Series 3

1. Key metrics of good performance: give to the following examples one or more key metrics how you would measure the performance. And why you would choose this metric / these metrics:
   a. An elevator manufacturer is analysing its whole supply chain about the supply chain risk the manufacturer has.
   b. An insurance company is analysing the claims for fraud cases.
   c. A watch manufacturer wants to reduce the amount of chemicals in the production process without loss of quality and efficiency.
   d. A cargo firm wants to improve the transport routes and efficiency. The payload-distance is typically measured in tonne-kilometre.
   e. A bank is analysing the non-compliant product sales for products which should not be sold to a person based on the person's risk profile.
   f. A machine manufacturing company is analysing how to improve the production process with regards of processing time, needed resources and required raw material.

2. Which data type do you have in the following examples:
   a. You would like to analyse the movement over time of book prices on Amazon for statistic books.
   b. You are filling out a multiple choice test with four possible answers.
   c. You are analysing tweeds on Twitter.
   d. For the pricing of an insurance product you are analysing the claims and customer data of an insurance company.
   e. You are supporting in a clinical trial as a statistical expert in testing a new drug against a control group. The probands are examined during the trials and the information collected and analysed afterwards.
   f. An online hotel booking web page is using the data when a customer has accessed the web page and with which device for determining the price range of offered rooms to that customer.

3. Sentiment Analysis: The purpose of this exercise is to illustrate the practicality of sentiment analyses.

    Open the R Console and run the script "Get_Sentiment_Functionality.R". This will install the required packages together with the function "sentiment_analysis". The input to the function is a vector of strings such as for instance

    ```
    some_txt<- c("The lecture is excited.","The lecture is very interesting.","The lecture
    is dull.","I like the different topics of the lecture. What I detest is the weather.")
    ```

    The output is a dataframe counting and classifying the sentiments in the individual quotes.

	positive	negative	net
1	1	0	1
2	0	0	0
3	0	1	-1
4	1	1	0

    Before reading the text, run the sentiment analysis and investigate the output. You can do that by running the script "Sample_Text_on_Coke_Life.R" and considering the variables "output" and "summary". The latter collects the totals of the dataframe "output".

    a) Before reading the text, do you reckon that the bloggers like Coke Life?

    b) By reading the randomly chosen blog, assess the quality of the sentiment analysis. Please note that certain comments might be rude.

    c) Based on b) or your own examples, find expressions or idioms for which "sentiment_analysis" provides an erroneous sentiment.