

“HFT Execution Algo VS VWAP in KRX”

By 손승한 / Saturday7



High Frequency Trade,

Low Latency,

그 전에 시장 미시 구조를 이해하기 위해서



https://github.com/Ginger-Tec/saturday7_KRX_HFT

실습 데이터 및 코드 제공 (발표 자료 포함)

시장 미시 구조 데이터 소개

A3, B6, G7

시장 미시 구조 데이터 소개

A3: Trade Tick Data

[`tick_direction`]

0: 장전, 장마감 동시호가

1: ask(매도)에 존재하는 가격이 체결.

(매수자가 상대 호가 주문으로 매수함)

-1: bid(매수)에 존재하는 가격이 체결.

(매도자가 상대 호가 주문으로 매도함)

time	tick_price	tick_qty	tick_direction
2022-10-04 09:00:03.936	55300	1332445	0
2022-10-04 09:00:03.993	55200	1	-1
2022-10-04 09:00:04.000	55300	20	1
2022-10-04 09:00:04.022	55200	100	-1
2022-10-04 09:00:04.063	55200	4	-1
2022-10-04 09:00:04.075	55200	9	-1
2022-10-04 09:00:04.084	55200	700	-1
2022-10-04 09:00:04.100	55200	2218	-1
2022-10-04 09:00:04.102	55100	1847	-1
2022-10-04 09:00:04.105	55000	1859	-1
2022-10-04 09:00:04.107	54900	1307	-1
2022-10-04 09:00:04.108	54800	769	-1
2022-10-04 09:00:04.108	54800	5	-1
2022-10-04 09:00:04.126	55200	2	1
2022-10-04 09:00:04.133	55200	1	1
2022-10-04 09:00:04.138	55200	2	1
2022-10-04 09:00:04.154	55200	1	1
2022-10-04 09:00:04.178	54800	7	-1
2022-10-04 09:00:04.181	54800	100	-1
2022-10-04 09:00:04.205	54800	1	-1
2022-10-04 09:00:04.213	54900	1	1
2022-10-04 09:00:04.213	55000	1	1

시장 미시 구조 데이터 소개

B6: Order Book Data (Level 2 Data)

data_preview: Out 48

data_preview: Out 48 ×

100 행 × 100 rows × 4 columns pd.DataFrame

CSV ↴

time	ask_price	ask_qty	bid_price	bid_qty
2022-10-04 09:00:01.004	[55400, 55500, 55600, 0, 0, 0, 0, 0, 0, 0]	[34672, 44325, 15163, 0, 0, 0, 0, 0, 0, 0]	[55300, 55200, 55100, 0, 0, 0, 0, 0, 0, 0]	[881, 3031, 1847, 0, 0, 0, 0, 0, 0, 0]
2022-10-04 09:00:01.835	[55300, 55400, 55500, 0, 0, 0, 0, 0, 0, 0]	[12078, 37251, 44548, 0, 0, 0, 0, 0, 0, 0]	[55200, 55100, 55000, 0, 0, 0, 0, 0, 0, 0]	[3032, 1847, 1859, 0, 0, 0, 0, 0, 0, 0]
2022-10-04 09:00:02.487	[55300, 55400, 55500, 0, 0, 0, 0, 0, 0, 0]	[12078, 37252, 44548, 0, 0, 0, 0, 0, 0, 0]	[55200, 55100, 55000, 0, 0, 0, 0, 0, 0, 0]	[3032, 1847, 1859, 0, 0, 0, 0, 0, 0, 0]
2022-10-04 09:00:03.963	[55300, 55400, 55500, 55600, 55700, 55800, 55900,...]	[19478, 37252, 44548, 15174, 17051, 11807, 17363,...]	[55200, 55100, 55000, 54900, 54800, 54700, 54600,...]	[3032, 1847, 1859, 1307, 4027, 245, 71602, 21741,...]
2022-10-04 09:00:04.749	사이 동안 체결은 무수히 발생함.	[55000, 55100, 55200, 55300, 55400, 55500, 55600,...]	[35643, 26, 19987, 13570, 36917, 45484, 14700, 16...]	[54800, 54700, 54600, 54500, 54400, 54300, 54200,...]
2022-10-04 09:00:05.725	[54700, 54800, 54900, 55000, 55100, 55200, 55300,...]	[10, 17556, 126, 56208, 1036, 3960, 29415, 38044,...]	[54600, 54500, 54400, 54300, 54200, 54100, 54000,...]	[71155, 21725, 20854, 21384, 21213, 22856, 31000,...]
2022-10-04 09:00:06.541	[54800, 54900, 55000, 55100, 55200, 55300, 55400,...]	[2928, 148, 56889, 201, 9109, 20553, 37974, 46538,...]	[54700, 54600, 54500, 54400, 54300, 54200, 54100,...]	[6421, 16926, 21725, 20854, 21384, 21213, 22856, ...]
2022-10-04 09:00:07.451	[54700, 54800, 54900, 55000, 55100, 55200, 55300,...]	[587, 7156, 2333, 56490, 942, 4825, 22459, 37298,...]	[54600, 54500, 54400, 54300, 54200, 54100, 54000,...]	[36784, 21825, 20854, 21384, 21193, 22856, 31001,...]
2022-10-04 09:00:08.203	[54800, 54900, 55000, 55100, 55200, 55300, 55400,...]	[12433, 2987, 56794, 2683, 5079, 22597, 33041, 46...]	[54700, 54600, 54500, 54400, 54300, 54200, 54100,...]	[97, 34568, 24123, 20854, 21384, 21193, 22856, 31...]
2022-10-04 09:00:09.048	[54700, 54800, 54900, 55000, 55100, 55200, 55300,...]	[1, 11910, 2644, 56645, 2633, 4963, 22597, 33055,...]	[54600, 54500, 54400, 54300, 54200, 54100, 54000,...]	[34700, 24123, 20854, 21384, 21193, 26282, 35835,...]
2022-10-04 09:00:09.862	[54700, 54800, 54900, 55000, 55100, 55200, 55300,...]	[261, 11547, 741, 56568, 2718, 5008, 23343, 33062,...]	[54600, 54500, 54400, 54300, 54200, 54100, 54000,...]	[26189, 24270, 20854, 21384, 21193, 26282, 35835,...]
2022-10-04 09:00:10.640	[54700, 54800, 54900, 55000, 55100, 55200, 55300,...]	[15212, 11930, 410, 56619, 2825, 5008, 23339, 330...]	[54600, 54500, 54400, 54300, 54200, 54100, 54000,...]	[25850, 24270, 20854, 21384, 21193, 26282, 35835,...]
2022-10-04 09:00:11.431	[54700, 54800, 54900, 55000, 55100, 55200, 55300,...]	[14929, 12153, 913, 56752, 2825, 5117, 16346, 329...]	[54600, 54500, 54400, 54300, 54200, 54100, 54000,...]	[23713, 24770, 20854, 21386, 21193, 26282, 35835,...]
2022-10-04 09:00:12.352	[54700, 54800, 54900, 55000, 55100, 55200, 55300,...]	[25883, 8228, 914, 56855, 2844, 5119, 16466, 3297...]	[54600, 54500, 54400, 54300, 54200, 54100, 54000,...]	[17136, 24770, 20854, 21386, 21193, 26302, 35835,...]
2022-10-04 09:00:13.177	[54700, 54800, 54900, 55000, 55100, 55200, 55300,...]	[28775, 7845, 910, 56861, 2934, 5085, 19026, 3284...]	[54600, 54500, 54400, 54300, 54200, 54100, 54000,...]	[67771, 24770, 20858, 21386, 21193, 26302, 35835,...]
2022-10-04 09:00:13.933	[54700, 54800, 54900, 55000, 55100, 55200, 55300,...]	[13898, 7779, 946, 56865, 2997, 5115, 19068, 3281...]	[54600, 54500, 54400, 54300, 54200, 54100, 54000,...]	[45672, 24820, 20858, 21386, 21194, 26302, 35835,...]
2022-10-04 09:00:14.691	[54700, 54800, 54900, 55000, 55100, 55200, 55300,...]	[8933, 7702, 978, 56876, 3128, 5116, 18256, 32828...]	[54600, 54500, 54400, 54300, 54200, 54100, 54000,...]	[68622, 24839, 20858, 21386, 21194, 26302, 35835,...]
2022-10-04 09:00:15.551	[54800, 54900, 55000, 55100, 55200, 55300, 55400,...]	[7598, 1144, 56931, 952, 5117, 18290, 32800, 4676...]	[54700, 54600, 54500, 54400, 54300, 54200, 54100,...]	[20683, 82425, 25140, 20861, 21419, 21195, 26302,...]
2022-10-04 09:00:16.461	[54800, 54900, 55000, 55100, 55200, 55300, 55400,...]	[3345, 1229, 56936, 1012, 5118, 19270, 35327, 467...]	[54700, 54600, 54500, 54400, 54300, 54200, 54100,...]	[9270, 78835, 26575, 20974, 21419, 21195, 26303, ...]
2022-10-04 09:00:17.311	[54800, 54900, 55000, 55100, 55200, 55300, 55400,...]	[557, 2048, 56938, 1012, 5415, 19372, 38727, 4677...]	[54700, 54600, 54500, 54400, 54300, 54200, 54100,...]	[16556, 78882, 26576, 20974, 21519, 21195, 26305,...]
2022-10-04 09:00:18.075	[54900, 55000, 55100, 55200, 55300, 55400, 55500,...]	[1999, 56991, 1007, 5415, 19374, 38709, 46871, 15...]	[54700, 54600, 54500, 54400, 54300, 54200, 54100,...]	[17087, 78894, 26576, 20975, 21519, 21195, 26305,...]
2022-10-04 09:00:18.531	[54900, 55000, 55100, 55200, 55300, 55400, 55500,...]	[1936, 57174, 1007, 5415, 19380, 38734, 45872, 15...]	[54800, 54700, 54600, 54500, 54400, 54300, 54200,...]	[869, 17511, 79401, 26580, 20976, 21519, 21205, ...]

시장 미시 구조 데이터 소개

G7: 체결 틱 기반 예상 호가 정보

data_preview: Out 89

data_preview: Out 89 ×

100 행 × 100 rows × 7 columns pd.DataFrame ↗

CSV ↘

time	tick_price	tick_qty	tick_direction	ask_price	ask_qty	bid_price	bid_qty
2022-10-04 09:00:03.936	55300.0	1332445.0	0.0	[55400, 55500, 0, ...	[37252, 44548, 0...	[55200, 55100, 55000, 0, 0, 0, 0, 0, 0]	[3032, 1847, 1859, 0, 0, 0, 0, 0, 0]
2022-10-04 09:00:03.993	55200.0	1.0	-1.0	[55300, 55400, 55...	[19478, 37252, 4...	[55200, 55100, 55000, 54900, 54800, 54700, 54600, 54500, 54400, 54300]	[3031, 1847, 1859, 1307, 4027, 245, 71602, 21741, 20854, 21366]
2022-10-04 09:00:04.000	55300.0	20.0	1.0	[55300, 55400, 55...	[19458, 37252, 4...	[55200, 55100, 55000, 54900, 54800, 54700, 54600, 54500, 54400, 54300]	[3031, 1847, 1859, 1307, 4027, 245, 71602, 21741, 20854, 21366]
2022-10-04 09:00:04.022	55200.0	100.0	-1.0	[55300, 55400, 55...	[19458, 37252, 4...	[55200, 55100, 55000, 54900, 54800, 54700, 54600, 54500, 54400, 54300]	[2931, 1847, 1859, 1307, 4027, 245, 71602, 21741, 20854, 21366]
2022-10-04 09:00:04.063	55200.0	4.0	-1.0	[55300, 55400, 55...	[19458, 37252, 4...	[55200, 55100, 55000, 54900, 54800, 54700, 54600, 54500, 54400, 54300]	[2927, 1847, 1859, 1307, 4027, 245, 71602, 21741, 20854, 21366]
2022-10-04 09:00:04.075	55200.0	9.0	5.0	[19458, 37252, 4...	[55200, 55100, 55000, 54900, 54800, 54700, 54600, 54500, 54400, 54300]	[2918, 1847, 1859, 1307, 4027, 245, 71602, 21741, 20854, 21366]	
2022-10-04 09:00:04.084	55200.0	700.0	시장가 체결에 의한 호가 변경 상황	[19458, 37252, 4...	[55200, 55100, 55000, 54900, 54800, 54700, 54600, 54500, 54400, 54300]	[2218, 1847, 1859, 1307, 4027, 245, 71602, 21741, 20854, 21366]	
2022-10-04 09:00:04.100	55200.0	2218.0	-1.0	[55300, 55400, 55...	[19458, 37252, 4...	[55100, 55000, 54900, 54800, 54700, 54600, 54500, 54400, 54300, 0]	[1847, 1859, 1307, 4027, 245, 71602, 21741, 20854, 21366, 0]
2022-10-04 09:00:04.102	55100.0	1847.0	-1.0	[55300, 55400, 55...	[19458, 37252, 4...	[55000, 54900, 54800, 54700, 54600, 54500, 54400, 54300, 0, 0]	[1859, 1307, 4027, 245, 71602, 21741, 20854, 21366, 0, 0]
2022-10-04 09:00:04.105	55000.0	1859.0	-1.0	[55300, 55400, 55...	[19458, 37252, 4...	[54900, 54800, 54700, 54600, 54500, 54400, 54300, 0, 0, 0]	[1307, 4027, 245, 71602, 21741, 20854, 21366, 0, 0, 0]
2022-10-04 09:00:04.107	54900.0	1307.0	1.0	[55300, 55400, 55...	[19458, 37252, 4...	[54800, 54700, 54600, 54500, 54400, 54300, 0, 0, 0, 0]	[4027, 245, 71602, 21741, 20854, 21366, 0, 0, 0, 0]
2022-10-04 09:00:04.108	54800.0	769.0	-1.0	[55300, 55400, 55...	[19458, 37252, 4...	[54800, 54700, 54600, 54500, 54400, 54300, 0, 0, 0, 0]	[3258, 245, 71602, 21741, 20854, 21366, 0, 0, 0, 0]
2022-10-04 09:00:04.108	54800.0	5.0	-1.0	[55300, 55400, 55...	[19458, 37252, 4...	[54800, 54700, 54600, 54500, 54400, 54300, 0, 0, 0, 0]	[3253, 245, 71602, 21741, 20854, 21366, 0, 0, 0, 0]
2022-10-04 09:00:04.126	55200.0	2.0	1.0	[55300, 55400, 55...	[19458, 37252, 4...	[54800, 54700, 54600, 54500, 54400, 54300, 0, 0, 0, 0]	[3253, 245, 71602, 21741, 20854, 21366, 0, 0, 0, 0]
2022-10-04 09:00:04.133	55200.0	1.0	1.0	[55300, 55400, 55...	[19458, 37252, 4...	[54800, 54700, 54600, 54500, 54400, 54300, 0, 0, 0, 0]	[3253, 245, 71602, 21741, 20854, 21366, 0, 0, 0, 0]
2022-10-04 09:00:04.138	55200.0	2.0	1.0	[55300, 55400, 55...	[19458, 37252, 4...	[54800, 54700, 54600, 54500, 54400, 54300, 0, 0, 0, 0]	[3253, 245, 71602, 21741, 20854, 21366, 0, 0, 0, 0]
2022-10-04 09:00:04.154	55200.0	1.0	1.0	[55300, 55400, 55...	[19458, 37252, 4...	[54800, 54700, 54600, 54500, 54400, 54300, 0, 0, 0, 0]	[3253, 245, 71602, 21741, 20854, 21366, 0, 0, 0, 0]
2022-10-04 09:00:04.178	54800.0	7.0	-1.0	[55300, 55400, 55...	[19458, 37252, 4...	[54800, 54700, 54600, 54500, 54400, 54300, 0, 0, 0, 0]	[3246, 245, 71602, 21741, 20854, 21366, 0, 0, 0, 0]
2022-10-04 09:00:04.181	54800.0	100.0	-1.0	[55300, 55400, 55...	[19458, 37252, 4...	[54800, 54700, 54600, 54500, 54400, 54300, 0, 0, 0, 0]	[3146, 245, 71602, 21741, 20854, 21366, 0, 0, 0, 0]
2022-10-04 09:00:04.205	54800.0	1.0	-1.0	[55300, 55400, 55...	[19458, 37252, 4...	[54800, 54700, 54600, 54500, 54400, 54300, 0, 0, 0, 0]	[3145, 245, 71602, 21741, 20854, 21366, 0, 0, 0, 0]
2022-10-04 09:00:04.213	54900.0	1.0	1.0	[55300, 55400, 55...	[19458, 37252, 4...	[54800, 54700, 54600, 54500, 54400, 54300, 0, 0, 0, 0]	[3145, 245, 71602, 21741, 20854, 21366, 0, 0, 0, 0]
2022-10-04 09:00:04.213	55000.0	1.0	1.0	[55300, 55400, 55...	[19458, 37252, 4...	[54800, 54700, 54600, 54500, 54400, 54300, 0, 0, 0, 0]	[3145, 245, 71602, 21741, 20854, 21366, 0, 0, 0, 0]

SUM: 7231.0 12 셀, 4 행 8:3

시장 미시 구조 데이터 소개

B6 + G7 = S7: 체결 정보가 반영된 호가 흐름

data_preview: Out 88

data_preview: Out 88

100 행 × 10 columns pd.DataFrame

CSV

time	ask_price	ask_qty	bid_price	bid_qty	tick_price	tick_qty	tick_direction	best_ask_price	best_bid_price	mid_price
2022-10-04 09:00:01.004	[55400, 5550...	[34672, 443...	[55300, 5520...	[881, 3031,...	NaN	NaN	NaN	55400	55300	55350.0
2022-10-04 09:00:01.835	[55300, 5540...	[12078, 372...	[55200, 5510...	[3032, 1847...	NaN	NaN	NaN	55300	55200	55250.0
2022-10-04 09:00:02.487	[55300, 5540...	[12078, 372...	[55200, 5510...	[3032, 1847...	NaN	NaN	NaN	55300	55200	55250.0
2022-10-04 09:00:03.936	[55400, 5550...	[37252, 445...	[55200, 5510...	[3032, 1847...	55300.0	1332445.0	0.0	55400	55200	55300.0
2022-10-04 09:00:03.963	[55300, 5540...	[19478, 372...	[55200, 5510...	[3032, 1847...	NaN	NaN	NaN	55300	55200	55250.0
2022-10-04 09:00:03.993	[55300, 5540...	[19478, 372...	[55200, 5510...	[3031, 1847...	55200.0	1.0	-1.0	55300	55200	55250.0
2022-10-04 09:00:04.000	[55300, 5540...	[19458, 372...	[55200, 5510...	[3031, 1847...	55300.0	20.0	1.0	55300	55200	55250.0
2022-10-04 09:00:04.022	[55300, 5540...	[19458, 372...	[55200, 5510...	[2931, 1847...	55200.0	100.0	-1.0	55300	55200	55250.0
2022-10-04 09:00:04.063	[55300, 5540...	[19458, 372...	[55200, 5510...	[2927, 1847...	55200.0	4.0	-1.0	55300	55200	55250.0
2022-10-04 09:00:04.075	[55300, 5540...	[19458, 372...	[55200, 5510...	[2918, 1847...	55200.0	9.0	-1.0	55300	55200	55250.0
2022-10-04 09:00:04.084	[55300, 5540...	[19458, 372...	[55200, 5510...	[2218, 1847...	55200.0	700.0	-1.0	55300	55200	55250.0
2022-10-04 09:00:04.100	[55300, 5540...	[19458, 372...	[55100, 5500...	[1847, 1859...	55200.0	2218.0	-1.0	55300	55100	55200.0
2022-10-04 09:00:04.102	[55300, 5540...	[19458, 372...	[55000, 5490...	[1859, 1307...	55100.0	1847.0	-1.0	55300	55000	55150.0
2022-10-04 09:00:04.105	[55300, 5540...	[19458, 372...	[54900, 5480...	[1307, 4027...	55000.0	1859.0	-1.0	55300	54900	55100.0
2022-10-04 09:00:04.107	[55300, 5540...	[19458, 372...	[54800, 5470...	[4027, 245,...	54900.0	1307.0	-1.0	55300	54800	55050.0
2022-10-04 09:00:04.108	[55300, 5540...	[19458, 372...	[54800, 5470...	[3258, 245,...	54800.0	769.0	-1.0	55300	54800	55050.0
2022-10-04 09:00:04.108	[55300, 5540...	[19458, 372...	[54800, 5470...	[3253, 245,...	54800.0	5.0	-1.0	55300	54800	55050.0
2022-10-04 09:00:04.126	[55300, 5540...	[19458, 372...	[54800, 5470...	[3253, 245,...	55200.0	2.0	1.0	55300	54800	55050.0
2022-10-04 09:00:04.133	[55300, 5540...	[19458, 372...	[54800, 5470...	[3253, 245,...	55200.0	1.0	1.0	55300	54800	55050.0
2022-10-04 09:00:04.138	[55300, 5540...	[19458, 372...	[54800, 5470...	[3253, 245,...	55200.0	2.0	1.0	55300	54800	55050.0
2022-10-04 09:00:04.154	[55300, 5540...	[19458, 372...	[54800, 5470...	[3253, 245,...	55200.0	1.0	1.0	55300	54800	55050.0
2022-10-04 09:00:04.178	[55300, 5540...	[19458, 372...	[54800, 5470...	[3246, 245,...	54800.0	7.0	-1.0	55300	54800	55050.0

체결 정보가 없다면, B6 호가 정보.

미시 구조 데이터의 피쳐링

- Order Book Imbalance
- Order Flow
- Order Flow Imbalance
- Order Active

Order Book Imbalance

시장 미시 구조의 피쳐링

		qty
order	price	
ask	55700	17251
	55600	15169
	55500	46766
	55400	35327
	55300	19270
	55200	5118
	55100	1012
	55000	56936
	54900	1229
	54800	520
bid	54700	9181
	54600	78835
	54500	26575
	54400	20974
	54300	21419
	54200	21195
	54100	26303
	54000	35840
	53900	84753
	53800	93254

- ask 호가의 qty와 bid 호가 qty의 비교합니다.
- 단순히 bid, ask의 qty 합을 빼거나 나눠볼 수 있습니다.
- 혹은 적절한 가중치를 줘서 bid, ask를 비교할 수 있습니다.

Order Flow

시장 미시 구조의 피쳐링

		qty
order	price	
ask	55700	17251
	55600	15169
	55500	46766
	55400	35327
	55300	19270
	55200	5118
	55100	1012
	55000	56936
	54900	1229
	54800	520
bid	54700	9181
	54600	78835
	54500	26575
	54400	20974
	54300	21419
	54200	21195
	54100	26303
	54000	35840
	53900	84753
	53800	93254

		qty
order	price	
ask	55700	17272
	55600	15169
	55500	46777
	55400	38727
	55300	19372
	55200	5415
	55100	1012
	55000	56938
	54900	2048
	54800	557
bid	54700	16556
	54600	78882
	54500	26576
	54400	20974
	54300	21519
	54200	21195
	54100	26305
	54000	35840
	53900	83752
	53800	93254

		qty_diff
order	price	
ask	55700	21
	55600	0
	55500	11
	55400	3400
	55300	102
	55200	297
	55100	0
	55000	2
	54900	819
	54800	37
bid	54700	7375
	54600	47
	54500	1
	54400	0
	54300	100
	54200	0
	54100	2
	54000	0
	53900	-1001
	53800	0

Order Flow Imbalance

시장 미시 구조의 피쳐링

		qty_diff
order	price	
ask	55700	21
	55600	0
	55500	11
	55400	3400
	55300	102
	55200	297
	55100	0
	55000	2
	54900	819
bid	54800	37
	54700	7375
	54600	47
	54500	1
	54400	0
	54300	100
	54200	0
	54100	2
	54000	0
	53900	-1001
	53800	0

- Order Flow의 불균형을 계산.

Order Active

시장 미시 구조의 피쳐링

전체 오더북 qty 대비 qty_diff

		qty_diff
order	price	
ask	55700	21
	55600	0
	55500	11
	55400	3400
	55300	102
	55200	297
	55100	0
	55000	2
	54900	819
	54800	37
	54700	7375
	54600	47
bid	54500	1
	54400	0
	54300	100
	54200	0
	54100	2
	54000	0
	53900	-1001
	53800	0

		qty
order	price	
ask	55700	17272
	55600	15169
	55500	46777
	55400	38727
	55300	19372
	55200	5415
	55100	1012
	55000	56938
	54900	2048
	54800	557
	54700	16556
	54600	78882
bid	54500	26576
	54400	20974
	54300	21519
	54200	21195
	54100	26305
	54000	35840
	53900	83752
	53800	93254

qty_diff 대비 체결량

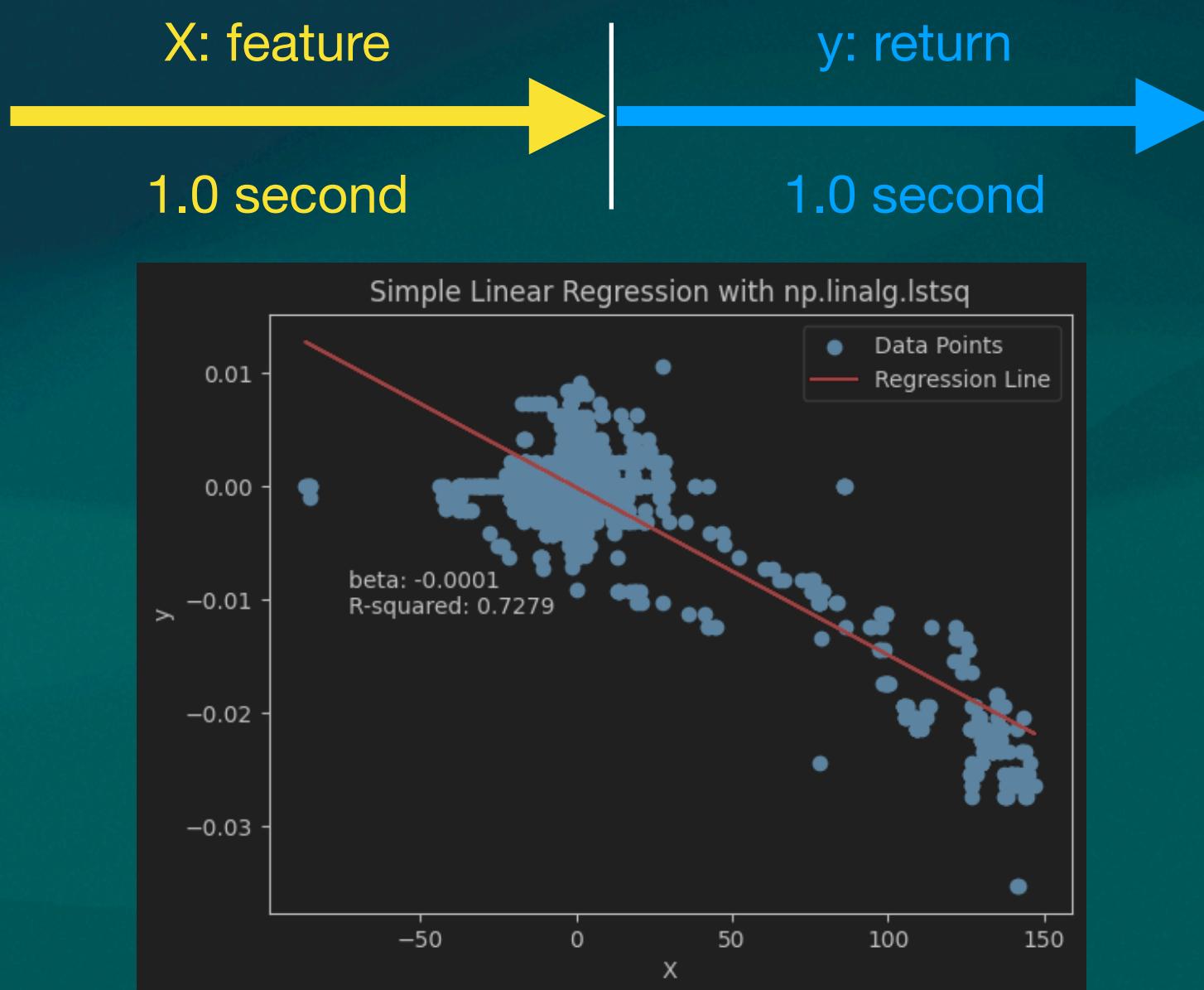
visualizer_order_book: Out 65	
visualizer_order_book: Out 65 ×	
time	tick_qty
2022-10-04 09:00:03.936	1332445
2022-10-04 09:00:03.993	1
2022-10-04 09:00:04.000	20
2022-10-04 09:00:04.022	100
2022-10-04 09:00:04.063	4
2022-10-04 09:00:04.075	9
2022-10-04 09:00:04.084	700
2022-10-04 09:00:04.100	2218
2022-10-04 09:00:04.102	1847
2022-10-04 09:00:04.105	1859
2022-10-04 09:00:04.107	1307
2022-10-04 09:00:04.108	769
2022-10-04 09:00:04.108	5
2022-10-04 09:00:04.126	2
2022-10-04 09:00:04.133	1
2022-10-04 09:00:04.138	2
2022-10-04 09:00:04.154	1
2022-10-04 09:00:04.178	7
2022-10-04 09:00:04.181	100
2022-10-04 09:00:04.205	1
2022-10-04 09:00:04.213	1
2022-10-04 09:00:04.213	1
2022-10-04 09:00:04.219	1000
2022-10-04 09:00:04.240	202
2022-10-04 09:00:04.313	23

qty_diff		
order	price	
ask	55700	21
	55600	0
	55500	11
	55400	3400
	55300	102
	55200	297
	55100	0
	55000	2
	54900	819
	54800	37
	54700	7375
	54600	47
bid	54500	1
	54400	0
	54300	100
	54200	0
	54100	2
	54000	0
	53900	-1001
	53800	0

피쳐 데이터의 통계 확인(feat. OLS)

y: forward return(mid price), n second

X: feature rolling sum, m second



$$y \sim X$$

$$\beta, R^2$$

X의 개수(feature 종류): 28 개

rolling n sec: 1~10 sec

분석종목 수: 약 350 종목

분석 영업일: 약 60일

총 분석 row: 약 7000만

피쳐 데이터의 통계 확인(feat. OLS)

analysis_feature2: Out 27

analysis_feature2: Out 27 ×

1-100 < > 73905800 rows × 11 columns pd.DataFrame ↗

analysis_feature2: Out 22

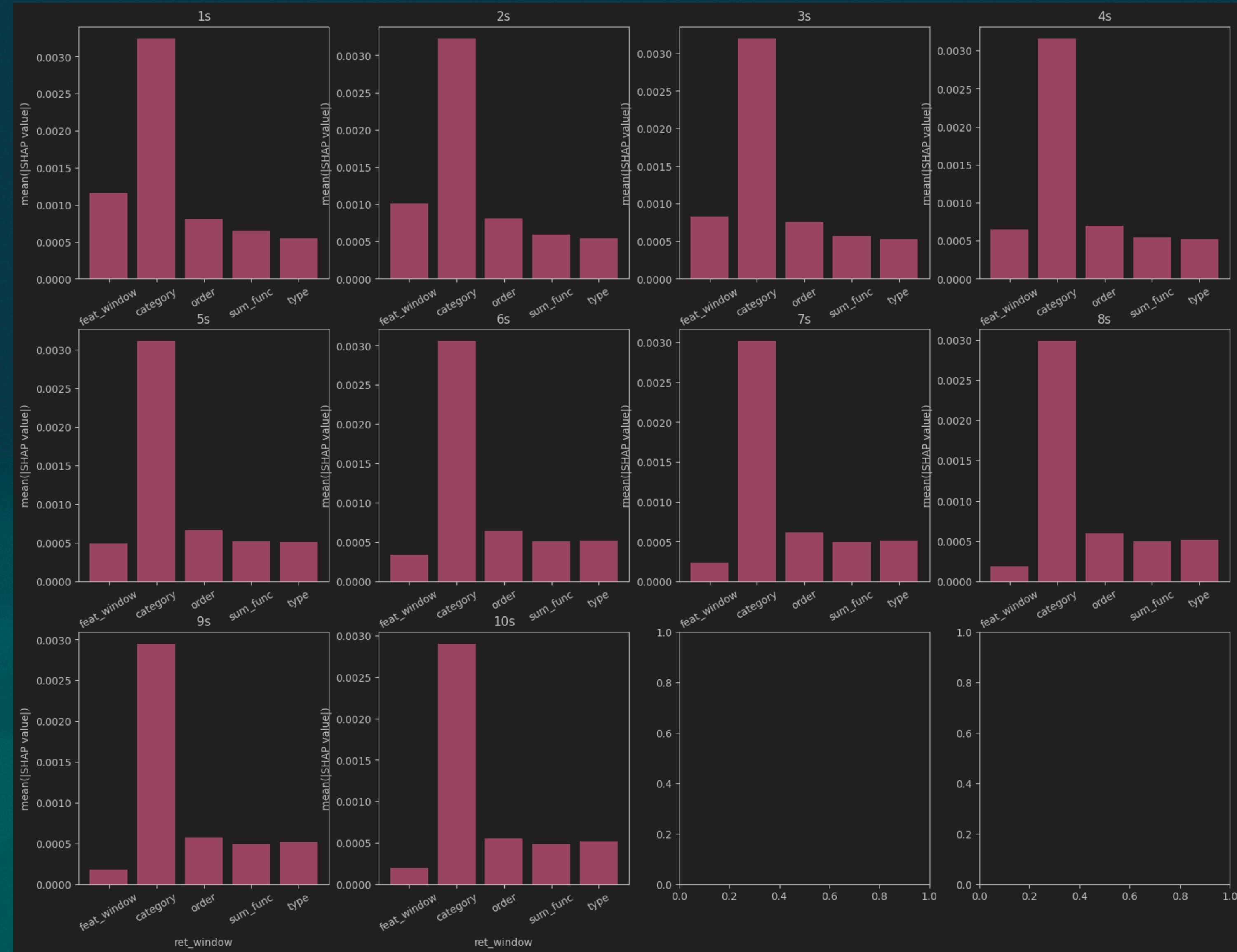
analysis_feature2: Out 22 ×

10 행 < > 10 열 × 6 행 pd.DataFrame ↗

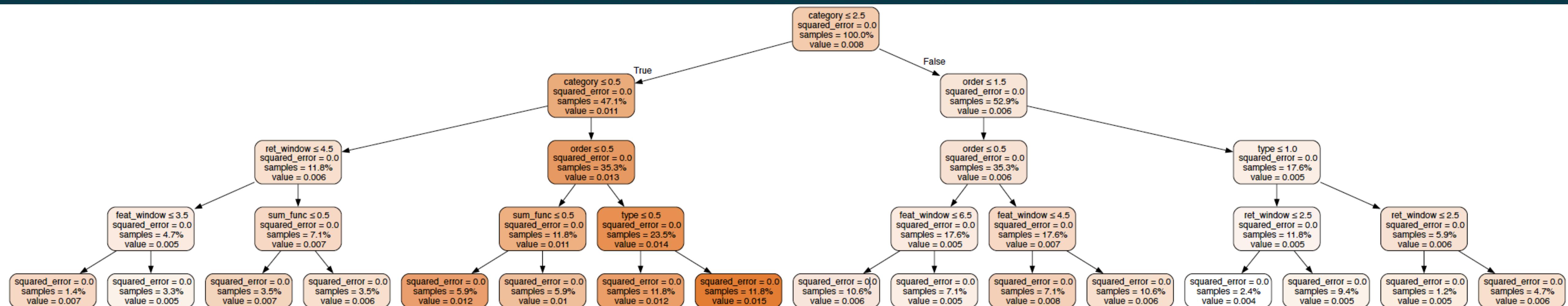
CSV ↓ ⌂

ymd	isin	ret_window	feat_window	feat_name	beta	R2	category	order	sum_func	type
9	20221209	KR7131290009	1s	1s	of_b_wt_ratio	0.008687	0.730144	of	wt	ratio
1917	20221121	KR7235980000	7s	6s	ofi_ba_wt_sub	-0.000148	0.727910	ofi	wt	sub
349	20221209	KR7131290009	1s	2s	of_b_wt_ratio	0.008651	0.727217	of	wt	ratio
8	20221110	KR7086900008	1s	1s	of_b_eq_ratio	0.003068	0.715781	of	eq	ratio
1951	20221121	KR7235980000	8s	6s	ofi_ba_wt_sub	-0.000148	0.714362	ofi	wt	sub
43	20221209	KR7131290009	2s	1s	of_b_wt_ratio	0.008674	0.709647	of	wt	ratio
2019	20221121	KR7235980000	10s	6s	ofi_ba_wt_sub	-0.000144	0.707217	ofi	wt	sub
383	20221209	KR7131290009	2s	2s	of_b_wt_ratio	0.008639	0.706827	of	wt	ratio
9	20221222	KR7215600008	1s	1s	of_b_wt_ratio	0.006338	0.701698	of	wt	ratio
1982	20221004	KR7298380007	9s	6s	of_a_eq_ratio	0.004750	0.701571	of	eq	ratio
9	20221018	KR7033290008	1s	1s	of_b_wt_ratio	0.001198	0.700652	of	wt	ratio
8	20221018	KR7033290008	1s	1s	of_b_eq_ratio	0.001200	0.700019	of	eq	ratio
1611	20221004	KR7237690003	8s	5s	ofi_ba_wt_sub	-0.000162	0.699067	ofi	wt	sub
1605	20221004	KR7237690003	8s	5s	of_a_wt	0.000176	0.698285	of	wt	none
1573	20221121	KR7235980000	7s	5s	of_b_wt_ratio	-0.012145	0.697242	of	wt	ratio
22	20221018	KR7033290008	1s	1s	oa_b_eq_ratio	-0.001193	0.694694	oa	eq	ratio
23	20221018	KR7033290008	1s	1s	oa_b_wt_ratio	-0.001190	0.	obi	wt	ratio
9	20221110	KR7086900008	1s	1s	of_b_wt_ratio	0.002676	0.	ofi	wt	sub
1985	20221121	KR7235980000	9s	6s	ofi_ba_wt_sub	-0.000140	0.	ba	wt	sub
2016	20221004	KR7298380007	10s	6s	of_a_eq_ratio	0.004675	0.	wt	eq	ratio
348	20221110	KR7086900008	1s	2s	of_b_eq_ratio	0.002957	0.	ofi	wt	sub
1949	20221121	KR7235980000	8s	6s	of_a_wt_ratio	0.003759	0.	ba	wt	sub
1948	20221004	KR7298380007	8s	6s	of_a_eq_ratio	0.004634	0.	a	eq	ratio
8	20221209	KR7131290009	1s	1s	of_b_eq_ratio	0.008445	0.	ba	wt	sub
1641	20221121	KR7235980000	9s	5s	of_b_wt_ratio	-0.011716	0.	ab	eq	ratio
23	20221209	KR7131290009	1s	1s	na_h_wt_ratio	-0.008207	0.	5s	wt	sub
								6s	eq	ratio
								7s	wt	sub
								8s	eq	ratio
								9s	wt	sub
								10s	eq	ratio

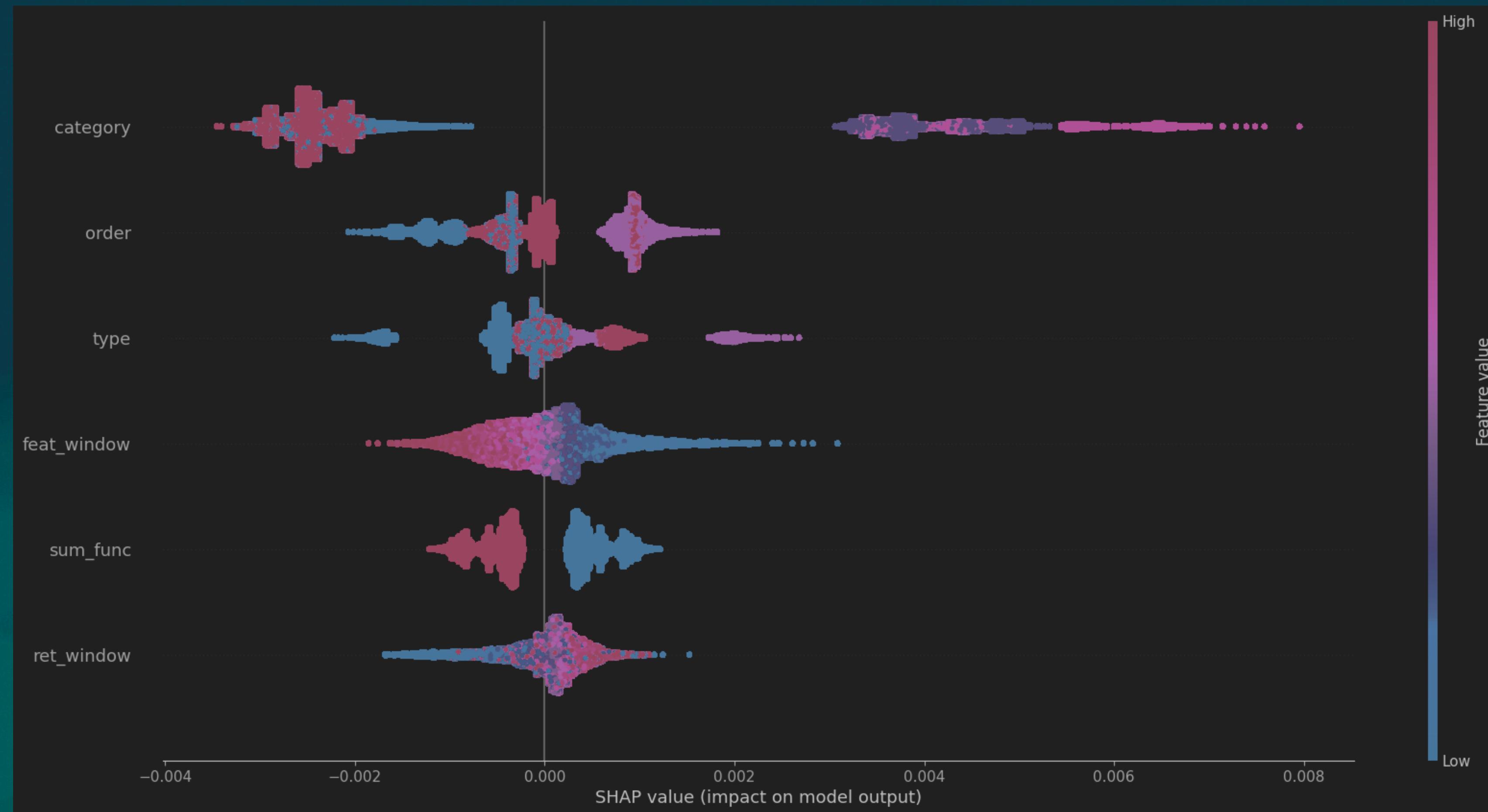
피쳐 데이터의 통계 확인(feat. shap)



피쳐 데이터의 통계 확인(feat. sklearn)



피쳐 데이터의 통계 확인(feat. shap)



틱 - 오더북 데이터 기반 백테스트

Python, Cython, Rust

데이터 타입

Raw 데이터 타입

- Order Book Delta
- OrderBookDeltas (L1/L2/L3)
- Quote Tick**
- Trade Tick**
- Bar

Bar 데이터 타입

- TIME (초, 분, 시, 일 ..)
- TICK (RUNS, IMBALANCE)
- VOLUME (RUNS, IMBALANCE)**
- VALUE (RUNS, IMBALANCE)

Price Type

- BID
- ASK
- MID
- LAST**

Volume bar, Dollar bar



Figure 1. Trade, time-based, volume and dollar bars depiction. Asterisks are plotted every time a candle is sampled

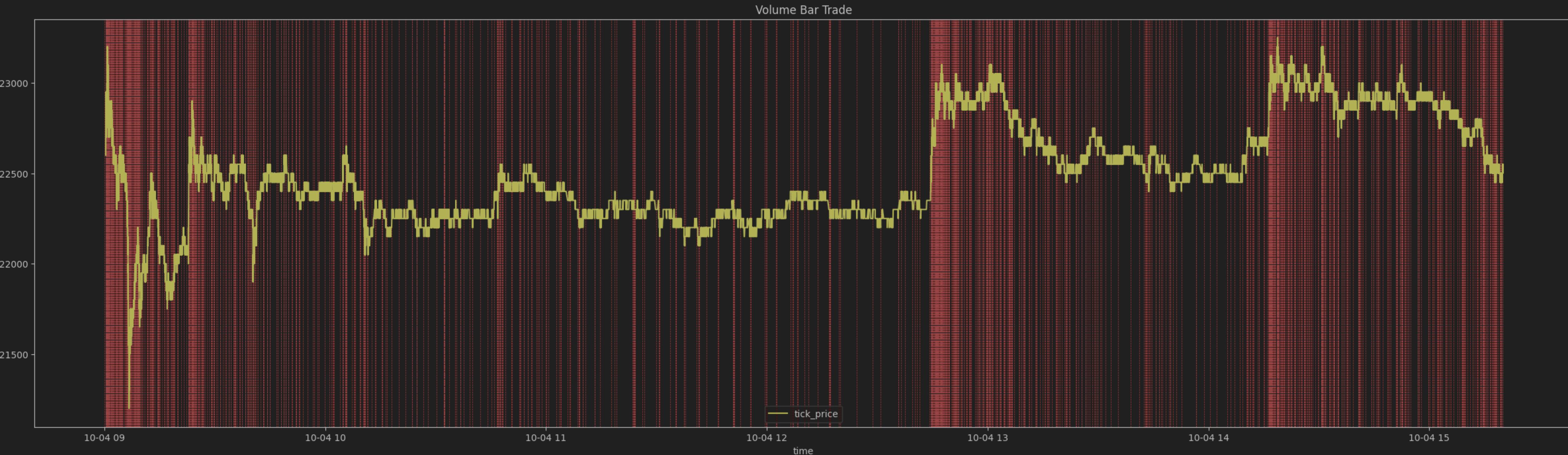
Volume bar 기반으로 VWAP 거래.

$$VWAP = \frac{\sum_{t=1}^T P_t V_t}{V_T}$$

t 시점의 체결 가격
t 시점에 발생한 거래량
총 거래량

$$vwap = \frac{\sum_{t=1}^T p_t v_t}{v_T}$$

t 시점에 집행한 가격
t 시점에 집행한 수량
총 집행 수량



Volume bar 기반으로 VWAP 거래.

$$VWAP = \frac{\sum_{t=1}^T P_t V_t}{V_T}$$

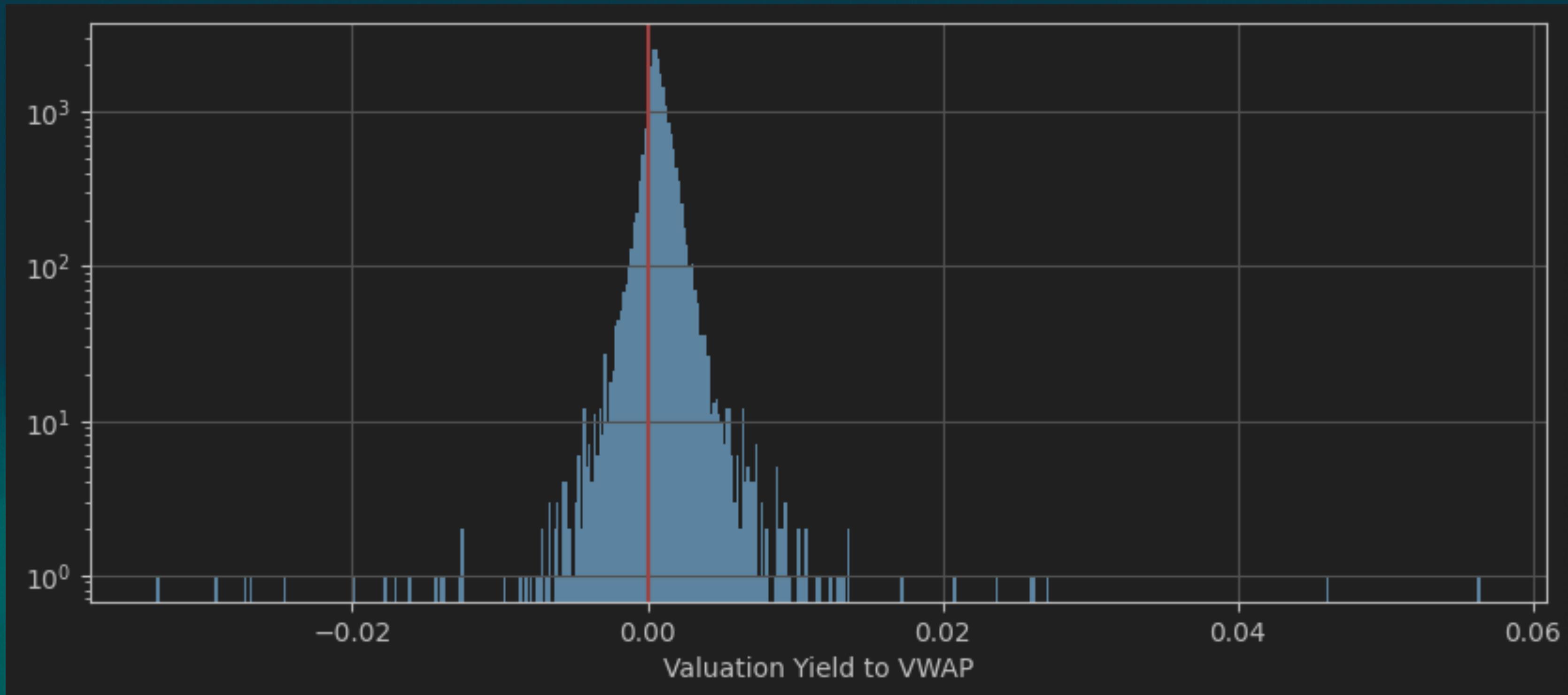
t 시점의 체결 가격
t 시점에 발생한 거래량

$$vwap = \frac{\sum_{t=1}^T p_t v_t}{v_T}$$

t 시점에 집행한 가격
t 시점에 집행한 수량

총 거래량
총 집행 수량

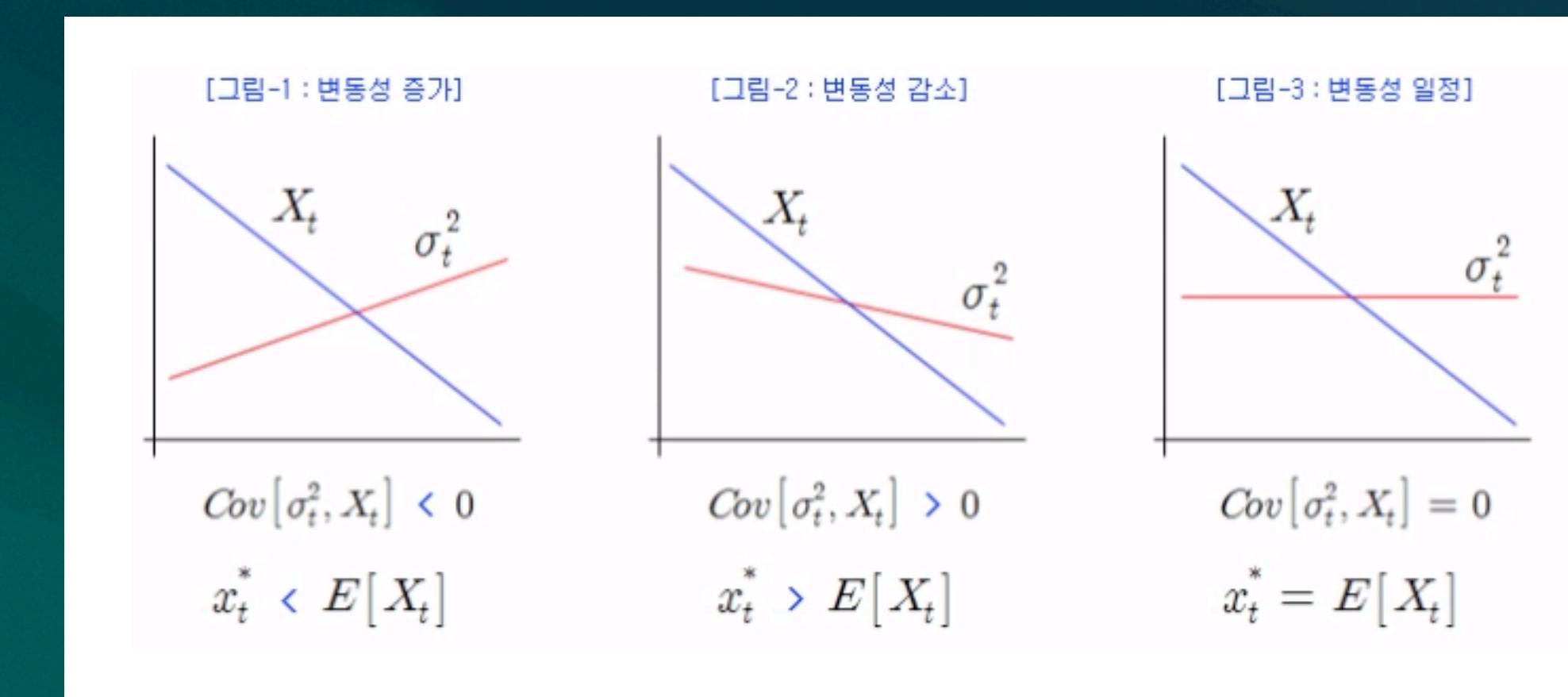
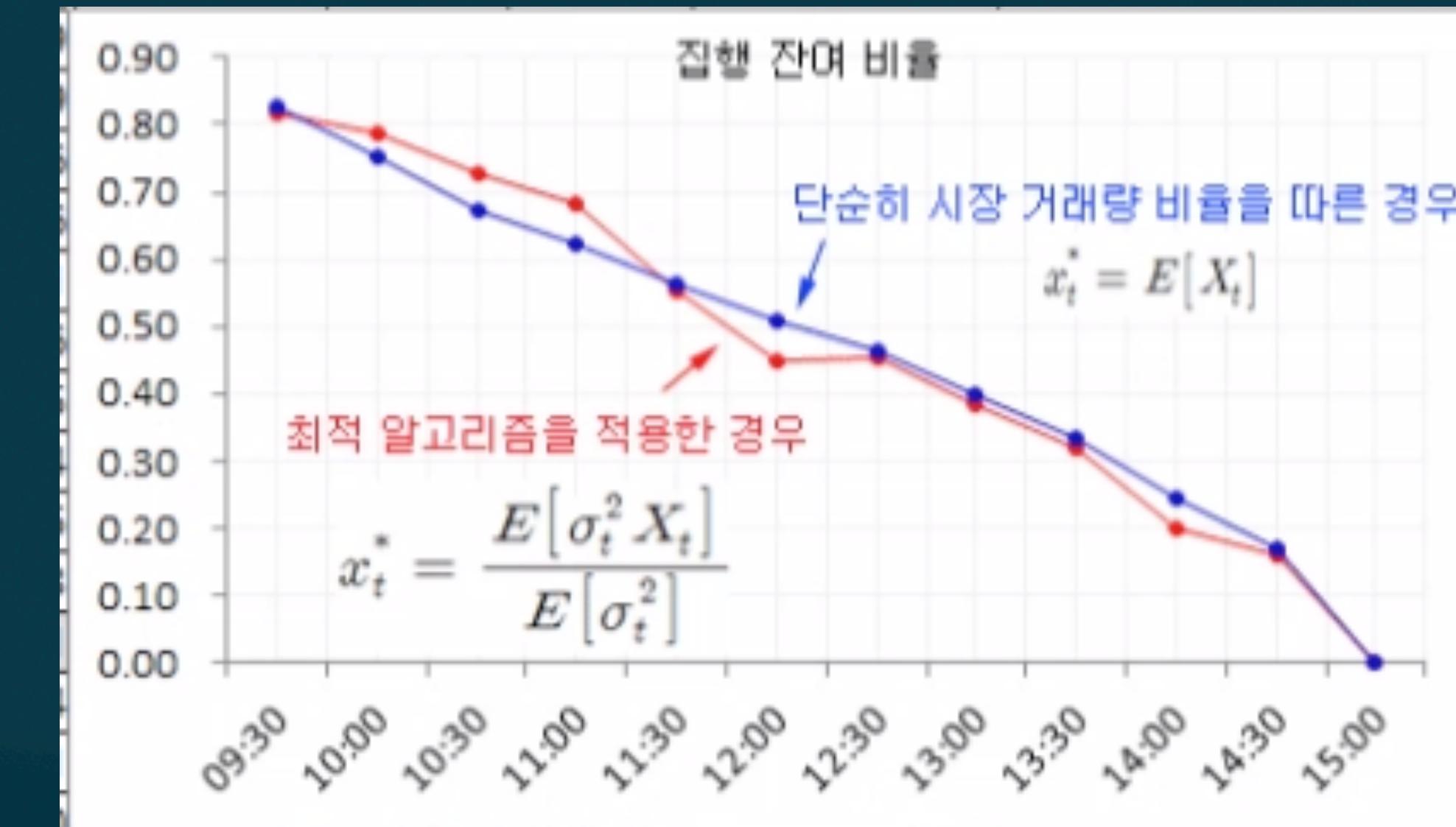
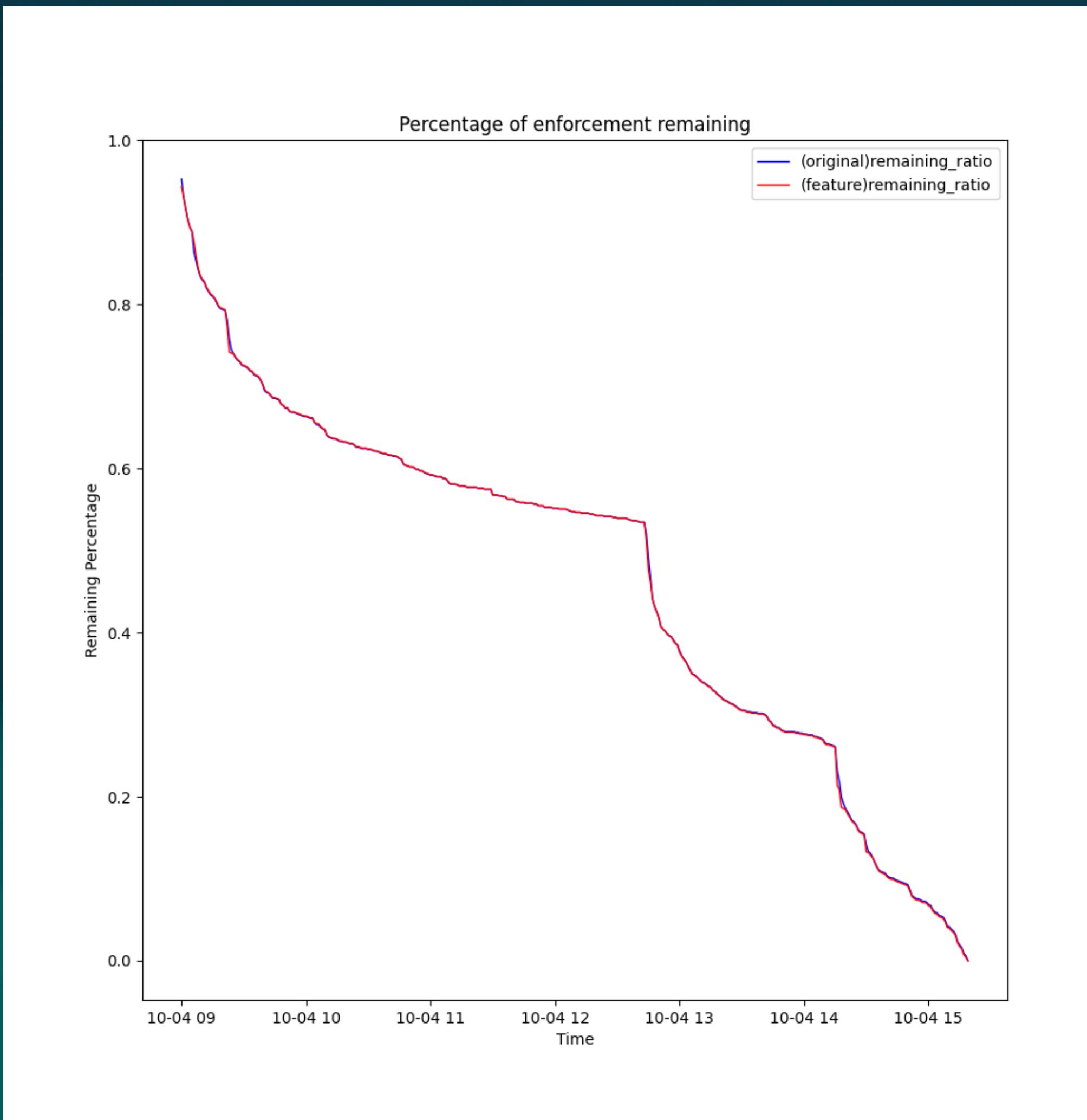
vwap - VWAP
평균 0.0729%



backtest_vwap2: Out 37

	vwap_v_bar_ratio
count	21797.00000
mean	0.000729
std	0.001375
min	-0.033268
25%	0.000239
50%	0.000630
75%	0.001159
max	0.056445
skew	3.651165
kurtosis	263.376335

Feature 정보를 활용한 volume bar 주문 수량 조절



Feature 정보를 활용한 volume bar 주문 수량 조절

feat_vwap - VWAP

평균 0.0876%

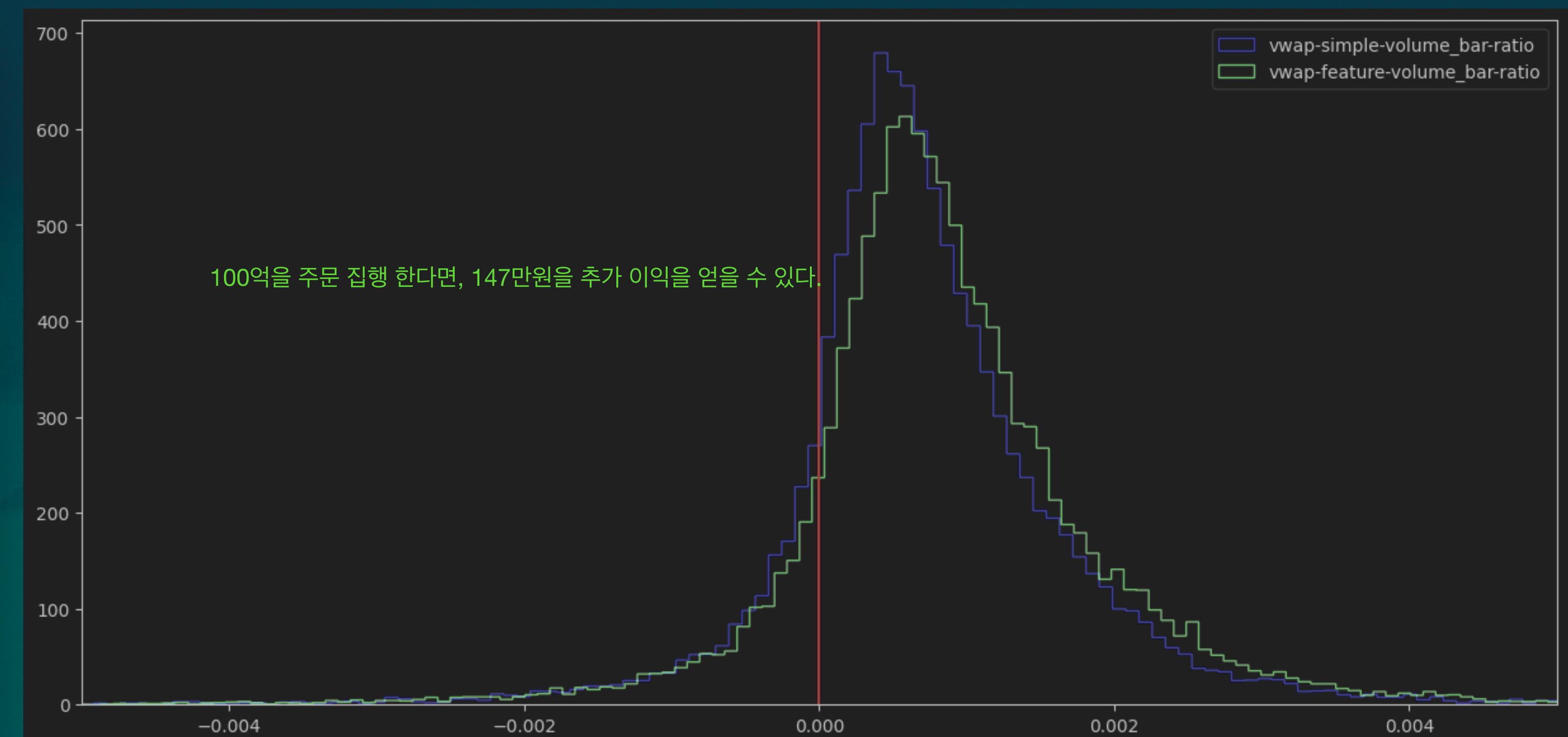
기존 vwap 보다 0.0147% 증가(개선)

● ● ● backtest_vwap_feature: Out 20

backtest_vwap_feature: Out 20 × :

◀ < 10 행 ▶ > ▶>

	vwap_v_bar_ratio
count	21797.000000
mean	0.000876
std	0.001435
min	-0.033122
25%	0.000336
50%	0.000767
75%	0.001342
max	0.051463
skew	2.903940
kurtosis	197.017476



마무리, 복습

1. 시장 미시 데이터 구조와 전처리 방법을 알아보았습니다.
2. 미시 데이터를 피쳐링하는 몇 가지 방법론을 알아봤습니다.
3. 피쳐를 평가하기 위해 선형 회귀를 사용할 수 있습니다.
4. 피쳐 종류에 따라 설명에 미치는 영향을 결정트리나 SHAP으로 분석해 봤습니다.
5. Time bar 가 아닌 Volume bar를 기준으로 거래할 수 있습니다.
6. Volume bar를 잘 활용하면 이론적 VWAP 가격에 체결시킬 수 있습니다.
7. Volume Bar의 주문 수량을 분석한 Feature로 잘 조절하면 기존 Volume bar 집행보다 더 나은 가격에 주문을 집행할 수 있었습니다.
→ 시장 미시 구조 feature를 다양하게 활용할 수 있다.

[감사한 분들]

SATURDAY7 멤버분들(한글 순) : {머니플로우님, 몬테카를로님, 시장의구조님, 시카리우스님, 오늘만큼은님, 프로집중러님, 효열님, 케빈님}

[분석 데이터 정보]

종목 : kospi 200 + kosdaq 150

기간 : 2022.10.04 ~ 2022.12.28

[python lib]

pandas==1.5.3

modin==0.23.1

numpy==1.24.4

pyarrow==13.0.0

numba==0.57.1

[사용한 머신]

Mac m1 pro, 32GB Ram, 10core

Workstation, 512GB Ram, 40core

참고 문헌

Enzo, Stephen. 2015, “Volume Weighted Average Price Optimal Execution”

Lopez de Prado Marcos, Advances in Financial Machine Learning (WILEY, 2018), 26

Advanced candlesticks for machine learning (ii): volume and dollar bars (<https://towardsdatascience.com/advanced-candlesticks-for-machine-learning-ii-volume-and-dollar-bars-6cda27e3201d>)

아마퀀트, 43. 주문 집행 알고리즘 (4) - VWAP 집행 알고리즘 (1) (<https://m.blog.naver.com/chunjein/220645305401>)

아마퀀트, 45. 주문 집행 알고리즘 (6) - Konishi의 최적 집행 알고리즘 적용 예시 (<https://m.blog.naver.com/chunjein/220656694991>)