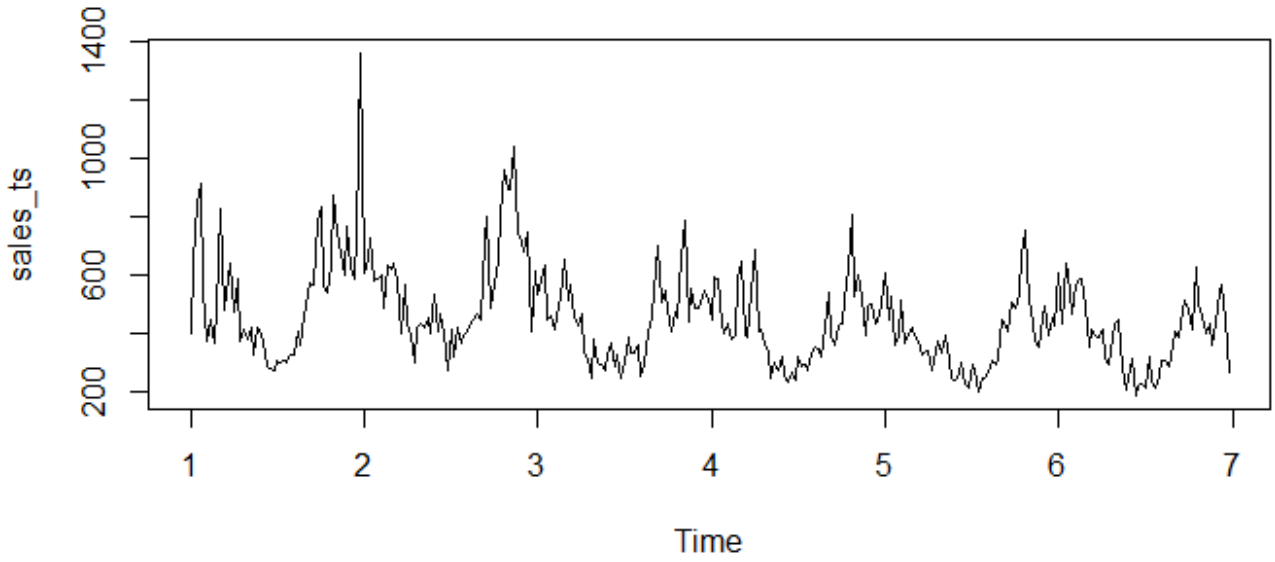
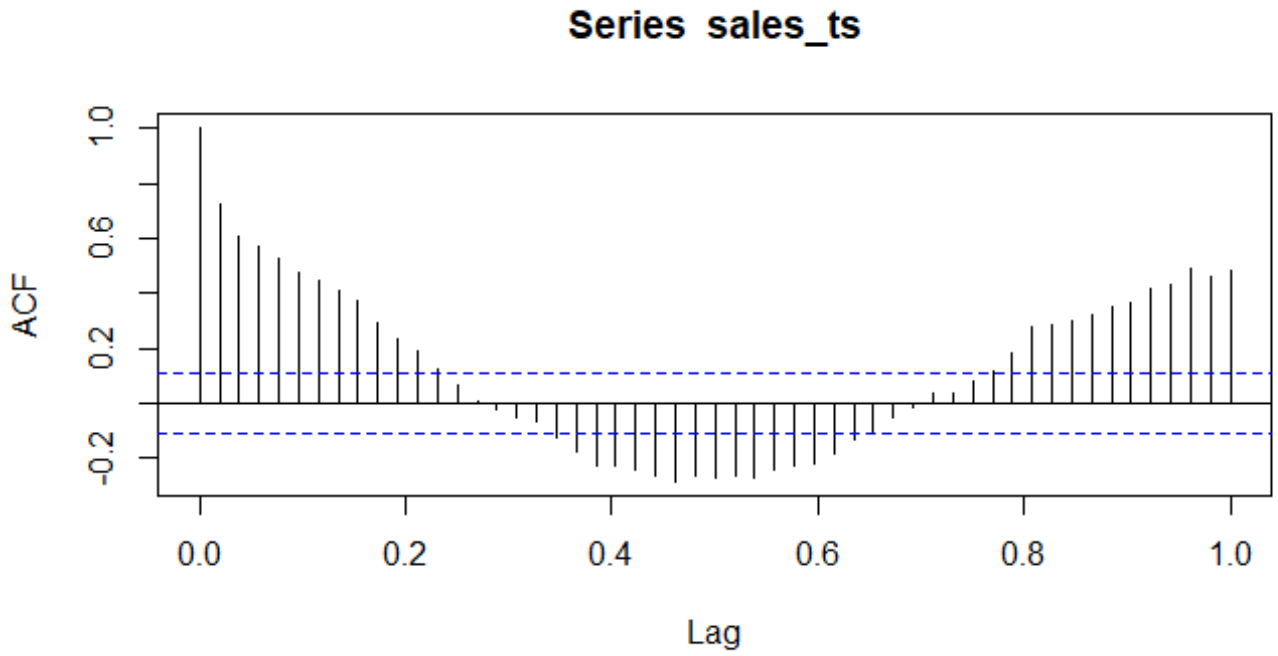
Homework 3

Liqi Zhu

Part 1

> g)





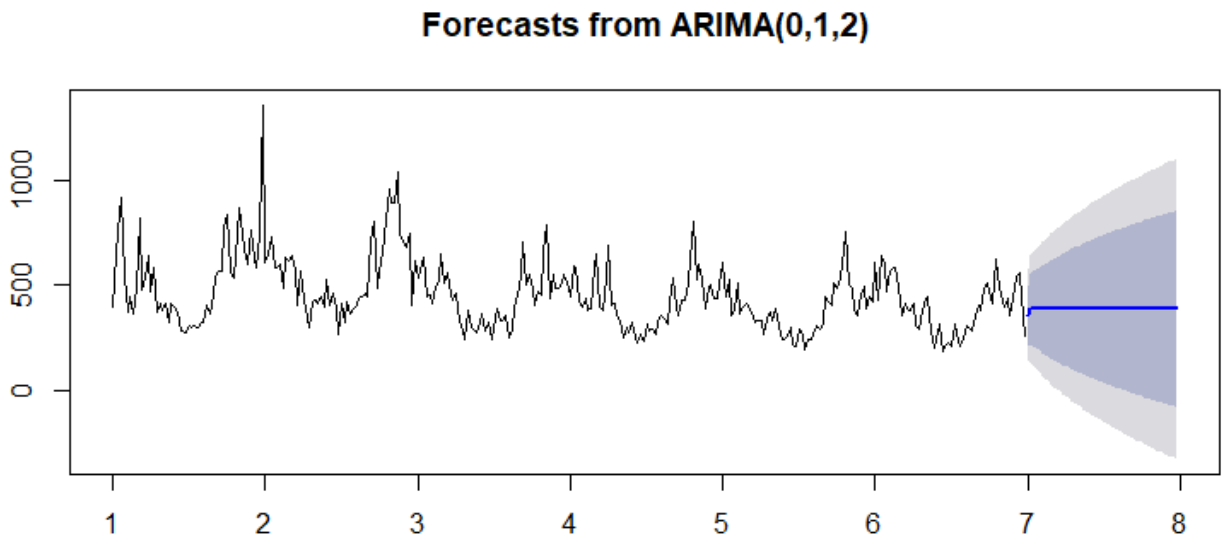
From the time series plot we can find that seasonality can be observed. And the large positive autocorrelations at lag(1.0) indicates seasonality as well. And since there are values over the Critical Values in ACF plot, elements of the time series are correlated with previous values, our error terms are correlated over time. So these plots indicate a forecasting model **is required**.

> h)

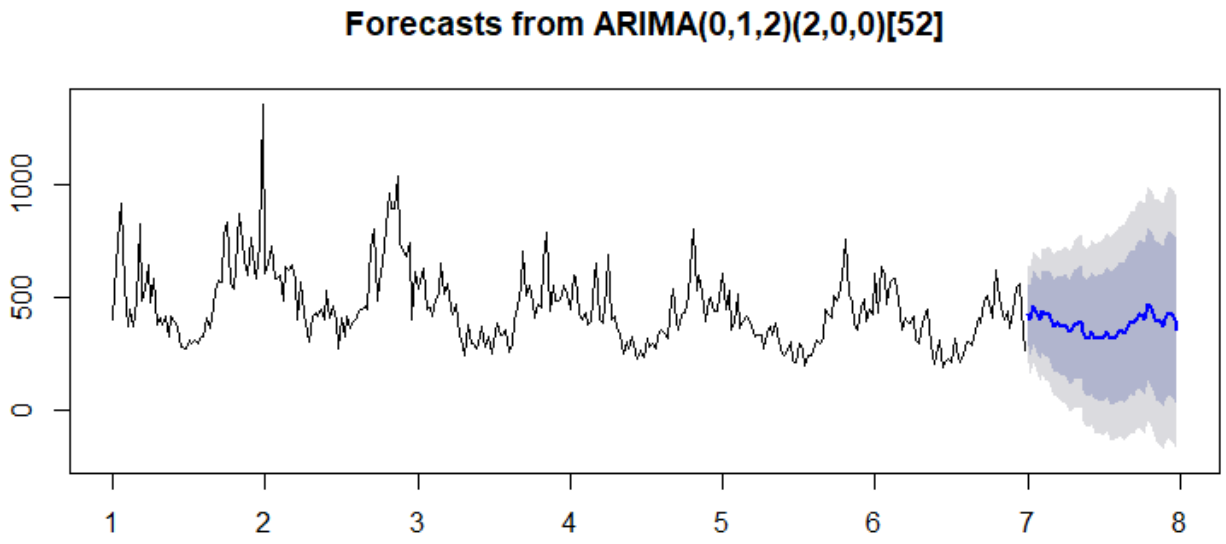
Auto.arima function returns best ARIMA model according to either AIC, AICc or BIC value. The function conducts a search over possible model within the order constraints provided.

> i)

Non-Seasonal plot



Seasonal plot



Performance

```
> performance
                    MSE      AIC
Seasonal       9226.053 3799.118
Non-seasonal  12766.316 3812.139
```

Seasional model has the better AIC and the better out of sample MSE, its plot also has the smaller confidence interval, thus I would choose the seasonal model.

Part 2

> e)

The price coefficient changes from **-0.329357** to **-0.221610** to **-0.16848**, coefficient interval changes from **[-0.34,-0.32]** to **[-0.236,-0.207]**,which is statistically significant, and from **[-0.236,-0.207]** to **[-0.267,-0.07]**, wich is not statistically significant.

> f)

I expect the price coefficient to be lower in absolute value than the former ones. But its significance level must be lower than each of the three. Its change from weekly aggregated must be less significant as well.

Part 3

> h)

Incorporating isFeature change coefficient interval changes from **[-0.34,-0.32]** to **[-0.314,-0.294]**, so it's **underestimated** and the change is statistically significant. **IsFeature is positively correlated with Units and negatively correlated with PricePerCan**.

> i)

Incorporating storeNum change coefficient interval from **[-0.294,-0.274]** to **[-0.325,-0.305]**, so it's **overestimated** and the change is statistically significant. **StoreNum1, compared to overall data in StorNum, is positively correlated with units and positively correlated with PricePerCan**.

> j)

Incorporating productNum change coefficient interval from **[-0.325,-0.305]** to **[-0.15,-0.11]**, so it's **underestimated** and the change is statistically significant. **ProductNum1, compared to overall data in ProductNum, is negatively correlated with units and positively correlated with pricePerCan.**

> k)

Standard errors are **0.004829** in 3b, **0.004869** in 3c, **0.004940** in 3d, **0.005329** in 3e, **0.0108629** in 3f and **0.0109025** in 3g, thus it keeps increasing. This would happen because when omitted variable are controlled, bias decreases and variance increases. And since the standard error indicate the precision of estimate, it will increase.