

```
----
title: Predicting Electricity Consumption at Dillman Road Wastewater Treatment Plant
author: Ahmed Ibrahim Yunus
date: "2024-02-21"
output: html_document
----
```

```
# Project: Predicting Electricity Consumption at Dillman Road Wastewater Treatment Plant
```

```
## Introduction
```

In this project, we aimed to develop a predictive model for electricity consumption at the Dillman Road Wastewater Treatment Plant based on influent volume. Electricity consumption is a critical factor in wastewater treatment plant operations, impacting costs and energy efficiency. By accurately predicting electricity consumption, plant operators can optimize resource allocation and improve overall sustainability.

```
## Setup
```

```
` `{r setup, include=FALSE}
# Set CRAN mirror
options(repos = c(CRAN = "https://cran.r-project.org"))
```

```
# Install required packages
```

```
install.packages(c("psych", "lm.beta", "lmtest", "sandwich", "car"))
```

```
# Load libraries
```

```
library(psych)
library(lm.beta)
library(lmtest)
library(sandwich)
library(car)
` `{r}
```

```
## Data Import and Exploration
```

```
` `{r data-import}
# Import dataset
DF <- read.csv("dillman_wwtp_daily_electricity_and_influent_data.csv")
```

```
# View structure and summary of the dataset
```

```
str(DF)
describe(DF)
```

```
# Correlation analysis
```

```
corr.test(DF, use = "complete.obs", method = "pearson", adjust = "bonferroni")
```

```
# Pairwise scatterplot
```

```
pairs.panels(DF, methods = "pearson", density = TRUE)
` `{r}
```

```
## Model Building and Analysis
```

```
` `{r model-building}
# Linear regression model
fit <- lm(Operation_bill ~ MG + billed_kWh + total_kWh, data = DF)
```

```
# Display model summary
```

```
summary(fit)
```

```
# Standardized coefficients
```

```
lm.beta(fit, complete.standardization = TRUE)
```

```
# Confidence intervals
```

```
confint(fit, level = 0.95)
```

```
# Heteroskedasticity-robust standard errors
```

```
vcm <- vcovHC(fit, type = "HC3")
```

```

coeftest(fit, vcov. = vcm)

# Variance inflation factor (VIF)
vif(fit)

# Partial regression plots
avPlots(fit)

# Residual plots
residualPlots(fit, fitted = TRUE, type = "rstudent", test = TRUE)

# Breusch-Pagan test for heteroskedasticity
bptest(fit, studentize = TRUE)

# Influence index plots
infIndexPlot(fit, vars = "cook")
infIndexPlot(fit, vars = "Studentized")
infIndexPlot(fit, vars = "hat")
```


Conclusion

The multiple linear regression model provides insights into the relationship between electricity consumption and influent volume at the Dillman Road Wastewater Treatment Plant. Further analysis and refinement may be necessary for practical implementation and predictive accuracy.


```

```

```


```