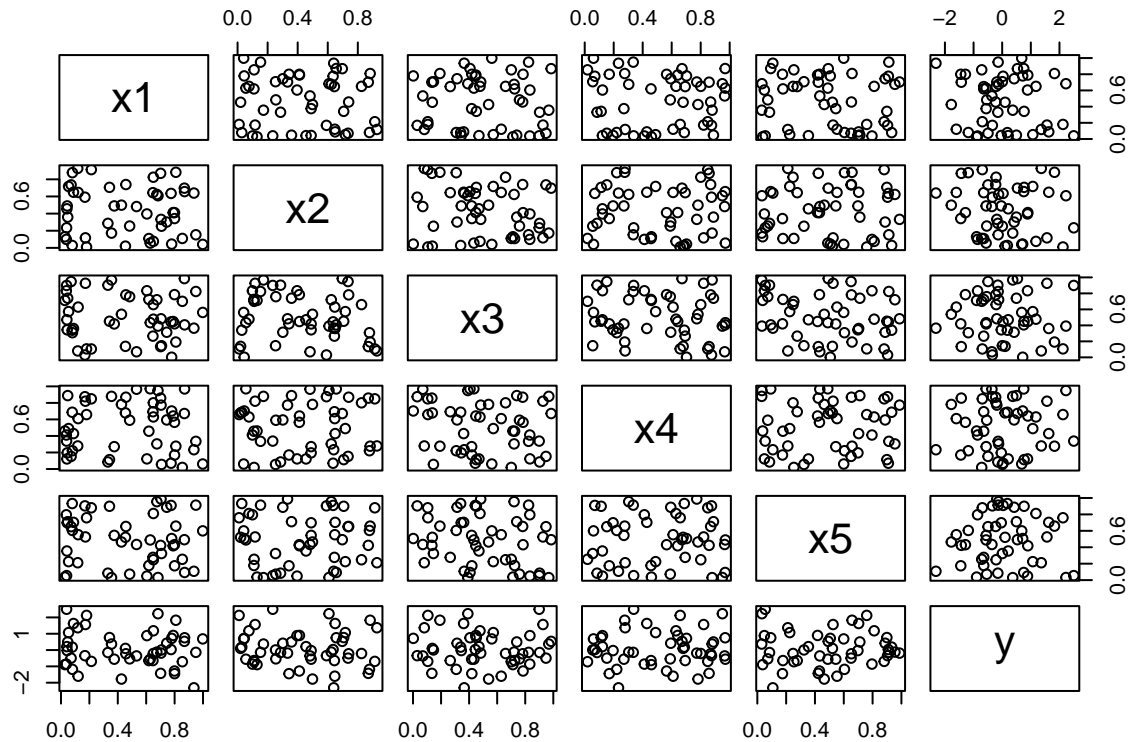


Assignment 3

Question 2.4)

```
library(broom)
```

```
mymat <- cbind(matrix(runif(5 * 50), ncol = 5), rnorm(50))
colnames(mymat) <- c(paste0("x", 1:5), "y")
pairs(mymat)
```



```
mylm <- lm(y ~ ., data = as.data.frame(mymat))
```

```
# Set up important quantities
```

```
X <- mymat[, 1:5]
```

```
Xbar <- apply(X, 2, mean)
```

```
hat_values <- hatvalues(mylm)
```

```
#Euclidean Distance
```

```
euc_dist <- sqrt(rowSums((X-Xbar)^2))
```

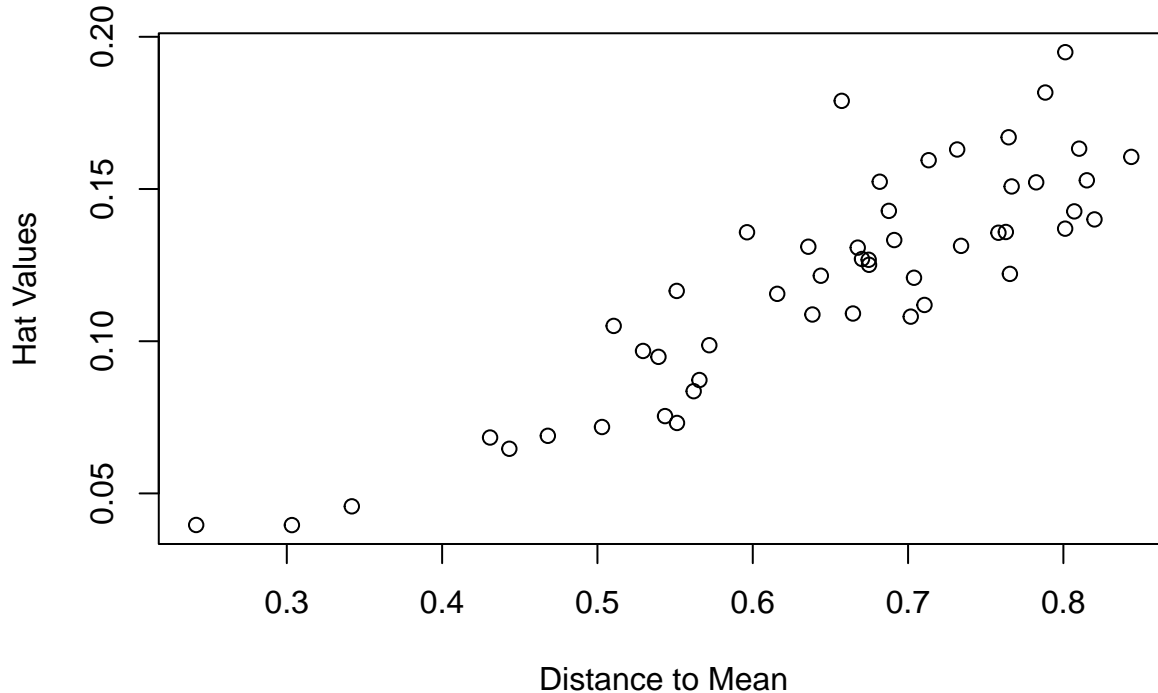
```
euc_dist
```

```
## [1] 0.5619560 0.6675887 0.7655592 0.6438253 0.5656482 0.4680709 0.5293498
## [8] 0.6572838 0.7017132 0.4308705 0.6876291 0.7316161 0.6910941 0.6357706
## [15] 0.7581654 0.5512143 0.7105529 0.8012151 0.6748615 0.7647424 0.5720281
## [22] 0.5435200 0.8200168 0.8011190 0.5963017 0.8101357 0.7824868 0.4432744
## [29] 0.8150880 0.3032791 0.7038192 0.6703323 0.7883414 0.3418472 0.6645039
```

```
## [36] 0.2416508 0.7132529 0.8437114 0.7341077 0.6746442 0.7629358 0.5104633
## [43] 0.6817572 0.8070003 0.6383903 0.7666798 0.5392851 0.5510380 0.6157534
## [50] 0.5029343
```

#Plotting hat values against distance to the mean

```
plot(euc_dist, hat_values, xlab = "Distance to Mean", ylab = "Hat Values")
```



Question 3)

```
n <- 100 # Number of observations
rho <- 0.5

errors <- numeric(n) # Initialize the errors vector
errors[1] <- 0 # Set the first error to 0

set.seed(500) # Set a seed for reproducibility

for (t in 2:n) {
  errors[t] <- rho * errors[t-1] + rnorm(1)
}

x <- 1:n # Predictor variable
y <- x + errors # Response variable

mymodel <- lm(y ~ x)

library(stats)

residuals <- resid(mymodel)
pacf(residuals)
```

Series residuals

