Welcome to smava challenge!

Exercises:

The Smava Test consists of three parts. The first part is an Excel test and you have to answer questions about marketing conversions (Exercise 1) and analyze some KPIs for loan offers (Exercise 2). The second part are SQL questions with a different level of complexity. In the third part you should briefly interpret some R statements. Take time to think about the questions and don't worry if you can't answer some questions. As you can see some exercises are marked with "nice to have" meaning that the questions are harder than the other ones and that we don't expect a perfect solution. The first part is the most important one, so try to answer the questions as good as possible.

Write your solution and comments in a Word or PowerPoint document and save the Excel files with your calculations. If you have any questions feel free to ask us.

Good Luck!

A: Excel

1. Marketing Conversion (1 Task, Details in the file)
2. Analysis of loan offers (2 Tasks, Details in the file)

B: SQL

1. What does ***UNION***? What is the difference between ***UNION*** and ***UNION ALL***? What does ***INTERSECT*** ? What are the prerequisites to use these statements?
2. Explain shortly the idea of transaction blocks.
3. What does the statement ***CREATE TRIGGER*** ? Explain shortly.
4. Write a statement to extract number of full months (result shown as integer) between two timestamps: ***date1*** and ***date2***.
5. Explain the results from the query below. What is the name of the function in the last column? What is the difference between this function and a regular aggregate function. How else can we use this function? Give some examples.
   SELECT depname, empno, salary, rank() OVER (PARTITION BY depname ORDER BY salary

   DESC) FROM empsalary;
6. Write a statement to create a multicolumn index on columns: ***loan_id***, ***bank_id*** in

   loan_request table.

7. Write a statement to find all records from table ***person*** for which ***person_id*** is NOT unique.
8. Can you write a simple SQL statement which selects the last_name and application_date of customers who have applied after January 12th 2012, whereas:
   - Column last_name is in table person
   - Column application_date is in table account
   - The connection between both tables is via the column account_id which exists in both tables
9. There are given two tables (details below). Write ONE query to show for each customer: the last successful loan_request (loan_id, creation_date, loan_amount), sum of loan_amount of all failed loan_requests, ratio of successful_to_failed applications. ***(nice to have)***

customer_account – contains the unique accounts of customers columns: ***account_id*** (type:int8, key column, IS_NULLABLE:NO), ***creation_date*** (type:timestamp), ***account_type*** (type: varchar, possible values: {CUSTOMER, EMPLOYEE})

loan_request – contains each loan request made through Smava columns: ***loan_id*** (int8, key column, IS_NULLABLE:NO) ***creation_date*** (timestamp), ***account_id*** (int8), ***loan_amount*** (float8) ***type*** (type:varchar, possible values {SUCCESS, FAILED})

C: R ***(nice to have)***

1. Can you write a R statement which gives you the average, median and variance of the variable main_income
2. Please interpret the following statements
   a. ba_level_postbank <- ba_level_postbank[,!(names(ba_level_postbank) == "signup_date")]
   b. ba_level_postbank$ec_property_owned_square_metres[is.na(ba_level_postbank$ec_property_owned_square_metres)] <- 0
   c. ba_level_postbank <- subset(ba_level_postbank, received_palimony >= 0)
   d. ba_level_postbank$lead_canal <- as.factor(ba_level_postbank$lead_canal)
   e. ba_level_postbank$effective_interest <- as.factor(ifelse(ba_level_postbank$ effective_interest <= 6, 1, ifelse(ba_level_postbank$ effective_interest <= 8, 2, 3)))