1. Starbucks

<table>
<thead>
<tr>
<th>Type</th>
<th>Mean</th>
<th>Median</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma_{0,t} )</td>
<td>4.67</td>
<td>3.10</td>
<td>0</td>
<td>60.32</td>
</tr>
<tr>
<td>( \sigma_{1,t} )</td>
<td>6.27</td>
<td>4.42</td>
<td>0.11</td>
<td>80.03</td>
</tr>
<tr>
<td>( \sigma_{2,t} )</td>
<td>4.86</td>
<td>3.86</td>
<td>0.67</td>
<td>101.25</td>
</tr>
<tr>
<td>( \sigma_{3,t} )</td>
<td>8.66</td>
<td>6.88</td>
<td>1.19</td>
<td>181.01</td>
</tr>
<tr>
<td>( \sigma_{5,t} )</td>
<td>4.92</td>
<td>3.98</td>
<td>0.65</td>
<td>117.05</td>
</tr>
<tr>
<td>( \sigma_{6,t} )</td>
<td>8.97</td>
<td>7.22</td>
<td>1.18</td>
<td>213.23</td>
</tr>
</tbody>
</table>

2. The time plot is given in Figure 1. Let \( x_t \) be the log volatility, the model selected is

\[
(1 - 0.104B - 0.070B^2 - 0.044B^3)(1 - B)x_t = a_t,
\]

where \( \sigma_a^2 = 6.82 \times 10^{-4} \). The model fits the series well based on the model checking statistics, but there are a few large outliers. The 1-step to 5-step ahead forecasts are \(-1.495, -1.497, -1.498, -1.498, -1.498\).

3. The monthly returns of Procter and Gamble (PG).

(a) The fitted logistic model is

\[
\text{logit}(p_t) = 0.44 + 0.1M_{t-1} - 0.32M_{t-2} - 0.12S_{t-1} - 0.02S_{t-2}.
\]

Based on the \( z - value \), the coefficient of \( M_{t-2} \) is significant so that the model has some predictive power, even it is not strong.

(b) The fitted model is

\[
\hat{M}_t = \begin{cases} 
1 & \text{if } o_t > 0 \\
0 & \text{if } o_t \leq 0,
\end{cases}
\]

where

\[
o_t = -0.03 + 5.18h_{1t} + 0.80h_{2t} + 0.13M_{t-1} - 5.19M_{t-2} - 0.24S_{t-1} + 0.27S_{t-2}
\]

where

\[
h_{1t} = \frac{\exp[-6.39 - 5.39M_{t-1} + 23.68M_{t-2} + 4.78S_{t-1} - 0.59S_{t-2}]}{1 + \exp[-6.39 - 5.39M_{t-1} + 23.68M_{t-2} + 4.78S_{t-1} - 0.59S_{t-2}]},
\]

\[
h_{2t} = \frac{\exp[-0.44 + 11.05M_{t-1} - 12.86M_{t-2} - 8.10S_{t-1} - 18.48S_{t-2}]}{1 + \exp[-0.44 + 11.05M_{t-1} - 12.86M_{t-2} - 8.10S_{t-1} - 18.48S_{t-2}]}.
\]
(c) The number of forecast errors for NNET is 55 whereas that for logistic regression is 41. However, based on the range of the probabilities of forecasts and with 0.5 as cut-off, logistic regression predicts every point as “up”. This might be be a good feature. One may change the cut-off probability (threshold) to refine the prediction.

4. TAQ data of Starbucks.
   - The 5-m log returns are shown in Figure 2.
   - Based on the Ljung-Box statistics, there are no serial correlations. We have $Q(10) = 14.48$ with $p$-value 0.15.
   - The 5-m realized volatilities (in percentages) are 0.7245, 0.7069, 0.9481, 0.7744, 0.9254, 0.8696, 0.7602.
   - The 1-m realized volatilities (in percentages) are 0.8965, 0.9121, 1.1317, 0.7168, 0.8699, 1.0194, 0.9889.

5. Again, consider the tick-by-tick trade data of Starbucks from December 20 to December 31, 2014.
   - See the upper plot of Figure 3.
   - See the lower plot of Figure 3. There exists diurnal pattern in the trading intensity.
Figure 2: Time plot of intraday 5-m log returns of Starbucks stock in December 22 to 31, 2014.

Figure 3: Trading intensity of Starbucks stock in December 22 to 31, 2014.