**Evaluations and Results:**

1. **Validating the models:**

* We performed residual analysis using Ljung Box test on all the models to identify which models consist of white noise.
* The models having white noise are valid models.
* Here, we have shown the Ljung box test for one model from male and female each:
* **Male:**

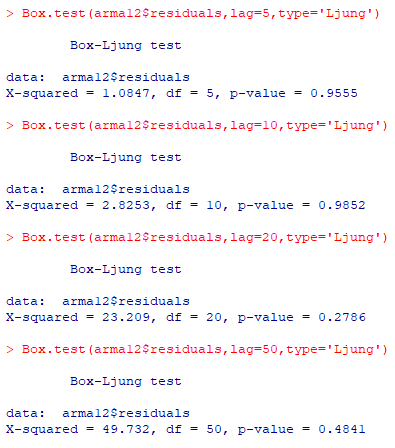
**Ljung Box Test for ARMA(3,6):**

A screenshot of a social media post

Description generated with very high confidence

Even though the lag is 5 or 100, the p-value is greater than 0.05, thus, we can say that, it has white noise and our model is qualified.

* **Female:**

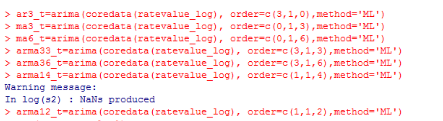


Whether the lag is 5 or 50, the p-value is greater than 0.05, thus, we can say that, it has white noise and our model is qualified.

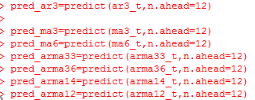
1. **Evaluation of models:**

* To evaluate which model is the best, we need to test all the models against the test data. But since we have applied first differencing on male data and second differencing on the female data, we need to first build ARIMA models on the original, non-stationary data to evaluate the best models.
* For simplicity we have shown the process for Test Window 1 for both male and female test data as follows:
* **Male:**

**Step 1: ARIMA models:**



**Step 2: Creating prediction variables for test data:**

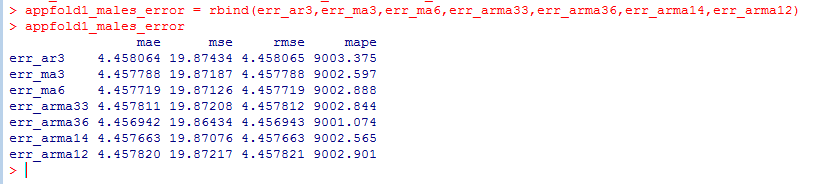


We have created prediction variable for each tasting model which we have created in Step 1. Here, ‘**n.ahead**’ denotes the number of records present in tasting data.

**Step 3: Generating errors:**

**A close up of a piece of paper

Description generated with high confidence**

****

Similarly, we calculated error for all the windows. Based on each fold, we chose the best model which has lowest MAE value.

|  |  |  |  |
| --- | --- | --- | --- |
| **Sr No.** | **Model** | **MAE value** | **Window** |
| 1 | ARMA(3,6) | 4.456942 | Window1 |
| 2 | ARMA(1,4) | 4.457438 | Window2 |
| 3 | ARMA(1,4) | 4.461118 | Window3 |
| 4 | ARMA(2,4) | 4.457336 | Window4 |

Equation for the ARMA(3,6) model is:

**A picture containing object, clock

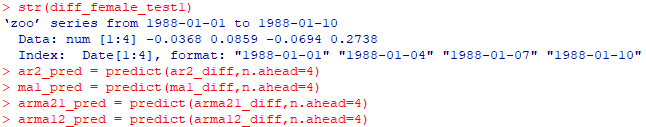
Description generated with very high confidence**

* **Female:**

**Step 1: ARIMA models:**

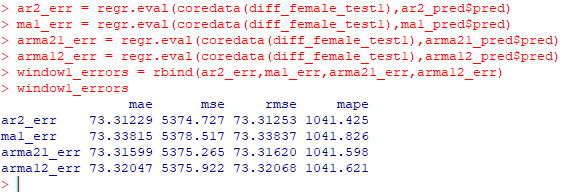
****

**Step 2: Creating prediction variables for test data:**

****

We have created prediction variable for each tasting model which we have created in Step 1. Here, ‘**n.ahead**’ denotes the number of records present in tasting data.

**Step 3: Generating errors:**



Similarly, we calculated error for all the windows. Based on each fold, we chose the best model which has lowest MAE value.

|  |  |  |  |
| --- | --- | --- | --- |
| **Sr No.** | **Model** | **MAE value** | **Window** |
| 1 | AR(2) | 73.31229 | Window1 |
| 2 | ARMA(1,3) | 73.39444 | Window2 |
| 3 | ARMA(1,3) | 73.51356 | Window3 |
| 4 | ARMA(1,2) | 73.23264 | Window4 |

Equation for the ARMA(1,2) model is:

****