**Question 1:Are there significant differences in electricity demand across different seasons?**

**Test Conducted: ANOVA (Analysis of Variance)**

**Result:**

* **F-statistic = 36624.80321157812**
* **p-value = 0.0**

**Interpretation:**

Since the p-value is 0.0 (which is less than the typical significance level of 0.05), we reject the null hypothesis, also the high F-statistic means large difference between means of different seasons This indicates that there are statistically significant differences in electricity demand across different seasons specially winter and summer.

**Question 2: Is there a statistically significant relationship between temperature and electricity demand?**

**Test Conducted: Pearson correlation and Spearman rank correlation**

**Pearson Correlation Result:**

* **r = 0.5507032707619595,**
* **p-value = 0.0**

**Spearman Correlation Result:**

* **rho = 0.47811995859977857**
* **p-value = 0.0**

**Interpretation:**

Both Pearson and Spearman correlation tests show a statistically significant relationship between temperature and electricity demand, as indicated by the p-values of 0.0. The Pearson correlation coefficient (r = 0.5507) suggests a moderate positive linear relationship, while the Spearman correlation coefficient (rho = 0.4781) suggests a moderate positive monotonic relationship.

**Question 5: How well does the model perform in predicting electricity demand during extreme temperature days compared to average days?**

**Test Conducted:** Independent t-test

**Result:**

* **t-statistic:** 9.154617822304882
* **p-value:** 5.937021962145192e-20

**Interpretation:** The very low p-value (much less than 0.05) indicates that there is a statistically significant difference in model errors between extreme temperature days and average temperature days. The positive t-statistic suggests that the model errors are higher on extreme temperature days compared to average temperature days. This means the model performs worse in predicting electricity demand during extreme temperature days.

**Question 6: What is the lag effect of temperature on electricity demand, and how many days (or hours) of historical data provide the most accurate predictions?**

**Test Conducted:** Cross-correlation analysis

**Result:** The cross-correlation plot (generated in the code) shows the correlation values for different lags (in hours). The lag with the highest correlation value indicates the time delay at which temperature most significantly affects electricity demand.

**Interpretation:** By examining the cross-correlation plot, we can identify the lag (in hours) where the correlation is highest. This lag represents the time delay at which temperature most significantly impacts electricity demand. The lag with the highest correlation value provides the most accurate prediction of electricity demand based on historical temperature data.

**Question 7: Does demand prediction accuracy vary significantly during holidays and weekends compared to regular weekdays?**

**Test Conducted:** Independent t-test

**Result:**

* **t-statistic:** -0.15978847952389535
* **p-value:** 0.8730486360107029

**Interpretation:** The high p-value (greater than 0.05) indicates that there is no statistically significant difference in prediction errors between holidays/weekends and regular weekdays. The near-zero t-statistic suggests that the model errors are very similar between holidays/weekends and regular weekdays. This means the model's prediction accuracy does not vary significantly during holidays and weekends compared to regular weekdays.

**Summary of Statistical Analysis**

**Model Performance on Extreme vs. Average Temperature Days**

* **Conclusion:** The model performs significantly worse on extreme temperature days compared to average temperature days.

**Lag Effect of Temperature on Electricity Demand**

* **Conclusion:** The lag with the highest correlation value indicates the time delay at which temperature most significantly impacts electricity demand.

**Prediction Accuracy on Holidays/Weekends vs. Regular Weekdays**

* **Conclusion:** There is no significant difference in prediction accuracy between holidays/weekends and regular weekdays.

A graph of a line

Description automatically generated with medium confidence