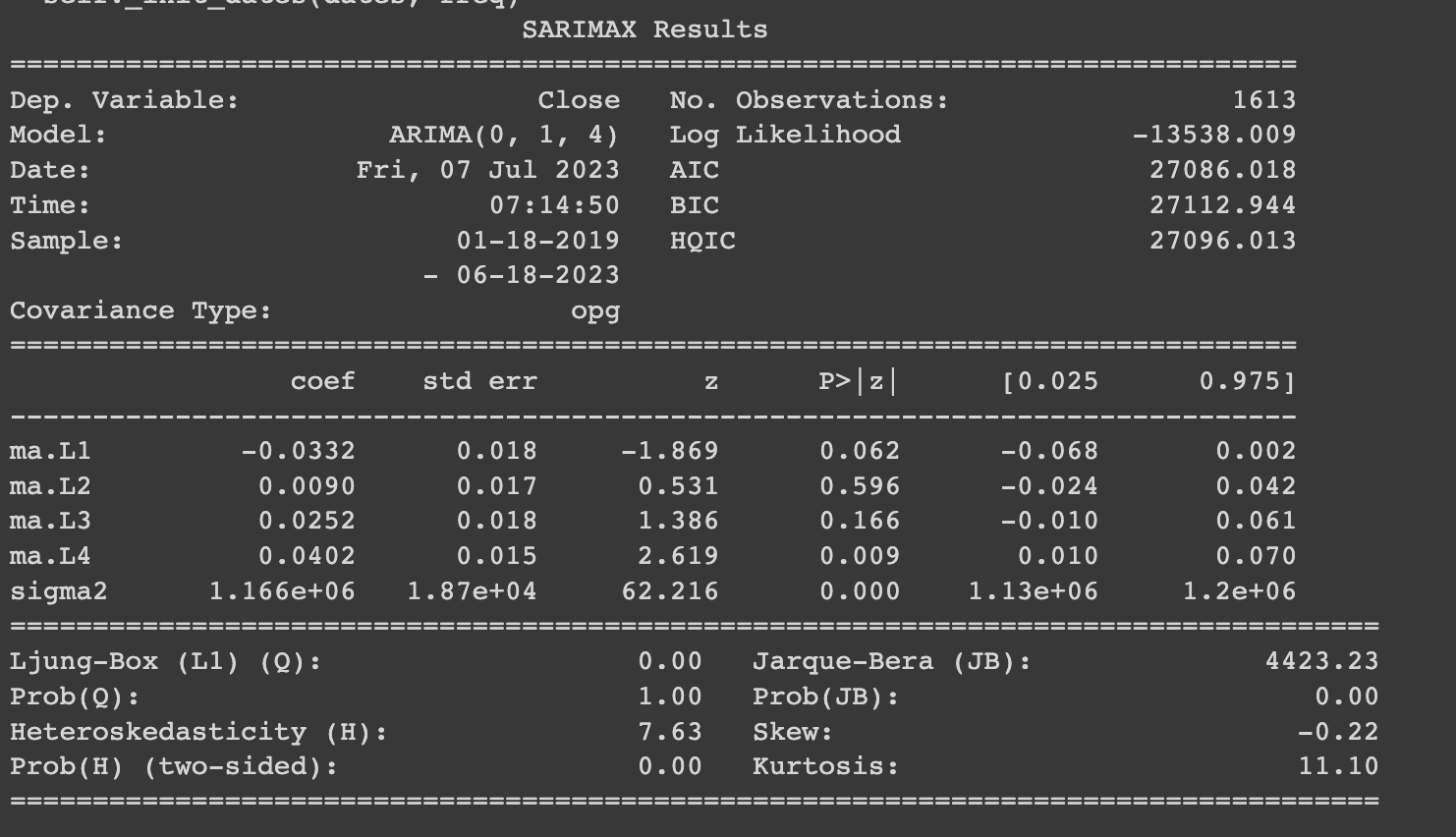
When Arima is set (0,1,4):



Check for outliers using z-scores, control threshold:

Around 97% of data contained: threshold = 2.6:

There are a total of 60 outliners:

Date

2019-01-18 3.378775

2020-03-12 -2.706076

2021-01-11 -2.794090

2021-01-13 3.178222

2021-01-21 -4.403923

2021-01-28 2.716201

2021-02-08 6.720684

2021-02-17 2.744868

2021-02-19 3.852611

2021-02-22 -3.149708

2021-02-23 -5.284137

2021-03-01 4.229059

2021-03-13 3.458115

2021-03-15 -3.304777

2021-03-22 -2.802202

2021-03-26 3.264588

2021-04-13 3.279620

2021-04-18 -4.168104

2021-04-26 4.712746

2021-04-30 3.617057

2021-05-04 -3.727139

2021-05-05 3.674174

2021-05-12 -7.023466

2021-05-15 -2.752404

2021-05-17 -2.723344

2021-05-19 -5.378008

2021-05-20 3.392382

2021-05-21 -2.955637

2021-05-24 3.491447

2021-05-28 -2.728965

2021-06-09 3.644713

2021-06-13 3.075469

2021-06-21 -3.664100

2021-06-25 -2.670553

2021-08-13 3.000163

2021-09-07 -5.429545

2021-09-20 -4.123415

2021-09-22 2.615526

2021-10-01 4.127709

2021-10-06 3.617667

2021-10-15 3.865012

2021-10-21 -3.515921

2021-11-08 4.039028

2021-11-16 -3.204603

2021-11-18 -3.160521

2021-11-26 -3.346522

2021-12-03 -2.722771

2021-12-04 -4.153466

2021-12-09 -2.680672

2021-12-13 -3.002220

2021-12-28 -2.846175

2022-01-21 -3.931172

2022-02-04 4.011108

2022-02-17 -3.234483

2022-02-28 4.967645

2022-03-04 -3.365487

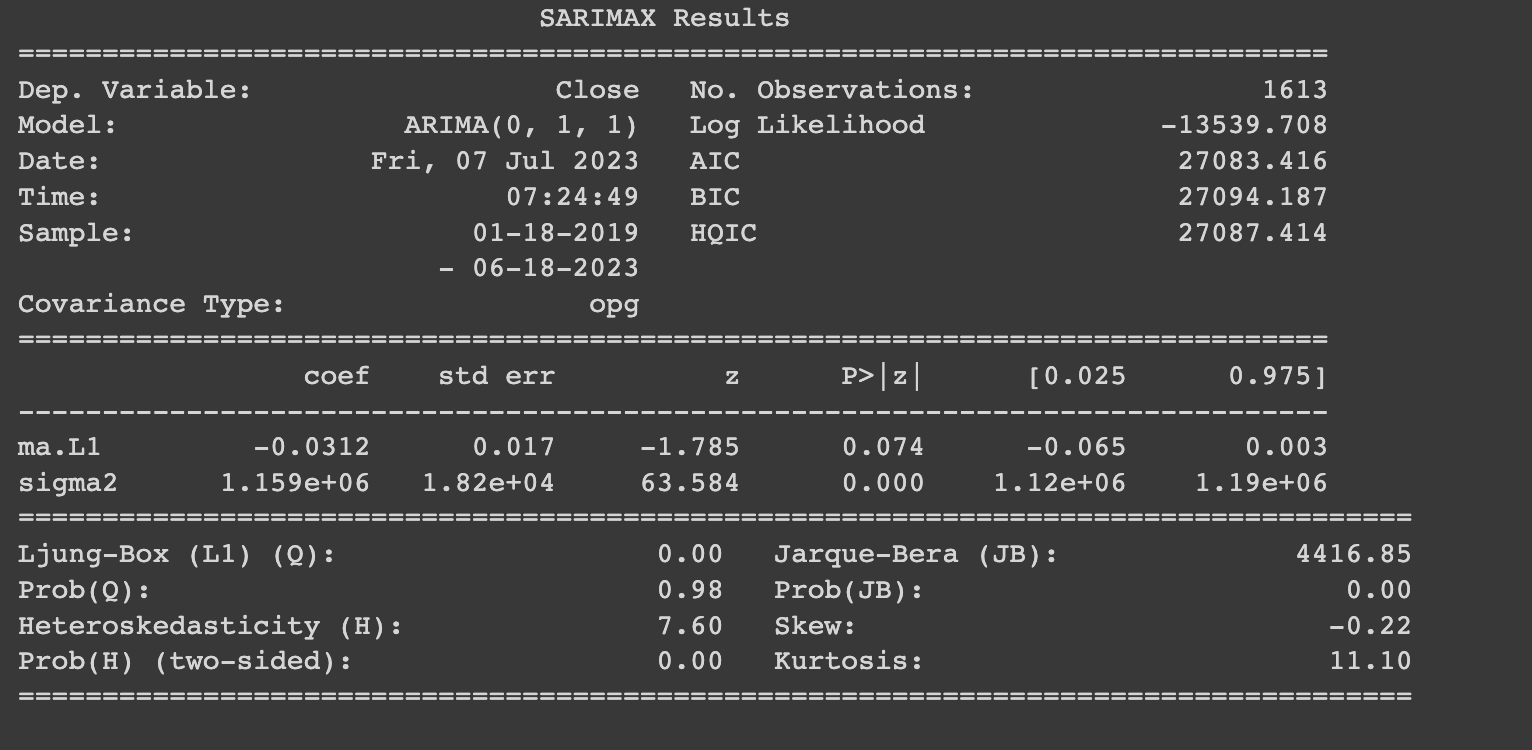
2022-03-09 3.039965

2022-05-05 -2.874069

2022-05-09 -3.415360

2022-06-13 -3.991323

When the Arima is set (0,1,1):



Check for outliers using z-scores, control threshold:

96.281% data contained: threshold = 2.6:

There are a total of 60 outliners:

Date

2019-01-18 3.374224

2020-03-12 -2.740830

2021-01-06 2.666492

2021-01-11 -2.656343

2021-01-13 3.079099

2021-01-21 -4.407331

2021-01-28 2.734104

2021-02-08 6.733668

2021-02-11 2.711678

2021-02-17 2.753044

2021-02-19 3.873526

2021-02-22 -3.062202

2021-02-23 -5.099178

2021-03-01 4.118123

2021-03-13 3.596822

2021-03-15 -3.214604

2021-03-22 -2.813415

2021-03-26 3.133313

2021-04-13 3.321307

2021-04-18 -4.183237

2021-04-26 4.603078

2021-04-30 3.835258

2021-05-04 -3.583754

2021-05-05 3.663919

2021-05-12 -6.994031

2021-05-15 -2.902355

2021-05-17 -2.731977

2021-05-19 -5.510659

2021-05-20 3.316396

2021-05-21 -3.135015

2021-05-24 3.551646

2021-06-09 3.568538

2021-06-13 3.219041

2021-06-21 -3.740470

2021-06-25 -2.790464

2021-08-13 3.069185

2021-09-07 -5.384454

2021-09-20 -4.137044

2021-10-01 4.058252

2021-10-06 3.619423

2021-10-15 3.942674

2021-10-21 -3.469161

2021-11-08 3.965932

2021-11-16 -3.217688

2021-11-18 -3.187685

2021-11-26 -3.421544

2021-12-03 -2.705811

2021-12-04 -4.174901

2021-12-09 -2.646307

2021-12-13 -3.107136

2021-12-28 -2.848210

2022-01-21 -3.960674

2022-02-04 4.016445

2022-02-17 -3.204587

2022-02-28 5.025704

2022-03-04 -3.129927

2022-03-09 3.011396

2022-05-05 -2.854492

2022-05-09 -3.544523

2022-06-13 -4.024561

AR, 101

证明时间序列的过程是什么样子的

写 paper：

证明时间序列的过程：

1. Background：
   1. Data source: Yahoo Finance;
   2. Reason to do this?:
2. Trend: Data Visualization: 数据长啥样，哪里出现高峰，哪里出现低峰，哪里出现小数据不正常现象
3. Seasonality: 不符合，此处【需要数据】
4. ACF 和 PACF，三者的图需要展示, how do we determine (p,d,q) by that graph?
5. 拟合度如何，有多确定这三者是时间序列，或者不符合时间序列，哪里可以发现这一规律
6. 白噪声检验【把数据沾过来】
7. 根据时间序列异常的部分 Irregularities and outliners: 在数据非常不同寻常的时间段，我们需要找这个的噪声，【新闻，Twitter】之类的，这些是猜想
8. How does this relate to the context and purpose of the investigation?