Autor:	Radosław Smoter
Numer sprawozdania:	4
Data oddania:	15.05.2022
Prowadzący:	dr inż. Krzysztof Schiff

Probabilistyka w zastosowaniach technicznych

Kraków 2022

1 Kody

1 dlf<-read.delim(”DownloadFestival.dat”, header=TRUE)

```
# Wczytaj data frame DownloadFestival.dat.
dlf <- read.delim("DownloadFestival.dat", header = T);
```

2 hist.day1<-ggploat(dlf,aes(day1))+opts(legend.position=”none”) +geom_histogram(aes(y=..density..),colour=”black”, fill=”white”)+labs(x=”Hygiene score on day 1”, y=”Density”) + stat_fucntion(fun=dnorm, args=list(mean=mean(dlf$day1, na.rm=TRUE), sd=sd(dlf$day1, na.rm=TRUE)), colour=”black”, size=1)

2.1 hist.day1

2.2 qqplot.day1<-qplot(sample=dlf$day1, stat=”qq”)

2.3 qqplot.day1

```
# opts jest zdeprecjonowana, zamiast tego uzywam theme; geom_histogram
wyswietla histogram funkcji gestosci, labs podpisuje osie, stat_function
dodaje do wykresu linie normy
hist.day1 <- ggplot(dlf, aes(day1)) + theme(legend.position = "none") +
geom_histogram(aes(y = ..density..), color = "black", fill = "white") +
labs(x = "Hygiene score on day 1", y = "Density") + stat_function(fun =
dnorm, args = list(mean = mean(dlf$day1, na.rm = T), sd = sd(dlf$day1,
na.rm = T)), color = "black", size = 1);
hist.day1;
# Utworz prosty wykres dla day1; stat jest zdeprecjonowany.
qqplot.day1 <- qplot(sample = dlf$day1);
qqplot.day1;
```

3 Wykonaj powyższe dla day2 i day3

```
# Powtorz dla day2.
hist.day2 <- ggplot(dlf, aes(day2)) + theme(legend.position = "none") +
geom_histogram(aes(y = ..density..), color = "black", fill = "white") +
labs(x = "Hygiene score on day 2", y = "Density") + stat_function(fun =
dnorm, args = list(mean = mean(dlf$day2, na.rm = T), sd = sd(dlf$day2,
na.rm = T)), color = "black", size = 1);
hist.day2;
qqplot.day2 <- qplot(sample = dlf$day2);
qqplot.day2;
# Powtorz dla day3.
hist.day3 <- ggplot(dlf, aes(day3)) + theme(legend.position = "none") +
geom_histogram(aes(y = ..density..), color = "black", fill = "white") +
labs(x = "Hygiene score on day 3", y = "Density") + stat_function(fun =
dnorm, args = list(mean = mean(dlf$day3, na.rm = T), sd = sd(dlf$day3,
na.rm = T)), color = "black", size = 1);
hist.day3;
```

```
qqplot.day3 <- qplot(sample = dlf$day3);
qqplot.day3;
```

4 rexam←read.delim(”rexam.dat”,header=TRUE)

4.1 rexam$uni<-factor(rexam$uni, levels=c(0:1),labels=c(”Duncetown University”, ”Sussex University”)) – wytłumacz na tym przykładzie co robi funkcja factor()

```
# factor zamienia wskazane wartosci na inne.
rexam <- read.delim("RExam.dat", header = T);
rexam$uni <- factor(rexam$uni, levels = c(0:1), labels = c("Duncetown University", "Sussex University"));
```

5 Wykorzystaj do danych z pliku rexam.dat funkcję stat.desc() i otrzymaj histogramy dla scores, computer literacy, numeracy oraz lectures attended.

```
# Utworz tabele podstawowych wlasnosci statystycznych. Tylko statystyki opisowe.
desc_rexam <- stat.desc(rexam, basic = F);
ggplot(data = desc_rexam, aes(exam)) + geom_histogram();
# Powtorz dla pozostalych kolumn.
ggplot(data = desc_rexam, aes(computer)) + geom_histogram();
ggplot(data = desc_rexam, aes(lectures)) + geom_histogram();
ggplot(data = desc_rexam, aes(numeracy)) + geom_histogram();
```

6 by(cbind(data=rexam$exam, data=rexam$numeracy), rexam$uni, describe)

```
# Przeksztalc dane dla wartosci kolumn uni, pokaz dla nich wlasnosci statystyczne.
by(cbind(data = rexam$exam, data = rexam$numeracy), rexam$uni, describe);
```

7

7.1 hist.numeracy.duncetown <- ggplot(dunceData, aes(numeracy)) + opts(legend, position = ”none” + geom_histogram(aes(y = ..density..), fill = ”white”, colour = ”black”, binwidth = 1) + labs(x = ”Numeracy Score”, y = ”Density”) + stat_function(fun = dnorm, args = list(mean = mean(dunceData&numeracy, na.rm = TRUE), sd = sd(dunceData$numeracy, na.rm = TRUE)), colour = ”blue”, size = 1)

7.2 hist.numeracy.duncetown

7.3 to samo co powyżej dla Susses University dla exam i numeracy i dla duncetown exam

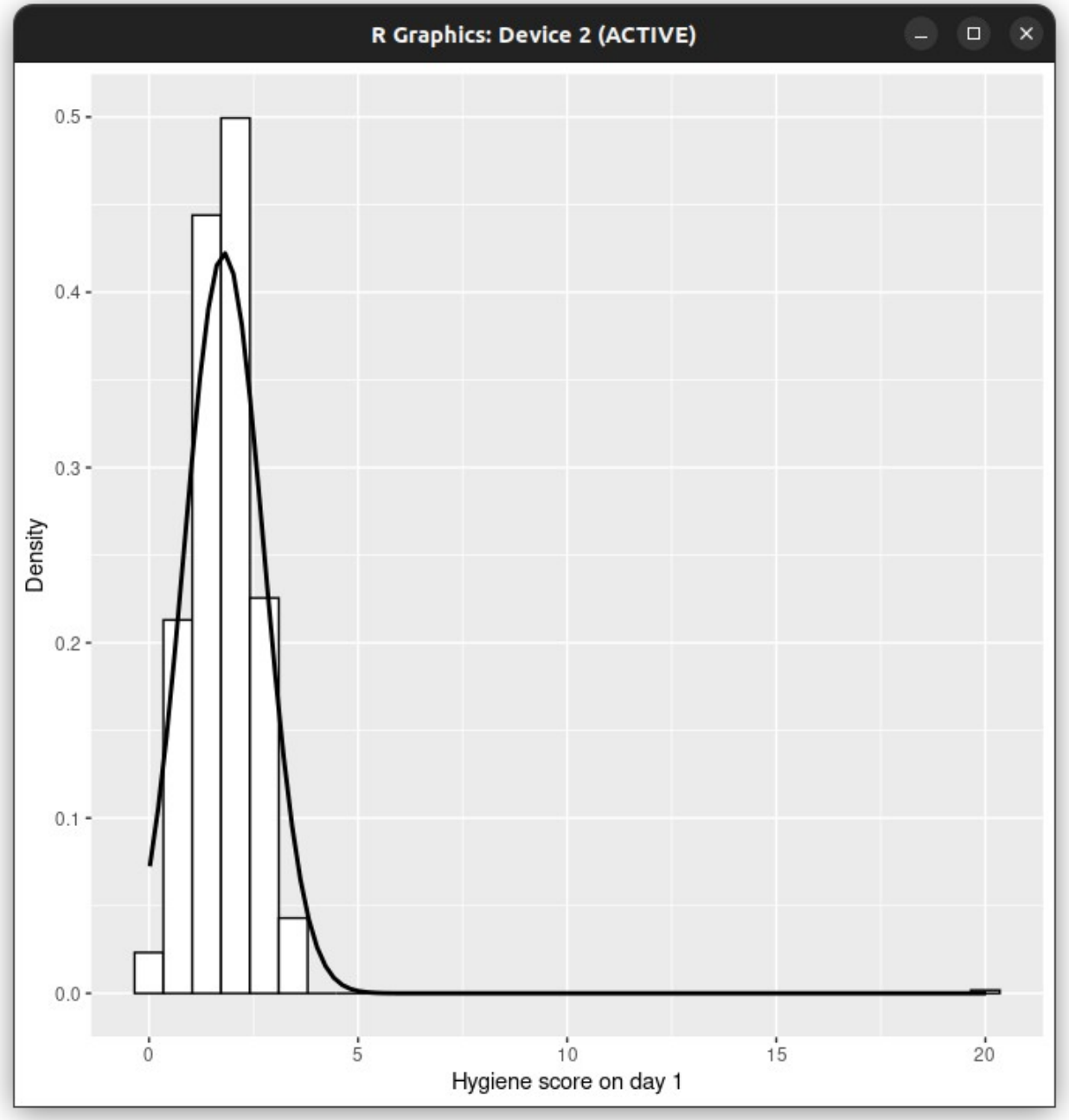
```
# Wyodrebnij dunceData z rexam.
dunceData <- rexam[rexam["uni"] == "Duncetown University", c("exam", "computer", "lectures", "numeracy", "uni")];
# Stworz wykres dla uniwersytetu Duncetown. Bez legendy. Dodaj histogram dla funkcji gestosci. Dodaj linie funkcji normy zalezna od sredniej.
hist.numeracy.duncetown <- ggplot(dunceData, aes(numeracy)) + theme(legend.position = "none") + geom_histogram(aes(y = ..density..),
```

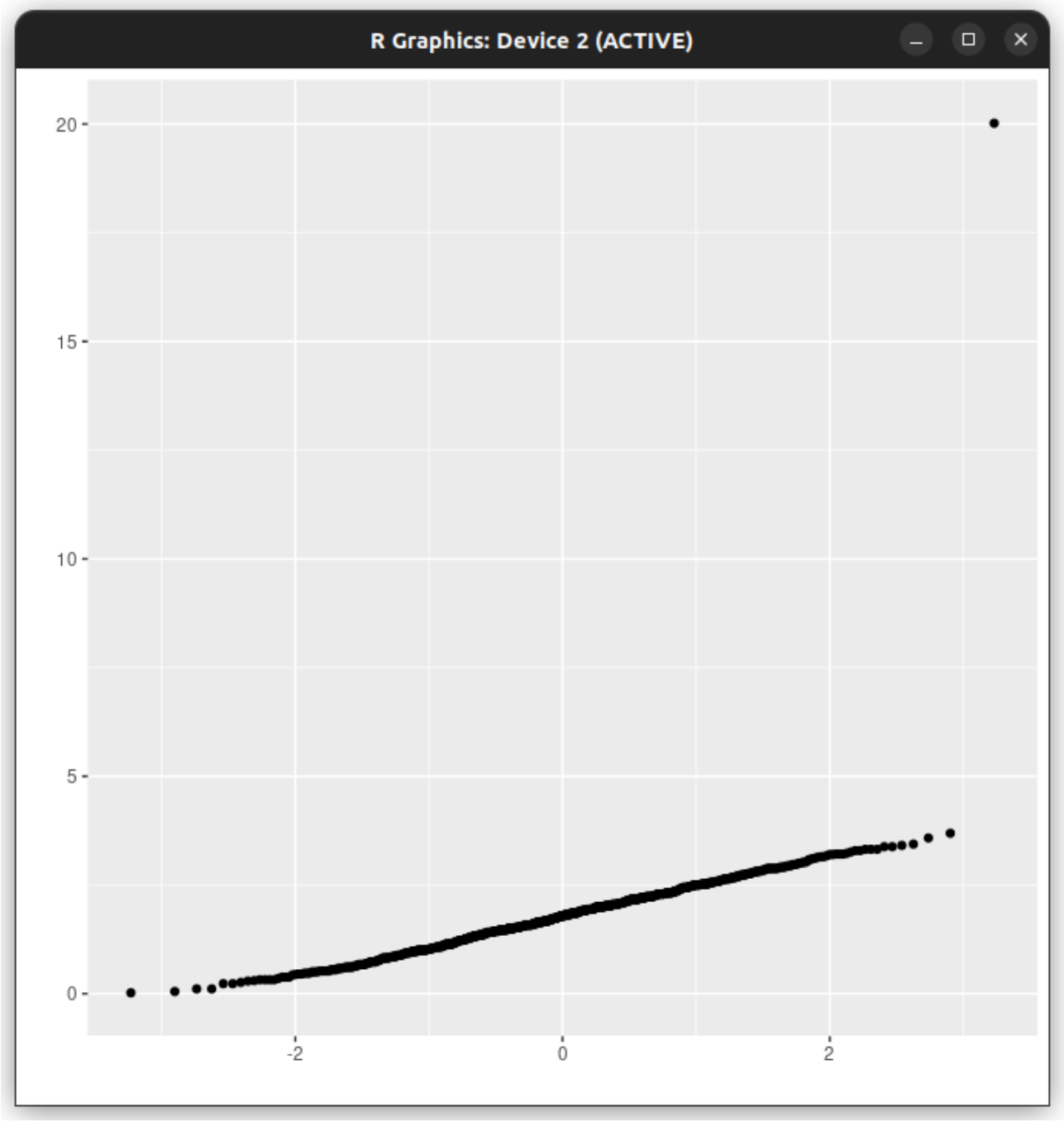
```
fill = "white", color = "black", binwidth = 1) + labs(x = "Numeracy
Score", y = "Density") + stat_function(fun = dnorm, args = list(mean =
mean(dunceData$numeracy, na.rm = T), sd = sd(dunceData$numeracy, na.rm =
T)), color = "blue", size = 1);
hist.numeracy.duncetown;
# Powtorz dla Sussex.
susData <- rexam[rexam["uni"] == "Sussex University", c("exam",
"computer", "lectures", "numeracy", "uni")];
hist.numeracy.sussex <- ggplot(susData, aes(numeracy)) +
theme(legend.position = "none") + geom_histogram(aes(y = ..density..),
fill = "white", color = "black", binwidth = 1) + labs(x = "Numeracy
Score", y = "Density") + stat_function(fun = dnorm, args = list(mean =
mean(susData$numeracy, na.rm = T), sd = sd(susData$numeracy, na.rm = T)),
color = "blue", size = 1);
hist.numeracy.sussex;
```

2 Wyniki

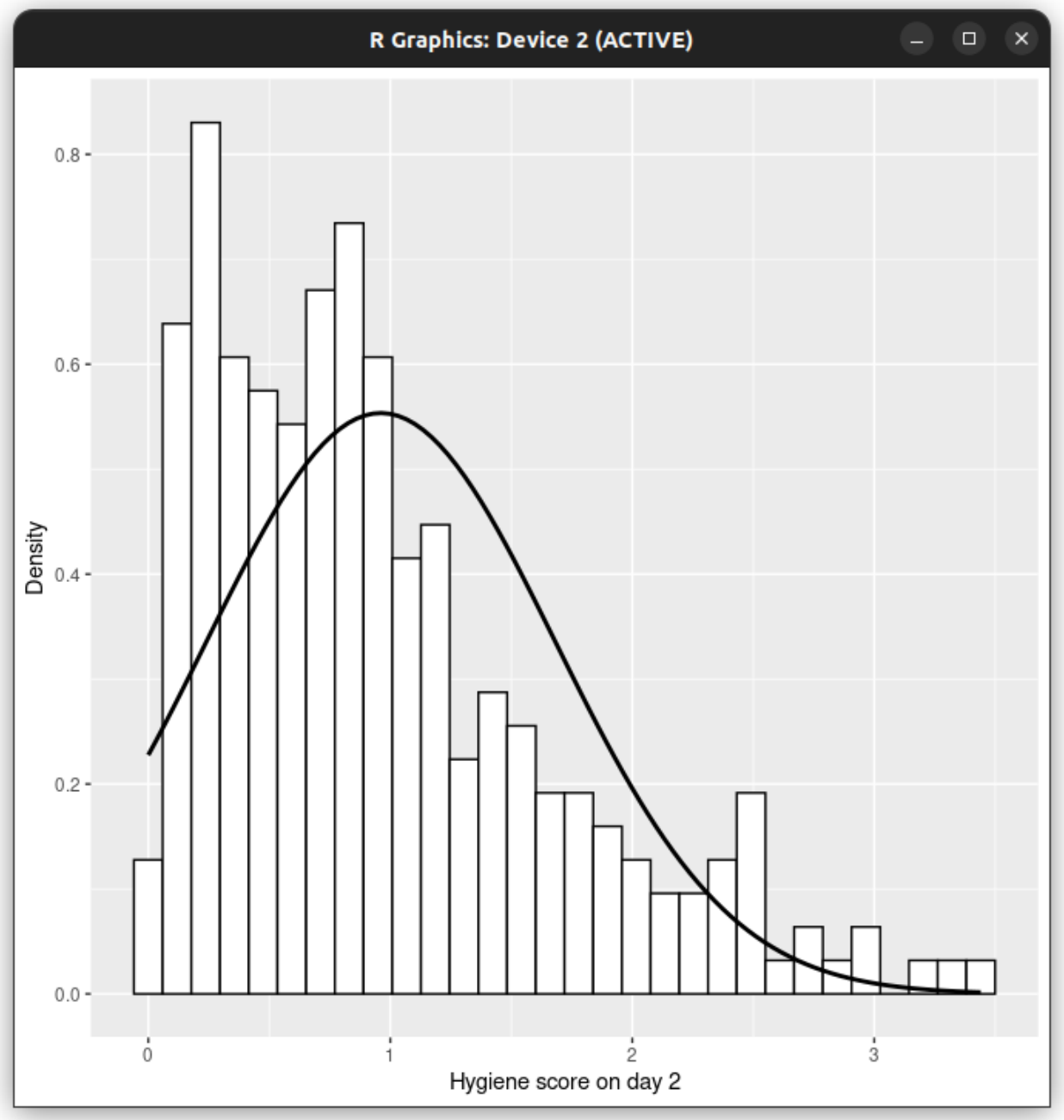
1.

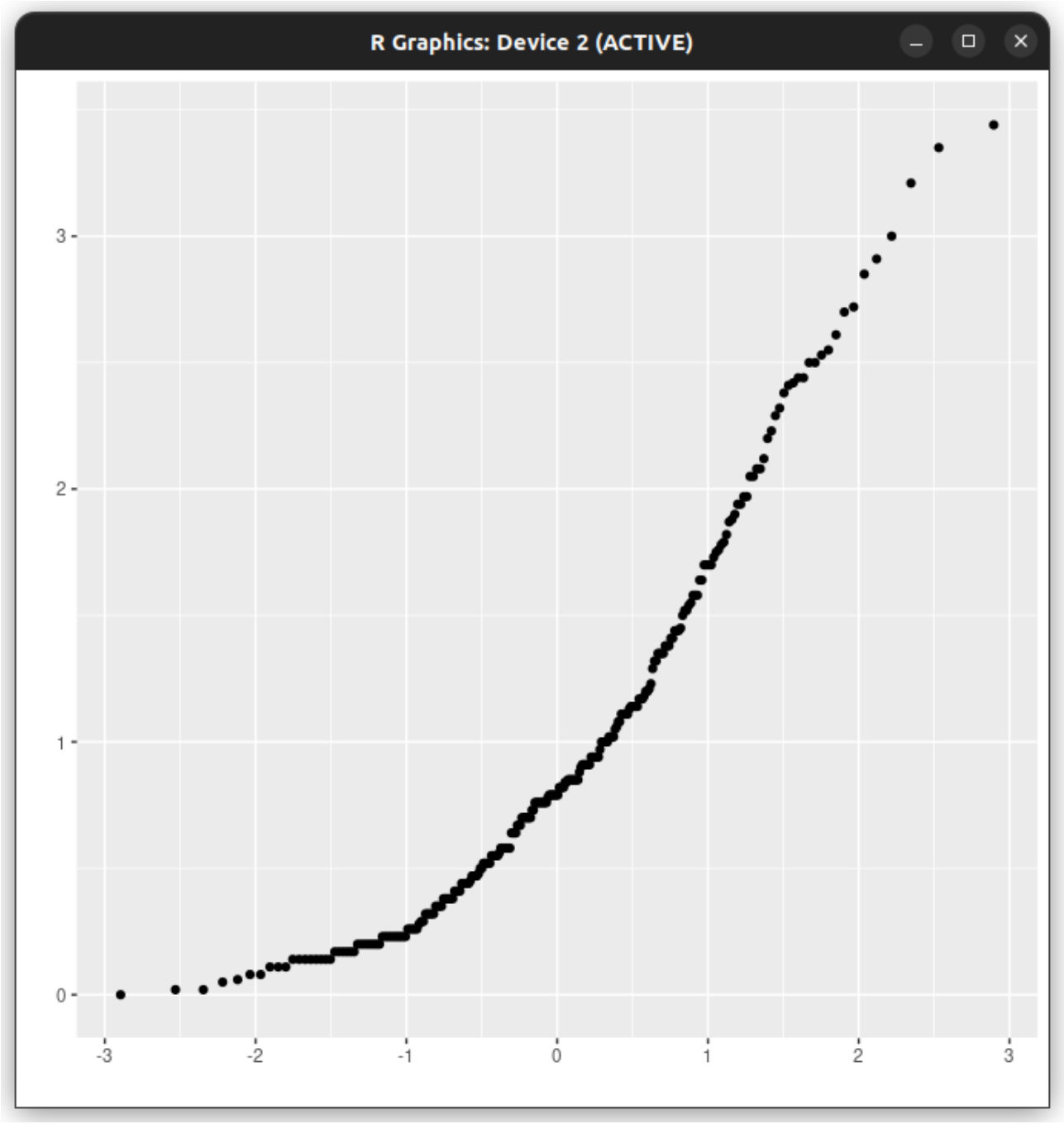
2.

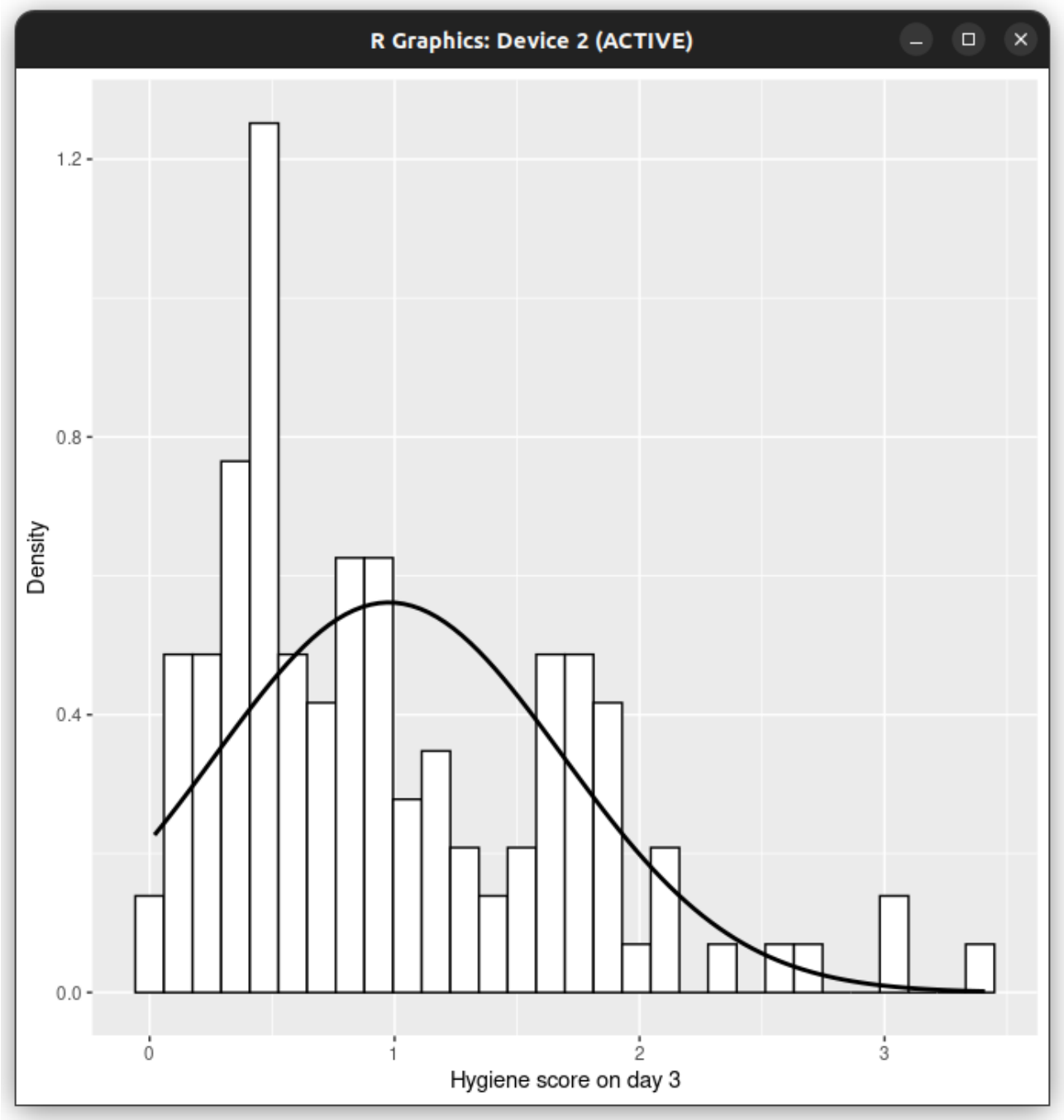

R Graphics: Device 2 (ACTIVE)
0.5
0.4
0.3
0.2
0.1
0.0
Density
0
5
10
15
20
Hygiene score on day 1

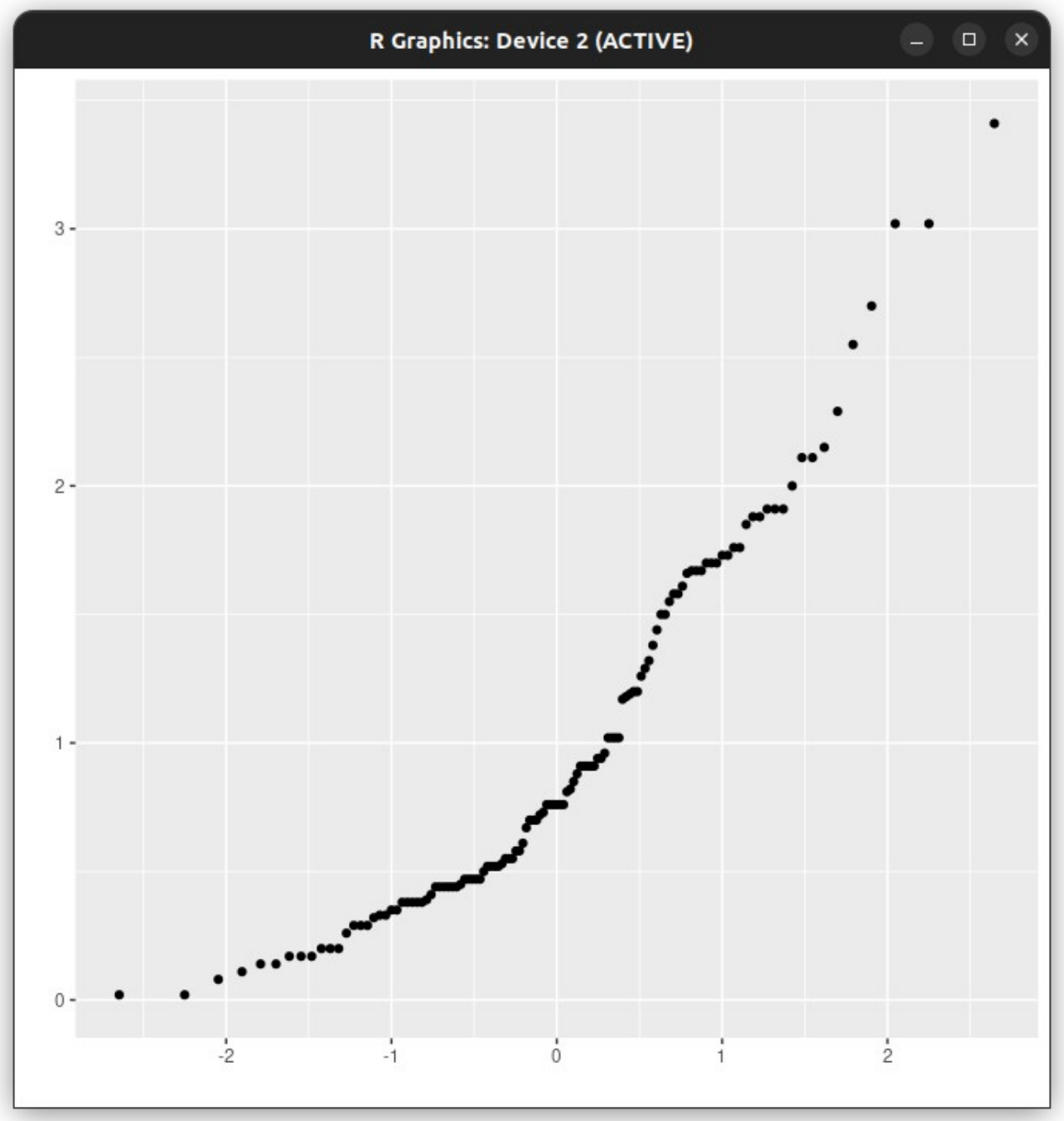

R Graphics: Device 2 (ACTIVE)
20
15
10
5
0
-2
0
2

3.


R Graphics: Device 2 (ACTIVE)
0.8
0.6
0.4
0.2
0.0
Density
0
1
2
3
Hygiene score on day 2


R Graphics: Device 2 (ACTIVE)
3
2
1
0
-3
-2
-1
0
1
2
3

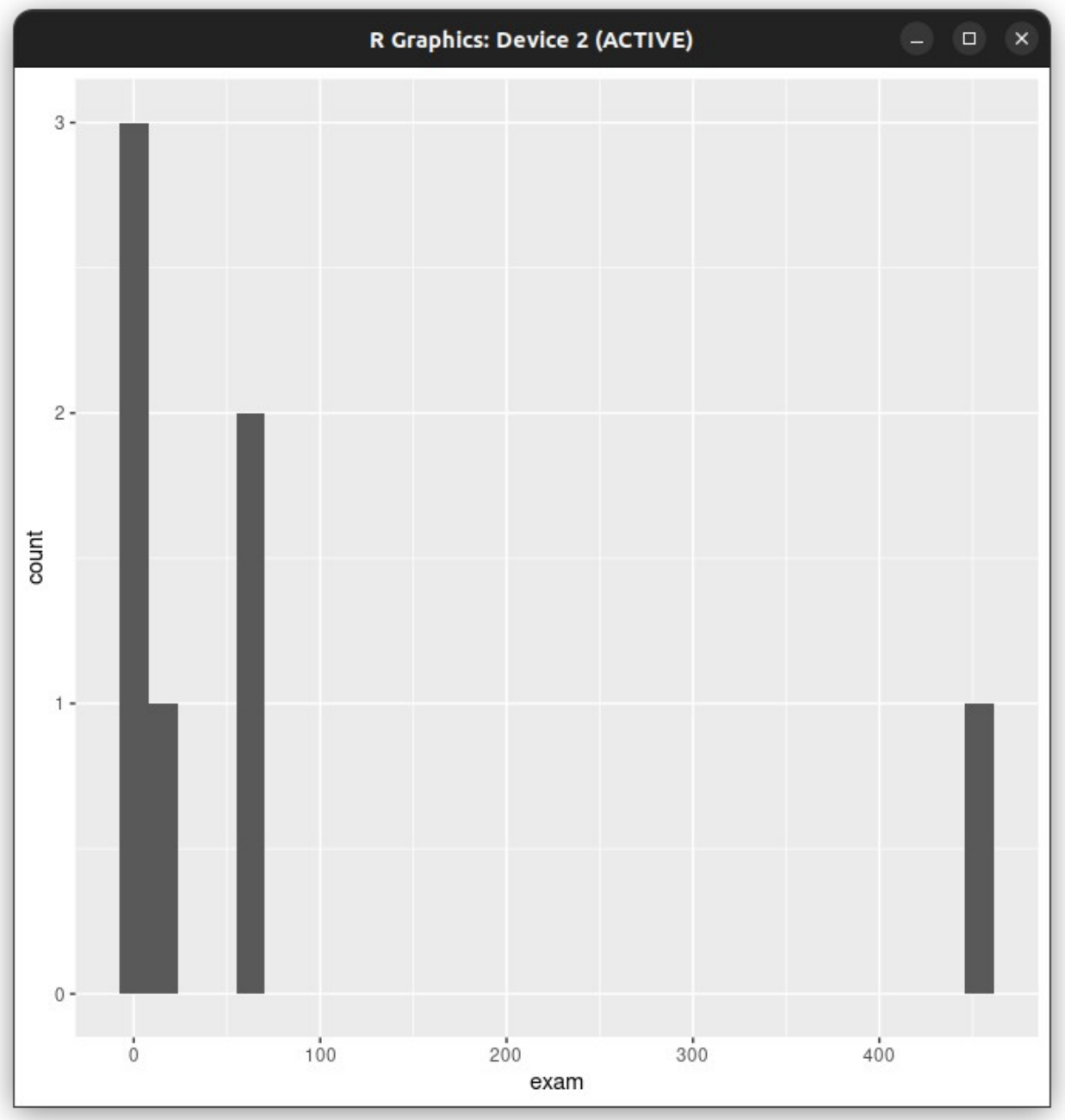

R Graphics: Device 2 (ACTIVE)
1.2
0.8
0.4
0.0
Density
0
1
2
3
Hygiene score on day 3

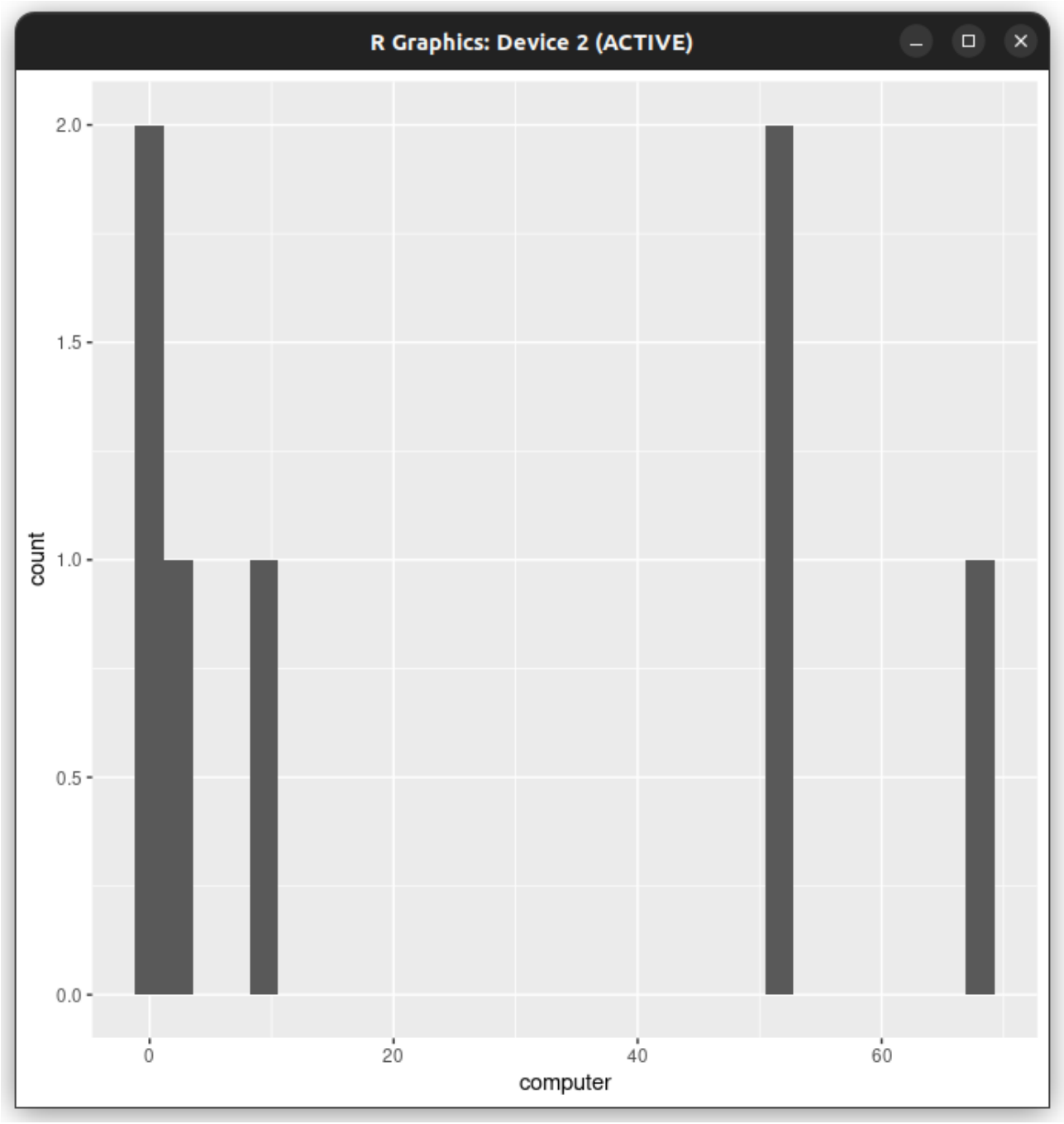


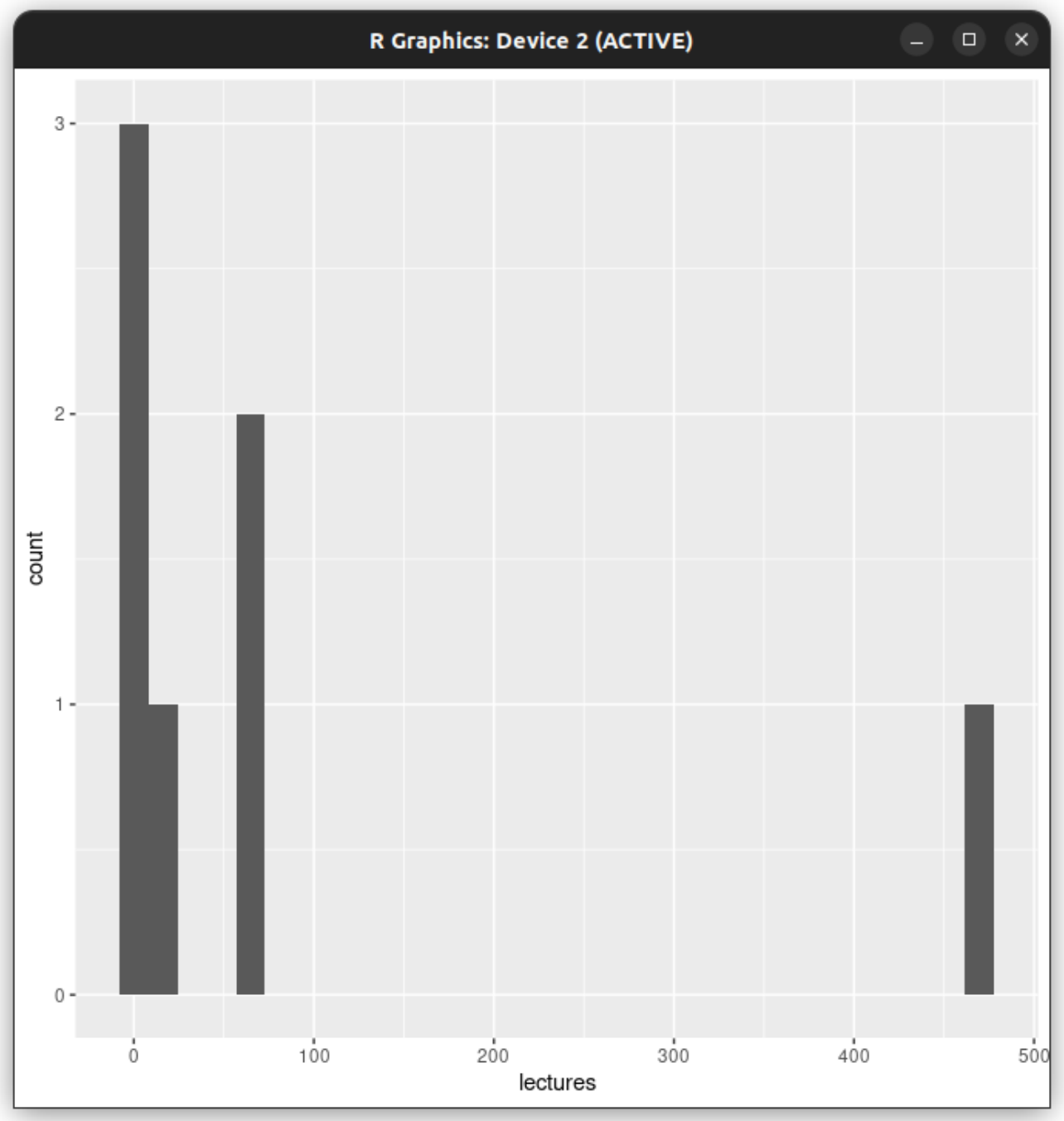
4.

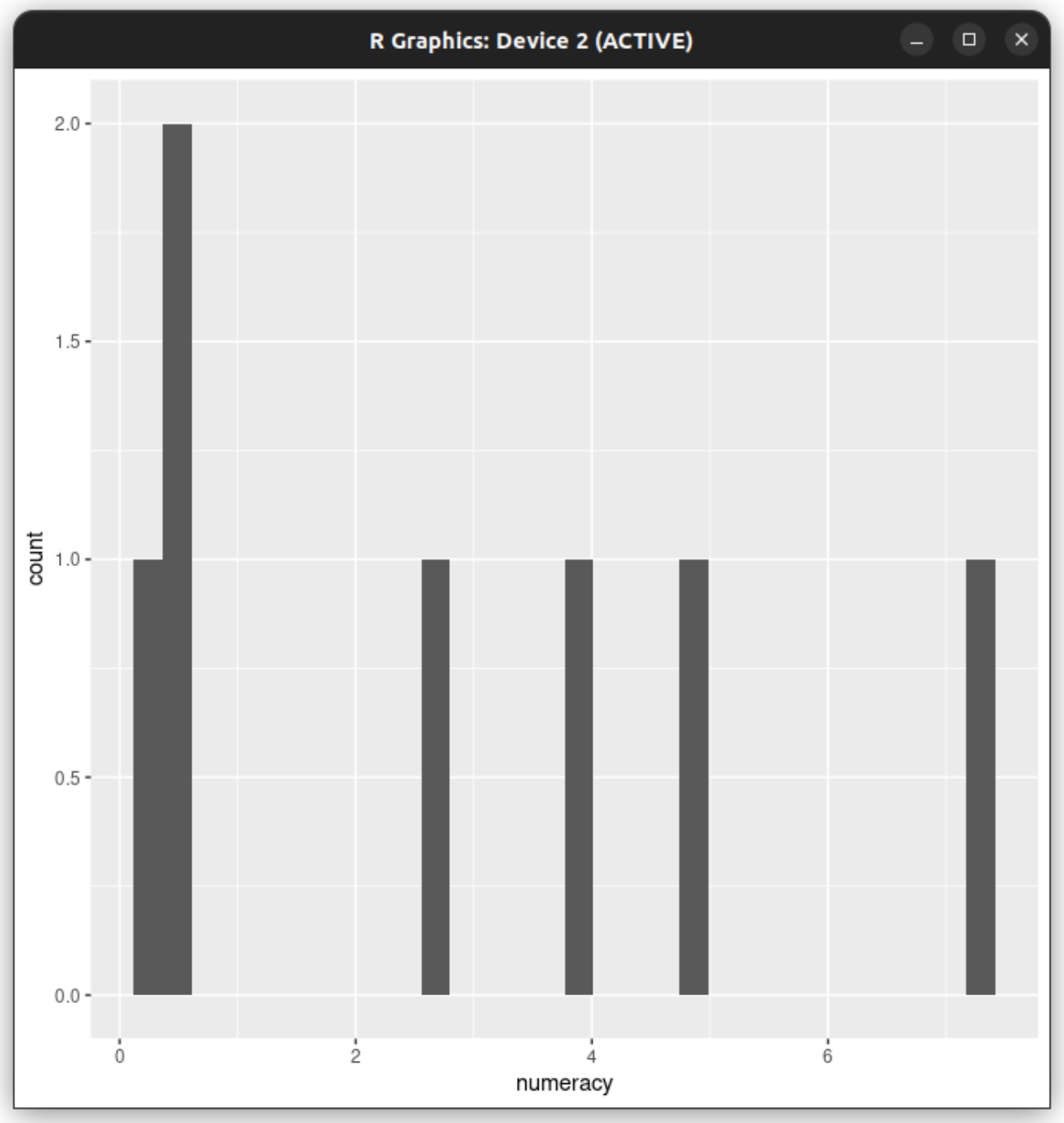
```
> head(rexam);
  exam computer lectures numeracy                 uni
1   18       54     75.0        7 Duncetown University
2   30       47      8.5        1 Duncetown University
3   40       58     69.5        6 Duncetown University
4   30       37     67.0        6 Duncetown University
5   40       53     44.5        2 Duncetown University
6   15       48     76.5        8 Duncetown University
>
```

5.




R Graphics: Device 2 (ACTIVE)
2.0
1.5
1.0
0.5
0.0
count
0
20
40
60
computer

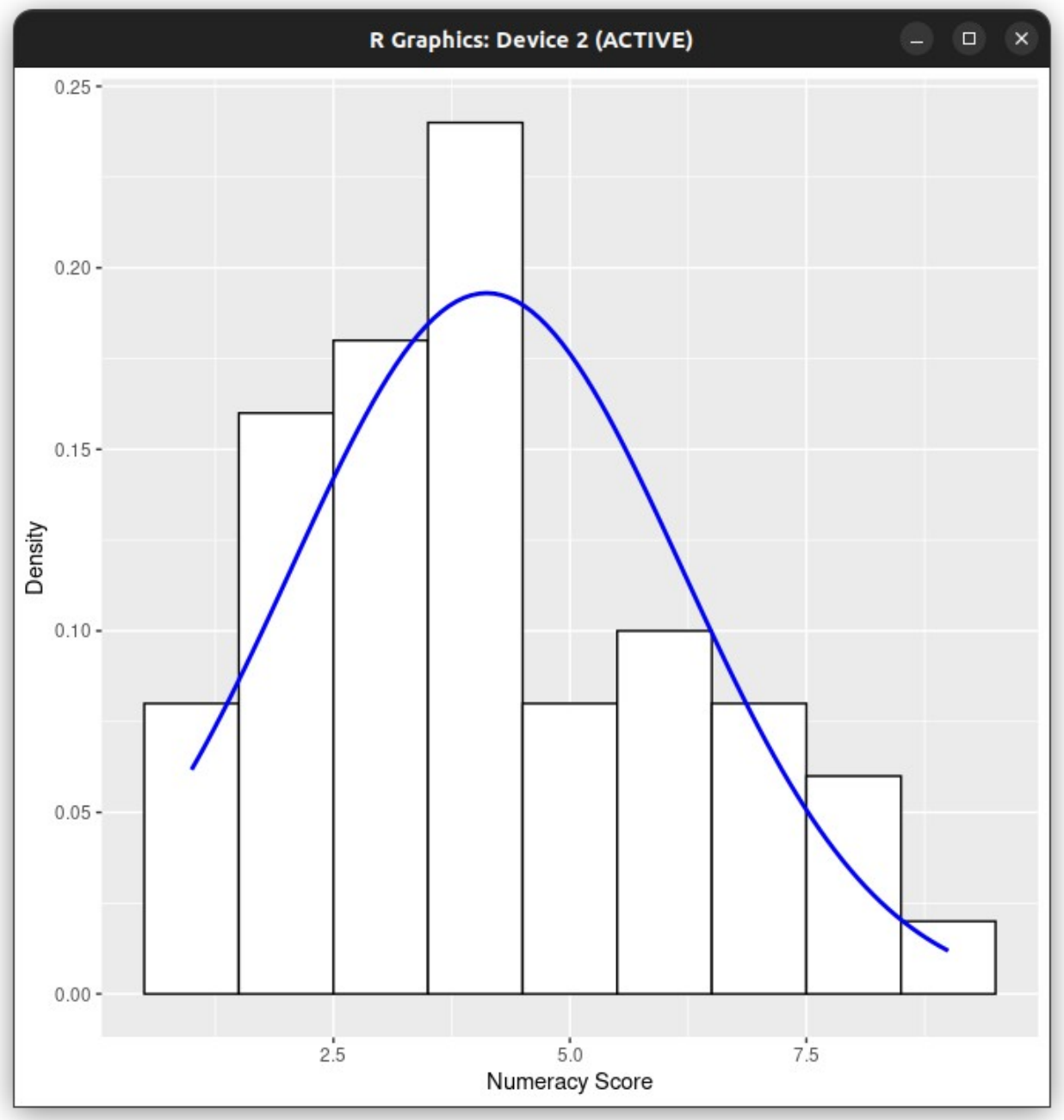

R Graphics: Device 2 (ACTIVE)
count
3
2
1
0
0
100
200
300
400
500
lectures

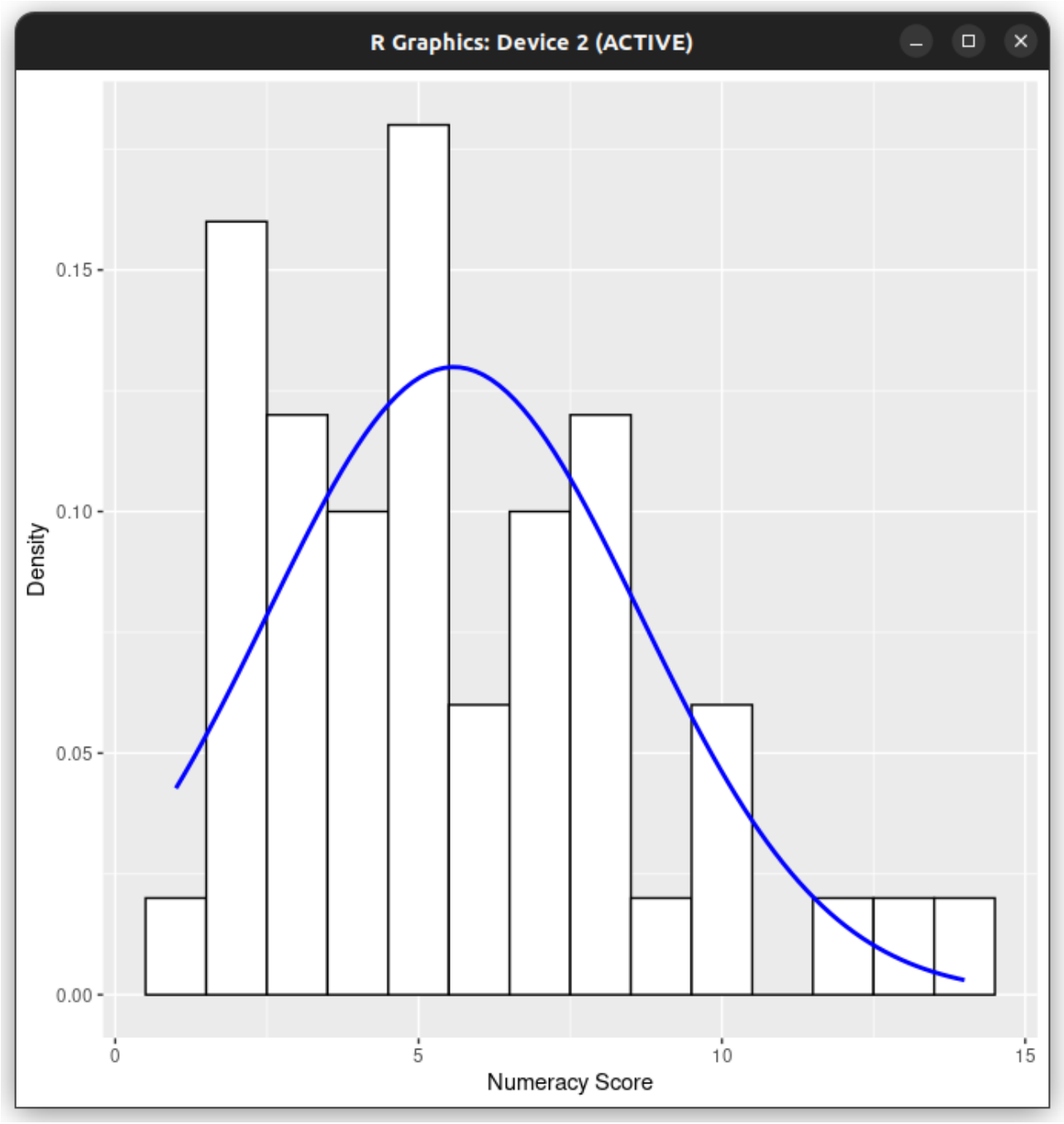


6.

```
> # Przeksztalc dane dla wartosci kolumn uni, pokaz dla nich wlasnosci statystyczne.
by(cbind(data = rexam$exam, data = rexam$numeracy), rexam$uni, describe);
INDICES: Duncetown University
  vars  n  mean    sd median trimmed   mad min max range skew kurtosis   se
1    1 50 40.18 12.59     38   39.85 12.60  15  66    51 0.29    -0.72 1.78
2    2 50  4.12  2.07      4    4.00  2.22   1   9     8 0.48    -0.65 0.29
------------------------------------------------------------
INDICES: Sussex University
  vars  n  mean    sd median trimmed  mad min max range skew kurtosis   se
1    1 50 76.02 10.21     75   75.70 8.90  56  99    43 0.26    -0.46 1.44
2    2 50  5.58  3.07      5    5.28 2.97   1  14    13 0.75    -0.01 0.43
> 
```

7.


R Graphics: Device 2 (ACTIVE)
0.25
0.20
0.15
0.10
0.05
0.00
Density
2.5
5.0
7.5
Numeracy Score


R Graphics: Device 2 (ACTIVE)
Density
0.15
0.10
0.05
0.00
0
5
10
15
Numeracy Score