Data Analysis Project Report

Team: I love data analysis
Peter Felber & Andreas Heindl & Jakob Hütter

January 7, 2025

1 Contributions

The following contributions were made by each team member:

- Peter Felber:
  - Data preprocessing tasks
  - Initial visualization development
- Andreas Heindl:
  - Statistical analysis implementation
  - Regression analysis
- Jakob Hütter:
  - Advanced visualizations
  - Report writing and documentation

2 Dataset Description

- Dataset name and source: Solar Power Generation Data by Ani Kannal from Kaggle
- Time period and sampling frequency: data has been collected over a period of 34 days with a sampling frequency of 15 minutes
- Key variables analyzed:
- Basic statistical properties:
  - Number of observations:
  - Missing values:
  - Key statistics:

3 Methods and Analysis

3.1 Data Preprocessing

- Cleaning procedures:
- Outlier handling:
- Missing value treatment:
- Data transformations:

3.2 Exploratory Data Analysis

- Distribution analysis:
- Time series patterns:
- Correlation analysis:
- Key visualizations:

3.3 Statistical Analysis

- Probability analysis:
- Law of Large Numbers demonstration:
- Central Limit Theorem application:
  - Q-Q plot analysis (if applicable):
- Regression analysis:
  - Model selection:
  - Model fitting and validation:
  - Cross-validation (if applicable):

4 Key Findings

4.1 Statistical Insights

- Distribution characteristics:
- Significant correlations:
- Probability analysis results:

4.2 Pattern Analysis

- Temporal patterns:
- Variable relationships:
- Identified anomalies:

4.3 Advanced Analysis Results

- Interactive visualization insights:
- Regression performance:
- Additional findings:

5 Summary and Conclusions

- Main insights:
- Limitations:
- Future analysis suggestions: