

很多期货交易者尤其是技术流交易者对交易数据是有着独特的兴趣的，虽然现在各种行情软件已经提供了针对交易数据的丰富的指标图表分析，我相信很多人跟我一样还是喜欢从原始数据的获取到数据的加工处理尽可能都经自己的手来，这样才能让自己对数据分析的结果更加有信心。

前面我已经发了几个文章大概介绍了如何在各个期货交易所官方网站浏览下载原始的交易数据，这些原始数据都是公开发布的并且可以合法下载的。我相信有朋友已经按照我介绍的方法去浏览下载查看过了，但凡下载查看过各个交易所的交易数据的朋友也一定会发现每个交易所发布的交易数据格式并不一致，即便同一个交易所发布的不同分类的交易数据之间也缺乏关联性，这对于我们进一步使用这些交易数据是比较头痛的。

本文就试图来针对交易所数据的整理介绍一下我得一些方法，限于时间有限内容比较多，本文就先做一个开头，先把我在处理交易数据的一些逻辑说一下。

首先，明确数据整理目标

对于交易所公开发布的数据，我主要关注和下载的数据有三部分，就是期货品种基础信息、每日收盘行情数据和每日收盘交易排名数据。基于这三部分数据，我的主要分析对象想必大家也就清楚了，就是两个：期货品种(虽然也可以深入分析期货合约，但是根据本人交易习惯暂时不细分到合约进行分析)和期货会员。明确了我们的分析

对象，同

时下载到了原始数

据，接下来我们数据整理的目标也就

出来了：

**将交易所原始数据按照期货品种和交易会员进行分类梳理，并建立起期货品种与交易会员之间的关系。**

第二，仔细查看分析原始数据格式和数据项

这里以郑州期货交易所的原始数据来加以介绍。我们通过郑州期货交易所可以下载到每日交易收盘的行情数据和每日交易排名数据的Excel文件格式，如下图(以2023年2月28日交易所发布的格式为例)：

郑州商品交易所郑商期货持仓排名表 (2023-02-28)									
品种: 焦炭CF					日期: 2023-02-28				
名次	会员简称	成交量(手)	增减量	会员简称	持仓总量	增减量	会员简称	持仓总量	增减量
1	海通期货	40,384	14,404	国泰君安	14,989	483	海通期货	13,600	751
2	中信期货	21,257	4,045	海通期货	14,804	661	国泰君安	13,555	481
3	国泰君安	20,739	6,509	中信期货	13,479	960	中泰期货	13,482	-81
4	东证期货	17,188	7,371	华安期货	10,450	703	东兴期货	11,308	-224
5	招商期货	16,341	6,515	东兴期货	10,354	-343	光大期货	11,227	-854
6	光大期货	15,344	3,664	光大期货	10,077	-284	华安期货	10,768	618
7	华安期货	15,313	3,784	中泰期货	8,971	-129	中信期货	8,729	-339
8	银河期货	12,224	3,299	东证期货	8,630	-110	华安期货	7,314	-139
9	东方红期货	8,086	2,925	华安期货	7,137	-358	银河期货	7,028	-486
10	方正中期	7,625	2,594	银河期货	6,932	10	东证期货	6,791	-450
11	华泰期货	7,536	2,551	德中期货	6,201	-62	永安期货	6,435	100
12	中泰期货	6,384	1,786	华泰期货	5,725	-92	瑞华期货	6,375	-348
13	永安期货	5,941	2,167	新湖期货	4,324	-37	广发期货	5,823	311
14	东兴期货	5,551	1,130	国证期货	4,820	-448	新湖期货	5,224	-77
15	海证期货	5,483	1,934	永安期货	4,902	-80	永安期货	5,197	-96
16	永安期货	5,140	2,043	广发期货	4,713	-226	中信建投	4,845	-207
17	新湖期货	5,044	2,909	中信建投	4,313	-191	方正中期	4,840	-309
18	中信期货	4,900	1,437	国证期货	3,858	-18	徽商期货	4,825	-448
19	中信期货	4,879	2,086	青盛期货	3,793	-10	宏源期货	4,170	183
20	中信建投	4,818	830	银河期货	3,744	-423	长城期货	4,118	-48
合计		238,370	77,458		153,258	-751		154,887	-1,785
品种: 棉花CF					日期: 2023-02-28				
名次	会员简称	成交量(手)	增减量	会员简称	持仓总量	增减量	会员简称	持仓总量	增减量
1	中信期货	133,648	-43,664	中信期货	101,689	1,435	中信期货	118,819	7
2	东证期货	71,903	-24,144	国泰君安	44,989	3,894	国泰君安	71,372	-5,934
3	海通期货	54,194	-5,482	永安期货	43,466	968	中信国际	59,020	-338
4	国泰君安	36,355	-1,447	中泰期货	38,955	2,109	永安期货	57,732	801
5	华泰期货	28,232	-8,724	银河期货	38,530	2,617	东证期货	54,937	1,007
6	中信期货	22,967	-1,258	东证期货	34,974	1,470	银河期货	47,433	839
7	华安期货	20,733	-4,726	浙商期货	33,984	509	中信期货	41,035	648
8	国信期货	18,063	-1,089	华泰期货	32,730	623	华泰期货	36,225	-115

观察上面的 Excel文件可以发现郑州期货交易所发布的日行情数据包含了交易日期、合约代码,昨结算,今开盘,最高价,最低价,今收盘,今结算,涨跌1,涨跌2,成交量(手),持仓量,增减量,成交额等有效数据项,日交易排名数据提供了品种、交易日、会员简称,成交量,增减量,持仓量,增减量,持仓量,增减量等数据项目。由于交易所发布数据的时候已经按照一定的格式进行了规范展示,这样在浏览的时候相对比较方便一些,但是如果用于统计则很不方便,这就需要进行下一步工作。

### 第三, 设计数据结构选取有用数据项计算基础指标

根据前面对交易所原始数据的观察, 结合我们对交易数据分析的需要, 我们需要设计新的数据结构。根据本文一开始我们明确的分析对象, 这里我们简单设计两个相互关联的数据结构。

### 3.1期货品种日交易行情数据结构

因为我们只关心品种的每日价格，这里可以选择对合约结算价进行加权运算得到一个品种的结算价，然后成交量、持仓量直接对合约相应字段合计即可得到。据此我们设计的期货品种日交易行情数据结构就是如下所示：

交易日	品种代码	品种名称	公司简称	净持仓	变化
2023/2/28	AP	苹果	国泰君安	14988	483
2023/2/28	AP	苹果	海通期货	14804	661
2023/2/28	AP	苹果	中信期货	13479	360
2023/2/28	AP	苹果	华闻期货	10650	705

从以上我们设计的期货品种行情数据结构和日净持仓数据结构可以看到，两个数据结构都有交易日和品种代码两个字段，这也就为后期连接两个数据结构对期货公司进行分析提供了条件。

## 4、说明

本文针对从郑州期货交易所下载的日交易行情数据和日持仓排名数据，为了能将其变成我们可分析用的基础数据，做了一个开端介绍，设计了两个初级的数据结构，利用这个结构我们有兴趣的朋友可以手动做一下数据处理得到的结果就是我们的数据基础，当然手动做这个工作必然会非常辛苦，真正处理还是需要通程序化工具来批量处理。

当然以上数据结构设计好了，将期货交易所的数据转换输出到这个数据结构还有很多细节需要处理，比如交易编码的提取，交易日期的设置，期货品种结算价格的计算等等，这些细节我也是通过写程序搞定了，这里就暂时不介绍了，以后有朋友非常有兴趣并且对软件处理数据感兴趣的时候我们再交流。