Data Description Report

We have attached a variable description file which contain all the information about all the variables used in this report. We have two datasets, Intraday data from EPEXSPOT and DayAhead data from EPEXSPOT and ENTSOE.

INTRADAY DATA

- The data has been collected from 01/Jan/2014 – 31/Dec/2018
- Dimension: 43824 rows 13 columns
- "***Last***" is the dependent variable in the intraday dataset, because the prices are similar to the DayAhead price in the DayAhead Dataset.
- Below is the summary of the Intraday dataset.

	DateTime	Low	High	Last	Weighted_Avg	Idx	ID3	Buy_Vol	Sell_Vol	Index_Base	Index_Peak
1	2014-01-01 00:00:00	16.80	35.15	21.00	22.07	22.07	NA	185.0	220.0	12.53	13.50
2	2014-01-01 01:00:00	9.00	32.96	9.00	17.94	17.94	NA	112.0	340.0	12.53	13.50
3	2014-01-01 02:00:00	2.00	16.00	3.00	9.36	9.36	NA	124.0	355.0	12.53	13.50
4	2014-01-01 03:00:00	1.00	12.10	5.00	5.23	5.23	NA	86.0	380.0	12.53	13.50
5	2014-01-01 04:00:00	2.00	13.00	11.00	8.63	8.63	NA	NA	280.0	12.53	13.50
6	2014-01-01 05:00:00	1.00	10.00	10.00	6.44	6.44	NA	NA	200.0	12.53	13.50
7	2014-01-01 06:00:00	4.00	8.50	8.50	5.94	5.94	NA	NA	240.0	12.53	13.50
8	2014-01-01 07:00:00	5.00	6.50	6.50	5.30	5.30	NA	NA	145.0	12.53	13.50
9	2014-01-01 08:00:00	7.00	12.25	10.00	9.42	9.42	NA	80.0	30.0	12.53	13.50
10	2014-01-01 09:00:00	8.00	18.00	12.00	15.21	15.21	NA	275.0	330.0	12.53	13.50
11	2014-01-01 10:00:00	6.50	18.00	9.50	10.49	10.49	NA	745.0	204.0	12.53	13.50

Showing 1 to 11 of 43,824 entries

Console Terminal

```
> summary(IntradayContinuousEPEXSPOT)
    DateTime                       Low                High              Last         Weighted_Avg
 Min.   :2014-01-01 00:00:00   Min.   :-98.00   Min.   :-10.00   Min.   :-69.9   Min.   :-47.47
 1st Qu.:2015-04-02 12:45:00   1st Qu.: 25.00   1st Qu.: 32.52   1st Qu.: 28.0   1st Qu.: 28.71
 Median :2016-07-02 00:30:00   Median : 34.20   Median : 44.00   Median : 38.8   Median : 39.12
 Mean   :2016-07-02 00:29:59   Mean   : 35.77   Mean   : 47.39   Mean   : 41.3   Mean   : 41.47
 3rd Qu.:2017-10-01 12:15:00   3rd Qu.: 45.30   3rd Qu.: 57.50   3rd Qu.: 51.5   3rd Qu.: 51.23
 Max.   :2018-12-31 23:00:00   Max.   :700.00   Max.   :950.00   Max.   :800.0   Max.   :979.51
                               NA's   :200      NA's   :201      NA's   :201     NA's   :200
      Idx               ID3             Buy_Vol          Sell_Vol        Index_Base       Index_Peak
 Min.   :-47.47   Min.   : -4.88   Min.   :   0.1   Min.   :   0.1   Min.   :  1.71   Min.   :  3.08
 1st Qu.: 28.62   1st Qu.: 40.54   1st Qu.: 177.3   1st Qu.: 186.0   1st Qu.: 30.75   1st Qu.: 33.03
 Median : 39.06   Median : 51.91   Median : 354.2   Median : 377.0   Median : 39.16   Median : 43.54
 Mean   : 41.40   Mean   : 52.51   Mean   : 460.4   Mean   : 473.9   Mean   : 41.39   Mean   : 45.70
 3rd Qu.: 51.19   3rd Qu.: 64.02   3rd Qu.: 627.3   3rd Qu.: 658.7   3rd Qu.: 49.89   3rd Qu.: 54.85
 Max.   :979.51   Max.   :500.30   Max.   :4406.9   Max.   :3234.0   Max.   :138.54   Max.   :188.86
 NA's   :48       NA's   :35952    NA's   :410      NA's   :486      NA's   :48       NA's   :73
>
```

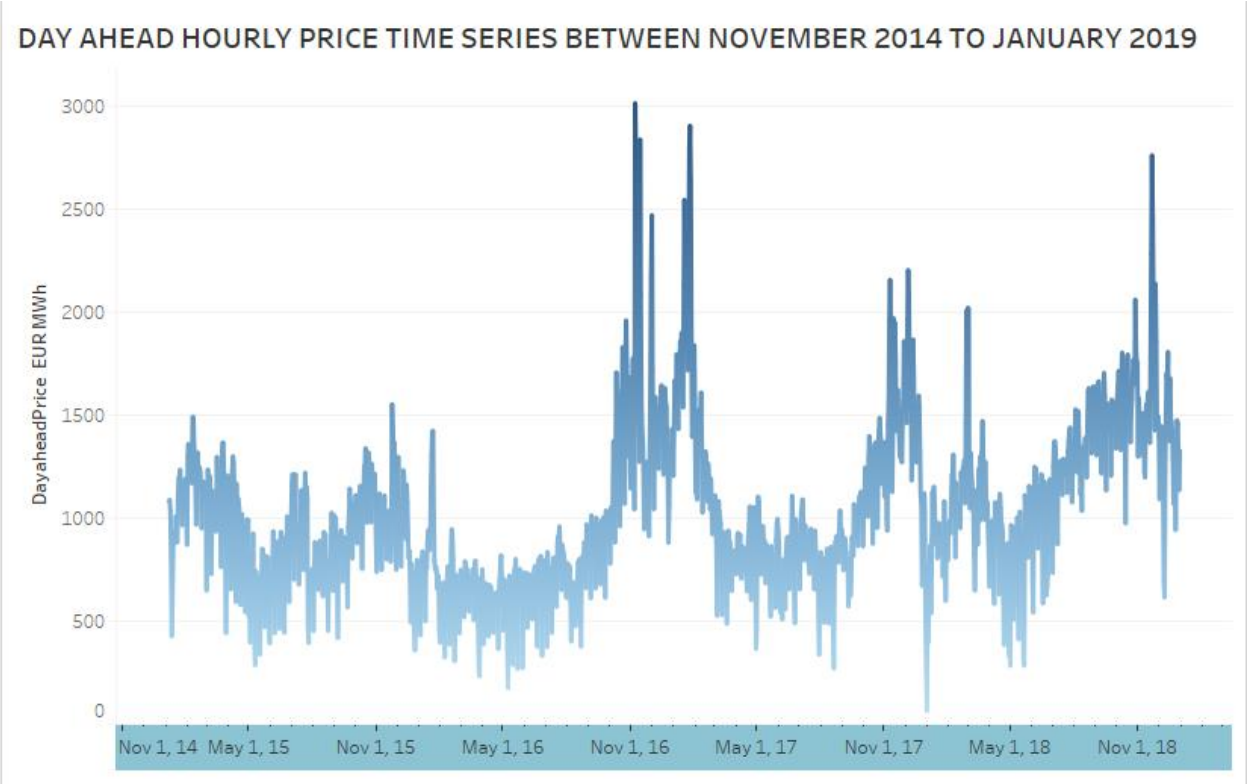
- The columns Low, High, Last and Weighted Average contain missing values (.45%)
- The Column ID3 has a lot of missing values about 75% of the data.
- Buy Volume and Sell Volume has around 1% missing values

DAYAHEAD DATA

- Timeline: The data has been collected from 01/Jan/2014 – 31/Dec/2018
- Dimension: The data has 22184 rows 28 columns
- The dependent variable in this dataset is DayaheadPrice_Eur
- We collected the data from two sources
  i) Epexspot
  ii) Entsoe
- Both the datasets were then merged to get the final dataset. Summary report are shown as below.

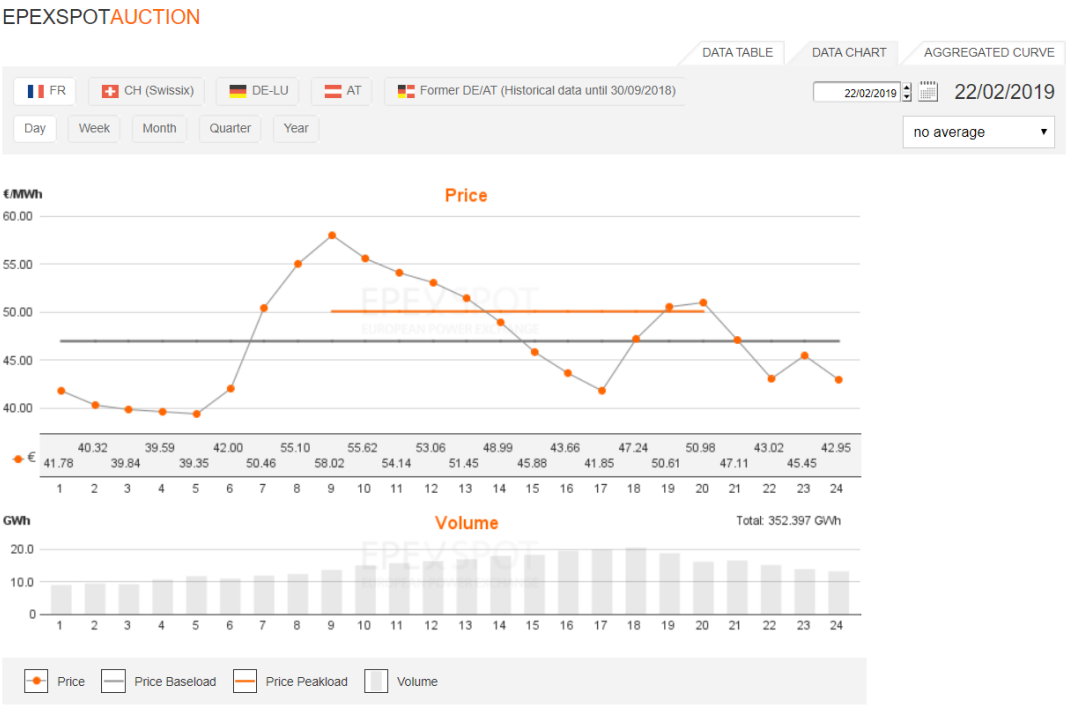
```
   fromdate             fromtime         Volume       MiddleNight     EarlyMorning     LateMorning
Min.   :2015-01-13   Min.   : 0.00   Min.   : 5594   Min.   : 6.40   Min.   :  2.56   Min.   :  6.35
1st Qu.:2016-01-13   1st Qu.: 5.00   1st Qu.:10624   1st Qu.:25.74   1st Qu.: 27.14   1st Qu.: 36.73
Median :2016-12-31   Median :11.00   Median :12203   Median :32.72   Median : 34.77   Median : 46.99
Mean   :2017-01-11   Mean   :11.47   Mean   :12402   Mean   :34.72   Mean   : 36.84   Mean   : 49.42
3rd Qu.:2017-12-31   3rd Qu.:17.00   3rd Qu.:13999   3rd Qu.:43.51   3rd Qu.: 45.61   3rd Qu.: 59.94
Max.   :2018-12-31   Max.   :23.00   Max.   :23214   Max.   :91.42   Max.   :108.89   Max.   :165.09
                                     NA's   :24
EarlyAfternoon      RushHour         OffPeak2         Night          OffPeak1         Business
Min.   :  3.27   Min.   :  7.10   Min.   : 12.32   Min.   :  9.69   Min.   :  7.55   Min.   : 6.14
1st Qu.: 32.13   1st Qu.: 35.05   1st Qu.: 36.13   1st Qu.: 32.84   1st Qu.: 34.92   1st Qu.:24.76
Median : 42.45   Median : 45.13   Median : 43.55   Median : 40.38   Median : 44.57   Median :31.35
Mean   : 43.82   Mean   : 49.40   Mean   : 46.55   Mean   : 43.46   Mean   : 47.54   Mean   :33.41
3rd Qu.: 53.90   3rd Qu.: 60.79   3rd Qu.: 55.85   3rd Qu.: 52.93   3rd Qu.: 57.80   3rd Qu.:41.52
Max.   :119.37   Max.   :303.60   Max.   :105.00   Max.   :121.10   Max.   :170.44   Max.   :89.28

   OffPeak         Morning          HighNoon       Afternoon         Evening          SunPeak
Min.   : 5.69   Min.   :  4.94   Min.   : 7.90   Min.   :  3.45   Min.   :  4.97   Min.   : -0.18
1st Qu.:26.62   1st Qu.: 34.64   1st Qu.:29.94   1st Qu.: 34.43   1st Qu.: 34.87   1st Qu.: 30.92
Median :33.90   Median : 44.68   Median :36.96   Median : 44.18   Median : 45.71   Median : 40.56
Mean   :35.78   Mean   : 46.62   Mean   :39.37   Mean   : 46.54   Mean   : 47.13   Mean   : 43.24
3rd Qu.:44.19   3rd Qu.: 56.20   3rd Qu.:47.90   3rd Qu.: 56.69   3rd Qu.: 57.09   3rd Qu.: 53.60
Max.   :99.23   Max.   :142.23   Max.   :98.65   Max.   :161.73   Max.   :131.64   Max.   :121.65

  BasePrice        BaseVolume     PeakPrice      PeakVolume       DateTime            AL_MW
Min.   :  9.69   Min.   :179799   Mode:logical   Mode:logical   Length:22184       Min.   :  0.0
1st Qu.: 32.84   1st Qu.:264217   NA's:22184     NA's:22184     Class :character   1st Qu.:250.0
Median : 40.38   Median :293539                                 Mode  :character   Median :250.0
Mean   : 43.46   Mean   :297660                                                    Mean   :243.4
3rd Qu.: 52.93   3rd Qu.:330483                                                   3rd Qu.:250.0
Max.   :121.10   Max.   :469001                                                   Max.   :250.0
                                                                                  NA's   :25
    GR_MW       DayaheadPrice _EUR_MWh Dayahead_Load Forecast ActualTotalLoad
Min.   :  0.0   Min.   : -4.90          Min.   :29650          Min.   :30426
1st Qu.:250.0   1st Qu.: 30.70          1st Qu.:45500          1st Qu.:45409
Median :250.0   Median : 41.00          Median :52700          Median :52613
Mean   :240.7   Mean   : 43.41          Mean   :54245          Mean   :54179
3rd Qu.:250.0   3rd Qu.: 53.01          3rd Qu.:62350          3rd Qu.:62253
Max.   :250.0   Max.   :829.79          Max.   :95150          Max.   :94492
NA's   :25      NA's   :24                                     NA's   :37
```

Exploratory Data Analysis.

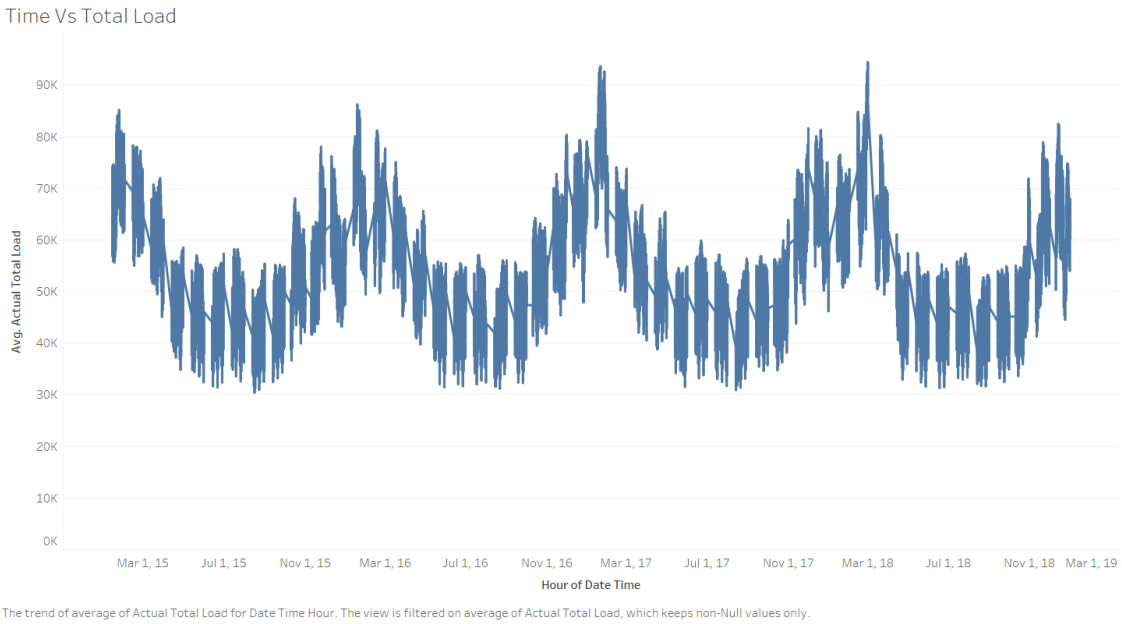


The above figure depicts the hourly price time series of the DA market, considering the period under study. The blue shaded between 7 November 2016 and 25 January 2017, which exhibited the highest prices during this four-year period. When applying statistical learning methods heavily dependent on the availability of sufficient historical data with high (or low) price regimes, it is expected that periods like this one will pose a challenge due to the non-existence of historical similarity.

DayAhead

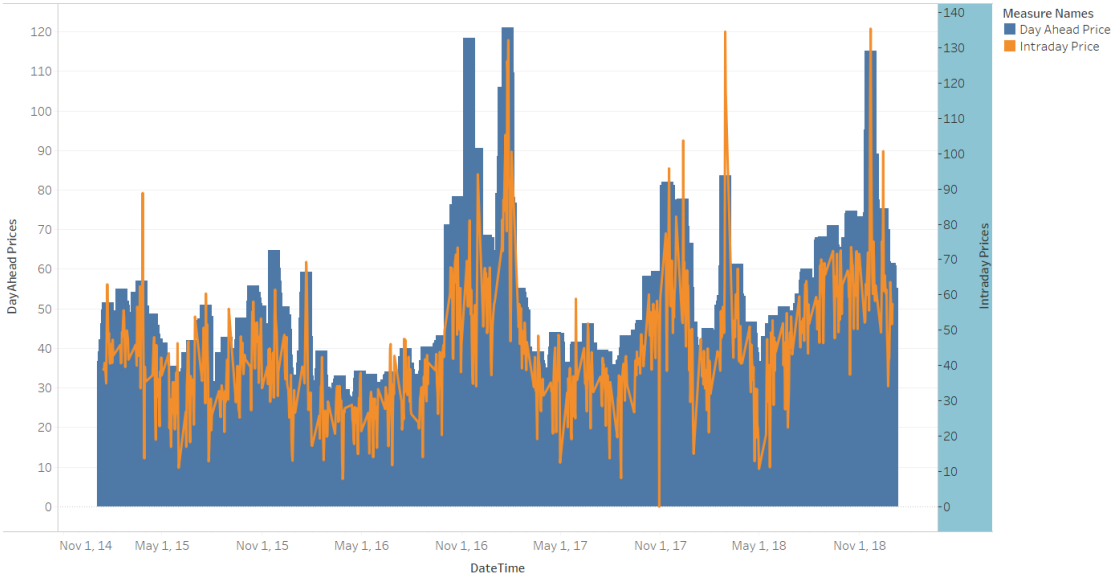


Changes in Price and volume over a period of 24 hrs. for 22/Feb/2019. We can observe that the maximum consumption and maximum price is during the peak hours (08:00 – 20:00).



The total load vs time graph tells us about the cyclic nature of load consumption. There is a peak during the winter months and troughs during the summer months. There seems to be an increasing trend in the load consumption.

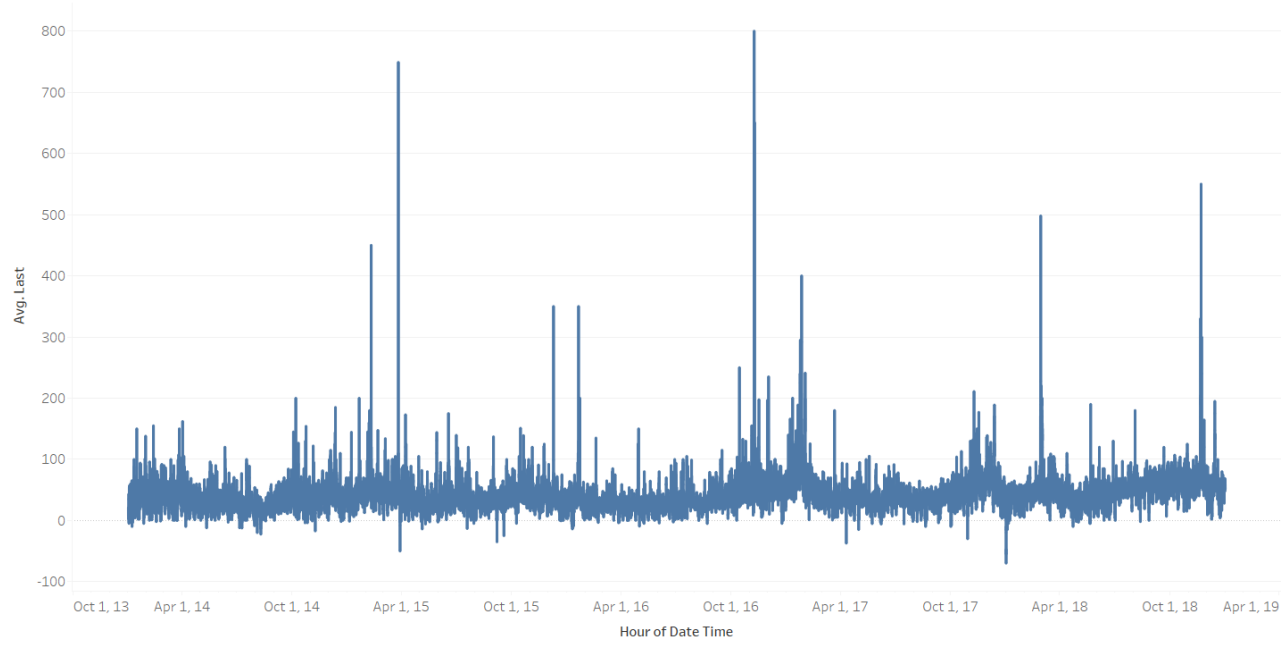
Intraday Vs DayAhead



The trends of Day Ahead Price and Intraday Price for DateTime (Dayahead.csv) Day. Color shows details about Day Ahead Price and Intraday Price. The view is filtered on average of DayaheadPrice EUR MWh, which keeps non-Null values only.

Comparison of DayAhead Price and Intraday price over the last 4yrs tells us that the overall price is almost the same for both.

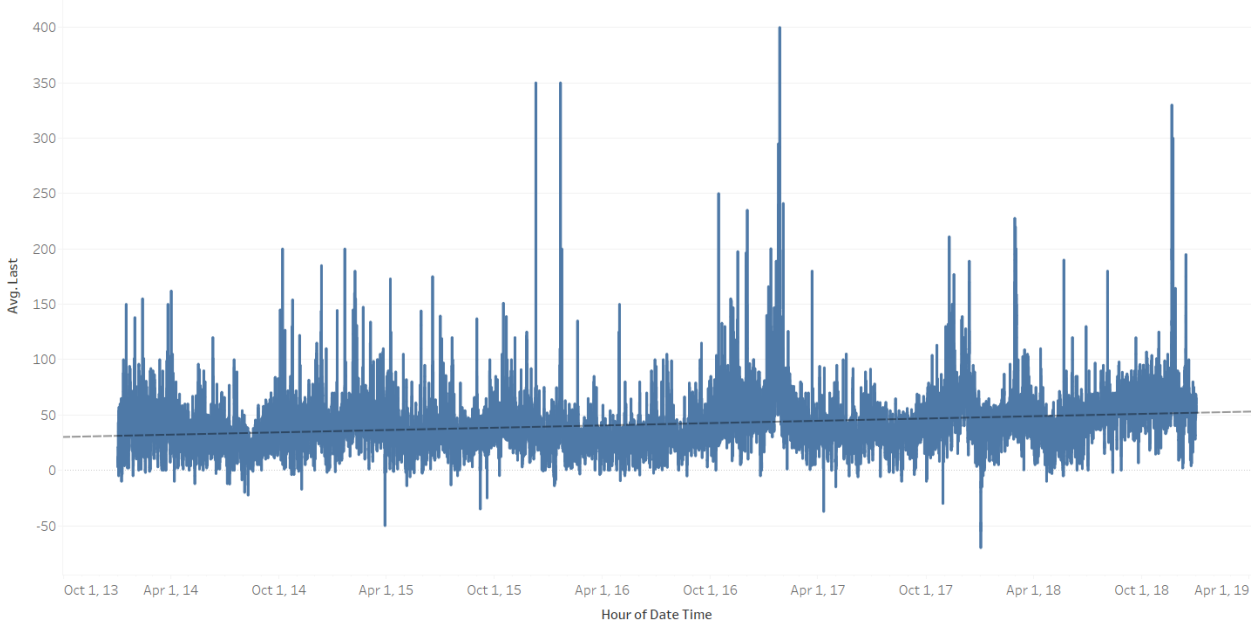
Time Vs Last Price



The trend of average of Last for Date Time Hour.

There seems to be an increasing trend when analyzing the Intraday last price with time. Also, there seems to be some outliers on specific days and specific time over the period of 5 years.

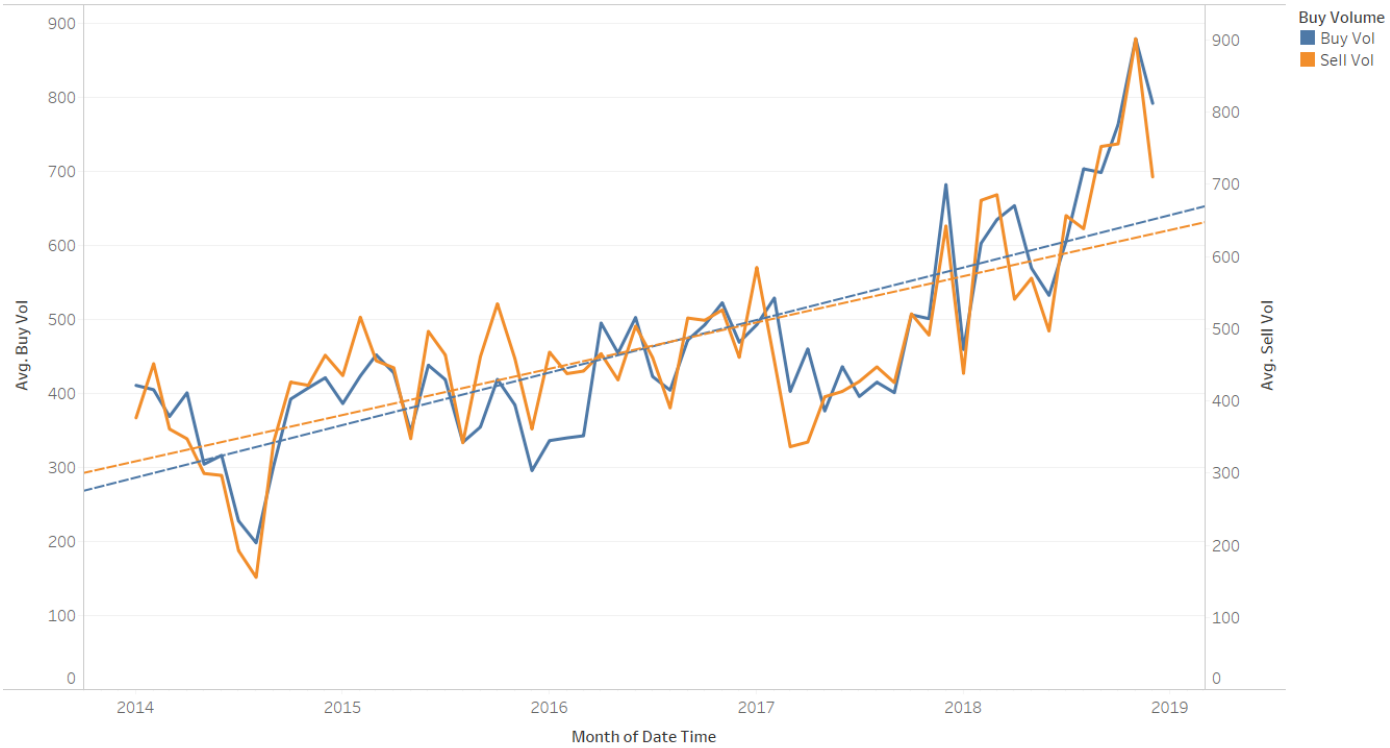
Time Vs Intraday Last Price



The trend of average of Last for Date Time Hour. The data is filtered on Date Time Hour, which keeps 43,813 of 43,819 members.

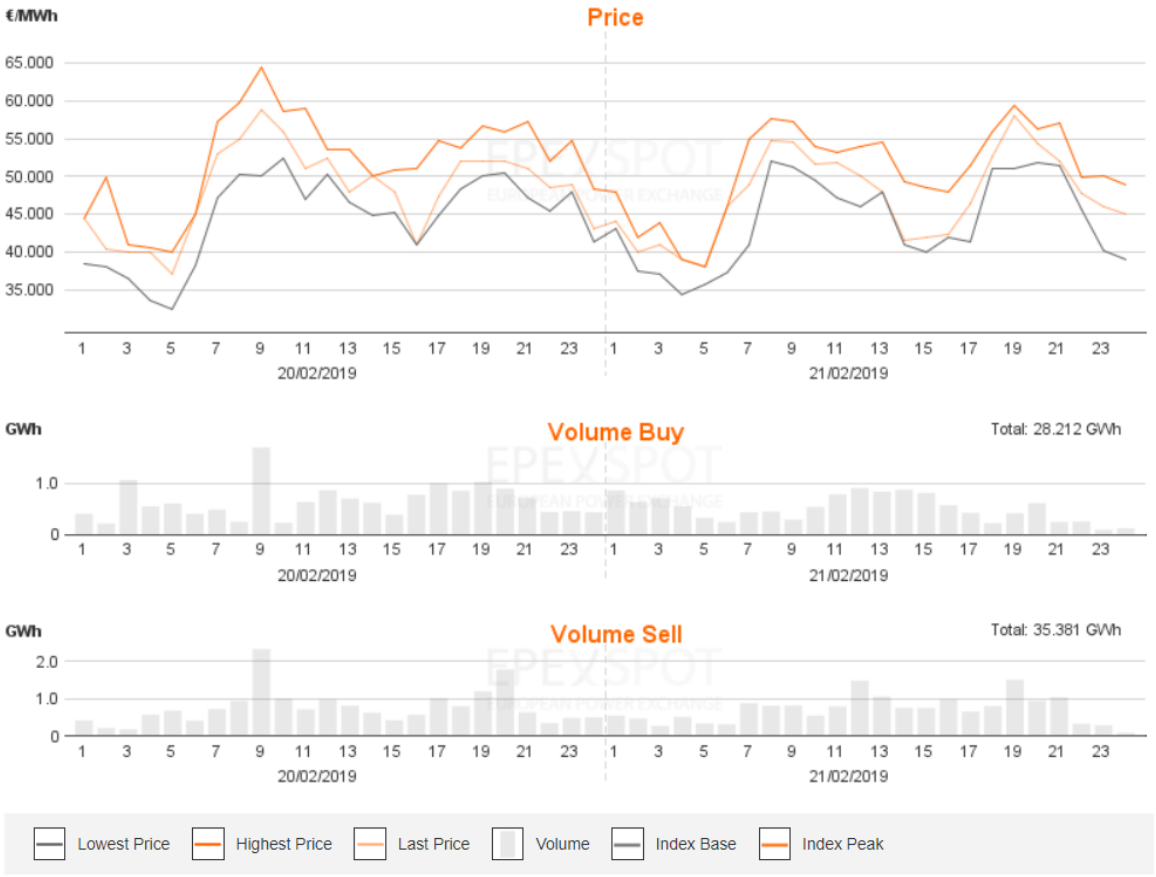
Intraday last price vs time after removing the outliers, the data is not biased and equally distributed.

Time Vs Buy & Sell



The trends of Buy Vol and Sell Vol for Date Time Month. Color shows details about Buy Vol and Sell Vol.

Relationship between Buying and selling units over a period of 5 years. There is a linear trend in both buying and selling units.

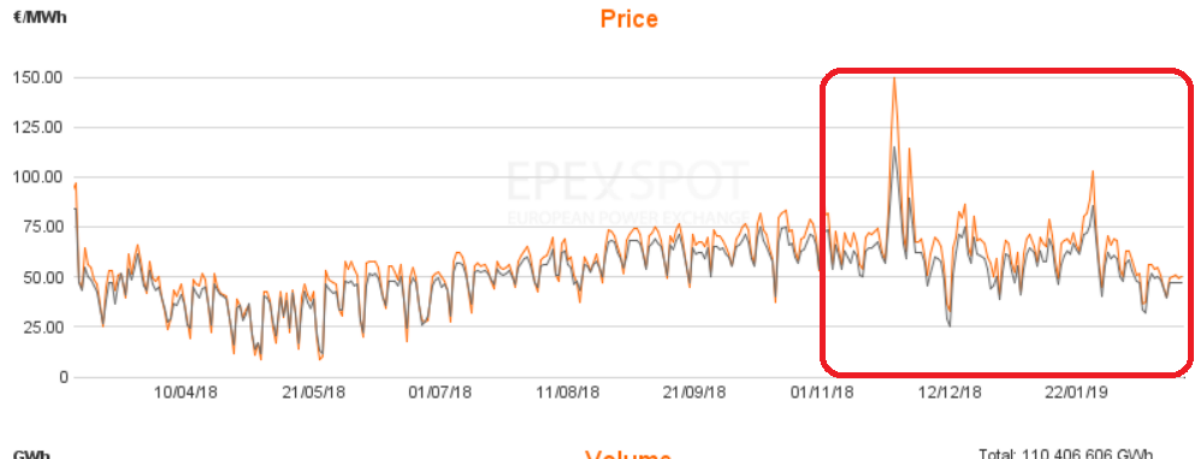


Relationship between the Intraday price and Volume bought and sold in recent days.

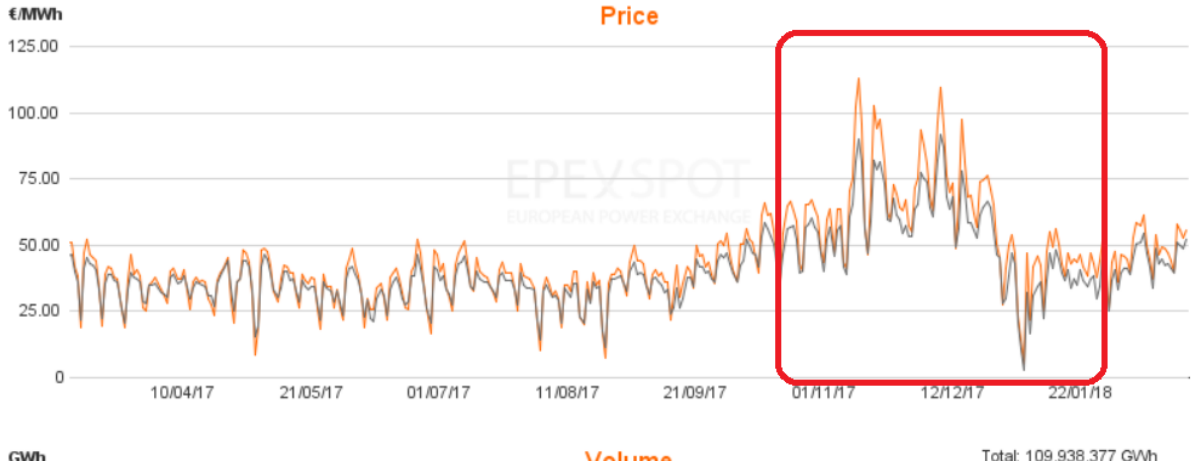
Seasonal Effects

Looking at the season change in the graphs for both intra day or interday we definitely see a price hike in **Nov & Dec** compared to other months.

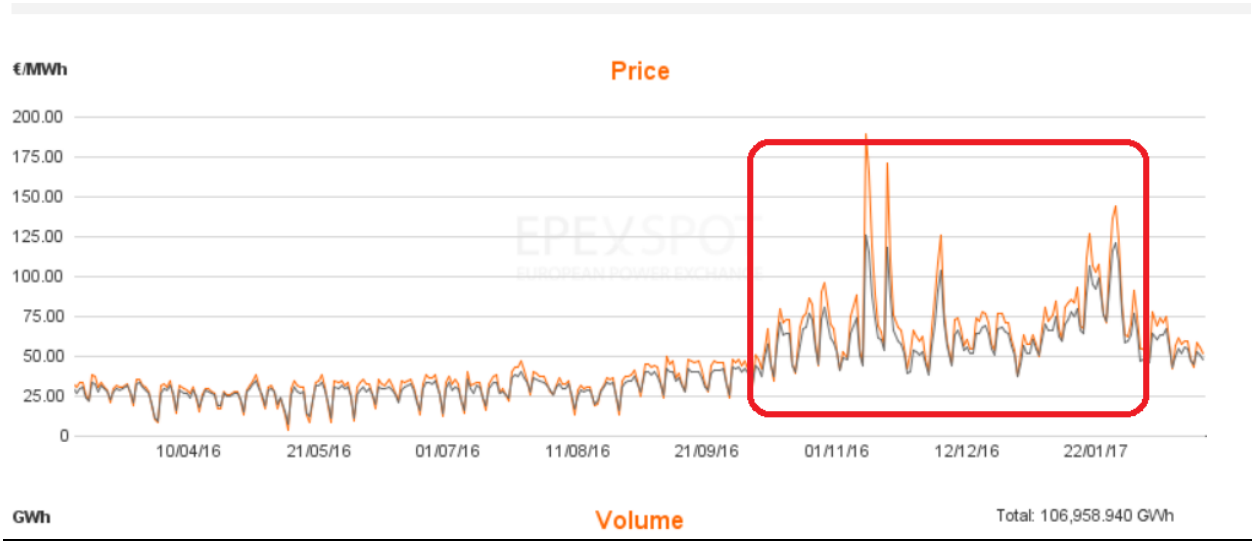
2018



2017



2016



Created By:

Ajay Parihar

Manjunagaraj Rudrappa

Patrick Dundon

Shivam Sarin

(Group 7)