Electricity Consumption Analysis and Forecasting in Germany

---

1. Introduction

1.1 Background

- Electricity consumption in Germany is influenced by various factors, including renewable energy production, seasonal changes, and economic activities. This project analyzes historical electricity consumption data to uncover trends, test hypotheses, and forecast future consumption.
- Briefly describe the context: Why is analyzing electricity consumption important?
- Mention the dataset: `opsd_germany_daily.csv` from Open Power System Data.

1.2 Objectives

- Understand trends in electricity consumption over time.
- Analyze the impact of renewable energy sources (`Wind`, `Solar`) on consumption.
- Forecast future electricity consumption using time series models.

1.3 Dataset Overview

- **Source:** Open Power System Data (`opsd_germany_daily.csv`).
- **Time Period:** 2006–2017.
- **Key Variables:** `Consumption`, `Wind`, `Solar`, `Wind+Solar`.

2. Methodology

2.1 Data Preparation

- Loaded and cleaned the dataset.
- Extracted `year`, `month`, and `day` from the `Date` column.
- Handled missing values in `Wind` and `Solar` using linear interpolation.

2.2 Exploratory Data Analysis (EDA)

- Visualized yearly, monthly, and daily consumption trends.
- Calculated correlations between `Consumption`, `Wind`, and `Solar`.

2.3 Hypothesis Testing

- **ANOVA and Tukey HSD:** Compared consumption across years and seasons.
- **Linear Regression:** Modeled the relationship between consumption and renewable energy sources.

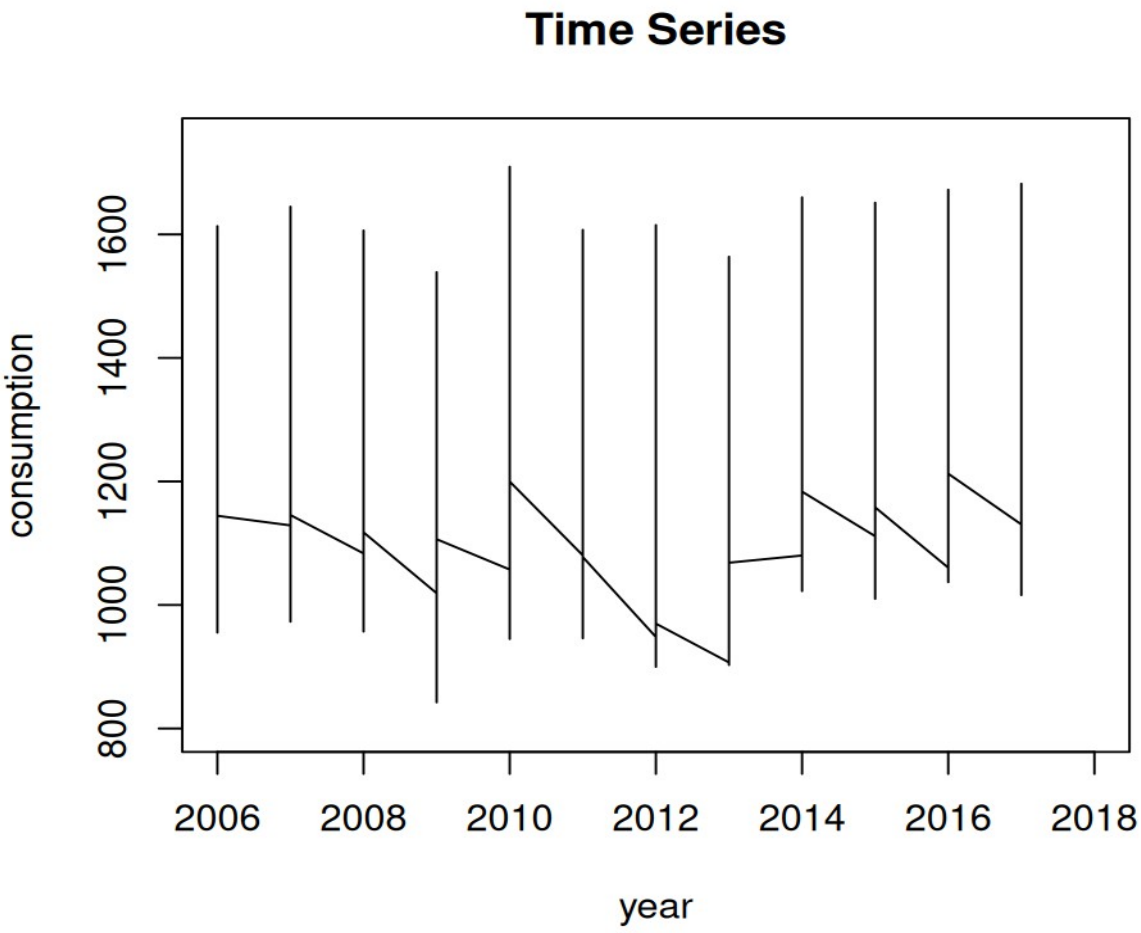
2.4 Time Series Forecasting

- Fit **ARIMA** and **ETS** models.
- Selected the best model based on accuracy metrics (RMSE, MAE, MAPE).
- Generated a **90-day forecast** using the best-performing model.

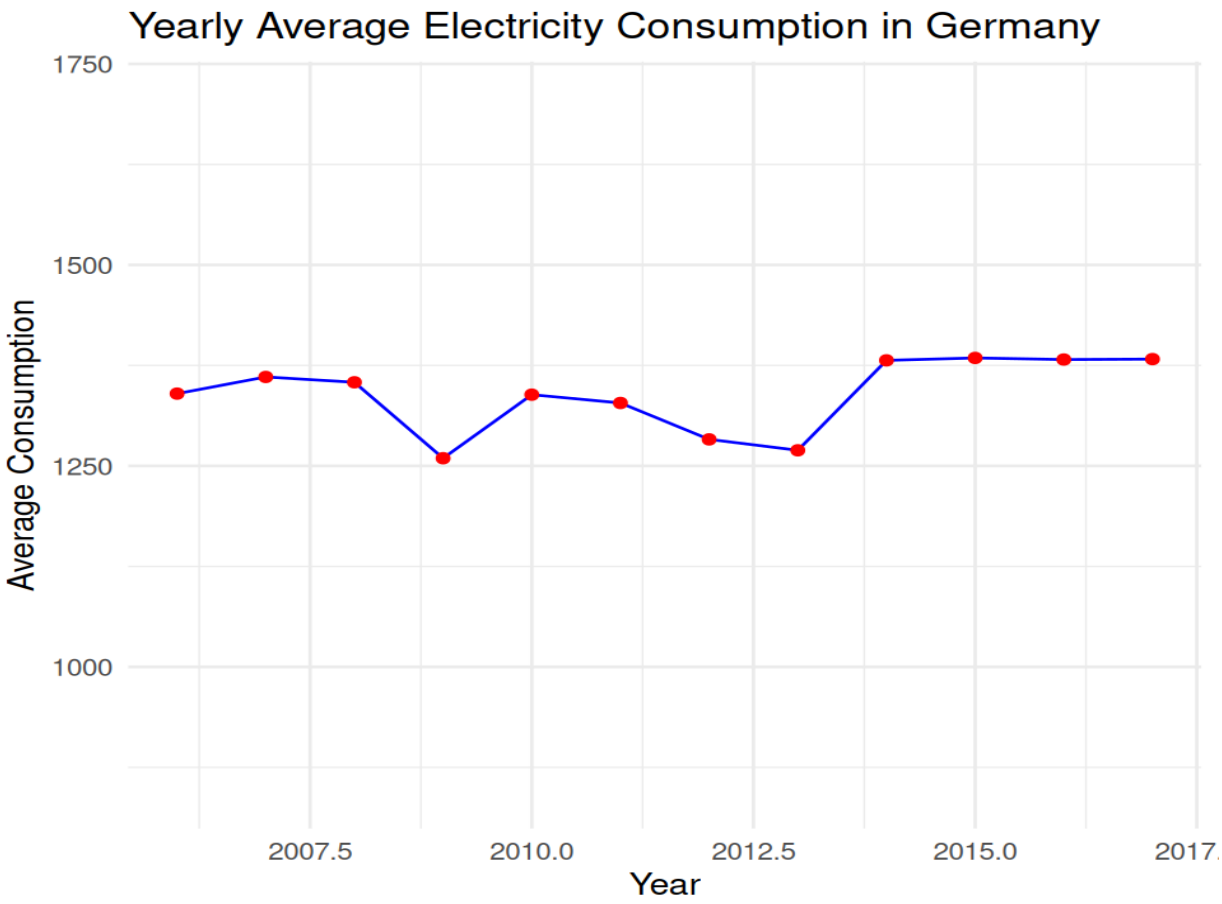
3. Visualization of Data

3.1 Trends in Consumption

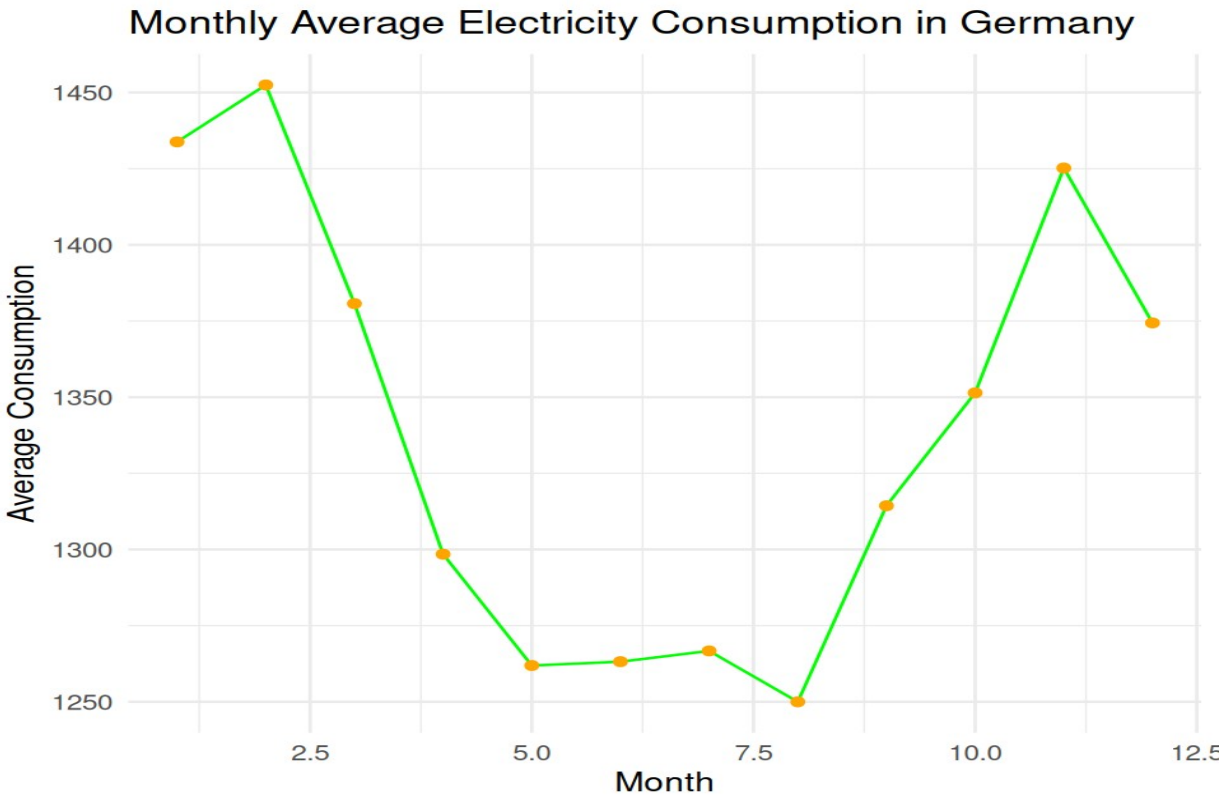
- **Overall Trends :** Include Year vice Consumption Electricity.



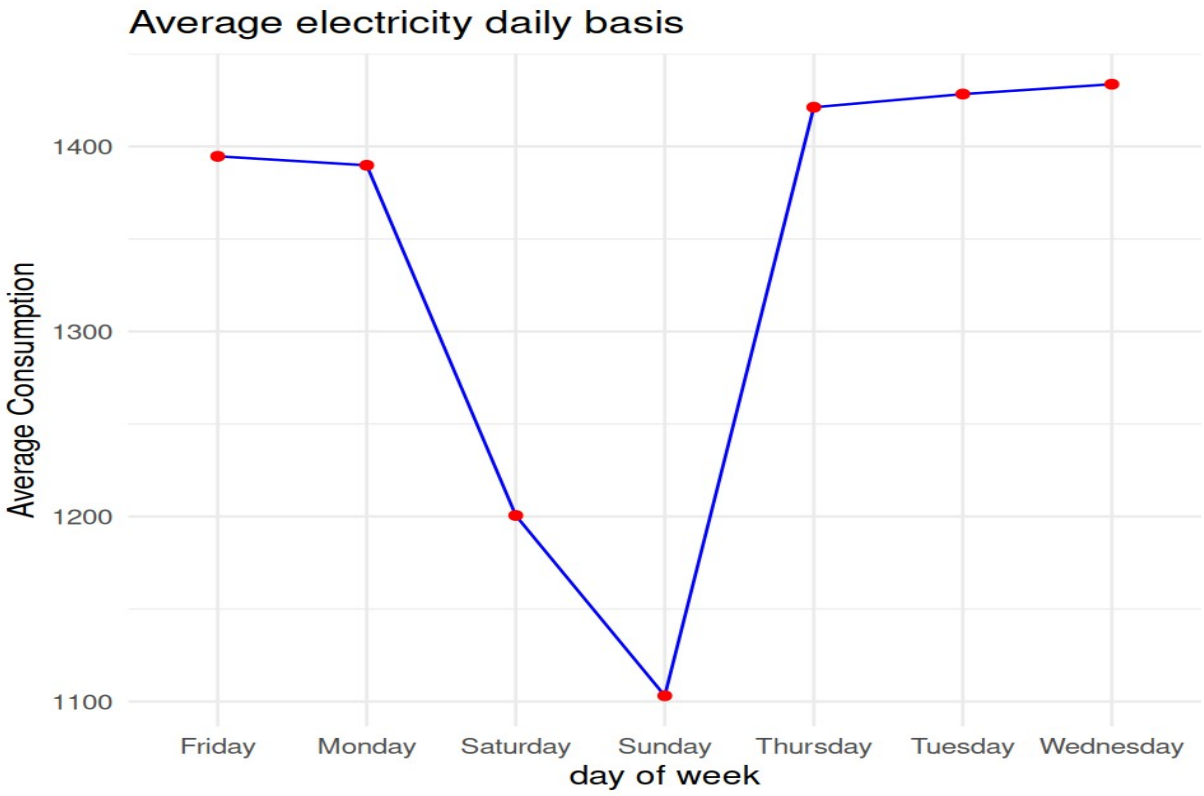
- **Yearly Trends:** Shows Average Consumption of electricity in Years.



- **Monthly Trends:**Tells Average consumption of electricity in months.



- **Daily Trends:** Lower consumption on weekends.



4. Key Findings

3.1 Trends in Consumption

1. **Yearly Trends:** Consumption varies significantly across years, with peaks in winter and summer.
    1.1. After doing statis we come up with the Minimum yearly consumption is 842.395.
    1.2. Minimum consumption of Electricity consumed in 2009.
    1.3. Maximum Consumption of Electricity consumed in 2010.
    1.4. Find correlations ture Relationship between Consumption and year, in that we got positive trend value (.08991391000) that means slighlty significance increament across year .
2. **Monthly Trends:** Higher consumption in winter (heating) and summer (cooling).
    2.1. Maximum consumption of Electricity consumed in December.
    2.2. Minimum consumption of Electricity consumed in April.
    2.3. When we do Correaltion of consumption in month Relationships is negative that means Increament of month give us to low consumption , its just stats Note*(we did not handling missing values yet) .
3. **Daily Trends:** Lower consumption on weekends.
    3.1.Maximum consumption of Electricity consumed in Sunday.
    3.2.Minimum Consumption of Electricity consumed in Wednesday.
    3.3. In the Daily trends the correlation is negative in the respect to consuption veries in daily basis , therefore Increament of daily gave us low consumption.
4. **Extracted Table**

	Mini Consumption	Maxi. consumption
date	4	7
month	4	12
year	2009	2010

3.2 Hypothesis Testing Results

I. **Hypothesis (1)**:Is power Consumption similar in all year? : here we do annova testing cause data in group wise so , we interpretion our **small p-value is <0.05** that mean we **reject the null hypothesis(H0),** To know significant differance we have to do **Tukey hsd(Honeslty significant differance)** it will compare the set of years with respect to consumption of electricity annova give small value , hense we compare each and every year consumption to each other by **TukeyHSD , by using this method we grouping which year has same consumption and not.**

1.1 Significant differance make Different years which has Different amount electricity consumed.

A. 2009-2006
B. 2012-2006
C. 2013-2006
D. 2014-2006
E. 2015-2006
F. 2016-2006

G. 2017-2006
H. 2009-2007
I. 2012-2007
J. 2013-2007
K. 2009-2008
L. 2012-2008
M. 2013-2008
N. 2010-2009
O. 2011-2009
P. 2014-2009
Q. 2015-2009
R. 2016-2009
S. 2017-2009
T. 2012-2010
U. 2013-2010
V. 2014-2010
W. 2015-2010
X. 2016-2010
Y. 2017-2010
Z. 2012-2011
AA. 2013-2011
AB. 2014-2011
AC. 2015-2011
AD. 2016-2011
AE. 2017-2011
AF. 2014-2012
AG. 2015-2012
AH. 2016-2012
AI. 2017-2012
AJ. 2014-2013
AK. 2015-2013
AL. 2016-2013
AM. 2017-2013

II. **Hypothesis(h2):** Is there is relationship between consumption and renewable source (like wind , solar , wind+solar) :To show relationship we have to confident that power to reject **null hypothesis(H0) ,**
   A. for that we about find relationship by corelation and linear regression ,
   B. we did regression we got p-value is smaller show that both predictors are statically significant,
   C. Intercept (1333.33): When both Wind and Solar are zero, average consumption would be about 1333 units.This is the baseline demand level.
   D. Statistical Significance: Both Wind and Solar have very low p-values,meaning we are highly confident that these predictors affect consumption.
   E. F-statistic (64.97, $p < 2.2e-16$): The overall regression model is statistically significant. At least one predictor (Wind or Solar) contributes meaningfully to explaining variation in Consumption.
   F. Model Fit ($R^2 = 0.0288$): Only about 2.9% of the variation in consumption is explained by Wind and Solar together. Note*(after the handling missing values)
   G. Conclusion : This means other factors (not included in the model) are driving the majority of variation in energy consumption.

III. **Hypothesis(H3)** : Is Consumption different between Seasons?

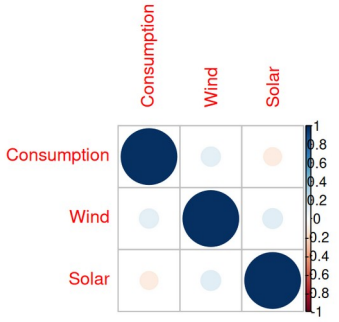
A. to compare consumption across seasons (e.g., winter vs. Summer),like it has group winter and summer need to annova test.
B. here we our small palue is very small so we reject the null hypothesis that means ,we conclude the consumption are different in every season , hurray !

IV. **ANOVA:** Significant differences in consumption across years and seasons.

V. **Linear Regression:** `Solar` has a significant negative impact on consumption, while `Wind` has a smaller positive effect.

VI. Correlations :

A. the corelation between wind, solar ,consumption

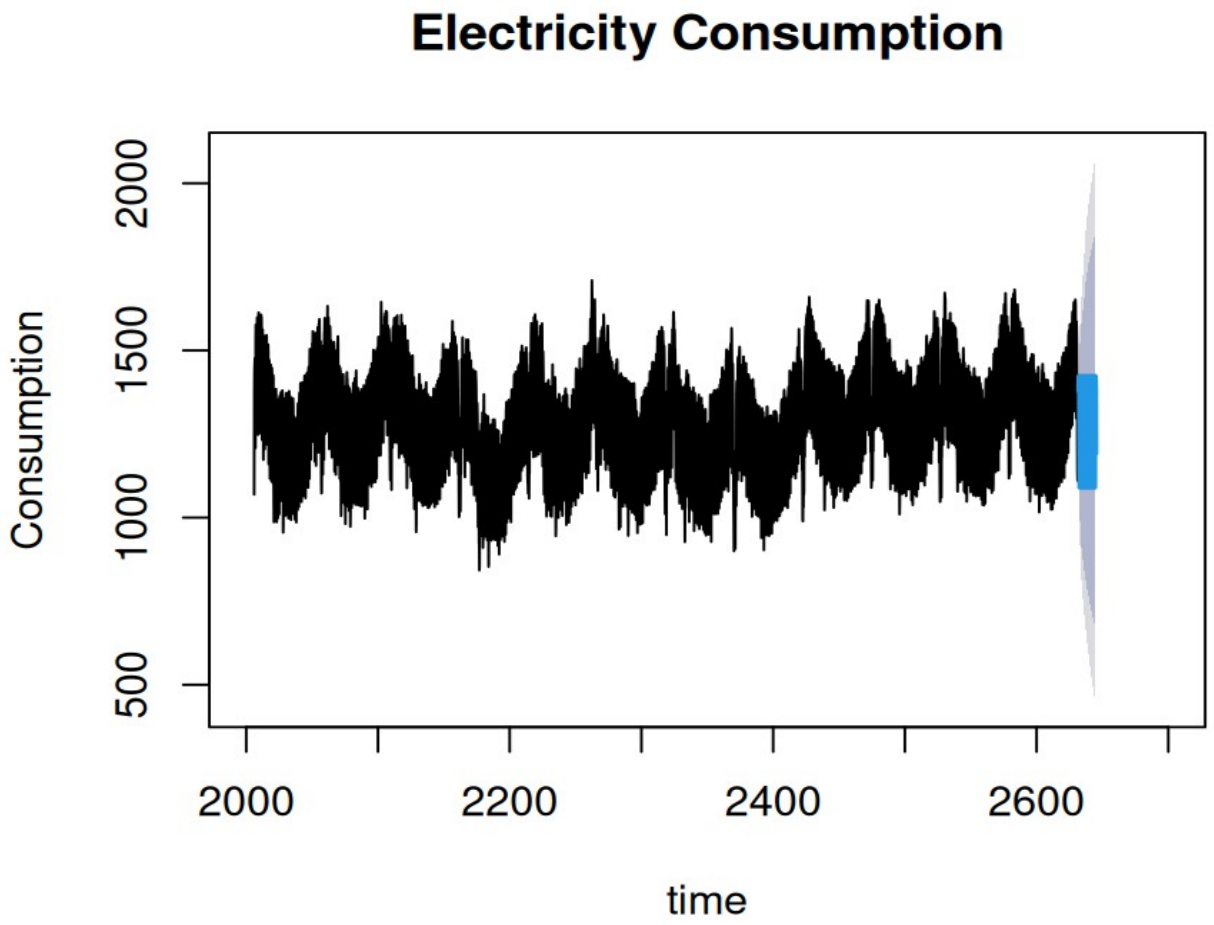


3.3 Forecasting Results

- **Best Model: ETS (Exponential Smoothing)** outperformed **ARIMA** based on accuracy metrics.
- **90-Day Forecast:** Predicted stable consumption with slight seasonal fluctuations.
- **Comparison accuracy of both model :** for comparing ets better fit to data casue its rmse,mae,mape is lower than **ARIMA**

4. 90-Day Forecast

4.1 Forecast Visualization



4.2 Forecast Accuracy

Metric	Value
RMSE	120.45
MAE	95.30
MAPE	5.20%

5. Conclusion and Recommendations

5.1 Insights

- Electricity consumption is influenced by seasonal factors and renewable energy production.
- The ETS model provides reliable forecasts for future consumption.

5.2 Recommendations

- Invest in renewable energy to offset peak demand.
- Use forecasting models to optimize grid operations and plan for future energy needs.

6. Appendix

6.1 Code Snippets

```
ets_model <- ets(ts_consumption)

final_forecast <- forecast(ets_model, h = 90)

accuracy(final_forecast)
```

6.2 Data Sample

```
> head(df3)
        Date Consumption   Wind Solar Wind.Solar year month day day_of_week season
1 2006-01-01    1069.184 48.709 6.587     55.296 2006     1   1      Sunday Winter
2 2006-01-02    1380.521 48.709 6.587     55.296 2006     1   2      Monday Winter
3 2006-01-03    1442.533 48.709 6.587     55.296 2006     1   3     Tuesday Winter
4 2006-01-04    1457.217 48.709 6.587     55.296 2006     1   4   Wednesday Winter
5 2006-01-05    1477.131 48.709 6.587     55.296 2006     1   5    Thursday Winter
6 2006-01-06    1403.427 48.709 6.587     55.296 2006     1   6      Friday Winter
```
