所有数据整合到一个时间轴上 从2015年1月1日开始 时间轴精度为分钟级 时区按US central time 24小时制

**天气数据**

天气数据有4个气象站 每个气象站每天会播报2或者4次数据

每次播报后几天的两种预测数据 HDD和CDD

（隔天的Delta可以基于累加值，当天预报值和Y10\_NORMAL计算出来）

**ECMEN**

0点公布的数据 转换成 1.40 （冬令时） 或 2.40 (夏令时)

12点公布的数据 转换成 13.40（冬令时）或14.40 （夏令时）

提取的数据包括

* 日期 INIT\_DAT
* 公布时间
* 公布的第一条HDD 的VERIF\_DATE
* 公布的第一条HDD Value
* 公布的第一个HDD的10Y\_NORMAL数据
* 公布的HDD总值 （14个Value值累加）
* 公布的第一条CDD 的VERIF\_DATE
* 公布的第一个CDD Value
* 公布的第一个CDD的10Y\_NORMAL数据
* 公布的CDD总值 （14个Value值累加）

**ECMOP**

0点公布的数据 转换成 0.55（冬令时） 或 1.55 (夏令时)

12点公布的数据 转换成 12.55（冬令时）或13.55（夏令时）

提取的数据包括

* 日期 INIT\_DAT
* 公布时间
* 公布的第一条HDD 的VERIF\_DATE
* 公布的第一条HDD Value
* 公布的第一个HDD的10Y\_NORMAL数据
* 公布的HDD总值 （9个Value值累加）
* 公布的第一条CDD 的VERIF\_DATE
* 公布的第一个CDD Value
* 公布的第一个CDD的10Y\_NORMAL数据
* 公布的CDD总值 （9个Value值累加）

**GFSEN**

0点公布的数据 转换成 0.28（冬令时） 或 1.28 (夏令时)

6点公布的数据 转换成 6.28（冬令时） 或 7.28 (夏令时)

12点公布的数据 转换成 12.28（冬令时）或13.28（夏令时）

18点公布的数据 转换成 18.29（冬令时） 或 19.29 (夏令时)

提取的数据包括

* 日期 INIT\_DAT
* 公布时间
* 公布的第一条HDD 的VERIF\_DATE
* 公布的第一条HDD Value
* 公布的第一个HDD的10Y\_NORMAL数据
* 公布的HDD总值 （15个Value值累加）
* 公布的第一条CDD 的VERIF\_DATE
* 公布的第一个CDD Value
* 公布的第一个CDD的10Y\_NORMAL数据
* 公布的CDD总值 （15个Value值累加）

**GFSOP**

0点公布的数据 转换成 23.11（冬令时） 或 0.11 (夏令时)

6点公布的数据 转换成 5.13（冬令时） 或 6.13 (夏令时)

12点公布的数据 转换成 11.13（冬令时）或12.13（夏令时）

18点公布的数据 转换成 17.13pm（冬令时） 或 18.13 (夏令时)

提取的数据包括

* 日期 INIT\_DAT
* 公布时间
* 公布的第一条HDD 的VERIF\_DATE
* 公布的第一条HDD Value
* 公布的第一个HDD的10Y\_NORMAL数据
* 公布的HDD总值 （15个Value值累加）
* 公布的第一条CDD 的VERIF\_DATE
* 公布的第一个CDD Value
* 公布的第一个CDD的10Y\_NORMAL数据
* 公布的CDD总值 （15个Value值累加）

**基本面数据**

Daily sd 每天公布时间为6.20 每个维度都需要 （delta之后可以计算）

weekly storage survey actual每周公布时间为9.30 每个维度都需要

**Daylight saving time in the United States**

Year Start End

2015 March 8 November 1

2016 March 13 November 6

2017 March 12 November 5

2018 March 11 November 4

2019 March 10 November 3

2020 March 8 November 1

2021 March 14 November 7

**价格数据**

首先把G列分开为年和月， 例如 1502表示15年2月交割， 1503表示15年3月交割。然后把价格表按照交割月份分为若干子表，再对每张子表分别按分钟进行统计。

输出的表格形式为

* Date交易时间 #A
* Time 时间 (分钟级别) #B
* Open price (每分钟第一次出现的价格)
* High price (每分钟内最高的价格)
* Low price (每分钟内最低的价格)
* Close price (每分钟最后一次出现的价格)
* Volume #H
* VWAP (按Volume求得平均价格)





