

title: 'תרגיל 1 - גיל שילה - 315440230, גילאי דעי - 205477599, עומר קשת - 314761321'

output:

html_document: default pdf_document: default —

```
membership_2015 <- read.csv(file = 'C:\\Users\\gilay\\Desktop\\study\\business\\third year\\stochastic\\ex_1\\Membership2015.csv')
membership_2016 <- read.csv(file = 'C:\\Users\\gilay\\Desktop\\study\\business\\third year\\stochastic\\ex_1\\Membership2016.csv')
membership_2017 <- read.csv(file = 'C:\\Users\\gilay\\Desktop\\study\\business\\third year\\stochastic\\ex_1\\Membership2017.csv')
```

A

```
membership_2017$today <- 2017
membership_2017$diff <- membership_2017$today - membership_2017$MemberSince
```

```
membership_2016$today <- 2016
membership_2016$diff <- membership_2016$today - membership_2016$MemberSince
mean_2016 <- mean(membership_2016$diff, na.rm=TRUE)
```

```
no_active_2016 <- membership_2016 %>% filter(Active == 0)
suspect_16 <- no_active_2016 %>% group_by(diff) %>% summarise(sum = length(diff))

x<-no_active_2016$ID
for(i in 1:length(x))
{membership_2017 <- filter(membership_2017,membership_2017$ID != x[i])}
mean_2017 <- mean(membership_2017$diff, na.rm=TRUE)
diff_17_16 <- (mean_2017 - mean_2016)

if(diff_17_16>0)
{print(paste0('increased by: ', diff_17_16))}else{print('not increased!')}
```

```
## [1] "increased by: 0.481370096464488"
```

ניתן לראות שבממוצע משך ההישארות של הלקוחות המנויים בכל שנה גדל בכ-0.5 בשנת 2017 לעומת שנת 2016.

B + C

```
no_active_2017 <- membership_2017 %>% filter(Active == 0)
x<-no_active_2017$ID
for(i in 1:length(x))
{no_active_2017 <- filter(no_active_2017,no_active_2017$ID != x[i])}

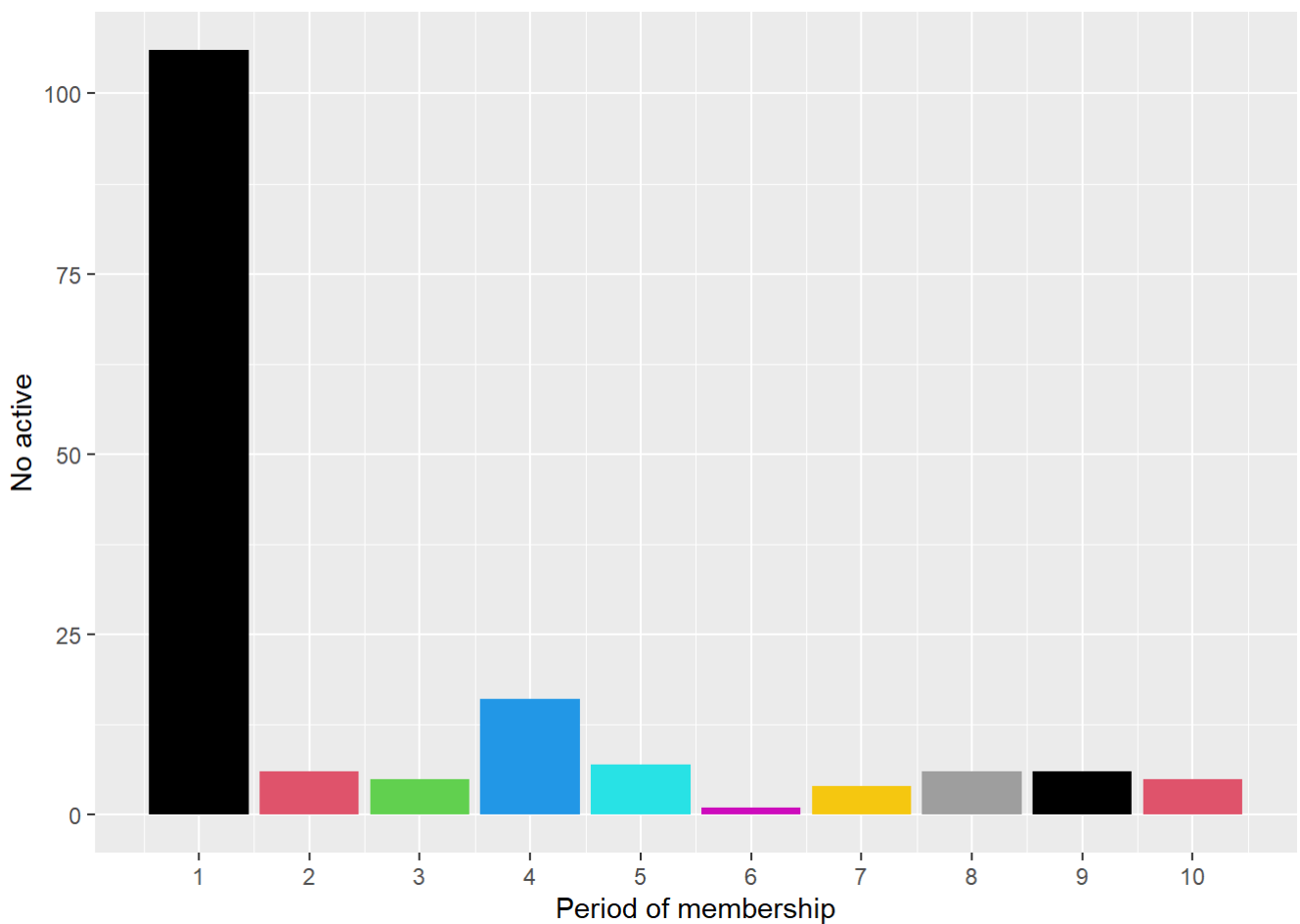
suspect_17 <- no_active_2017 %>% group_by(diff) %>% summarise(sum = length(diff))
```

```
plot_16 <- ggplot(suspect_16,aes(x=diff,y=sum)) +
  scale_x_continuous(breaks = seq(0, 10, by = 1)) + geom_col(fill = suspect_16$diff) + xlab(
'Period of membership') + ylab('No active')
```

```
print(suspect_16)
```

```
## # A tibble: 10 x 2
##   diff  sum
##   <dbl> <int>
## 1     1  106
## 2     2    6
## 3     3    5
## 4     4   16
## 5     5    7
## 6     6    1
## 7     7    4
## 8     8    6
## 9     9    6
## 10    10    5
```

```
plot_16
```



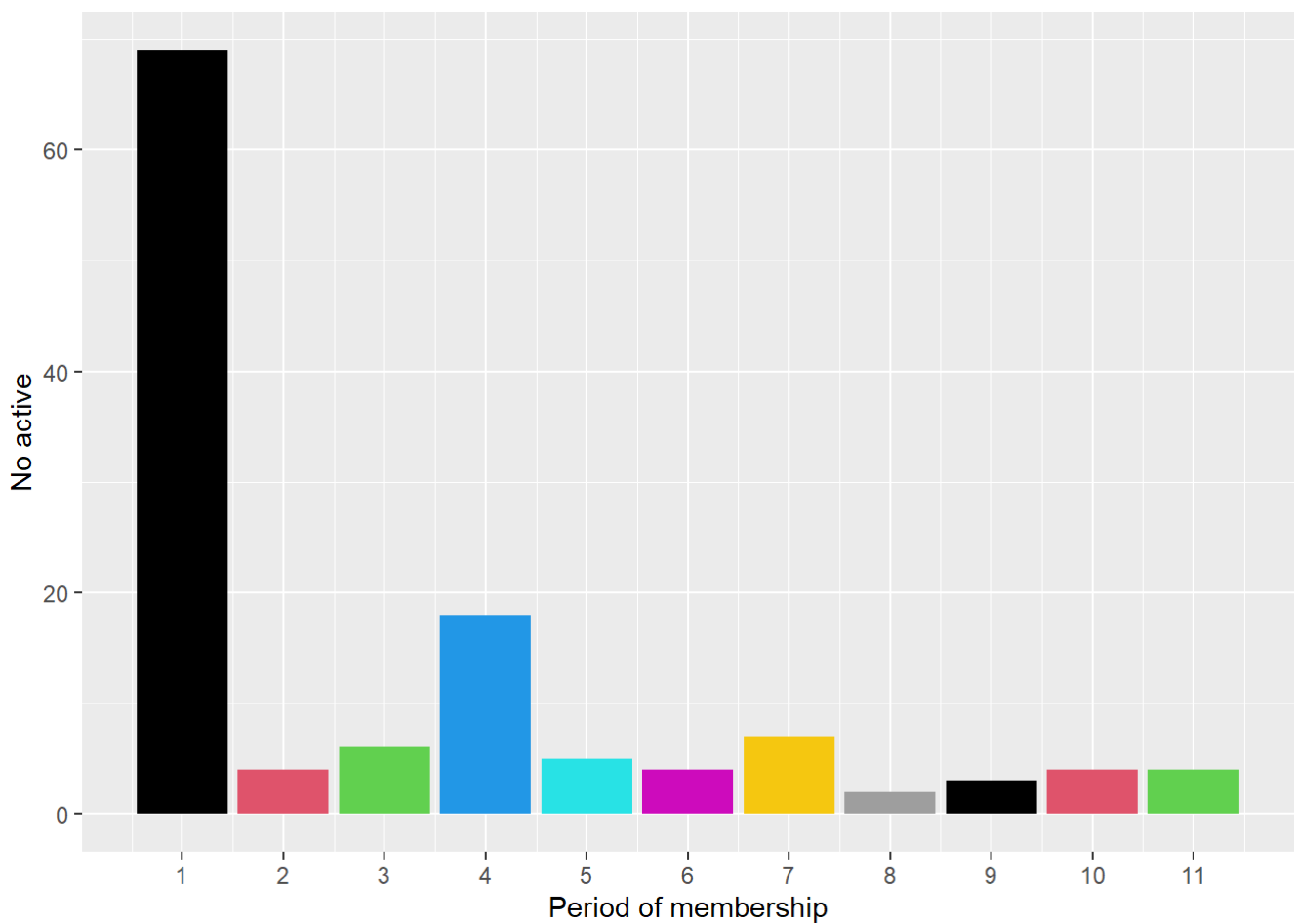
גרף זה מציג את כמות הלקוחות שעוזבים את החדר כושר לפי הותק, נכון לשנת 2016. ניתן לראות שבשנה הראשונה עוזבת כמות גדולה מאוד של לקוחות ביחס לשנים האחרות.

```
plot_17 <- ggplot(suspect_17,aes(x=diff,y=sum)) +
  scale_x_continuous(breaks = seq(0, 11, by = 1)) + geom_col(fill = suspect_17$diff) + xlab(
'Period of membership') + ylab('No active')
```

```
print(suspect_17)
```

```
## # A tibble: 11 x 2
##   diff    sum
##   <dbl> <int>
## 1     1    69
## 2     2     4
## 3     3     6
## 4     4    18
## 5     5     5
## 6     6     4
## 7     7     7
## 8     8     2
## 9     9     3
## 10    10     4
## 11    11     4
```

```
plot_17
```



גם בשנת 2017 ניתן לראות שהתצפית החריגה היא בשנה הראשונה, שבה עוזבים הכי הרבה לקוחות.

C

מהגרפים של הסעיפים הקודמים אנו למדים שצריך להתמקד במאמצים של השארת הלקוחות בשנה הראשונה שלהם. ניתן לראות שבשנים הראשונות יש עזיבה של הלקוחות החדשים ולאחר מכן התייצבות יחסית של כמות העוזבים בכל שנה.