



# A DATA SCIENCE APPROACH TO FORECAST ELECTRICITY CONSUMPTION IN AUSTRALIA

David Valido Ramos (z5516338), Katelyn Kemp (z5459347), Nick Mutton (z5549371),  
Senarath Seelanatha (z5595581), Shanjay Perinpanathan (z5339723), Waseem Alashqar  
(z5514810).

School of Mathematics and Statistics  
UNSW Sydney

April 2025

SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS OF  
THE CAPSTONE COURSE ZZSC9020

---

## Abstract

---

---

# Contents

---

Chapter 1	Introduction	1
Chapter 2	Literature Review	2
2.1	Forecasting electricity demand . . . . .	2
2.2	Weather in forecasting electricity demand . . . . .	3
2.3	Defining extreme temperature . . . . .	3
2.4	Modelling electricity demand in extreme weather . . . . .	4
Chapter 3	Material and Methods	5
3.1	Software . . . . .	5
3.2	Description of the Data . . . . .	5
3.3	Pre-processing Steps . . . . .	6
3.4	Data Cleaning . . . . .	6
3.5	Assumptions . . . . .	7
3.6	Modelling Methods . . . . .	7
Chapter 4	Exploratory Data Analysis	8
4.1	Overview of datasets( <b>This section would be going to previous chapter</b> ) . . . . .	9
4.1.1	Air temperature dataset . . . . .	9
4.1.2	Electricity demand dataset . . . . .	9
4.1.3	Forecast demand dataset . . . . .	10
4.1.4	Weather forecast dataset . . . . .	11
4.2	Univariate Analysis . . . . .	11
4.2.1	Temperature dataset . . . . .	11
4.2.2	Electricity Demand . . . . .	12
4.3	Bivariate Analysis . . . . .	14
4.3.1	Data correlation analysis . . . . .	14
4.3.2	Forecast error . . . . .	16
4.4	Summary of Key Findings . . . . .	17
Chapter 5	Analysis and Results	18
Chapter 6	Discussion	19
Chapter 7	Conclusion and Further Issues	20
Appendix		21

---

# CHAPTER 1

## Introduction

---

The ability to accurately forecast electricity demand is vital for future planning by informing decision making that will, among other things, manage costs, and ensure a reliable power supply. According to the client for this project, one of the greatest drivers of variability in electricity demand is the weather, therefore it is important forecasts consider this to improve accuracy. A summary of literature related to electricity forecasting, including incorporating temperature effects is presented in section 2.

The client has provided data related to electricity demand and forecast demand in NSW (sourced from AEMO's Market Management System database), along with air temperature data for Bankstown (sourced from the Australian Data Archive for Meteorology) that will form the basis for this project's forecast modelling.

Initial team analysis of forecast data against actual demand data revealed current electricity forecasts struggle to perform during hotter temperatures. Furthermore, there is expected to be an increase in extreme temperature events due to climate change (BOM, 2024).

The goal of this project is to improve electricity demand forecasting in periods of extreme temperatures, especially for extended periods of extreme temperature. The analysis will involve short-term forecasting (up to 40 hours) to improve on AEMO's forecast specifically, through greater consideration of temporal weather effects including extreme high or low temperatures (e.g. heatwaves).

Initially, time series modelling will be used to understand temporal effects and behaviours of temperature and electricity demand. Other machine learning techniques will also be explored, evaluated and implemented as required. This project plan details software, methods, tasks and team roles.

---

## CHAPTER 2

### Literature Review

---

#### 2.1 Forecasting electricity demand

Energy forecasts play a crucial role in planning and maintaining the energy sector, and ensuring forecasting accuracy helps to manage imbalances in energy production and consumption, reduce power system costs and improve operational safety (Koukaras et al., 2024). Energy forecasting therefore has a broad impact on a wide variety of stakeholders including residential customers, power generators, retailers, traders, industrial and commercial customers, system operators, and financial investors (Ghalekhondabi et al., 2016).

There are many risks in inaccurate energy forecasting. Over forecasting has cost and resource implications for providers, as well as environmental impacts. Under forecasting can cause outages, as well as having down the stream increased costs from inconsistent supply (Suganthi and Samuel, 2012). Shortages are also linked to political instability (Rakpho and Yamaka, 2021).

In the modern world where energy consumption is almost continuous, energy forecasting has become quite complex. There is no universal method of forecasting energy demand, and it is the data and the business need that tends to determine which technique is most useful (Pinheiro, Madeira & Francisco, 2023). Despite this, the forecast interval, which often indicates the goal of the forecast, also influences the kinds of models used.

Forecasting models are typically categorised into short-, medium-, and long-term, and while there is not a unanimous definition of what constitutes these time periods, researchers generally agree that short-term is a few minutes up to a few days (Ahmad and Chen, 2018) or two weeks (Klyuev et al., 2022), medium-term as one month to one year, and long-term as one year to ten years (Ahmad and Chen, 2018). The Australian Energy Market Operator (AEMO) has a more general classification with short-term as up to five years, and long term as longer than this (AEMO, 2022).

Short-term intervals tend to be the most accurate and are important for the effective management of electricity demand and help to reduce peak loads (need reference). Short-term forecast methods can be broadly categorised into two categories – mathematical algorithms such as time-series analysis and logistic regression, and artificial intelligence (AI) algorithms such as machine learning, deep learning and ensemble learning models (Deng et al., 2022). For short-term forecasting, AI methods are becoming more popular as they can consider the non-linear nature of power demand. Short term forecasting is also generally more interested in the accuracy of the forecast rather than the interpretability of the results which makes these

‘black box’ approaches appropriate (Klyuev et al., 2022). Other studies have found that machine learning models tend to outperform traditional models such as ARIMA in short-term forecasting (Divina et al., 2019).

Medium- and long-term forecasting supports the planning and maintenance of the electrical network such as smart grid eco-systems (Ahmad & Chen, 2018). Furthermore, long-term forecasting is more strategic and is necessary for the development of energy systems, planning capital construction at production or infrastructure facilities (Klyuev et al., 2022). These forecast intervals typically use econometric models, system dynamics, and grey prediction, with a focus on policy adjustments, economic indicators (such as GDP and CPI), and population trends (Mystakidis et al., 2024).

## 2.2 Weather in forecasting electricity demand

Temperature is a primary driver of electricity demand, shaping heating and cooling loads that dictate energy consumption. Research consistently identifies it as the dominant weather factor in electricity demand prediction, especially during peak periods. Liu et al. (2021) demonstrate that extreme temperatures lead to increased residential electricity consumption, finding that for each additional day in which the mean temperature exceeds 30 °C, there is an 16.8% increase in monthly residential electricity consumption. Similarly, for each additional day below -6 °C there is a 6% increase in monthly residential electricity consumption. This underscores temperature’s critical role in accurate demand forecasting, as it directly influences consumption patterns during extreme events.

Extreme temperatures can lead to significant errors in electricity demand forecasts, often underestimating demand. During Winter Storm Uri in Texas in February 2021 (Añel, 2024), minimum extreme cold temperatures of -34 °C and high winds of 260 km/h impacted 170 million people. Due to this extreme weather event, electricity demand unexpectedly increased from 40 GW to over 70 GW, resulting on blackouts that affected more than 4 million people. The economic cost of the power outages and disruption has been estimated between 26.1 and 130 billion U.S. dollars.

## 2.3 Defining extreme temperature

There are various approaches to define extreme weather events. The Intergovernmental Panel on Climate Change (IPCC, 2021) uses a percentile-based definition, which ensures a global standardised approach that determines the relative threshold of extreme weather events by region. An extreme heat event is recorded if the maximum temperature during a day is higher than the 90th percentile of historical weather records. Similarly, an extreme cold event is recorded if the minimum temperature during a day is lower than the 10th percentile of historical weather records. Other methods involve an absolute threshold (e.g. 35 °C is a hot day).

Other weather variables, particularly humidity and “feels like” temperature, enhance forecasting accuracy. Maia-Silva et al. (2020) found that using humidity-related measures, such as dew point and heat index, improves prediction accuracy, especially in high-energy-consuming regions, with improvements up to 8-9%. This highlights the need to consider composite weather indices, as air temperature alone underestimates demand during humid heatwaves by as much as 10-15%.

AEMO uses historical weather data from BOM for training their demand forecasting models (AEMO, 2024 Forecasting Assumptions). A weather station per region is selected based on data availability and correlation with regional consumption or demand. The Bureau of Meteorology (BOM) defines a heatwave as maximum temperatures unusually hot for over 3 days compared to local climate, using the Excess Heat Factor (EHF) for monitoring. The EHF combines the observed temperature in the past month, and the average temperature for a 3-day period to measure its severity.

## 2.4 Modelling electricity demand in extreme weather

We have narrowed our focus on a modification of short-term energy demand forecasts. Based on the literature, it appears that it is best suited for reducing demand-forecast variability. Temperature forecasts are reasonably accurate for the forward 1-3 days with diminishing accuracy up to eight days (Floehr, 2010), and energy suppliers have means of modulating their energy outputs based on short-term forecasts (Hydro, 2020) (IEA, 2023). Short-term energy models can be effectively categorised into two groups: those that prioritise explainability, and those that focus on forecast accuracy.

Explainable models consist mostly of traditional statistical and econometric models, such as regression (Papalexopoulos and Hesterberg, 1990) (Ertuğrul, Tekin and Tekin, 2020), time-series such as ARIMA (Tarmanini et al., 2023) (Ediger and Akar, 2007), as well as decision trees (Kopyt et al., 2024) (Wang et al., 2018). They also have natural extrapolations to medium-to-long term models, that are also econometric-based due to their relationship with longitudinal factors such as policy changes, modifications to the energy grid, or economic factors (such as GDP and population) (Ardakani and Ardehali, 2014).

Models that focus on forecast accuracy are generally black box machine learning models such as Neural Networks (Manno, Martelli and Amaldi, 2022) (Kuo and Huang, 2018), Support Vector Machines (Ahmad et al., 2014) (Ahmad et al., 2020), and ensemble methods, such as Random Forests (Divina et al., 2019) and XGBoost (Abbasi et al., 2019).

However, the best performing models tend to be some form of hybrid model using a combination of either explainable or machine learning models, such as NN-ARIMA or CNN-LTSM due to their stability and potential to reduce overfitting (Deng et al., 2022). This feature gives credence to our supplementation of AEMO’s forecast to improve extreme temperature forecasts. Hybrid models have also been shown to produce superior forecasts during temporal extreme temperature events (Phyo and Byun, 2021).

Short-term energy demand forecasting factors include current temperature and other climate measures such as temperature forecast, humidity, dew point, air pressure, windspeed, solar radiation etc (primarily focused on capturing the “feels like” influence of weather), historical energy demand (Suganthi and Samuel, 2012), and time factors such as time of day, day of the week, month, and season (Boroojeni et al., 2017). For extreme temperature, additional factors such as rolling average temperature (Gutiérrez et al., 2013) and consecutive days over/under an average temperature limit are also included (Zhang et al., 2022). Unsupervised learning techniques, like k-means clustering, are also utilised to reduce dimensionality and identify time associations (Singh and Yassine, 2018).

---

## CHAPTER 3

### Material and Methods

---

#### 3.1 Software

Software	Library(s)	Purpose
Python	Pandas	Reading, manipulating, cleaning and analysing datasets.
	Numpy	Manipulating data and mathematical calculations.
	Matplotlib, Seaborn	Visualising data to understand trends and patterns.
	Scikit-learn	Implementing and evaluating machine learning algorithms.
PowerBI	-	Summarising and visualising data.
RMarkdown	-	Writing final report.

#### 3.2 Description of the Data

Table xx describes the data that will be used in analysis. In addition to the datafiles provided by the client, historical weather forecasts including temperature, humidity and wind speed were sourced from OpenWeather, a global company specializing in environmental data products. Forecast weather data rather than actual weather data was used to ensure inputs into the forecast model were realistic.



Data	Description
<b>Electricity demand</b> Use for both training and testing models.	Electricity demand from 2010 to 2021. Well-structured and low complexity with no duplicates and no null values. Variables: Date-time, totalDemand, regionID Format: CSV, Storage: Github, Size: 6 Mb, Rows: 196,513
<b>Air temperature</b> Use to train the behaviour of temperature as a forecasting input.	Provides air temperature data in Bankstown from 2010 to 2021. Well-structured with no null values. It is medium complexity due to uneven time increments and duplicate rows. Variables: Date-time, location, temperature Format: CSV, Storage: Github, Size: 7 Mb, Rows: 220,326
<b>Forecast demand</b> Use as a baseline forecast model and improve it through the inclusion of more forecasting inputs.	Provides forecasted demand data from 2010 to 2021. Well-structured with no null values. It is high complexity due to uneven time increments and duplicate rows. Variables: Date-time, forecastDemand, totalDemand, regionID, preDispatchSeqNo, periodID, lastChange Format: CSV, Storage: Github, Size: 722 Mb, Rows: 10,906,019
<b>Forecast weather indicators</b>	Provides previous forecast weather data for Bankstown from October 7 2017. Variables: temperature, humidity, wind speed Format: CSV, Storage: XXX

### 3.3 Pre-processing Steps

What did you have to do to transform the data so that they become useable?

### 3.4 Data Cleaning

Data was found to be complete for Electricity Demand and Air temperature. Some forecast data were missing for forecast intervals >12 hours. To ensure complete data was used, the forecast model was trained and tested on 12 hour forecast intervals. There were no missing values in the Forecast Weather Indicators data, however the available data begins on October 7 2017.

Consequently, the relevant data used to train and test the forecast model was between October 7 2017 and 17 March 2021 with a 12 hour forecast interval. No further missing values were present in the data. Additional data cleaning steps performed on all datasets are detailed below: 1.Date/time variables were formatted consistently (i.e. d/m/y H:M).

2.Date/time variables were rounded to the nearest 30 minute increment to provide consistent 30-minute intervals.

3. Duplicate date/time rows were removed to ensure each date/time row was unique. After each dataset was cleaned and checked, they were merged into one clean dataset, joined on the unique date/time variable.

### 3.5 Assumptions

### 3.6 Modelling Methods

---

## CHAPTER 4

### Exploratory Data Analysis

---

This section presents an exploratory analysis of the temperature, forecasted demand, and actual electricity demand data. Emphasis is placed on understanding how demand responds to temperature variations and where forecast discrepancies are most pronounced. The data is manipulated and visualised with Python.

Begin with importing the modules required and clean datasets.

```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt

### Temperature dataset
df_temperature = pd.read_csv('temperature_nsw.csv',
                             names = ['location', 'date_time', 'temperature'], skiprows = 1)
df_temperature.date_time = pd.to_datetime(df_temperature.date_time,
                                          format = "%d/%m/%Y %H:%M")

### Demand dataset
df_demand = pd.read_csv('totaldemand_nsw.csv',
                        names = ['date_time', 'total_demand', 'region_id'], skiprows = 1)
df_demand.date_time = pd.to_datetime(
    df_demand.date_time, format = "%d/%m/%Y %H:%M")

### Forecast dataset
df_forecast = pd.read_csv('forecastdemand_nsw.csv',
                           names = ['id', 'region_id', 'period_id', 'forecast_demand',
                                    'date_time_forecast', 'date_time_prediction'],
                           skiprows = 1)
df_forecast.date_time_forecast = pd.to_datetime(
    df_forecast.date_time_forecast, format = "%Y-%m-%d %H:%M:%S")
df_forecast.date_time_prediction = pd.to_datetime(
    df_forecast.date_time_prediction, format = "%Y-%m-%d %H:%M:%S")
```

## 4.1 Overview of datasets(This section would be going to previous chapter)

### 4.1.1 Air temperature dataset

```
print("Locations = {}".format(set(df_temperature.location)))
print("Date Min = {} | Date Max = {}".format(
    df_temperature.date_time.min(), df_temperature.date_time.max()))
print("Temp Min = {} | Temp Max = {}".format(
    df_temperature.temperature.min(), df_temperature.temperature.max()))
print(df_temperature.head())
print("\nRows = {}".format(len(df_temperature)))
print("Average Datetime frequency: {}".format(df_temperature.date_time.diff().mean()))
print("NaN values \n-----\n{}".format(df_temperature.isnull().sum().to_string()))

## Locations = {'Bankstown'}
## Date Min = 2010-01-01 00:00:00 | Date Max = 2021-03-18 00:00:00
## Temp Min = -1.3 | Temp Max = 44.7
##      location      date_time  temperature
## 0  Bankstown 2010-01-01 00:00:00          23.1
## 1  Bankstown 2010-01-01 00:01:00          23.1
## 2  Bankstown 2010-01-01 00:30:00          22.9
## 3  Bankstown 2010-01-01 00:50:00          22.7
## 4  Bankstown 2010-01-01 01:00:00          22.6
##
## Rows = 220326
## Average Datetime frequency: 0 days 00:26:45.453761488
## NaN values
## -----
## location      0
## date_time     0
## temperature   0
```

### 4.1.2 Electricity demand dataset

```
print("Regions = {}".format(set(df_demand.region_id)))
print("Date Min = {} | Date Max = {}".format(
    df_demand.date_time.min(), df_demand.date_time.max()))
print("Demand Min = {} | Demand Max = {}".format(
    df_demand.total_demand.min(), df_demand.total_demand.max()))
print(df_demand.head())
print("\nRows = {}".format(len(df_demand)))
print("Average Datetime frequency: {}".format(
    df_demand.date_time.diff().mean()))
print("NaN values \n-----\n{}".format(
    df_demand.isnull().sum().to_string()))
```

```

## Regions = {'NSW1'}
## Date Min = 2010-01-01 00:00:00 | Date Max = 2021-03-18 00:00:00
## Demand Min = 5074.63 | Demand Max = 14579.86
##
##      date_time  total_demand region_id
## 0 2010-01-01 00:00:00      8038.00    NSW1
## 1 2010-01-01 00:30:00      7809.31    NSW1
## 2 2010-01-01 01:00:00      7483.69    NSW1
## 3 2010-01-01 01:30:00      7117.23    NSW1
## 4 2010-01-01 02:00:00      6812.03    NSW1
##
## Rows = 196513
## Average Datetime frequency: 0 days 00:30:00
## NaN values
## -----
## date_time      0
## total_demand   0
## region_id      0

```

#### 4.1.3 Forecast demand dataset

```

print("Regions = {}".format(set(df_forecast.region_id)))
print("Forecast Date Min = {} | Forecast Date Max = {}".format(
    df_forecast.date_time_forecast.min(), df_forecast.date_time_forecast.max()))
print("Predict Date Min = {} | Predict Date Max = {}".format(
    df_forecast.date_time_prediction.min(),
    df_forecast.date_time_prediction.max()))
print("Forecast Demand Min = {} | Forecast Demand Max = {}".format(
    df_forecast.forecast_demand.min(), df_forecast.forecast_demand.max()))
print(df_forecast.head())
print("\nRows = {}".format(len(df_forecast)))
print("NaN values \n-----\n{}".format(df_forecast.isnull().sum().to_string()))

## Regions = {'NSW1'}
## Forecast Date Min = 2009-12-30 12:31:49 | Forecast Date Max = 2021-03-17 23:31:33
## Predict Date Min = 2010-01-01 00:00:00 | Predict Date Max = 2021-03-18 00:00:00
## Forecast Demand Min = 4422.46 | Forecast Demand Max = 14736.66
##
##      id region_id  ...  date_time_forecast  date_time_prediction
## 0 2009123018    NSW1  ... 2009-12-30 12:31:49      2010-01-01
## 1 2009123019    NSW1  ... 2009-12-30 13:01:43      2010-01-01
## 2 2009123020    NSW1  ... 2009-12-30 13:31:36      2010-01-01
## 3 2009123021    NSW1  ... 2009-12-30 14:01:44      2010-01-01
## 4 2009123022    NSW1  ... 2009-12-30 14:31:35      2010-01-01
##
## [5 rows x 6 columns]
##
## Rows = 10906019
## NaN values

```

```
## -----
## id                0
## region_id         0
## period_id         0
## forecast_demand   0
## date_time_forecast 0
## date_time_prediction 0
```

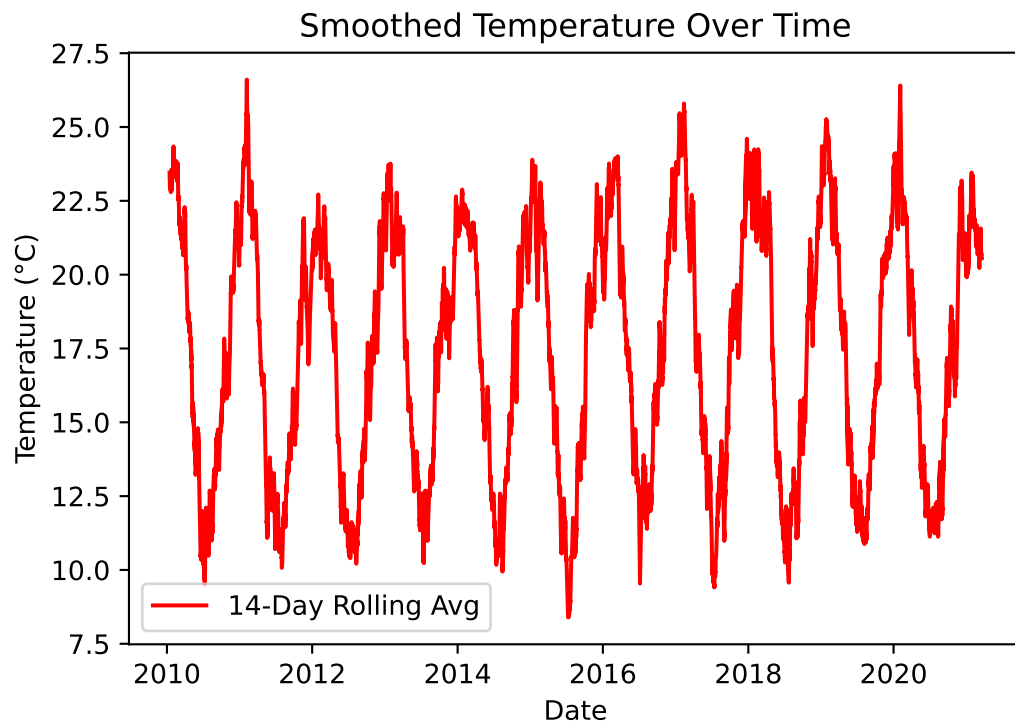
#### 4.1.4 Weather forecast dataset

## 4.2 Univariate Analysis

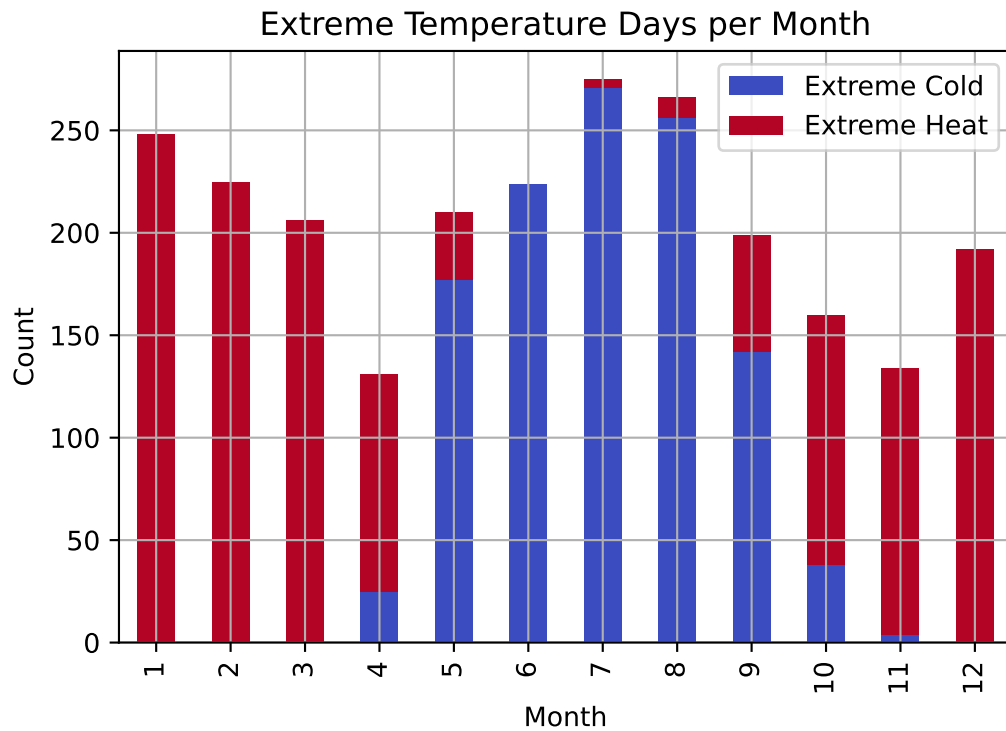
We begin the analysis by focusing on the individual distributions and characteristics of each dataset. This stage provides context on the seasonal variability of the data.

### 4.2.1 Temperature dataset

Temperature data over time reveals an expected cyclical pattern corresponding to seasonal changes consistently across all years.

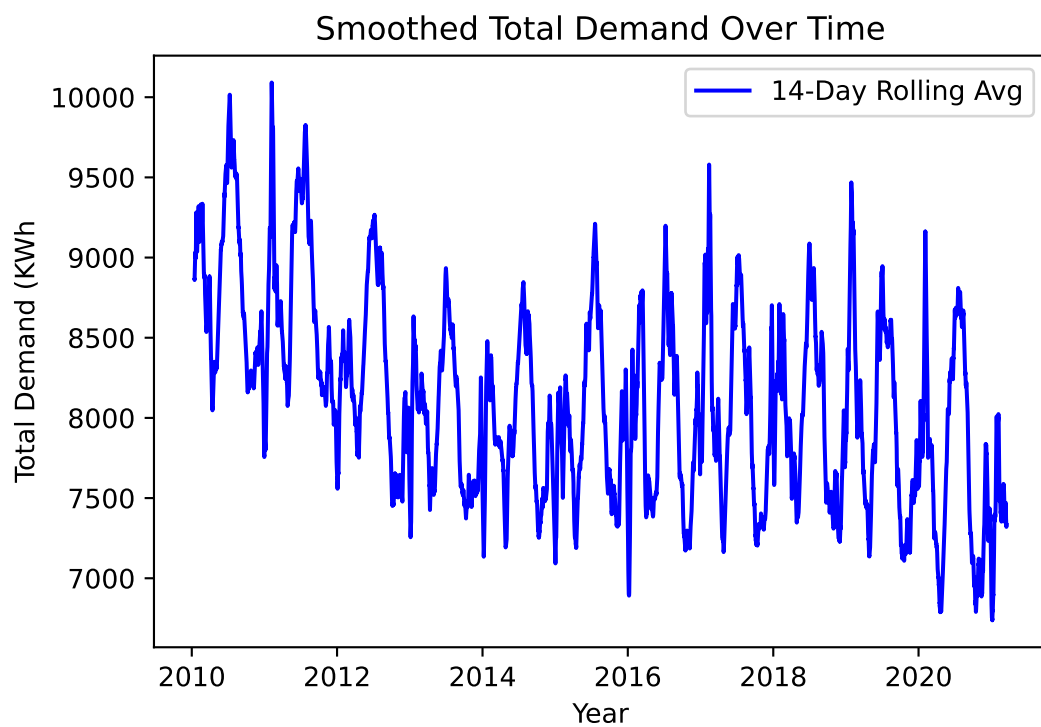


Extreme temperature events occur consistently across the year, with a lower count (but still significant) in April and November. Extreme heat days are defined as days with temperatures above the 90th percentile in the past year, and extreme cold days are defined as days with temperatures below the 10th percentile in the past year.



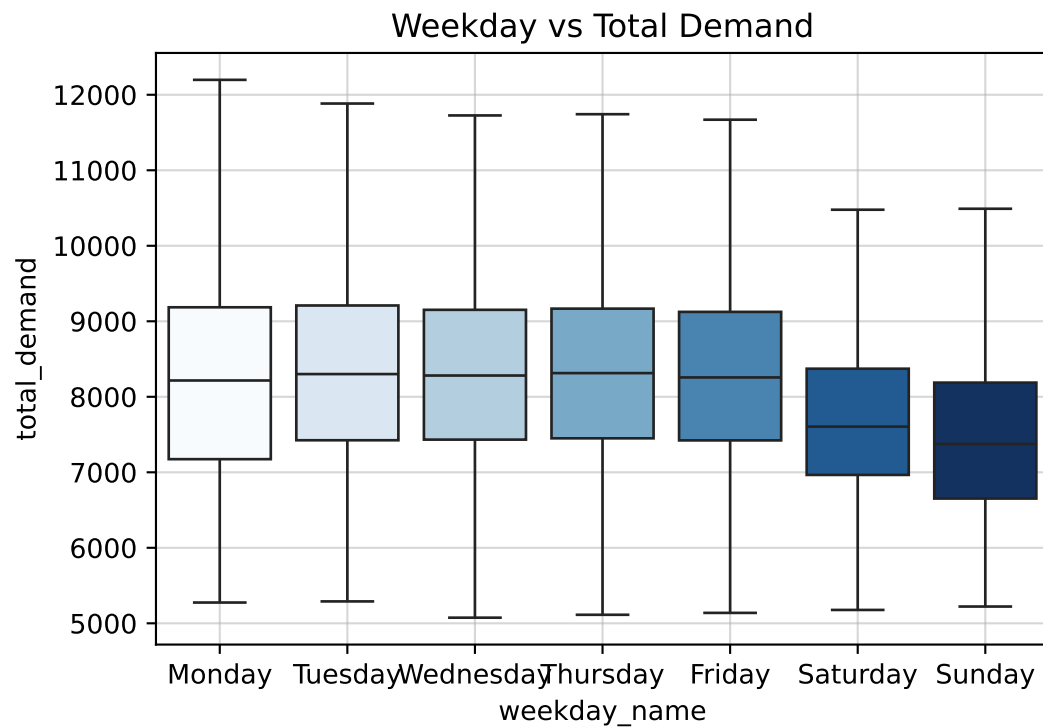
#### 4.2.2 Electricity Demand

Initial exploration of electricity demand data shows high fluctuations with an overall downward trend. Comparing with (Figure x), while both variables are cyclical, it can be observed a more irregular behaviour in electricity demand compared to temperature.

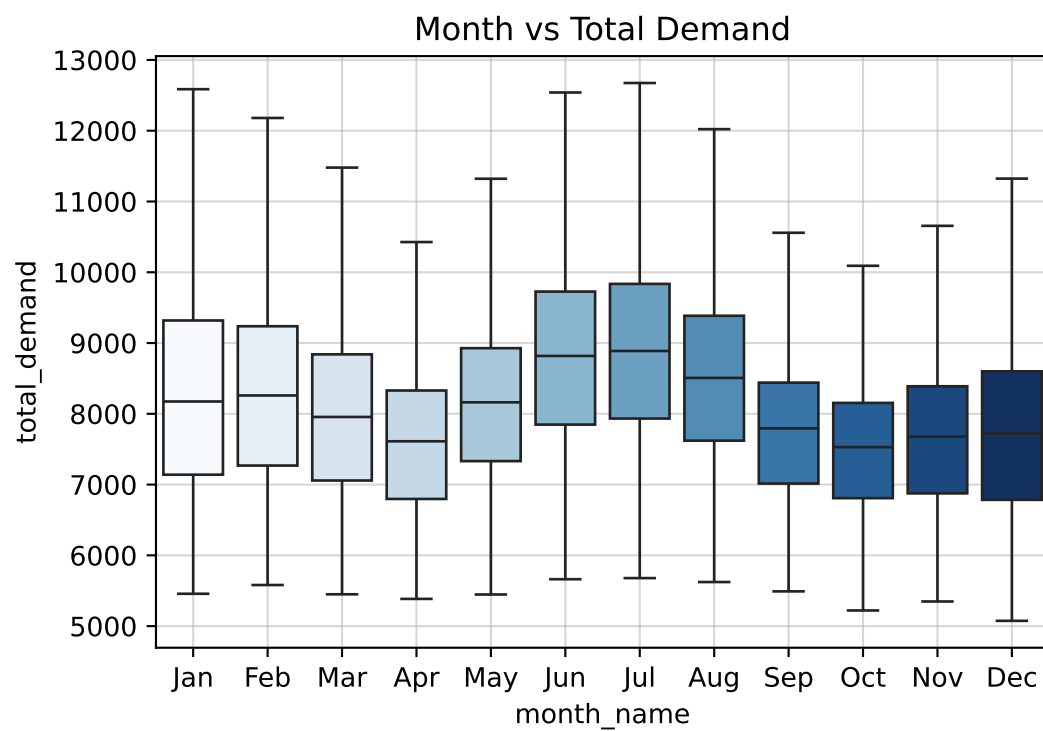


is stable during the weekdays and declines consistently on the weekends.

Demand



Higher electricity demand during winter and summer months, likely due to higher likeliness of extreme temperature events.





### 4.3 Bivariate Analysis

In this next section, we will examine for relationships between the datasets to find if and how temperature affects both the electricity demand and its forecast. To proceed with this analysis, all the datasets are merged following the data cleaning processes mentioned in Section x.x.

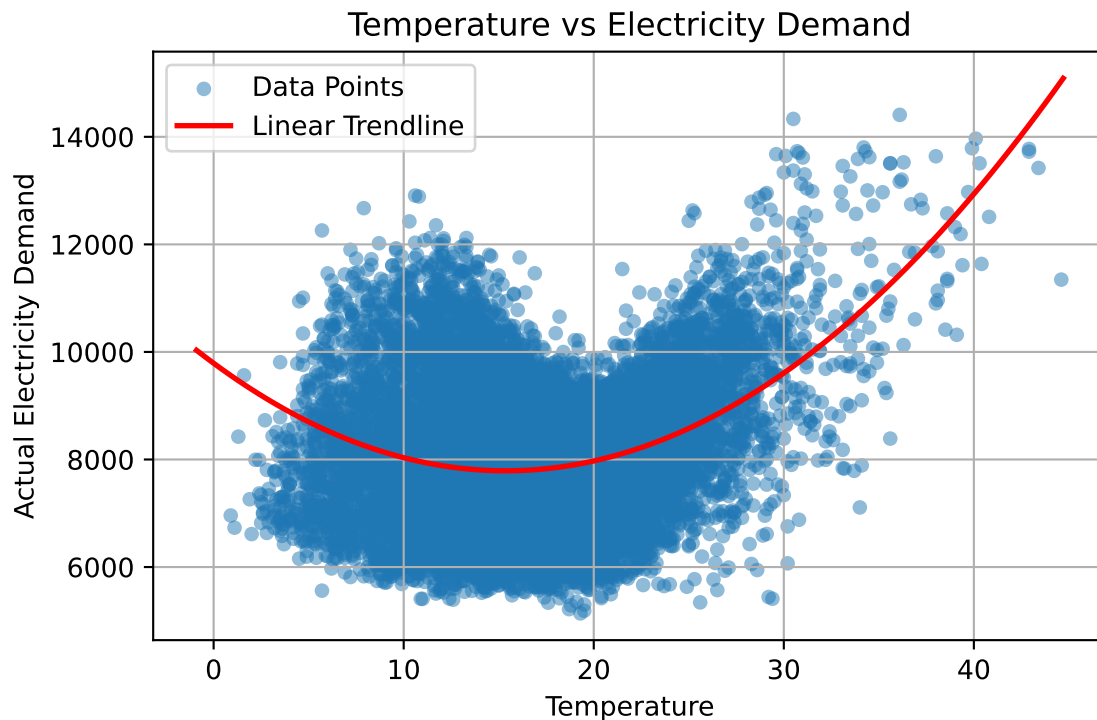
```
#Forecast_interval is the number of hours between prediction and it's forecast
interval = 60*60 #sets the interval in seconds
df_forecast["forecast_interval"] = df_forecast.date_time_prediction - df_forecast.date_time_forecast
df_forecast.forecast_interval = df_forecast.forecast_interval.apply(lambda x: x.total_seconds()//interval)

#Rounding forecast time to intervals of 30mins to match df_demand and only have one instance
interval_min, interval_max = 23 , 25 #sets a window for forecast periods
df_forecast_near24hour = df_forecast.loc[(df_forecast.forecast_interval > interval_min) & (df_forecast.forecast_interval < interval_max)]
df_forecast_near24hour["date_time_forecast_rounded"] = df_forecast_near24hour.date_time_forecast.round("h")
df_forecast_near24hour_1instance = df_forecast_near24hour.loc[df_forecast_near24hour.groupby("date_time_forecast_rounded").count().forecast_interval < 2]

#Merge forecast data with demand data
df_forecast_near24hour_1instance_with_demand = pd.merge(df_forecast_near24hour_1instance, df_demand, on="date_time_forecast_rounded")
df_forecast_near24hour_1instance_with_demand["forecast_error"] = df_forecast_near24hour_1instance_with_demand["forecast"] - df_forecast_near24hour_1instance_with_demand["demand"]
df_merged = pd.merge(df_forecast_near24hour_1instance_with_demand, df_temperature, on="date_time_forecast_rounded", left_on="date_time_forecast_rounded", right_on="date_time")
df_merged["forecast_error_relative"] = df_merged.forecast_error/df_merged.demand
```

#### 4.3.1 Data correlation analysis

As per the literature, the plot of temperature against electricity demand reveals a distinct U-shaped correlation. This pattern reflects energy usage behavior: demand increases during colder periods due to heating needs and rises again during hotter periods as cooling systems are used more intensively. The lowest demand levels generally occur in temperate conditions where neither heating nor cooling is heavily used.



Same

pattern emerges when plotting temperature against the forecasted electricity demand.

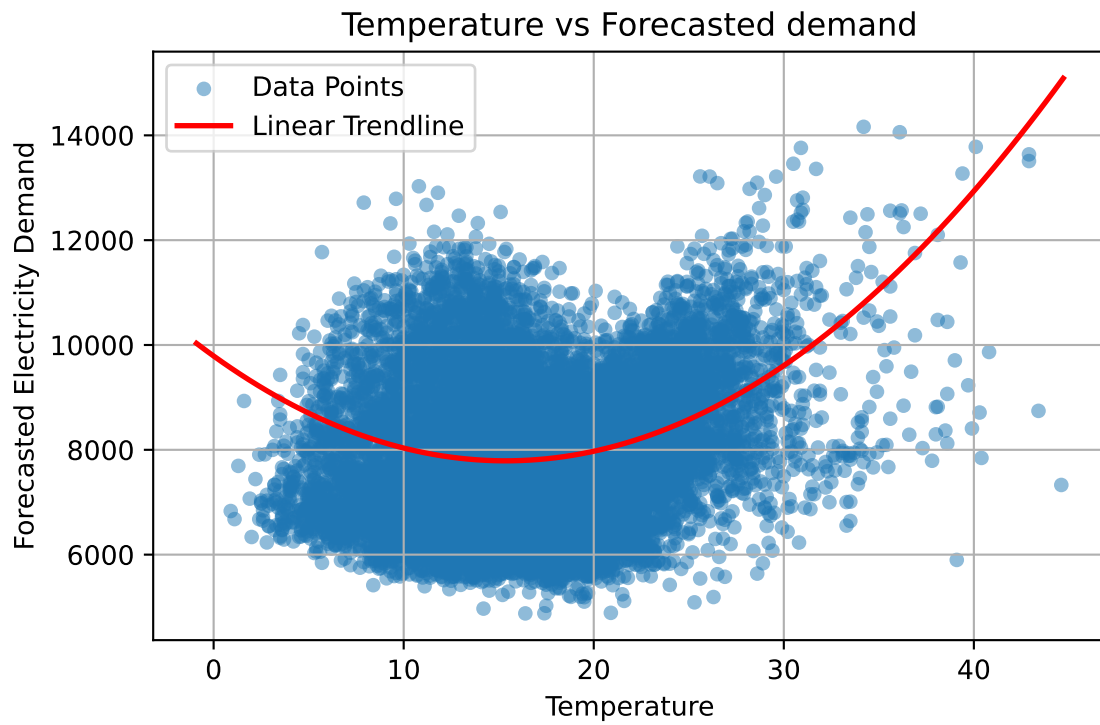
```
#Temperature vs Forecast Demand
coeffs_forecast = np.polyfit(df_merged.temperature, df_merged.forecast_demand, deg=2)
trendline_forecast = np.poly1d(coeffs_demand)

x_trend = np.linspace(df_merged.temperature.min(), df_merged.temperature.max(), 100)
y_trend = trendline_demand(x_trend)

df_sampled = df_merged.sample(frac=0.1, random_state=42)

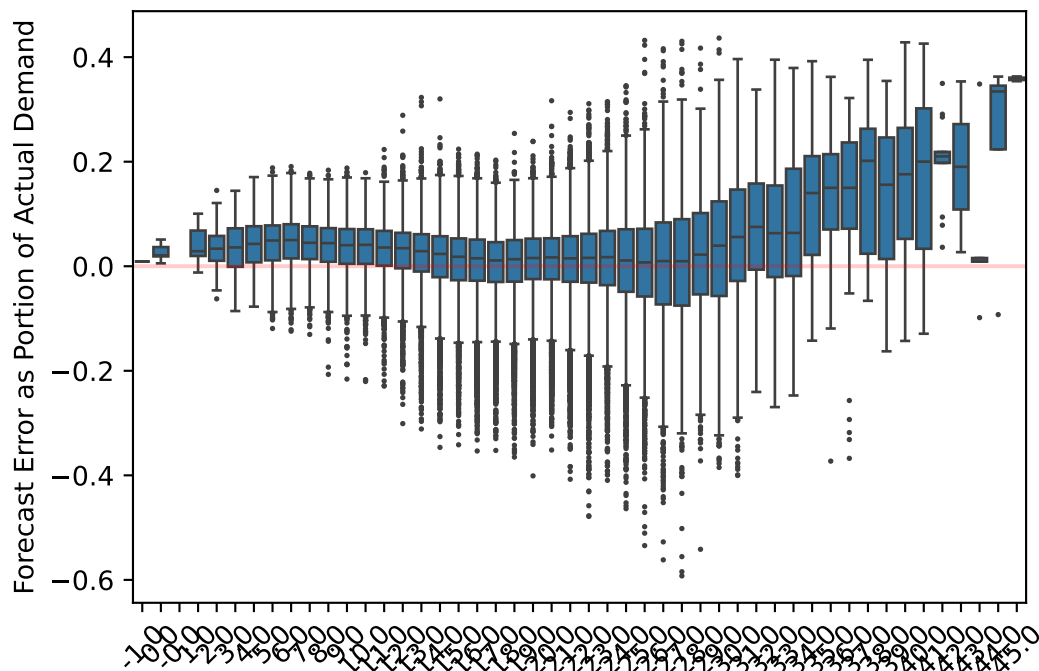
plt.figure(figsize=(6, 4))
plt.scatter(df_sampled.temperature, df_sampled.forecast_demand, alpha=0.5, s=20, label='Data Points')
plt.plot(x_trend, y_trend, color='red', linewidth=2, label='Linear Trendline')

plt.xlabel('Temperature')
plt.ylabel('Forecasted Electricity Demand')
plt.title('Temperature vs Forecasted demand')
plt.legend()
plt.grid(True)
plt.tight_layout()
plt.show()
```



#### 4.3.2 Forecast error

Forecast error notably increases as temperature increases above mild temperatures. This suggests are current forecasting models struggle to predict demand especially under extreme heat temperatures. In the contrary, forecasts during extreme cold temperatures, while still showing an increased forecast error, show to be relatively more accurate than hot temperatures.



## 4.4 Summary of Key Findings

- U-shaped relationship between temperature and electricity demand: Electricity demand increases during both extreme cold and extreme heat conditions, with the lowest demand observed during temperate conditions. This pattern is consistent with expected heating and cooling behavior and is evident in both actual and forecasted demand data.
- Forecasting models capture the general seasonal trend: Forecasted electricity demand shows a similar U-shaped relationship with temperature, indicating that the models are broadly aligned with seasonal usage patterns.
- Forecast error increases with temperature, especially during extreme heat: Forecast accuracy deteriorates significantly at higher temperatures, suggesting that current models underperform during periods of extreme heat. In comparison, performance during extreme cold is better, though still less accurate than under mild conditions.
- Model improvement needed for extreme temperature scenarios: These findings highlight the need for enhancing demand forecasting models, particularly in handling extreme heat events where demand becomes more volatile and difficult to predict.

---

## CHAPTER 5

### Analysis and Results

---

---

## CHAPTER 6

### Discussion

---

Put the results you got in the previous chapter in perspective with respect to the problem studied.

---

## CHAPTER 7

### Conclusion and Further Issues

---

What are the main conclusions? What are your recommendations for the “client”? What further analysis could be done in the future?

---

## Appendix

---