Data Science I

Project Report: Predicting Energy Prices

Submission Deadline: January 25 2026, 21:00 UTC

University of Oldenburg
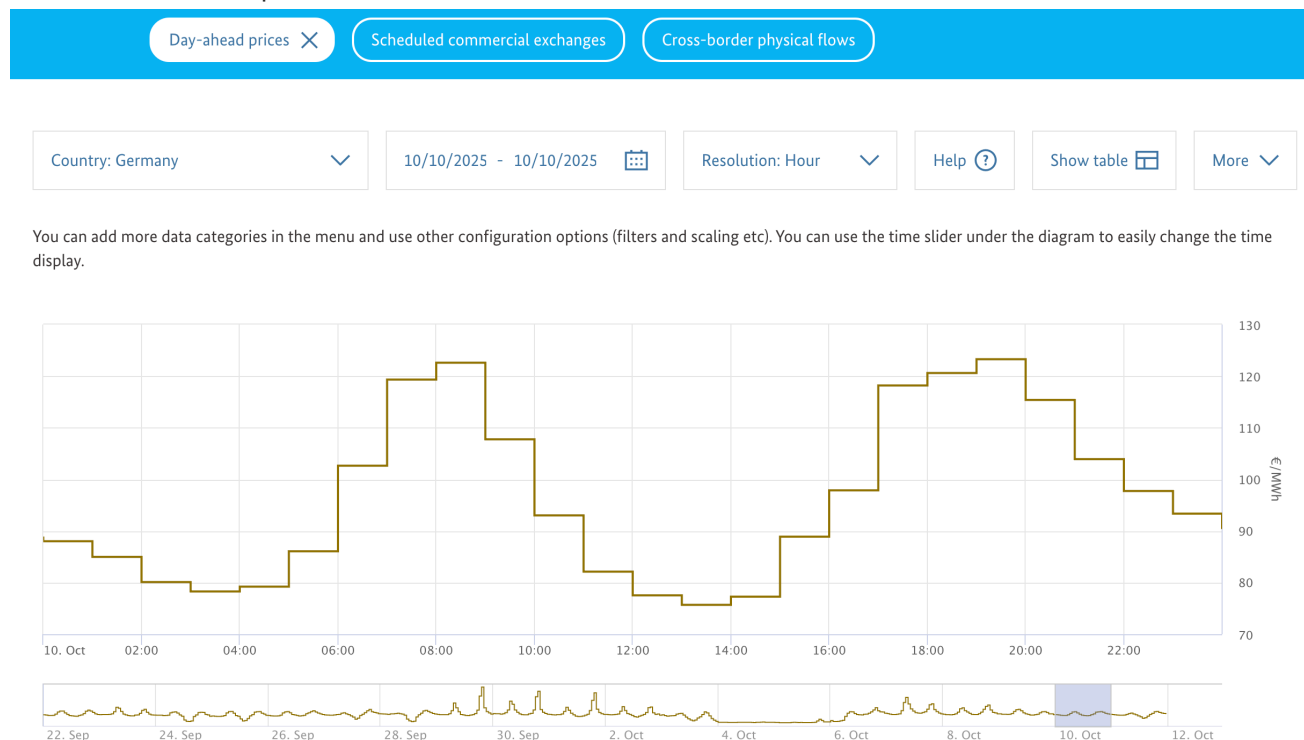**Winter 2025/2026**
Instructors: Jannik Schröder, Wolfram "Wolle" Wingerath

Submitted by: <your names here>

Introduction

Day-ahead prices are the prices set for electricity delivery on the next day, determined through a daily auction process based on anticipated demand and supply. These prices provide a forecast that helps producers and consumers plan their generation and consumption, reflecting expected market conditions. Predicting these prices is crucial for energy companies and traders to optimize their operations, manage risks, and improve market efficiency.

Over the course of this semester project, you will build a model to predict hourly day-ahead energy prices for Germany on **January 27, 2026** as reported by the SMARD.de information platform (see Figure 1 below for an example).

Report Structure & Semester Overview

This module is designed in a way that you can reuse and repurpose work from the exercise sheets when working on the project report. By building this semester project report, you can thus apply the knowledge you pick up in the lectures in a practical context. There are 4 phases overall that are loosely aligned with the topics of the lectures and exercise sessions.

Note: We may update the instructions with clarifications or additional resources during the semester, so **please check Stud.IP for new versions** before starting to work on a new phase. You can find the publication date of the project report instructions on the top-right corner.

Phase 1: Gathering Domain Knowledge & Data Sources `/ 25`

For building a model that predicts energy prices, you are required to gather not only historical data on energy prices, but also gain a strong understanding of the energy sector's dynamics. This entails collecting data for recent years on energy prices as well as demand and production capacities, weather conditions, supply metrics, and any significant events impacting energy prices. Additionally, acquiring domain knowledge is crucial for understanding the factors that influence energy markets.

You can think of your report as a story about data discovery, where the background information sets the stage for understanding the significance and narrative of your analytical findings. It should therefore include important facts on the domain that you are exploring.

- Research the energy market's structure, including how prices are set and the role of different energy sources.
- Understand regulatory policies, market reforms, and technological advancements that could impact energy prices.
- Review energy sector publications, expert analyses, and videos or other resources.

Possible data sources include energy market databases, weather archives, government publications on energy, and news articles reporting on the energy sector. Combining comprehensive data collection with in-depth domain knowledge will lay a solid foundation for developing accurate predictive models and meaningful analysis.

Phase 2: Data Cleaning & Exploratory Data Analysis (EDA) `/ 25`

This phase is critical for uncovering insights from your collected data and setting the stage for predictive modeling.

- Prepare your data for analysis by handling missing values, outliers, and ensuring data quality.
- Use statistical summaries and visualizations to understand the basic characteristics of the data. Explore the distribution of key variables and identify patterns or anomalies.

- Investigate the relationships between different variables, especially how various factors such as weather conditions, demand, and significant events correlate with energy prices. Utilize statistical methods like Pearson's or Spearman's rank correlation coefficient to quantify the strength and direction of these relationships. This step is crucial for identifying potential predictors for your models.

This initial analysis will not only help you understand the data better but also guide your choice of features for the predictive modeling phase. It's essential to document any insights or interesting findings during this stage, as they could be valuable for the final report and presentation.

Phase 3: Visualization & Storytelling   / 25

Effective visualization is key to interpreting the data and predictive modeling results comprehensively.

- Create visualizations that illustrate the historical trends in energy prices and highlight the key factors influencing price fluctuations.
- Use graphs to compare actual vs. predicted energy prices to evaluate the model's performance visually.
- Explain your findings by connecting the dots between visuals, forecasts, and insights, making it easier to understand how energy prices change and what that means for people and businesses.

Visualizations should not only aim to support the narrative of your findings but also uncover insights that might not be immediately apparent from the raw data or statistical metrics alone. A well-crafted visualization can make complex data more accessible and insights more compelling.

Phase 4: Predictive Modeling   / 25

Develop predictive models to forecast average hourly energy prices on **January 27, 2026**. You may consider various modeling techniques such as:

- Linear Regression
- Time Series Analysis (e.g. ARIMA)
- Machine Learning models (e.g. Random Forest, Neural Networks)

Evaluate different models based on reasonable performance metrics (e.g. RMSE, MAE) to select the best-performing model.

Evaluation Criteria

Your project will be evaluated based on the following criteria:

- **Accuracy of Predictions:** The precision of your model's energy price forecasts.

- **Innovativeness:** The novelty of your approach in predicting energy prices.
- **Data Analysis Depth:** The thoroughness of your initial data analysis and insights.
- **Model Complexity and Justification:** The complexity of your model and the rationale behind choosing it.
- **Presentation:** The clarity and effectiveness of your data visualization and presentation.

Submission Guidelines

Please submit the project report in your normal exercise groups in the same way as you submit the exercise sheets (as both **notebook and PDF file**). So please don't forget to …

- … choose a **file name** according to the scheme `2025_ds1_project_report_group01` (please **use your group name as a suffix**) and to
- … include the **execution output** in your submission!