Exercise 1:

Identify which type of machine learning (supervised or unsupervised? What type of task?) could be used in these cases:

a) When crossing the alps using the Brenner Autobahn, there is the option to pay electronically in advance. When approaching the toll station, the barrier automatically opens when the number plate was recognised. The recognition happens automatically by a digital camera system.

b) Diagnose whether a patient suffers from cancer or not.

c) The owner of an internet site wants to protect his system against various violations of the terms of service (bot programs, manipulation of timestamps, etc.)

d) An online shopping portal wants to determine products that are automatically offered to registered customers upon login.

e) We want to sort our news into different groups.

f) We want to sort our Email into Spam/Non-Spam.

g) In a supermarket, products that are often bought together are said to be placed side by side on a shelf to increase the sales.

h) We want to extract a list of skills from XING.

i) We want to know our top customers (i. e. highest sales, logistics, etc.).

Exercise 2:

Suppose we observe 6 data pairs and want to describe the underlying relationship between y_i and x_i

x	0.56	0.22	1.7	0.63	0.36	1.2
y	160	150	175	185	165	170

Calculate the β coefficients manually (+ calculator):

a) Assuming a standard linear relationship:

$$y_i = \beta_0 + \beta_1 x_i$$

b) Assuming a non-linear relationship (polynomial degree 2):

$$y_i = \beta_0 + \beta_1 x_i + \beta_1 x_i^2$$

Exercise 3:

Let the 2D feature vectors in the following figure be with two different class labels (triangles and circles). Classify the point (7,6) - represented by a square in the picture - with a k-nearest neighbor classifier. Distance function should be the L_1 norm (Manhattan distance):

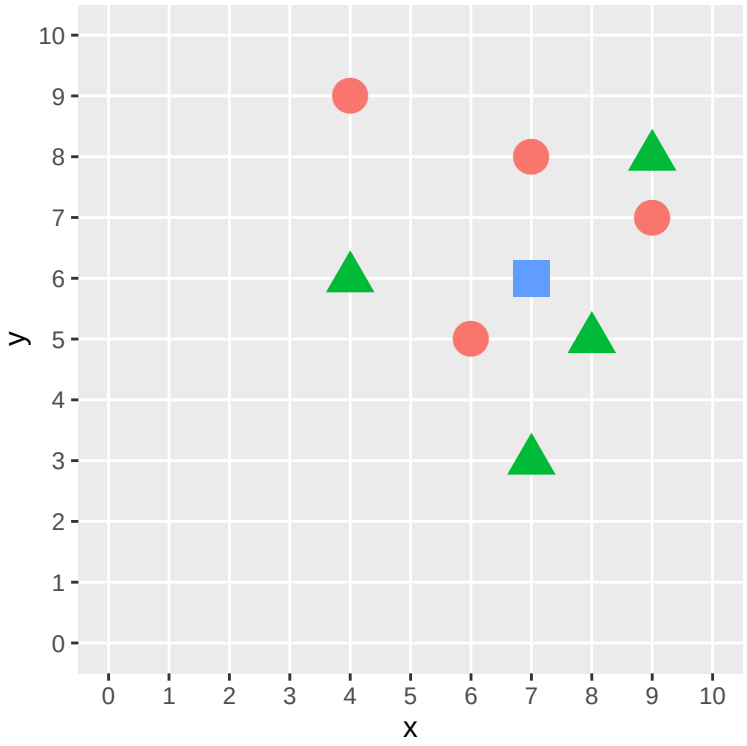
$$d_{\text{manhattan}}(x, \tilde{x}) = \sum_{j=1}^{p} |x_j - \tilde{x}_j|$$

As a decision rule, use the unweighted number of the individual classes in the k-next Neighbor Quantity, i. e. the point is assigned to the class that represents most k-nearest neighbors.

a) $k = 3$

b) $k = 5$

c) $k = 7$



Exercise 4:

How in mlr3 a learner can be constructed and what it represents can be found at `https://mlr3book.mlr-org.com/learners.html`.

a) How does a learner in mlr3 compare to what you've learned in the videos?

b) Pick an mlr3 learner of your choice. What are the different settings for this learner?
(Hint: Use `mlr_learners$keys()` to see all available learners)

Exercise 5:

We want to predict the age of an abalone using its longest shell measurement and its weight.

See: `http://archive.ics.uci.edu/ml/datasets/Abalone` for more details.

a) Plot `LongestShell`, `WholeWeight` on the x- and y-axis and color points with `Rings`

Using the mlr3-package:

b) Fit a linear model

c) Fit a k-nearest-neighbors model

d) Plot the prediction surface of lm and of knn (Hint: Use `autoplot()`)

Hint: See the official book manual of the mlr3 package for usage:

`https://mlr3book.mlr-org.com/index.html`

Exercise 6*[1]:

Let $i = 1, \ldots, n$ and $y^{(i)} \in \mathbb{R}$ be the ordered data of interest. Show that the median, which is then given by

$$y_{\text{med}} = \begin{cases} \frac{y^{(\frac{n}{2})} + y^{(\frac{n}{2}+1)}}{2}, & \text{if } n \text{ is even} \\ y^{(\frac{n+1}{2})}, & \text{if } n \text{ is odd} \end{cases},$$

is the best constant model with L1 loss.
Hint: Rearrange and combine terms of the empirical risk $\mathcal{R}_{\text{emp}}$ associated to the L1 loss L into new summands, s.t. you can use the fact that for $a, b \in \mathbb{R}, S_{a,b} : \mathbb{R} \to \mathbb{R}_0^+, c \mapsto |a - c| + |b - c|$ it holds that $c^* \in [a, b]$ minimizes $S_{a,b}$, to show that y_{med} minimizes each of these summands newly formed.

[1] This is a bonus exercise.