573\_HW\_1

Matthew Stoebe

2024-11-03

# Load the Johnson & Johnson data  
data("JohnsonJohnson")  
  
jj\_data <- data.frame(  
 Year = as.numeric(time(JohnsonJohnson)),  
 EPS = as.numeric(JohnsonJohnson)  
)  
  
jj\_data$log\_EPS <- log(jj\_data$EPS)  
  
jj\_data$t <- jj\_data$Year  
jj\_data$Quarter <- as.factor(cycle(JohnsonJohnson))  
jj\_data$Q1 <- ifelse(jj\_data$Quarter == 1, 1, 0)  
jj\_data$Q2 <- ifelse(jj\_data$Quarter == 2, 1, 0)  
jj\_data$Q3 <- ifelse(jj\_data$Quarter == 3, 1, 0)  
jj\_data$Q4 <- ifelse(jj\_data$Quarter == 4, 1, 0)  
  
model <- lm(log\_EPS ~ t + Q1 + Q2 + Q3 + Q4 - 1, data = jj\_data)  
  
summary(model)

##   
## Call:  
## lm(formula = log\_EPS ~ t + Q1 + Q2 + Q3 + Q4 - 1, data = jj\_data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.29318 -0.09062 -0.01180 0.08460 0.27644   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## t 1.672e-01 2.259e-03 74.00 <2e-16 \*\*\*  
## Q1 -3.283e+02 4.451e+00 -73.76 <2e-16 \*\*\*  
## Q2 -3.282e+02 4.451e+00 -73.75 <2e-16 \*\*\*  
## Q3 -3.282e+02 4.452e+00 -73.72 <2e-16 \*\*\*  
## Q4 -3.284e+02 4.452e+00 -73.77 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.1254 on 79 degrees of freedom  
## Multiple R-squared: 0.9935, Adjusted R-squared: 0.9931   
## F-statistic: 2407 on 5 and 79 DF, p-value: < 2.2e-16

beta <- coef(model)["t"]  
cat("Estimated average annual increase in logged earnings per share:", beta, "\n")

## Estimated average annual increase in logged earnings per share: 0.1671722

alpha3 <- coef(model)["Q3"]  
alpha4 <- coef(model)["Q4"]  
  
delta\_alpha <- alpha4 - alpha3  
cat("Change from Q3 to Q4 in logged earnings:", delta\_alpha, "\n")

## Change from Q3 to Q4 in logged earnings: -0.2687577

percentage\_change <- (delta\_alpha)/abs(alpha3)  
cat("Percentage change from Q3 to Q4:", percentage\_change, "%\n")

## Percentage change from Q3 to Q4: -0.0008189384 %

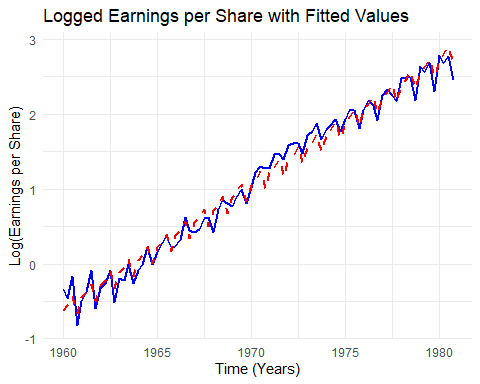
The average Log earnings rate decreases slightly from Q3 to Q4

model\_with\_intercept <- lm(log\_EPS ~ t + Q1 + Q2 + Q3 + Q4, data = jj\_data)  
summary(model\_with\_intercept)

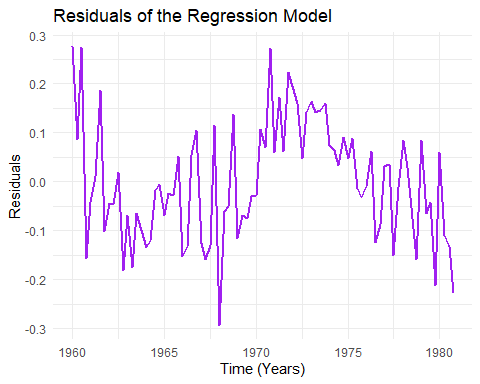
##   
## Call:  
## lm(formula = log\_EPS ~ t + Q1 + Q2 + Q3 + Q4, data = jj\_data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.29318 -0.09062 -0.01180 0.08460 0.27644   
##   
## Coefficients: (1 not defined because of singularities)  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -3.284e+02 4.452e+00 -73.771 < 2e-16 \*\*\*  
## t 1.672e-01 2.259e-03 73.999 < 2e-16 \*\*\*  
## Q1 1.705e-01 3.873e-02 4.403 3.31e-05 \*\*\*  
## Q2 1.986e-01 3.871e-02 5.132 2.01e-06 \*\*\*  
## Q3 2.688e-01 3.870e-02 6.945 9.50e-10 \*\*\*  
## Q4 NA NA NA NA   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.1254 on 79 degrees of freedom  
## Multiple R-squared: 0.9859, Adjusted R-squared: 0.9852   
## F-statistic: 1379 on 4 and 79 DF, p-value: < 2.2e-16

jj\_data$fitted <- fitted(model)  
  
# Plot the data and fitted values  
library(ggplot2)  
  
ggplot(jj\_data, aes(x = t, y = log\_EPS)) +  
 geom\_line(color = "blue", size = 1) +  
 geom\_line(aes(y = fitted), color = "red", size = 1, linetype = "dashed") +  
 labs(title = "Logged Earnings per Share with Fitted Values",  
 x = "Time (Years)",  
 y = "Log(Earnings per Share)") +  
 theme\_minimal()

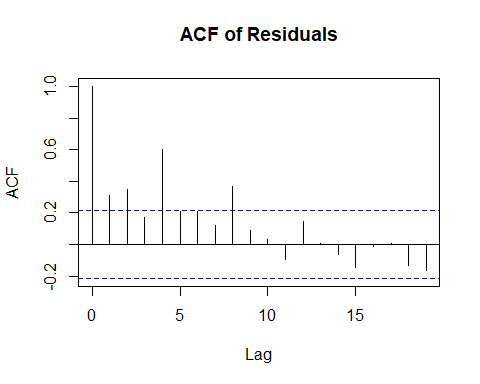
## Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.  
## ℹ Please use `linewidth` instead.  
## This warning is displayed once every 8 hours.  
## Call `lifecycle::last\_lifecycle\_warnings()` to see where this warning was  
## generated.



jj\_data$residuals <- resid(model)  
  
ggplot(jj\_data, aes(x = t, y = residuals)) +  
 geom\_line(color = "purple", size = 1) +  
 labs(title = "Residuals of the Regression Model",  
 x = "Time (Years)",  
 y = "Residuals") +  
 theme\_minimal()



acf(jj\_data$residuals, main = "ACF of Residuals")

 The Residuals do not look like white noise. in between 1970 and 1975 where we seem to systematically under predict. This model is OK but not great