

前面提到我们

针对期货交易所公开发布的交易数据进行整理的目标是

将期货交易所发布的原始数据按照期货品种和交易所会员进行分类梳理，并建立起期货品种与交易会员之间的关系。

为此，我们在前面讲了大概思路，需要熟悉原始数据的结构，明确需要梳理出来的结果数据结构，接下来就是想办法把原始数据进行分解关联汇总按照我们需要的结构进行输出。

由于数据梳理汇总统计过程涉及很多细节，我想了半天也没想好怎么顺序的把这个过程讲明白，这里我干脆采用结果反推的方式做一个简单介绍，看到结果结合前面介绍过的原始数据格式，再加上我的简单介绍希望能让大家明白这一步数据梳理的技术技巧。

首先看看行情数据梳理结果，我对行情数据梳理结果分为合约行情数据和品种行情数据两类，其中合约行情数据梳理结果如下图：

合约代码	合约名称	合约单位	最新价	涨跌额	今收盘	今开盘	最高价	最低价	成交量(手)	持仓量	持仓量(手)	交割月份
CU1201	铜	5	38900	-100	38900	39000	39100	38800	100	100	100	12-01
CU1202	铜	5	38800	-100	38800	38900	39000	38700	100	100	100	12-02
CU1203	铜	5	38700	-100	38700	38800	38900	38600	100	100	100	12-03
CU1204	铜	5	38600	-100	38600	38700	38800	38500	100	100	100	12-04
CU1205	铜	5	38500	-100	38500	38600	38700	38400	100	100	100	12-05
CU1206	铜	5	38400	-100	38400	38500	38600	38300	100	100	100	12-06
CU1207	铜	5	38300	-100	38300	38400	38500	38200	100	100	100	12-07
CU1208	铜	5	38200	-100	38200	38300	38400	38100	100	100	100	12-08
CU1209	铜	5	38100	-100	38100	38200	38300	38000	100	100	100	12-09
CU1210	铜	5	38000	-100	38000	38100	38200	37900	100	100	100	12-10
CU1211	铜	5	37900	-100	37900	38000	38100	37800	100	100	100	12-11
CU1212	铜	5	37800	-100	37800	37900	38000	37700	100	100	100	12-12

比较之下我们只需要将日期单独作为一列放，将合约代码前两位分离出来作为品种代码，并且关联品种代码写入名称即可，注意这里品种代码名称对应表在做数据梳理之前可以手动制作一份，因为数据有限所费工时很少。

三大交易所(包括金融期货交易所以及广州期货交易所就是五大，目前重点关注三大)每天的合约行情数据梳理出来大概有1000条左右的数据，由于合约持续时间一般就几个月，对于我们分析较长时间短的行情非常不利，所以有必要合并生成品种行情数据，三大交易所所有商品一共也就几十个，下面是对上海期货交易所商品合约行情进行合并之后形成的品种行情数据例子：

会员多空持仓成交数据例子

A1	A	B	C	D	E	F	G	H	I	J
1	日期	品种代码	合约代码	会员简称	多仓	多仓增减	空仓	空仓增减	成交量	成交量增减
2	20230331	ag	ag2304	中银期货	80	0	0	0	0	0
3	20230331	ag	ag2304	兴证期货	66	-5	0	0	5	0
4	20230331	ag	ag2304	国信期货	0	0	68	-1	1	0
5	20230331	ag	ag2304	华安期货	150	-1	500	-1	2	0
6	20230331	ag	ag2304	平安期货	0	0	0	0	31	21
7	20230331	ag	ag2304	一德期货	68	0	0	0	0	0
8	20230331	ag	ag2304	华联期货	100	0	100	0	0	0
9	20230331	ag	ag2304	中泰期货	0	0	64	28	84	51
10	20230331	ag	ag2304	银河期货	0	0	134	-24	74	-113
11	20230331	ag	ag2304	中信建投	248	-2	468	49	53	-25
12	20230331	ag	ag2304	华闻期货	0	0	46	14	179	2
13	20230331	ag	ag2304	华泰期货	0	0	76	-27	218	-30
14	20230331	ag	ag2304	广发期货	0	0	66	-2	29	-11
15	20230331	ag	ag2304	大地期货	10	0	0	0	0	0
16	20230331	ag	ag2304	兴业期货	0	0	0	0	37	27
17	20230331	ag	ag2304	东证期货	34	-21	0	0	103	-10
18	20230331	ag	ag2304	中行期货	0	0	0	0	320	320
19	20230331	ag	ag2304	创元期货	8	0	0	0	0	0
20	20230331	ag	ag2304	中信期货	892	-31	0	0	319	-247
21	20230331	ag	ag2304	国投安信	0	0	1474	0	0	0
22	20230331	ag	ag2304	海通期货	92	0	272	-38	52	41
23	20230331	ag	ag2304	新湖期货	804	0	800	0	0	0
24	20230331	ag	ag2304	申万期货	16	0	0	0	0	0
25	20230331	ag	ag2304	国贸期货	80	-3	76	0	187	37
26	20230331	ag	ag2304	山金期货	202	-4	0	0	4	0
27	20230331	ag	ag2304	华金期货	0	0	1320	0	0	0
28	20230331	ag	ag2304	金瑞期货	994	-1	0	0	1	0
29	20230331	ag	ag2304	长江期货	0	0	50	-12	78	12
30	20230331	ag	ag2304	南华期货	0	0	82	1	1	0

合约持仓成交数据看上去并没有什么特点，下面看看期货交易所下载的持仓成交排名数据Excel文件样子，大家可以对照一下，看看哪些地方需要如何处理：

A1	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	日期	品种代码	品种名称	主力合约	加权价	会员简称	会员持仓	会员持仓	品种持仓	品种持仓	品种持仓	品种持仓	品种持仓	品种持仓	品种持仓	品种持仓
2	20230331	ag	白糖	ag2306	5314.36	一德期货	ag2306	4	10457	96	10457	96	0	0	192	0
3	20230331	ag	白糖	ag2306	5314.36	上海中期	ag2312	2	341	51	1284	52	1043	1	57	0
4	20230331	ag	白糖	ag2306	5314.36	东吴期货	ag2308	2	-514	-36	999	-230	1512	-194	424	0
5	20230331	ag	白糖	ag2306	5314.36	东吴期货	ag2308	3	8120	-27	5251	36	1131	63	99	0
6	20230331	ag	白糖	ag2306	5314.36	东吴期货	ag2308	4	6381	-68	21281	-126	12900	-68	17926	9139
7	20230331	ag	白糖	ag2306	5314.36	东方财富	ag2306	4	-11016	-690	0	0	11016	680	31220	8440
8	20230331	ag	白糖	ag2306	5314.36	五海期货	ag2305	1	7	6	267	0	260	0	0	0
9	20230331	ag	白糖	ag2306	5314.36	东吴期货	ag2306	1	10409	1263	10409	1263	0	0	1263	0
10	20230331	ag	白糖	ag2306	5314.36	东证期货	ag2308	7	10594	-4930	47853	-2696	37259	2234	179313	45923
11	20230331	ag	白糖	ag2306	5314.36	中信建投	ag2308	6	-1092	-2110	9388	-330	10480	1780	23786	6317
12	20230331	ag	白糖	ag2306	5314.36	中信期货	ag2306	7	4376	924	47028	1741	42652	817	214004	56713
13	20230331	ag	白糖	ag2306	5314.36	中国国贸	ag2306	1	-6841	-324	0	0	6843	324	324	0
14	20230331	ag	白糖	ag2306	5314.36	中泰期货	ag2307	7	570	-53	18659	62	17989	115	39745	14796
15	20230331	ag	白糖	ag2306	5314.36	中糖期货	ag2312	2	-304	-20	182	1	486	21	22	0
16	20230331	ag	白糖	ag2306	5314.36	中融汇信	ag2306	1	-6548	136	0	0	6548	-136	136	0
17	20230331	ag	白糖	ag2306	5314.36	中行期货	ag2308	3	2507	2	3532	-8	1025	-10	344	320
18	20230331	ag	白糖	ag2306	5314.36	中财期货	ag2306	2	9448	-98	9448	-98	0	0	98	0

上图就是我处理的期货商品合并持仓成交数据，除了对持仓成交进行了合并，同时关联了行情数据得到了价格信息，到此我相信你应该也知道该怎么用这个数据了吧

。

后记，这一篇想了好久都不知道怎么写，今天周末正好有空，干脆就通过结果数据和原始数据的对比方式来粗略描述一下数据梳理中的一些关键点，具体工作的开展如果你准备纯手工来做这个数据我就建议不要去弄了，一天的商品种行情数据接近100条，合约行情数据接近1000条，商品品种此仓成交数据接近5000条，商品合约持仓排名数据数万条，手工玩出来数据也失去了时间价值。如果你熟悉程序开发，或者会一些数据处理工具的话可以试验一下，最后分享一个利用该数据我做出来的数据展现效果：