The data set used this paper are published by the whatever it is

The data is relative to the canadian housing reslate data from 2007 to 2023, by using this data, we could futhury sytudent waht is going on Canadian housing makret.

This paper is fousing on study how canadian housing market will respond to thepolicy rate change if Bank of Canadian remine the policy rate high as 4.5%.

Author using the housing reslce data m and ther avalidalt macro data to consturion one ARIMA model and VAR model to conditional forecast if canaidna hosuing go over 45.

```
autoplot(Can_month_housing_sell.ts)
```
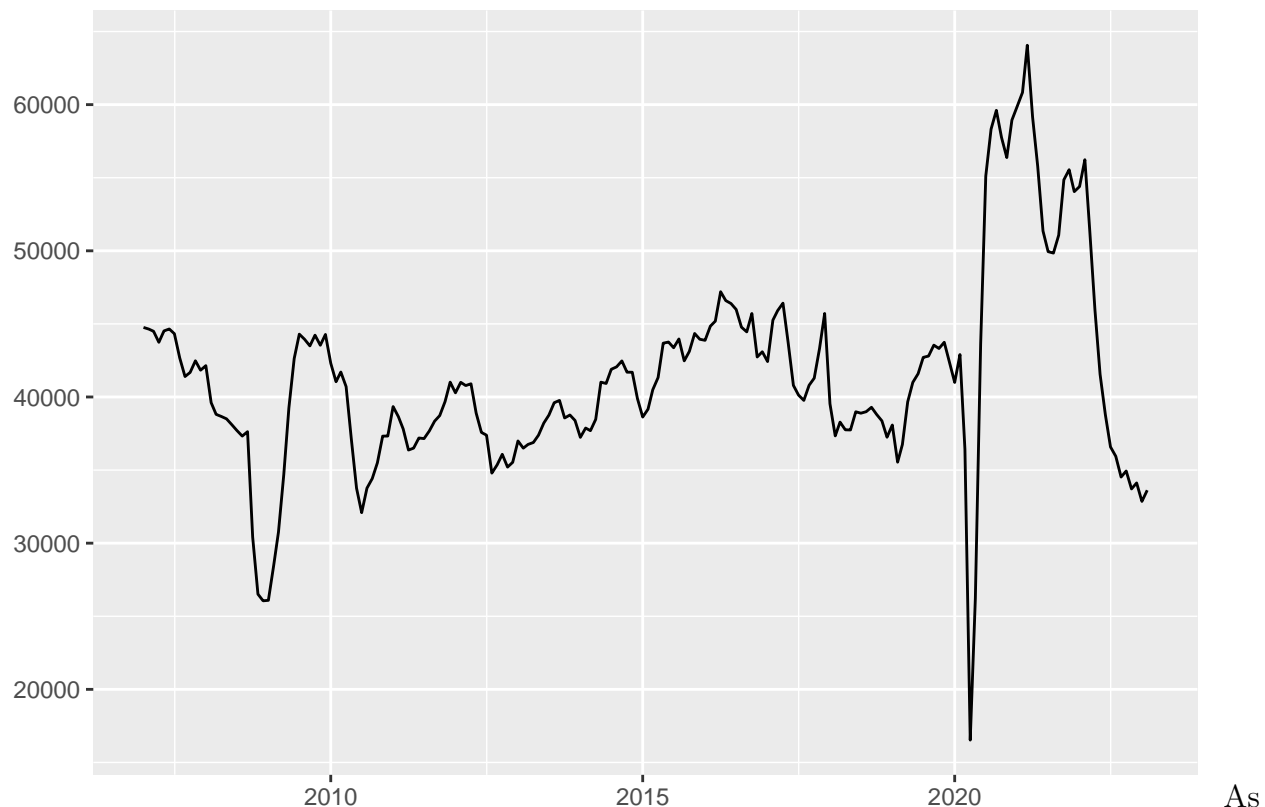

As shown in this graph, we can easily identify the trend of housing with the business cycle. The first drop can be explained as the dramatic impact of the 2008 financial crisis on the housing market. Another significant jump occurred around the beginning of 2020, when the first COVID-19 case was discovered in Canada, and COVID-19 became a global pandemic. Furthermore, as COVID-19 cases increased, the Canadian federal and provincial governments implemented COVID-19 restrictions to curb the spread, and the Bank of Canada lowered its policy rate to support the Canadian economy. As a result, there was a significant spike in housing sales. However, once the COVID-19 situation began to improve, and the war in Korea pushed up global energy prices, inflation in Canada reached an 8% 40-year high, resulting in a rising policy rate. As a result, the housing market began to cool down, and housing sales started to decline.

Part One: Constructing the ARIMA Model

1

To better analyze the data, the author constructed four ARIMA models using different time periods. The first model used the entire data set from 2007 to 2023, while the second focused on the period from 2016 to 2023, during the COVID-19 pandemic. The third model covered the period from 2009 July, when the 2008 recession ended, to 2023, and the fourth model covered the period from 2007 to 2019 January, with the aim of avoiding the impact of COVID-19.

After doing het different airma function 12 times, we got 12 different arime function while most of them are repace with self.

```
##                      Model
## 1           ARIMA(2,1,0)
## 2           ARIMA(0,1,1)
## 3 ARIMA(2,1,2)(0,0,1)[12]
## 4           ARIMA(0,1,0)
```

next,we will going to evaluate the different model to find out witch one is the best. first, let compare their AIC and BIC

By fit the all different ARIMA cobmine into the data that without covid-19 impact to see which mode is working better on fit the model wihtout covid 19

```
##                              AIC       BIC
## ARIMA(2,1,0)            2515.189 2524.098
## ARIMA(0,1,1)            2514.656 2520.596
## ARIMA(2,1,2)(0,0,1)[12] 2505.988 2523.807
## ARIMA(0,1,0)            2514.656 2520.596
```

As the output show, the ARIMA(2,1,2)(0,0,1)[12], it may relace to this combination is generate base on the data from 2007 to 2019, which are not suprised at all.

compare the AIC and BIC model base on the data that including the covid influence.

```
##                              AIC       BIC
## ARIMA(2,1,0)            3582.467 3592.255
## ARIMA(0,1,1)            3583.642 3590.168
## ARIMA(2,1,2)(0,0,1)[12] 3579.123 3598.699
## ARIMA(0,1,0)            3583.642 3590.168
```

Unsprised, the ARIMA(2,1,0) is best accrding to it AIC value, due to its fit by from the entire data set.

Futher, let's do the cross validation of four model on the data that inculding the covid and without the cov

the cross valdation

Furthermore, let's perform cross-validation on four models using data that includes and excludes the COVID period. The cross-validation process involves training the models using data up to 2016 and testing them using data from 2016 to 2019. Assuming that the relationship between housing sales and other macroeconomic factors remains the same in

the post-COVID period, we can focus on the market's behavior during normal times and ignore the impact of COVID. Therefore, this cross-validation test starts with the training set from 2007 to 2015 and uses the dataset from 2016 to 2018 for testing to evaluate the model's goodness.

```
##                    Model        ME     RMSE      MAE        MPE     MAPE
## 1          ARIMA(2,1,0) -141.3249 1391.234 1084.889 -0.3118674 2.496156
## 2          ARIMA(0,1,1) -143.7842 1412.804 1095.866 -0.3038535 2.513827
## 3 ARIMA(2,1,2)(0,0,1)[12] -189.4028 1515.188 1175.828 -0.4147509 2.701633
## 4          ARIMA(0,1,0) -161.5556 1387.015 1087.222 -0.3269795 2.508217
```

Based on the provided error metrics, the best model appears to be ARIMA(0,1,1) as it has the lowest MAE (1977.793) and the second-lowest MAPE (5.568452). Lower error metrics indicate better model performance. however, ARIMA(2,1,0) has the lowest AIC in both with or without covid influence.it suggests that this model might be the best balance between model complexity and goodness of fit.Given this new information, The ARIMA(2,1,0 to consider the ARIMA(2,1,0) model as the best