Bar Chart

Bar Chart

A simple tool to visualize the relative or absolute frequencies of observed values of a variable is a bar chart. A bar chart can be used for nominal and ordinal variables, as long as the number of categories is not very large. It consists of one bar for each category. The height of each bar is determined by either the absolute frequency or the relative frequency of the respective category and is shown on the y-axis.

If the variable is measured on an ordinal level, then it is recommended to arrange the bars on the x-axis according to their ranks or values.

Bar Chart - Example

Suppose there are ten people in a supermarket queue. Each of them is either coded as "F" (if the person is female) or "M" (if the person is male). The collected data may look like

$$M, F, M, F, M, M, M, F, M, M.$$

There are now two categories in the data: male (M) and female (F).

We use **a1** to refer to the male category and **a2** to refer to the female category. Since there are seven male and three female students, we have 7 values in category a1, denoted as n1 = 7, and 3 values in category a2, denoted as n2 = 3. The number of observations in a particular category is called the **absolute frequency**. It follows that n1 = 7 and n2 = 3 are the absolute frequencies of a1 and a2, respectively. Note that n1 + n2 = n = 10, which is the same as the total number of collected observations. We can also calculate the **relative frequencies** of a1 and a2 as

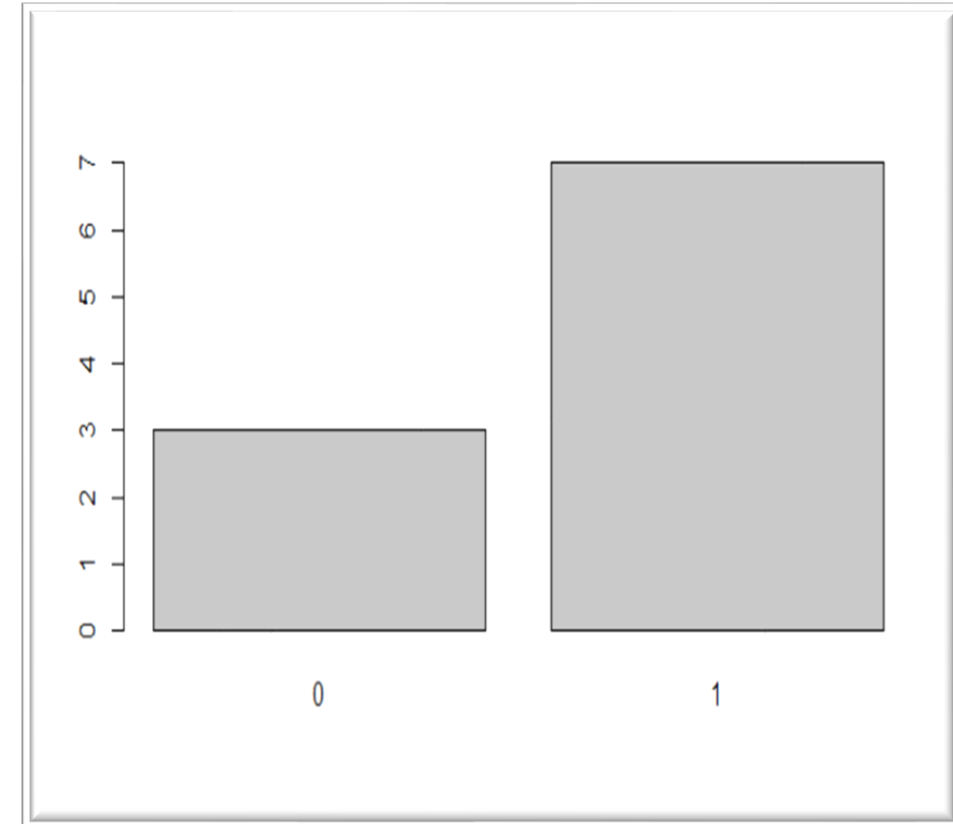$$f_1 = f(a_1) = \frac{n_1}{n} = \frac{7}{10} = 0.7 = 70\,\%$$

and

$$f_2 = f(a_2) = \frac{\overline{n_2}}{n} = \frac{3}{10} = 0.3 = 30\,\%$$

respectively. This gives us information about the proportions of male and female customers in the queue.

R Code

Data: $\mathbf{M, F, M, F, M, M, M, F, M, M.}$

Use 0, 1 to represent F and M respectively

```
> queue <- c(1, 0, 1, 0, 1, 1, 1, 0, 1, 1)
> queue
 [1] 1 0 1 0 1 1 1 0 1 1
> table(queue)
queue
0 1
3 7
> barplot(table(queue))
```

Bar diagrams

Visualizes the relative or absolute frequencies of observed values of a variable.

It consists of one bar for each category.

The height of each bar is determined by either the absolute frequency or the relative frequency of the respective category and is shown on the y-axis.

Width of the bar is immaterial or arbitrary.

Bar Diagram

Frequency distribution

Class	Frequency	Relative frequency
A_1	f_1	f_1/n
A_2	f_2	f_2/n
...	...	...
A_{k-1}	f_{k-1}	f_{k-1}/n
A_k	f_k	f_k/n

Bar Diagrams

```
> barplot(table(x))    #Bar plot with absolute frequency


> barplot(table(x)/length(x))    #Bar plot with relative
                                         frequency
```

Bar diagrams

barplot **Creates a bar plot with vertical or horizontal bars.**

Usage

```
barplot(height, ...)
```

```
barplot(height, width = 1, space = NULL,
names.arg = NULL, legend.text = NULL, beside
= FALSE, horiz = FALSE, density = NULL, angle
= 45, col = NULL, border = par("fg"), main =
NULL, sub = NULL, xlab = NULL, ylab = NULL,
xlim = NULL, ylim = NULL, xpd = TRUE, log =
"", axes = TRUE, axisnames = TRUE, cex.axis =
par("cex.axis"), cex.names = par("cex.axis"),
inside = TRUE, plot = TRUE, axis.lty = 0,
offset = 0, add = FALSE, args.legend = NULL,
...)
```

Bar plot

```
> help("barplot")
```

height   either a vector or matrix of values describing the bars which make up the plot. If height is a vector, the plot consists of a sequence of rectangular bars with heights given by the values in the vector.

width   optional vector of bar widths. Re-cycled to length the number of bars drawn. Specifying a single value will have no visible effect unless xlim is specified.

space   the amount of space (as a fraction of the average bar width) left before each bar. May be given as a single number or one number per bar.

names.arg   a vector of names to be plotted below each bar or group of bars. If this argument is omitted, then the names are taken from the names attribute of height if this is a vector, or the column names if it is a matrix.

legend.tex   a vector of text used to construct a legend for the plot, or a logical indicating whether a legend should be included.

…   …

…   …

Bar Diagram

```
barplot(x, width = 1, space = NULL,…)
```

```
> barplot(table(x))
```
#Bar plot with absolute frequency

```
> barplot(table(x)/length(x))
```
#Bar plot with relative frequency
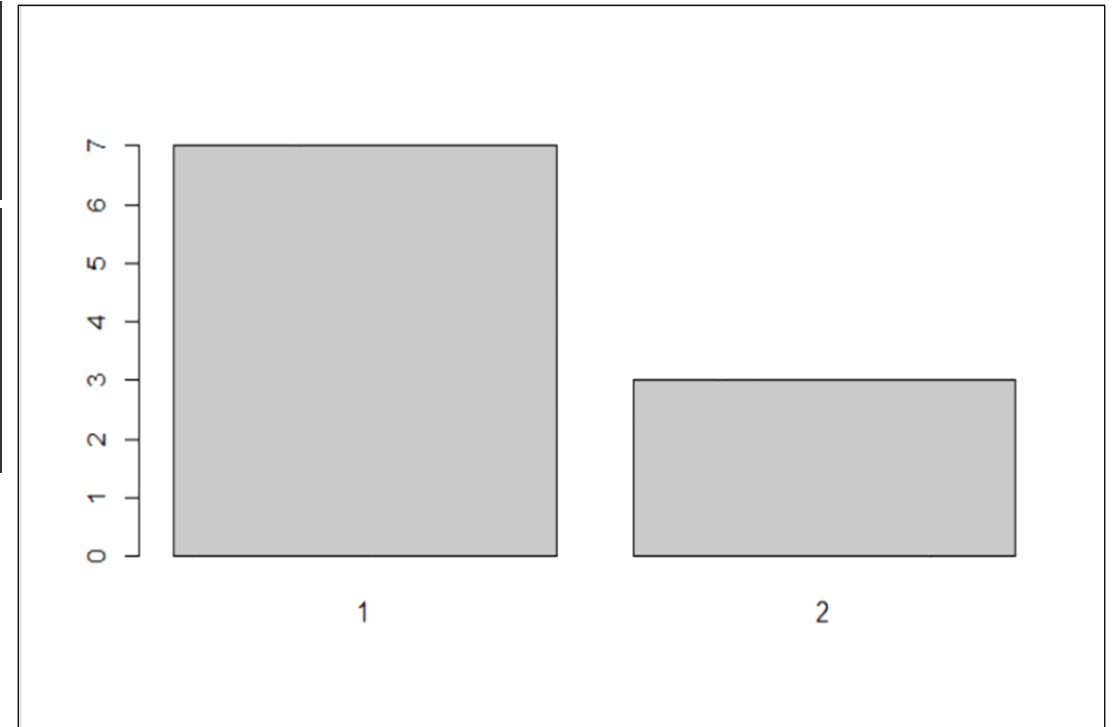
Bar Diagram – Example 1

Code of qualification of 10 persons by using, say 1 for graduate (G) and 2 for nongraduate (N).

G, N, G, N, G, G, G, N, G, G

1, 2, 1, 2, 1, 1, 1, 2, 1, 1

```
> qualif <- c(1, 2, 1, 2, 1, 1, 1, 2, 1, 1)
> qualif
 [1] 1 2 1 2 1 1 1 2 1 1

> table(qualif)
qualif
1 2
7 3

> barplot(table(qualif))
```

Bar Diagram – Example 2

There are three salespersons in a shop. They are denoted as 1, 2 and 3. Which salesperson serves the first 100 customers is recorded as follows:

1,1,2,1,2,3,2,2,3,3,3,1,2,3,2,2,3,1,1,3,3,1,2,1,3,3,3,2,2,2,2,1,2,2,1,1,

1,3,2,2,1,2,3,2,2,1,2,3,3,2,1,2,2,3,1,1,2,1,2,3,2,3,2,2,3,1,2,3,3,3,2,1,

1,1,2,1,1,2,1,2,3,3,1,2,3,3,2,1,2,3,2,1,3,2,2,2,2,3,2,2

```
> SalesP <- c(c(1,1,2,1,2,3,2,2,3,3,3,1,2,3,2,2,3,1,1,3,3,1,2,1,3,3,3,2,2,2,
2,1,2,2,1,1,1,3,2,2,1,2,3,2,2,1,2,3,3,2,1,2,2,3,1,1,2,1,2,3,2,3,2,2,3,1,2,3,
3,3,2,1,1,1,2,1,1,2,1,2,3,3,1,2,3,3,2,1,2,3,2,1,3,2,2,2,2,3,2,2))
> SalesP
 [1] 1 1 2 1 2 3 2 2 3 3 3 1 2 3 2 2 3 1 1 3 3 1 2 1 3 3 3 2 2 2 2 1 2 2 1
[36] 1 1 3 2 2 1 2 3 2 2 1 2 3 3 2 1 2 2 3 1 1 2 1 2 3 2 3 2 2 3 1 2 3 3 3
[71] 2 1 1 1 2 1 1 2 1 2 3 3 1 2 3 3 2 1 2 3 2 1 3 2 2 2 2 3 2 2
```
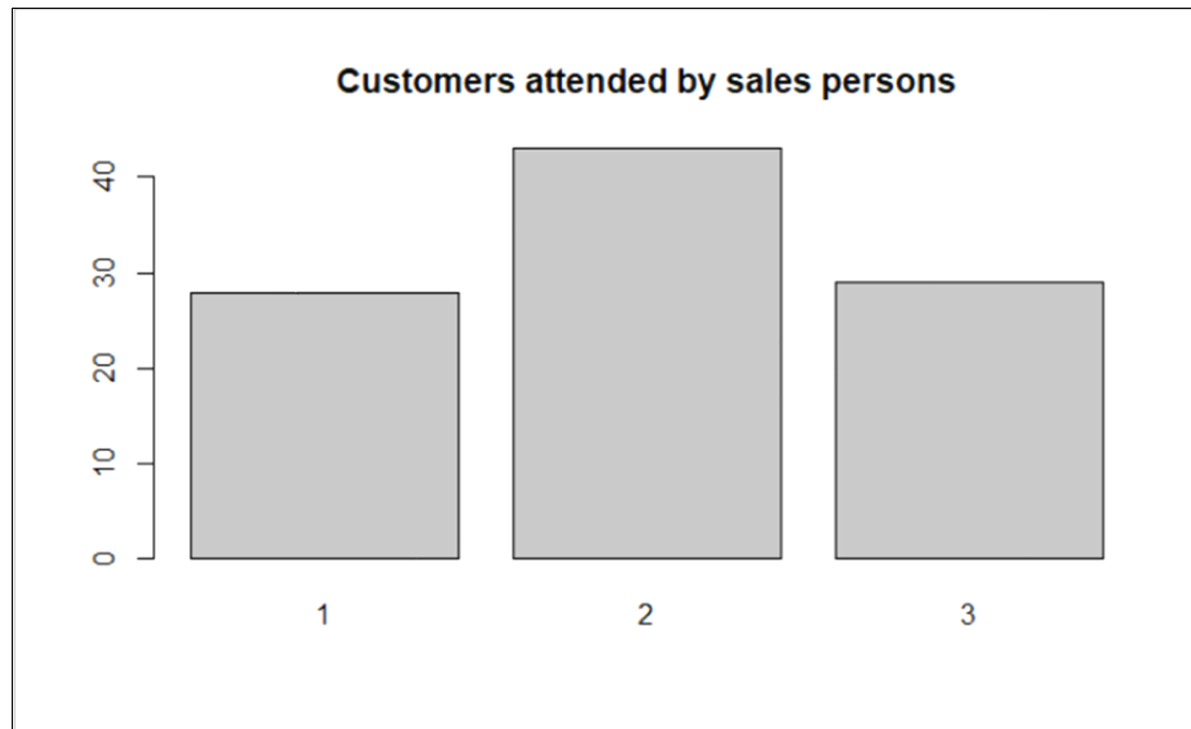
Example 2 – contd.

```
> barplot(table(SalesP))      # barplot of absolute frequencies
> barplot(table(SalesP)/length(SalesP)) # barplot of relative frequencies
```
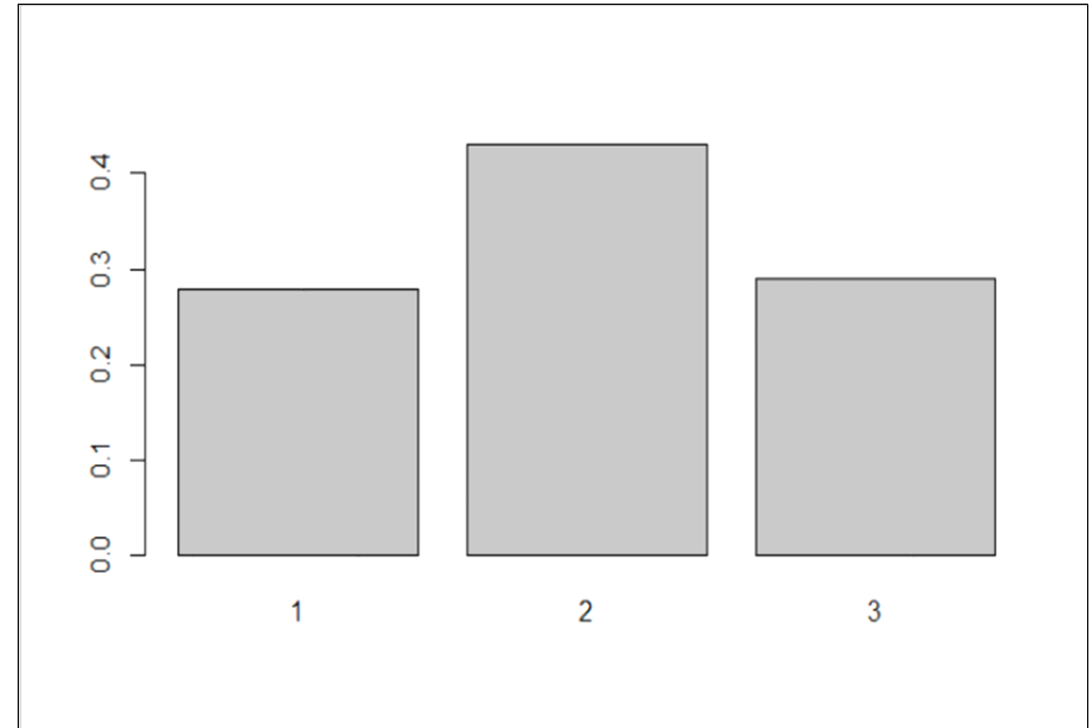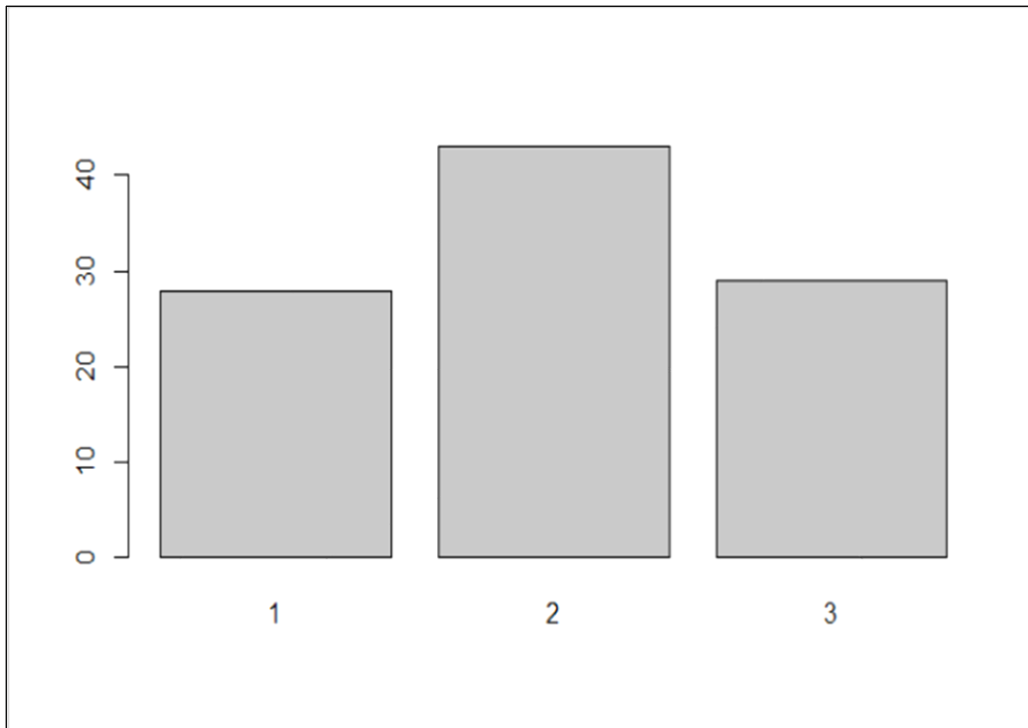
Example 2 – contd.
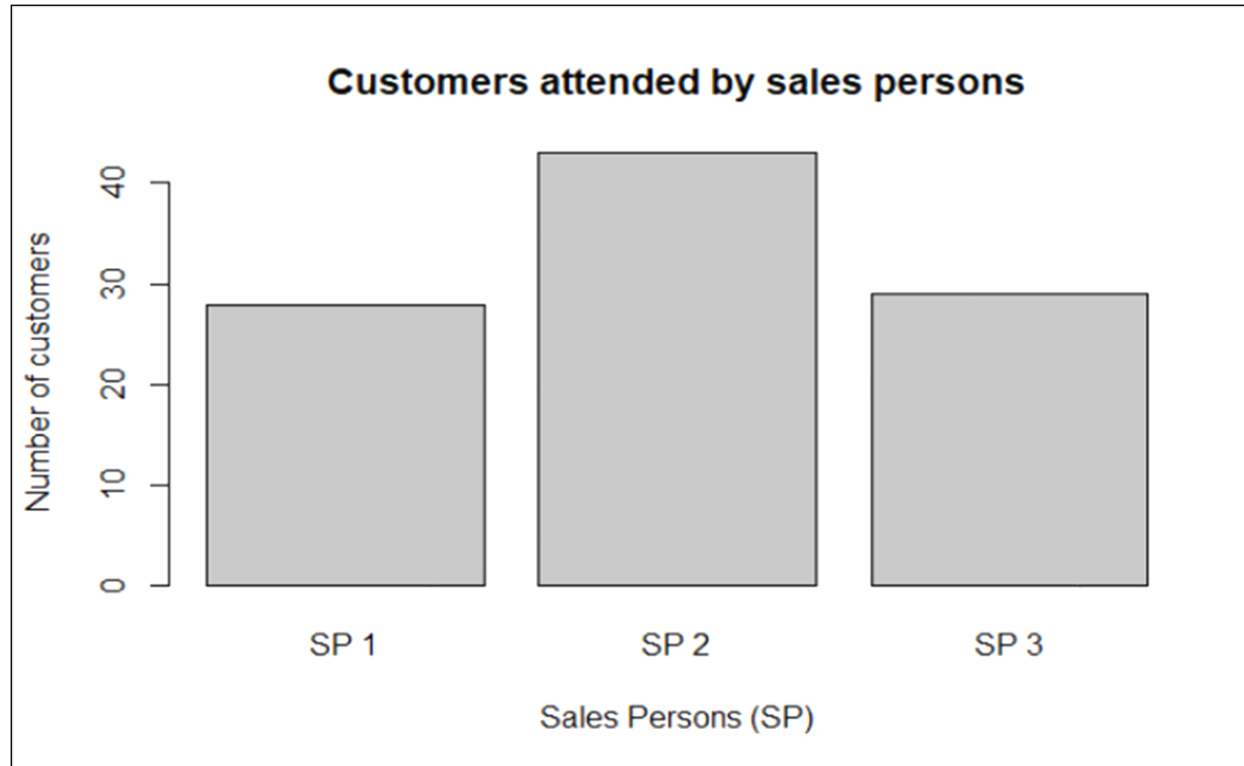
If we want to give a title to the grap

```
> barplot(table(SalesP), main = "Customers attended by sales persons")
```


Customers attended by sales persons

Example 2 – contd.

If we want to further add legends and axis titles to the graph

```
> barplot(table(SalesP), main = "Customers attended by sales persons", names.arg=
c("SP 1", "SP 2", "SP 3"), xlab = "Sales Persons (SP)", ylab = "Number of custome
rs")
```

Example 2 – contd.

If we want to further add colours in the bars to the graph

```
> barplot(table(SalesP), main = "Customers attended by sales persons", names.arg
 = c("SP 1", "SP 2", "SP 3"), xlab = "Sales Persons (SP)", ylab = "Number of cust
omers", col= c("red","blue","orange"))
```