

Figure 1: Pairs output showing correlation between Numeric Variables

**Summary**

With a correlation value of <0.1, we find little to no pattern between correlation pairs (PE, PP) and (PE, PS) as the observations were randomly scattered. PP has a weak positive linear relationship with PS of correlation value 0.35 temp has a low correlation value of 0.0046, -0.08957, 0.138 with PS, PP (Negative relationship), PE respectively. Also, larvae have a negative relationship with PE, PP, PS and temp with correlation value of -0.81(strong correlation), -0.52555(moderately correlated), -0.22(weak correlation), and -0.0097(extremely weak correlation) respectively.

Call:

lm(formula = larvae ~ PE + PP + PS + temp, data = data)

Residuals:

Min 1Q Median 3Q Max

-10.1717 -1.8030 -0.3335 1.3810 10.3551

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 52.549730 1.756252 29.922 <2e-16 \*\*\*

PE -0.059817 0.001789 -33.436 <2e-16 \*\*\*

PP -0.206932 0.010990 -18.828 <2e-16 \*\*\*

PS -0.002371 0.007457 -0.318 0.7509

Temp 0.273804 0.113041 2.422 0.0163 \*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.568 on 195 degrees of freedom

Multiple R-squared: 0.8932, Adjusted R-squared: 0.891

F-statistic: 407.5 on 4 and 195 DF, p-value: < 2.2e-16

**Least squares regression equation**

larvae = 52.54973 - 0.059817(PE) - 0.206932(PP) - 0.002371(PS) + 0.273804(temp)

**Test Hypothesis**

H0: βPE = βPP = βPS = βtemp = 0, Given all predictors have been fitted, they are all not useful

Given all predictors are fitted, at a 5% significance level, the significant predictors of fish larvae density include PE, PP, and temp, their p-values of 2e-16, 2e-16 and 0.0163 respectively gives enough significant statistical evidence to reject the null hypothesis (H0) which says all predictors are not useful.

Call:

lm(formula = larvae ~ PE + PP + temp, data = data)

Residuals:

Min 1Q Median 3Q Max

-10.1650 -1.7976 -0.3581 1.3403 10.4370

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 52.519088 1.749580 30.018 <2e-16 \*\*\*

PE -0.059837 0.001784 -33.544 <2e-16 \*\*\*

PP -0.208167 0.010257 -20.294 <2e-16 \*\*\*

temp 0.272595 0.112717 2.418 0.0165 \*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.562 on 196 degrees of freedom

Multiple R-squared: 0.8931, Adjusted R-squared: 0.8915

F-statistic: 545.9 on 3 and 196 DF, p-value: < 2.2e-16

**Least squares regression equation for new model**

larvae = 52.519088 - 0.059837(PE) - 0.208167(PP) + 0.272595(temp)

Estimate Std. Error 2.5 % 97.5 %

(Intercept) 52.51908805 1.749579675 49.06866985 55.96950625

PE -0.05983675 0.001783849 -0.06335475 -0.05631874

PP -0.20816699 0.010257325 -0.22839589 -0.18793810

temp 0.27259525 0.112717147 0.05030112 0.49488938

The 95% CI for βPE is (-0.063, -0.056) indicates that provided other predictors are held constant, a unit increase in PE (polyethylene microplastics) will result in a decrease in the average larvae between -0.056 and -0.063 µg/m3. The 95% CI for βPP is (-0.228, -0.188) indicates that provided other predictors are held constant, a unit increase in PP (polypropylene microplastics) will result in a decrease in the average larvae between -0.188 and -0.228 µg/m3. The 95% CI for βtemp is (0.05,0.495) indicates that provided other predictors are held constant, a unit increase in temp (water temperature) will result in an increase in the average larvae between 0.05 and 0.495 Celsius.

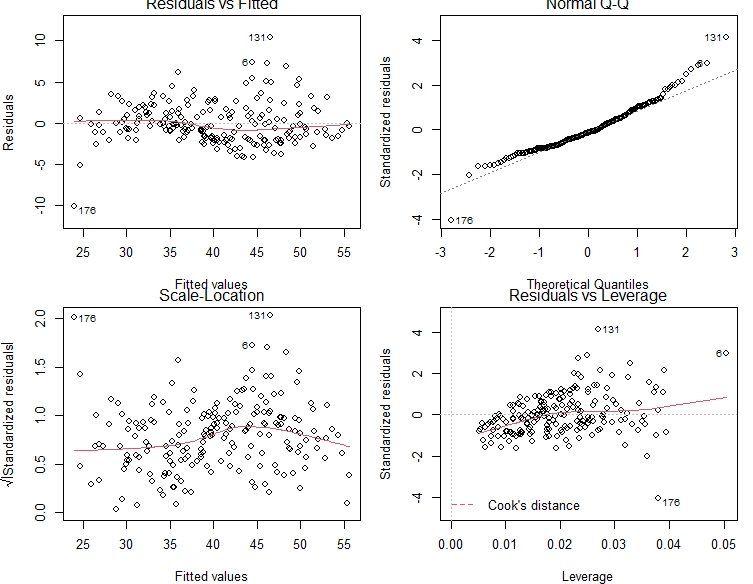


Figure 2: Diagnostic Plot

It can be observed from the residual vs fitted plot that linearity is fairly reasonably because, the red line seems to move closely with the dashed line. The heteroskedasticity while moving towards the right on the x-axis looks like there is a spread of the residuals which appears to be increasing at some points. Finally, points 131, 60 and 176 may be outliers, with large residual values.

The normal q-q plot shows that the residuals are basically normally distributed but a little skewness can be observed on the right of the plot.

The Residuals vs Leverage plot indicates that the outliers present in the data set do not cause significant changes to the model when altered.

**Confidence Level**

Fit lwr upr

1 28.93013 28.06894 29.79132

2 29.74792 28.38541 31.11042

Given PE, PP and PS are held constant with values 300, 50, and 80.5 respectively, at a 95% confidence level, a temperature of 17.5 and 20.5 according to the model will give a mean fish larvae density with values which ranges between (28.06894 and 29.79132) and (28.38541and 31.11042) respectively

**Prediction Level**

fit lwr upr

1 28.93013 23.80537 34.05489

2 29.74792 24.51552 34.98031

According to the model, a temperature of 17.5 and 20.5 will have a 95% prediction interval for the mean fish larvae density with values ranging from (23.80537 and 34.05489) and (24.51552 and 34.98031) respectively provided PE, PP and PS remains constant with values 300, 50, and 80.5 respectively.

**Report**

**Aim:** To get the effects of microplastics on aquatic organisms

**Method:** Regression Analysis was performed Using R

**Result:** It is found out that the percentage of polyethylene and polypropylene microplastics (PE and PP) on aquatic organisms, and the temperature of the water are useful in predicting the larvae density (per 100 m3), while the percentage of the polystyrene microplastics (PS) do not have a significant effect on the larvae density (per 100 m3)

The final model is E(larvae) = 52.519088 - 0.059837(PE) - 0.208167(PP) + 0.272595(temp)

**Conclusion:** The average larvae density will decrease when polyethylene and polypropylene microplastics (PE and PP) increases on aquatic organisms provided other predictors remain constant when either PE or PP is being studied. However, provided other predictors are held constant, the average larvae density is higher for increased temperature at 95% CI. The model using polyethylene, polypropylene microplastics (PE and PP) and temp (water temperature) explains about 89.1% of the variability in larvae density.

**Appendix: R Code**

# Import the data

data <- read.table(file.choose(),header = TRUE)

data

# Plot the correlation pairs figure

source("Rfunctions.R")

pairs(data[2:6],

lower.panel=panel.smooth,

upper.panel = panel.cor)

# Fit the model

mod <- lm(larvae~PE+PP+PS+temp,data)

summary(mod) # Summary table of the model

# larvae = 52.54973 - 0.059817(PE) - 0.206932(PP) - 0.002371(PS) + 0.273804(temp)

# H0: βPE = βPP = βPS = βtemp = 0, Given all predictors have been fitted, they are all not useful

mod2 <- lm(larvae~PE+PP+temp,data) # Fit new model with new derived predictors

print(summary(mod2)) # summary table of new model

# larvae = 52.519088 - 0.059837(PE) - 0.208167(PP) + 0.272595(temp)

# Get 95% confidence intervals for the regression coefficients

betaCI(mod2)

# Get diagnostic plots

par(mfrow=c(2,2))

plot(mod2)

par(mfrow = c(1,1))

# Estimating mean fish larvae density and the 95% CI

temp <- c(17.5,20.5)

PE <- c(300,300)

PP <- c(50,50)

newd <- data.frame(PE,PP,temp)

newd

predict(mod2, newdata = newd, interval = 'confidence','level' = 0.95)

predict(mod2, newdata = newd, interval = 'prediction','level' = 0.95)