MATH 584: Homework 3

Zachary Denis André Scialom
A20497400

11 March 2022

Contents

I have made one script for each question to make it easier to read.

1 Question A

To compute the different PnL processes, the code is mainly based on two dataframes called: df and df_{pnl}.
The df dataframe contains ten columns corresponding respectively to:

1. Ticker 1 of the pair studied
2. Ticker 2 of the pair studied
3. (1: pair open),(0,pair closed)
4. μ
5. σ
6. C^{ij}
7. Holding time increment
8. Shares in asset i
9. Shares in asset j
10. (1:long position),(0,short position)

The second dataframe contains the vector ζ_t and the vector of changes in price (price[t+1]-price[t]). Therefore, the absolute PnL is computed thanks to the dot product between the two vectors. The long side PnL and the short side PnL are derived similarly with other specific dataframes.

```
[0, 52671.402146602006, -112218.76217430412, -33684.975502960486, -92143.92802663037, -81820.24240963432, -14448.00656321898, -7805.737911058926,
281133.2702971509, 351423.093833164, 232480.9125622374, 81505.02077057504, 154930.32322303538, 125634.34409586614, -213985.99121081928, -184689
.40796084367, -367503.67693028355, -391148.2137005293, -318108.60454328556, -286928.8998145255, 266003.830875471, 587713.0915669026, 819522.801161943,
988552.6579112608, 973580.2773967204, 827973.5093312578, 630682.4018241649, 972150.9603613797, 1286222.60585359, 1497324.0204799317, 1576221
.2105693528, 2010189.895858982, 1952494.8350720427, 2096766.0142122065, 2449962.795971236, 2271758.426135828, 2382611.416831604, 2066877.4131438052,
2289998.1403705133, 1849265.7653094062, 1583050.9955193156, 1682315.5280864928, 1563126.0924616875, 1329551.3550474208, 1251922.2219789517, 1201501
.5498979848, 1426240.8264998167, 1545780.9488549342, 1792933.9850267014, 2051331.230583132, 1751452.203913706, 1789471.0802640019, 1697368.8824432516,
2111066.626612665, 2481816.7442482435, 2600417.436876397, 2731464.587493779, 2747501.1530255056, 2755591.590742769, 2758026.8907238687, 2778263
.6469661677, 2778990.185698071, 2610156.8684952105, 2497612.2542254888, 2521397.0477682883, 2509350.9597791093, 2501110.365472387, 2702803.2760215537,
2703391.2617793325, 2759286.240943364, 2652178.9238827433, 2641334.8975821533, 2733460.8603091524, 2696408.625443953, 2672670.565963604, 2982428
.44046878, 3064016.415011491, 3213354.307157436, 3340991.8465368454, 3148565.275583593, 3254807.468083258, 3187329.7220778484, 3307657.71703128,
```

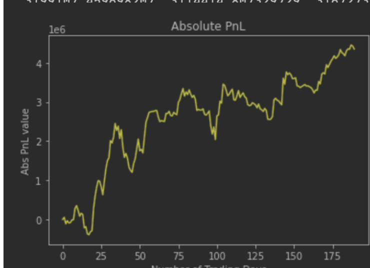

Figure 1: Absolute PnL (Values + Plot)

```
print('Number of opened pairs closed by failure condition: ',np.sum(tab_failure_num))
print('Number of opened pairs closed successfully: ',np.sum(tab_success_pair))

Number of opened pairs closed by failure condition:  163
Number of opened pairs closed successfully:  206
```

Figure 2: Number of success pairs and fail pairs

Figure 3: Long PnL, Short PnL and Absolute PnL

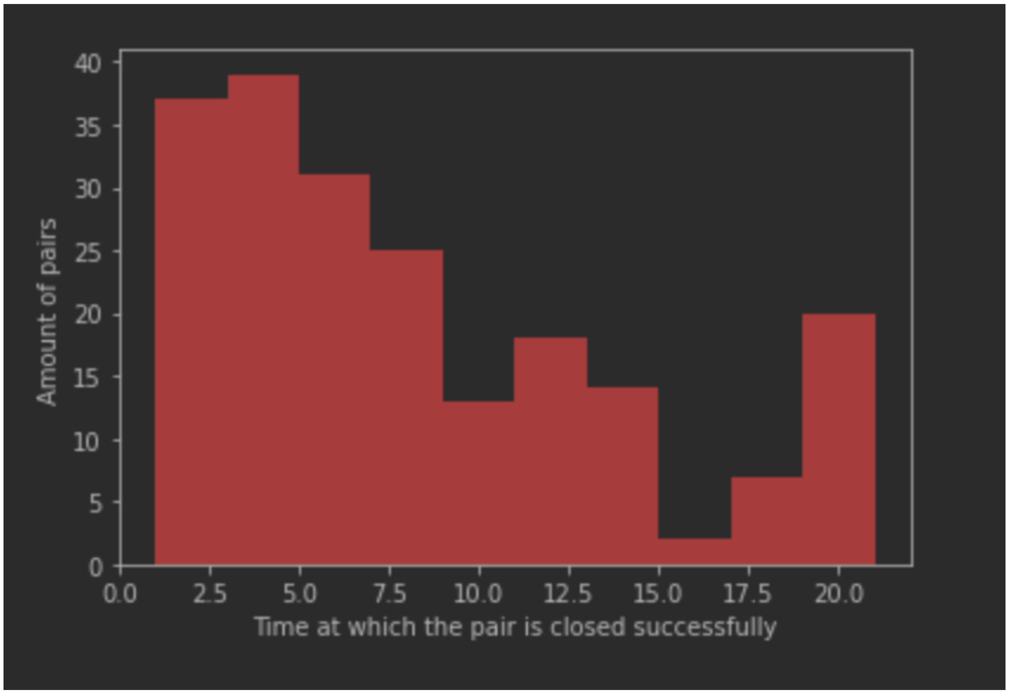

Figure 4: Histogram of the holding times: Reaching target

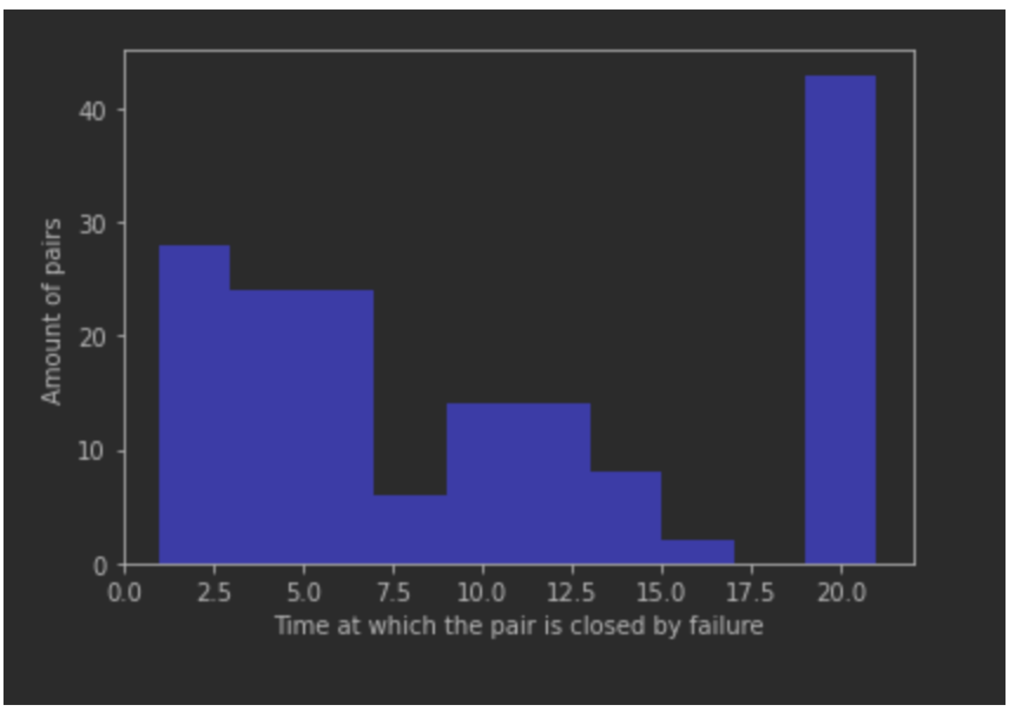

Figure 5: Histogram of the holding times: Stop-Loss condition

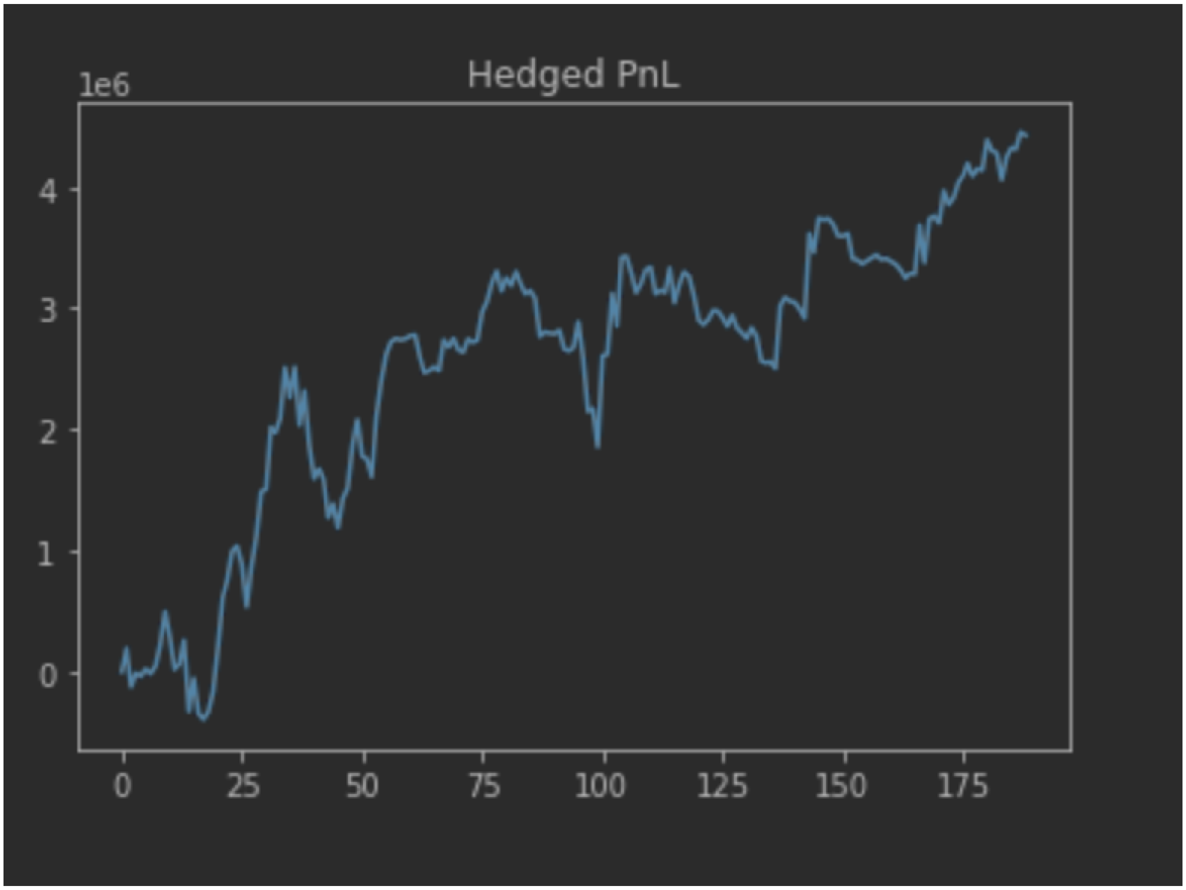

Figure 6: Hedged PnL

```
Absolute PnL Sharpe ratio:  2.118350876749855
Long PnL Sharpe ratio:  1.339658249390018
Short PnL Sharpe ratio:  0.9384284771937909
Hedged PnL Sharpe ratio:  1.8052228006992885
```

Figure 7: Different Sharpe Ratios obtained

The hedged PnL has a Sharpe ratio that is smaller than the one of the absolute PnL which makes sense as we are market-neutral, so the risk taken is less important.

2 Question B

To compute the absolute PnL when transaction costs are considered, another column has been added to the dataframe that enables to compute the PnL. This column contains the variation of the number of shares between time t and t+1. Indeed, once a pair is opened, μ, σ and C^{ij} do not change but the shares in assets i and j are calculated again: it is rebalancing. Thus, we must apply the T-cost to this difference in the number of shares used to re balance.
Obviously, it also must be taken into account when we first open the pair. The T-cost is applied directly to the number of shares (not any difference) in the beginning.
In the end, to the absolute PnL at time t, we subtract the sum of all the shares involved in an opening or in rebalancing multiplied by the T-cost.

```
[0, 52523.46953276581, -113335.74982319541, -36150.78760016443, -96127.6881477672, -87301.22319565913, -21670.569751811265, -16668.307813972664,
270744.1317278212, 339703.9296137577, 219705.56682273082, 67730.38307346139, 140135.76406034772, 109584.5396927243, -231367.50096545005, -203382
.80816115977, -387505.56313398, -412336.263196514, -340559.206433983, -310722.5719546228, 240209.94280728462, 560810.2663622928, 790516.9834494285,
957092.1479832752, 940489.1174645645, 794063.8364367543, 596210.616057965, 937021.2538223212, 1250119.0062325243, 1460470.3147414955, 1538990
.8275812906, 1972451.5722014317, 1914468.858153089, 2058575.4199682798, 2411296.4816978388, 2231975.8910847544, 2341156.8216735, 2021567.3618415964,
2240414.6124968915, 1794871.6725972097, 1523973.182321935, 1618469.2318598595, 1494499.7621047585, 1255713.9081545372, 1172755.651721241, 1116786
.884916196, 1336030.0022020615, 1450088.283388573, 1691728.5795037546, 1944632.6318660157, 1639363.2961462198, 1671921.2477611944, 1574199.1678869165,
1982178.5049532233, 2348301.1311273095, 2463963.8308159057, 2593301.97364724, 2608642.578766152, 2616115.460165673, 2617951.4557391526, 2637640
.9656518707, 2637856.633066482, 2468398.646935558, 2354293.987837553, 2376521.0857484867, 2363125.2083380013, 2353876.4564004517, 2554539.8818924744,
2554121.1076373127, 2608953.398936185, 2500751.4660890084, 2488796.439473863, 2579880.750221952, 2541834.0067721196, 2517118.779598741, 2825938
.104359873, 2907447.606657349, 3056585.9876979785, 3184068.4706691024, 2991280.0961306486, 3097163.593282246, 3029293.559762407, 3149179.8192690797,
```

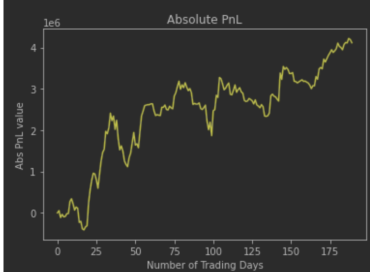

Figure 8: Absolute PnL (Values + Plot) with Tcost = 0.001

Figure 9: Long PnL, Short PnL and Absolute PnL with Tcost = 0.001

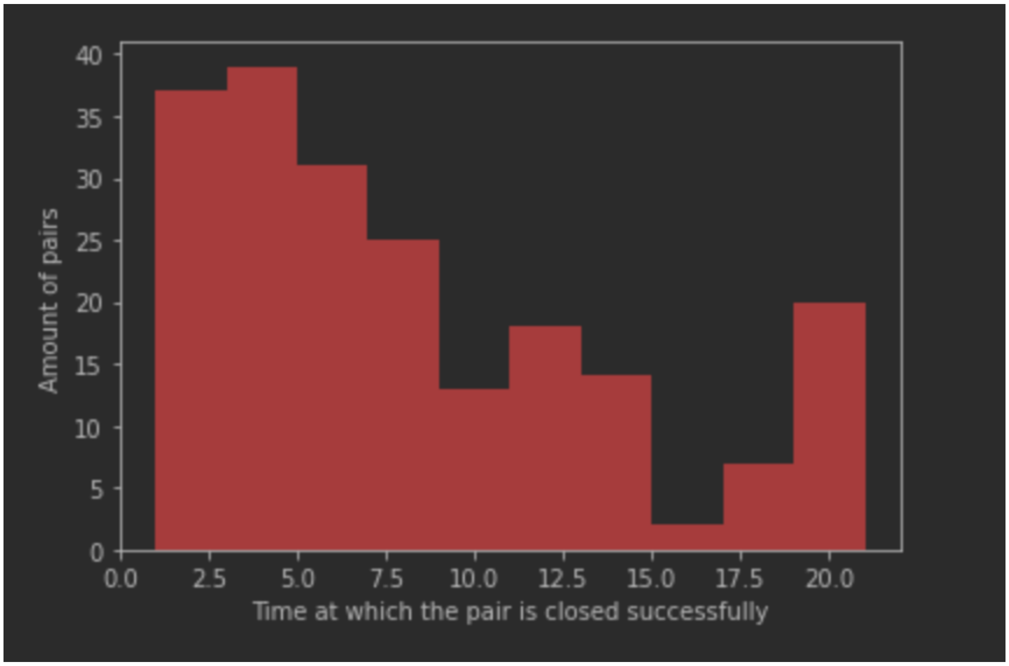

Figure 10: Histogram of the holding times: Reaching target with Tcost = 0.001

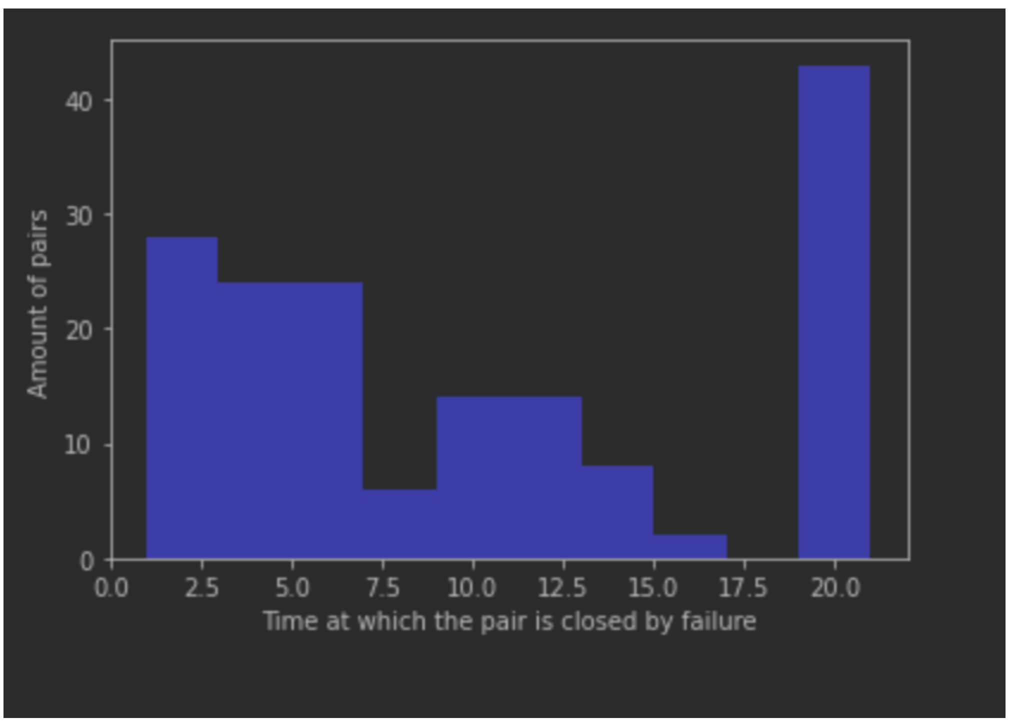

Figure 11: Histogram of the holding times: Stop-Loss condition with Tcost = 0.001

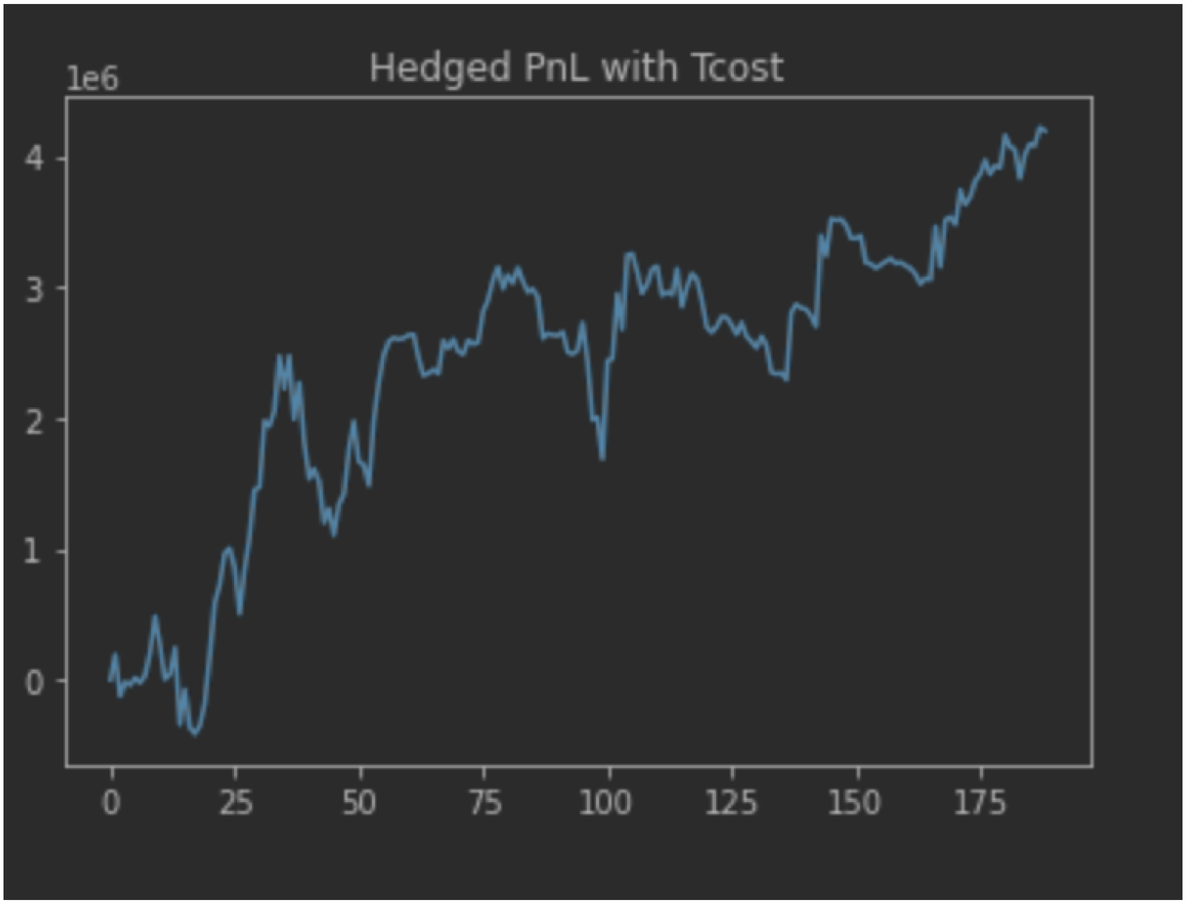

Figure 12: Hedged PnL with Tcost = 0.001

```
Absolute PnL Sharpe ratio:  2.00465643437137
Long PnL Sharpe ratio:  1.151607360230623
Short PnL Sharpe ratio:  0.9179181483117581
Hedged PnL Sharpe ratio:  1.7099175229701205
```

Figure 13: Different Sharpe Ratios obtained with Tcost = 0.001

Obviously, in presence of transaction costs, the different Sharpe Ratios genuinely diminish and this is confirmed by the results presented above.

3 Question C

The aim of this question was to focus on the performances obtained by taking log-prices. One main difference in the approach of this strategy is C^{ij} no longer represents the number of shares of j to short per one share of i, but the negative of the dollar amount to invest in asset j, per one dollar invested in asset i.

Therefore, it is now important to figure out on which asset the maximum capital available will be placed.

We can see C^{ij} as: $C^{ij} = \frac{S_j}{S_i}$. If $C^{ij} > 1$, then the maximum capital must be placed on asset j and vice-versa.

As before, we must apply the right sign to the right asset according to the position taken in the asset.

```
[0, 23546.40050678764, -173929.48441649208, -93006.74683007308, -144191.71364629563, -154634.62218897327, -86592.1838909535, -98139.2991013323,
249048.07198687003, 340545.6683844596, 245483.69421038099, 34962.055331521406, 88774.14977962617, 139923.91814540935, -252394.20231892398, -202553
.5766947359, -398712.11776421336, -399924.4823347654, -392209.9488883585, -291980.876335641, 519734.5632739167, 775521.1417749258, 996762.6257746933,
1024560.3065141338, 1029124.1564318256, 915322.8758791289, 698369.9613715073, 970003.4092082293, 1282777.4691035214, 1461516.096749252, 1504182
.9362715492, 1930059.3379284372, 1883566.0714718732, 2083702.084320621, 2459975.0111640384, 2261579.7353739906, 2310157.1579907457, 2037096.4244477763,
 2231601.9657807145, 1837597.4446837734, 1633551.9786108893, 1723359.6860850344, 1674608.4794327444, 1449233.5576229468, 1393688.1186438992, 1358128
.2587705273, 1545944.073976169, 1624944.1378394577, 1833798.7517386794, 2112941.1378309685, 1841009.5778350946, 1886515.1877154685, 1840777.0981773403,
 2164063.270465441, 2546068.9715289753, 2623120.7648953674, 2750358.9182341844, 2763519.684004348, 2798625.295178151, 2794298.4433298814, 2821481
.471758605, 2811804.782874773, 2651179.7557711336, 2548815.17241068, 2549424.977104436, 2564160.1444510873, 2509573.377184978, 2698916.3046612046,
2674919.0678596054, 2796009.2713054274, 2710977.1138906702, 2739808.515353017, 2869634.6447374127, 2787766.0322163734, 2749004.9717848655, 3070852
.483094788, 3124065.05561252, 3232740.8955280147, 3346055.9723197715, 3165418.254717256, 3245856.879870049, 3183039.5696970457, 3280471.1261056405,
```

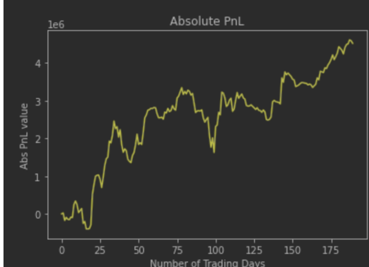

Figure 14: Absolute PnL (Values + Plot) with log-prices

```
print('Number of opened pairs closed by failure condition: ',np.sum(tab_failure_num))
print('Number of opened pairs closed successfully: ',np.sum(tab_success_pair))
```

```
Number of opened pairs closed by failure condition:  156
Number of opened pairs closed successfully:  199
```

Figure 15: Number of success pairs and fail pairs with log-prices

Figure 16: Long PnL, Short PnL and Absolute PnL with log-prices

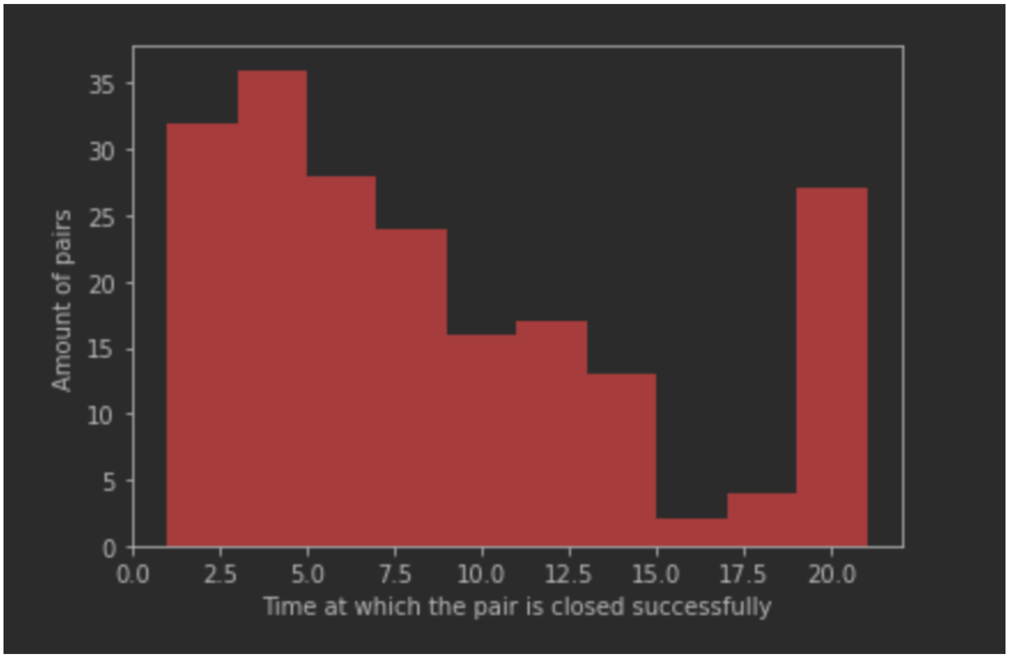

Figure 17: Histogram of the holding times: Reaching target with log-prices

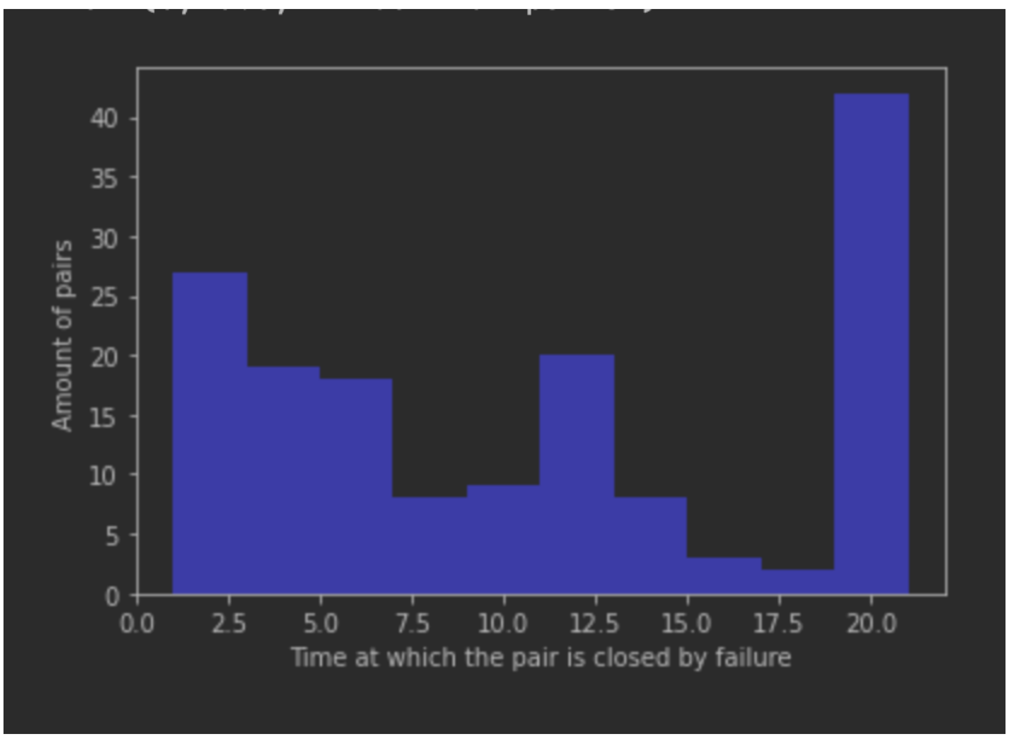

Figure 18: Histogram of the holding times: Stop-Loss condition with log-prices

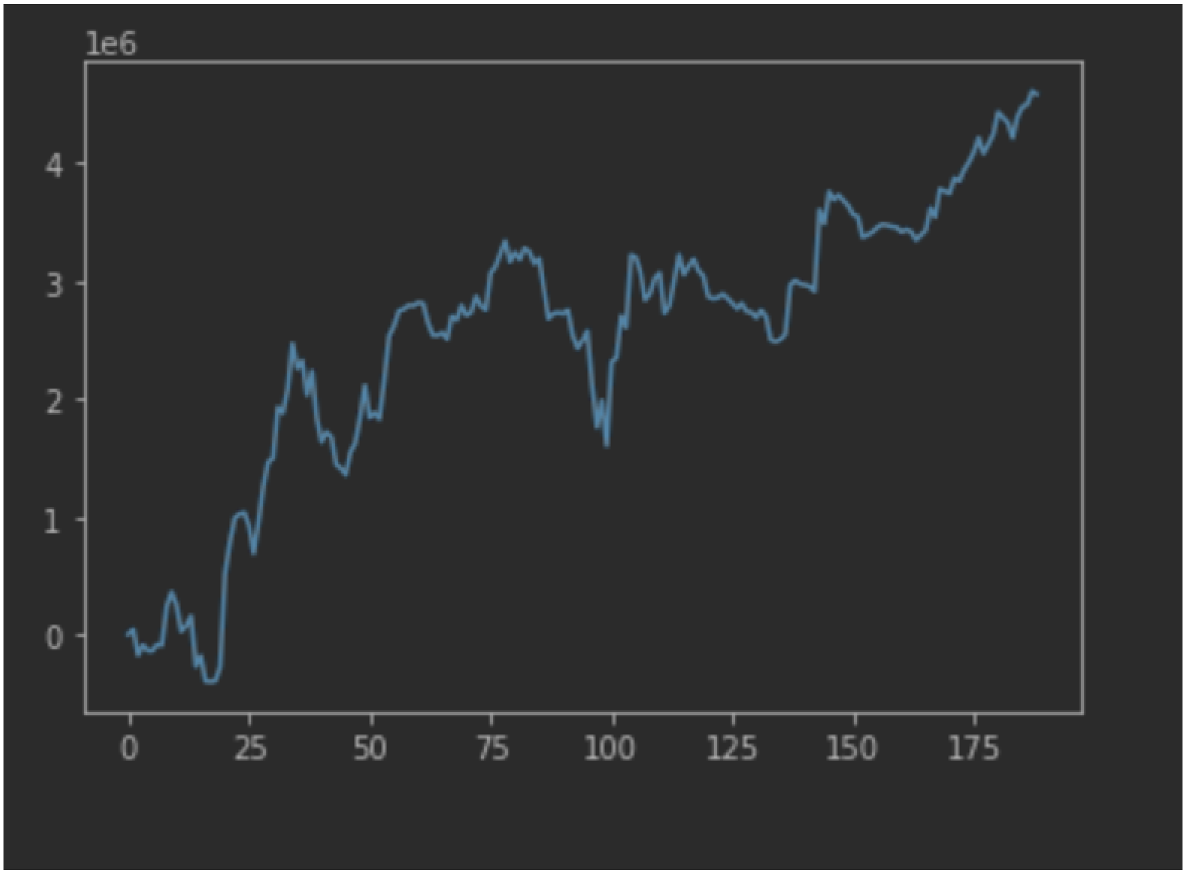

Figure 19: Hedged PnL with log-prices

```
Absolute PnL Sharpe ratio:  2.0813033930986604
Long PnL Sharpe ratio:  1.4415815139677552
Short PnL Sharpe ratio:  0.8022305081450141
Hedged PnL Sharpe ratio:  2.0882869633896624
```

Figure 20: Different Sharpe Ratios obtained with log-prices

It is hard to comment the effect of log-prices. Before implementing the strategy, I would expect to obtain a better Sharpe Ratio for the absolute PnL. However, this is not the case and both Sharpe ratios (question a and c) are really similar. Therefore, if no mistakes have been realized, log-prices may still lead to a better hedged portfolio.