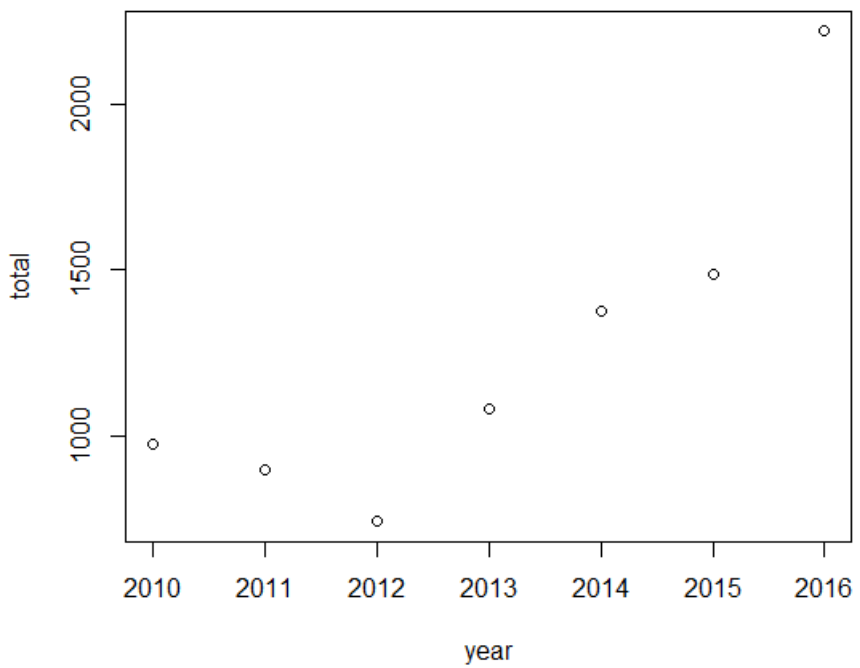
16

a)
```
> data <- read.csv(file = "https://raw.githubusercontent.com/galaamn/source-code-on-statistics/master/Data%20Sets/Controversy/data.csv")
```

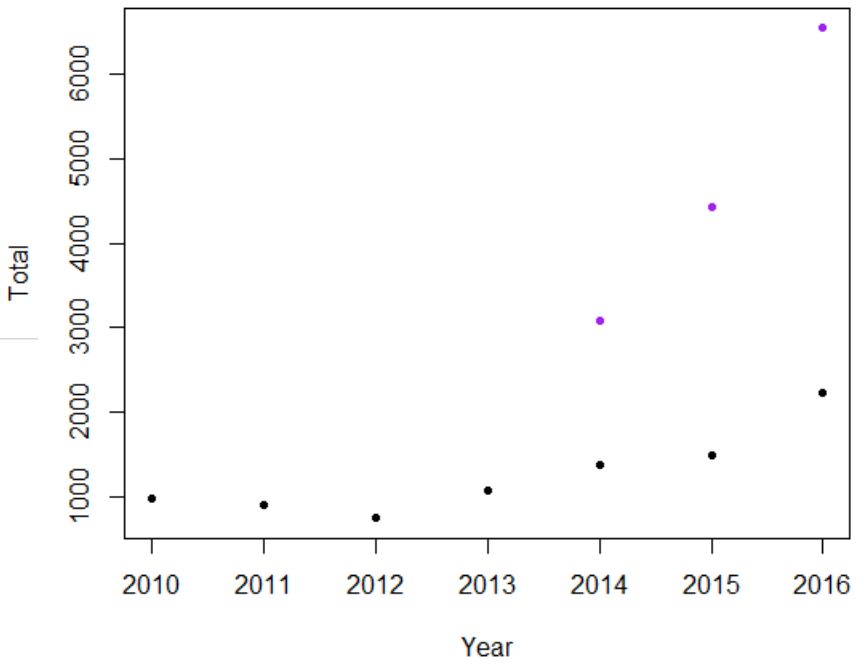
b)
```
> data <- read.csv(file = "https://raw.githubusercontent.com/galaamn/source-code-on-statistics/master/Data%20Sets/Controversy/data.csv")
> data <- data[1:2]
> print(data)
  year total
1 2010   978
2 2011   897
3 2012   742
4 2013  1080
5 2014  1377
6 2015  1487
7 2016  2220
```

c)



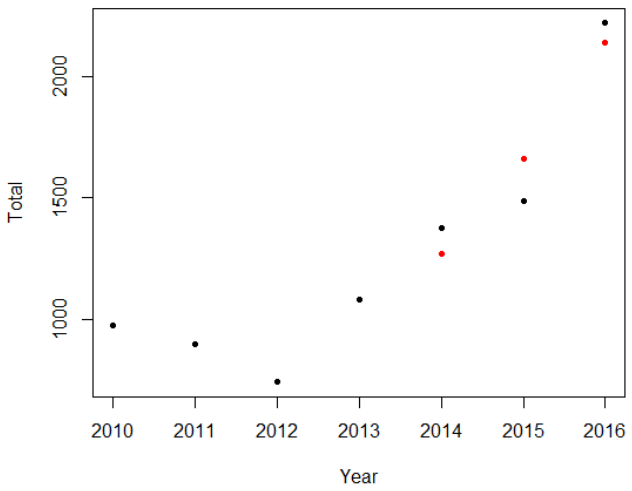
d)

```
data <- read.csv(file = "https://raw.githubusercontent.com/galaamn/source-code-on-statistics/master/Data%20Sets/Controversy/data.csv")
data <- data[1:2]
print(data)
plot(data)
partial_autocorrelations <- acf(x = data$total, lag.max
                                = 5, type = "partial",plot = TRUE)
print(partial_autocorrelations)
fit.ar <- ar.ols(x=data$total, order.max=1)
forecast <- predict(fit.ar, n.ahead = 3)$pred
plot(x = c(data$year, 2014:2016), y = c(data$total,
                                        forecast), xlab="Year", ylab="Total",
     pch = 20, col = c("black","Purple")[rep.int(x=1:2,
                                       times=c(length(data$year),3))])
```



e)

```
> plot(data)
> fit <- lm(formula = log2(total) ~ poly(x=year, degree = 3),
+           data = data, subset = year >=2010)
> forecast <- 2 ** predict(object = fit, newdata =
+                            data.frame(
+                              year = 2014:2016
+                            ))
> plot(x = c(data$year, 2014:2016), y = log2(c(data$total,
+                                              forecast)), xlab = "Year", ylab = "Total",
+      pch = 20, col = c("black", "red")[rep.int(x = 1:2,
+                                                times = c(length(data$year), 3))])
> plot(x = c(data$year, 2014:2016), y = c(data$total,
+                                         forecast), xlab = "Year", ylab = "Total",
+      pch = 20, col = c("black", "red")[rep.int(x = 1:2,
+                                                times = c(length(data$year), 3))])
```



f) Авторегрессийн загвар илүү үнэмшилтэй байна.

16.5

a)

```
The downloaded binary packages are in
        C:\Users\ldava\AppData\Local\Temp\RtmpodZa9t\downloaded_packages
> data <- epitools::expand.table(HairEyeColor[,,2]
+ install.packages("epitools")
```

b)

```
n <- nrow(data)
set.seed(0)
ind <- sample.int(n = n, size = round(0.2 * n))
training.dataset <- data[-ind,]
testing.dataset <- data[ind,]
```

c)

```
data <- epitools::expand.table(HairEyeColor[,,2])
n <- nrow(data)
set.seed(0)
ind <- sample.int(n = n, size = round(0.2 * n))
training.dataset <- data[-ind,]
testing.dataset <- data[ind,]
lassifier <- klaR::NaiveBayes(formula = class ~ ., data =
                                training.dataset)

```

d)

```
data <- epitools::expand.table(HairEyeColor[,,2])
n <- nrow(data)
set.seed(0)
ind <- sample.int(n = n, size = round(0.2 * n))
training.dataset <- data[-ind,]
testing.dataset <- data[ind,]
lassifier <- klaR::NaiveBayes(formula = class ~ ., data =
                                training.dataset)
test <- predict(classifier, testing.dataset)
table("true" = testing.dataset$class, "classifier" =
        test$class)
```

16.4

```
data <- read.csv(file = "https://www.datahub.io/machine-learning/optdigits/r/1.html", header = TRUE)
data$class <- as.factor(data$class)
tapply(X = data$input64, INDEX = data$class, FUN = var)
data$input64 <- NULL
set.seed(0)
subset <- sample.int(n = nrow(data), size =
                       round(nrow(data) * 0.8), replace = FALSE)
training.dataset <- data[subset,]
testing.dataset <- data[-subset,]
classifier <- klaR::NaiveBayes(formula = class ~ ., data =
                                 training.dataset)
test <- predict(classifier, testing.dataset)
table("true" = testing.dataset$class, "classifier" =
        test$class)
```