Class 1a

Exercise 1

a) Make sure you have all files for this problem saved to the directory where you saved this Rmd file. This means there should be a data set file "ice cream sales.csv". Run the command `list.files()` to list all files in your working directory:

b) Now load the file "ice cream sales.csv" into a data frame with name `dat` by using the function `read.csv`.

```
# enter your code here ...
```

Check your problem set after each chunk.

c) Show the column names of `dat`.

```
# enter your code here ...
```

Remember: You can type any time `stats()` in your R console to see your progress in the problem set.

d) Show the column "p" of dat (p stands for price). Use the $ syntax.

```
# enter your code here ...
```

e) Make scatterplot of prices on the x-axis against quantities on the y-axis using the command `plot`.

```
# enter your code here ...
```

Exercise 2

Could such a positive relationship between price and sold quantity be created from a 'normal' market with a decreasing demand function? We will first analze such a model theoretically and then simulate it.

a) Let p be a vector of possible prices consisting of all integer numbers from 0 and 100

```
# enter your code here ...
```

b) Let q be the vector of demanded quantities for each price p, given the demand function

   q = a - b * p

```
# with
a = 100
b = 2

# enter your code here ...
```

c) Let pi be a vector that contains for each price p the firm's profit, assuming the product has constant variable costs of c=11.

```
c = 11

# enter your code here ...
```

d) Plot the profits pi (on the y-axis) against the vector of prices p (on the x-axis)

```r
# enter your code here ...
```

e) Let ind.opt be the index of pi that has the highest profits

```r
# enter your code here ...
```

f) Let p.opt.grid be that price from your vector p that yields the highest profits

```r
# enter your code here ...
```

g) Just uncomment the following code to draw a vertical line in your plot at the position of p.opt.grid

```r
# abline(v=p.opt.grid)
```

h) Let `p.opt` be the profit maximizing price from all possible prices. This means compute `p.opt` by using the analytical formula for profit maximizing prices. The formula uses the parameters a, b and c.

```r
# enter your code here ...
```

i) Just uncomment the following code to add a blue vertical line to your plot at the position of p.opt

```r
# abline(v=p.opt, col="blue")
```

Exercise 3

We will now extend our model from exercise 2 by a random shock and simulate data from our model for

```r
T = 100
```

different periods.

a) Let eps be a vector of T normally distributed random variables with mean 0 and standard deviation 10

```r
# enter your code here ...
```

b) We assume that the intercept a of our demand function is given as follows:

```r
a0 = 100
a  = a0 + eps
```

This means eps is a random 'demand shock' that yields to a different demand function in each period. Sometimes demand at a given price p is higher, sometimes it is lower. You can think of eps containing factors like the weather. We can loosely interpret the resulting intercept of the demand function 'a' as a measure of 'market size' for every period.

Plot "a" to see how the market size differs between periods.

```r
# enter your code here ...
```

The slope of the demand function shall for simplicity be the same in all periods:

```r
b = 1
```

c) The variable costs c shall also be a random variable that differs between periods. Let c be a vector of T random variables that are independently drawn from a uniform distribution between 5 and 10.

```r
# enter your code here ...
```

d) Let p be the vector of profit maximizing prices (see exercise 2) for each of the T periods. (To be precise: you shall compute profit maximizing prices given that eps and c are known in each period.)

```r
# enter your code here ...
```

e) Compute a vector q of outputs for every period using the demand function from exercise 2:

```r
# enter your code here ...
```

f) Show a plot with prices on the x-Axis and output q on the y-Axis.

```r
# enter your code here ...
```

g) Compute the correlation between p and q

```r
# enter your code here ...
```

h) Run a linear regression of output against prices using the command: lm(q~p)

```r
# enter your code here ...
```

i) Look at the scatter plot and the estimated intercept and slope from h). Does the p-q relationship represent a noisy version of the demand function: q = a-b*p?

```r
# Enter "yes" or "no"

# enter your code here ...
```

j) Compute the correlations between the demand shock eps and output q, and the correlation between eps and prices p.

```r
# enter your code here ...


# enter your code here ...
```

If there is a large demand shock eps, e.g. good weather when everybody likes ice cream, the demand curve shifts upwards. This means:

- eps increases: output q is higher for every given price p
- eps increases: a profit maximizing firm will set higher prices p

The positive correlation of the price p with the unobserved demand shock eps can cause the positive correlation between p and q. This means that the observed relationship between q and p is in general *not* a noisy version of the demand function. This issue makes estimating a demand function much more tricky (since in real data we don't observe eps). We will discuss this so called "endogeneity problem" and solutions to it in detail in the course and subsequent problem sets.