Data Science for Marketing Analytics – Assignment 2

Questions to resolve in Assignment 1 and 2

The performance of the category beachwear is likely to be affected by the weather. Consequently, wehkamp.nl is looking for insights how the behavior of their customers is influenced by weather conditions and changes therein. Of course, the behavior of customer is also affected by other factors like customer background variables (age, gender), their search behavior, their browsing behavior, and so on. Your task is to create a data set (Assignment 1) that can be analyzed (Assignment 2) to generate insights in how weather conditions in addition to other factors drive customer behavior for the category beachwear.

For assignment 1, you created a data set that contains relevant key performance indicators (KPIs), in addition to factors that potentially affect these KPIs. Furthermore, you prepared this data so that it is ready analyses: this is the topic of Assignment 2.

In assignment 2, your task is to analyze the wehkamp.nl data, based on the data set that you extracted from their database, and the weather variables that you added.

You can analyze the data on two levels:

1. the daily level
2. the session level

On Nestor you find an sql query that will extract a session-level data set from the database ("Sessionleveldata.sql"). You need to add weather variables to this data set.

Using statistics that you learned in other courses, such as: t-tests (t.test in R), ANOVA (aov in R), regression (lm in R), and the machine learning techniques that you learn in this course, you are asked to find out how the weather variables affect the behavior of customers of wehkamp.nl, and to create a machine learning model that predicts for individual customers whether they will make a purchase.

The deliverables for Assignment 2 are:

(1) An analysis of the effect of weather variables on the behavior of customers of wehkamp.nl
(2) A machine-learning model that can be used to predict conversion on the individual customer level.
(3) Appendix: SQL query (in case you made modifications) & R code.
(4) A max 2 minutes video about the managerial insights.