Data Transformation with Campusbier Sales Orders

Exercise 1: Pick columns with select

Find solutions for the following tasks in which you have to select a subset of the total set of columns:

- 1. Select all columns ending in "_at"!
- 2. Select all columns that contain information about the delivery address!
- 3. Select all columns, which contain a date or time value!
- 4. Select the first and the last three columns!
- 5. Select all columns, which do not contain customer information or information to the delivery address!

Exercise 2: Reduce rows with filter

Find a solution to constrain the rows of the data set as described:

- 1. Keep only orders from the year 2022!
- 2. Remove all orders paid with PayPal!
- 3. Remove the top 1% of orders with the highest turnover!
- 4. Keep only orders from outside the 49 zip code area in the result!
- 5. Select all orders in which a discount code was used!

Exercise 3: Create new columns with mutate and transmute.

- 1. Create a new column with the day of the week on which the order was placed! The column should contain a speaking name of the weekday (Monday, Tuesday etc.)!
- 2. Determine all customers who accept marketing mails and who use PayPal at the same time! Create a new column marketing_paypal which contains TRUE if both are true!
- 3. Calculate a new column turnover_rank that assigns a rank number to each order based on its turnover. The order with the highest turnover should get rank 1.
- 4. Add a column that indicates whether an order has been fully paid and fulfilled! Look at the columns financial_status and fulfillment_status for this!
- 5. Calculate the time span between order receipt and final order completion and store the value in a new column!

Exercise 4: Summarize data with group_by and summarize

- 1. Create an overview for the frequency of the payment methods used! Sort the result in descending order!
- 2. Which top 10 customers generated the most revenue?
- 3. Show the sum of the turnover for all years contained in the data set! In addition to the turnover, also show the number of orders in each year!
- 4. From which postal code areas do how many orders come?
- 5. In which zip code area is the average purchase value per order the highest?