**Analytics Engineer - Challenge**

****Context

Let’s assume you work for a BNPL company like Kueski. The BNPL product is a payment solution that facilitates to a customer the purchase of any product, mainly through an ecommerce platform, by providing them a credit that shall be paid in installments (partialities), mainly with 0% interest.

You, as a part of the Analytics Engineering team of this company, are required to provide support for taking raw data and create useful data models that will constitute the data warehouse that will allow us to answer several business questions with ease.

## Questions

Answer the following considering the sample data provided

1. Design and develop a data model (we strongly recommend you to use Kimball methodology).
2. Show the ERD diagram and take advantage of your developed model for solving the following questions:
   1. How many loans have been disbursed in December 2021?
   2. Does the risk score is helpful for deciding what application is disbursed or not? What about the fraud score?
   3. What are the top 5 states by loans disbursed?
   4. How much Revenue by Interest did the company get by week?
   5. Do the customers prefer to obtain a loan with a specific term (tenure) or it is indistinct?
   6. Do the customers prefer to get a loan to buy things from any particular merchant?
3. Develop one dataset that contains the following “recurrency” definitions and provide any insight you may find using data visualization

Definition 1. “recurrence\_1” :

The number of loans disbursed in the client's history at the disbursement date, including the one being disbursed. (1,2,3,4,...., n)

Definition 2. “recurrence\_2” :

A loan is said to be recurrent if, at the disbursement date, the customer has paid **any loan** in its totality or has paid 3 or more installments of its **first loan disbursed**. (TRUE, FALSE)

Finally, let’s assume that the BNPL Product Director wants you to present the most relevant findings you did based on the list of questions he shared with you. Please, create a short presentation (2-3 slides) in order to explain what was the approach you took to achieve both definitions and to share any findings you did based on the data provided.

For data visualization you can use any programming language or BI tool you are familiar with. Tableau public can be downloaded from <https://public.tableau.com/en-us/s/>.

## DATA

Two datasets are provided: loan\_data.csv and repayment\_data.csv. The description of each datasets, as well as the description of some of its columns, is described below:

* loan\_data: dataset including all the loan applications of the BNPL startup.
  + customer\_id: customer unique identifier.
  + loan\_id: loan application unique identifier (primary key)
  + application\_date: timestamp at the beginning of the application.
  + term: number of installments requested to pay the loan.
  + disbursed\_date: disbursement date of the loan. (null if the application did not convert to a loan)
  + loanamount: the total price of the buyed items.
  + paid\_date: date in which the customer finished paying the loan and its obligations were fulfilled. (null if the application did not convert to a loan or if there is still a balance pending)
  + is\_disbursed: TRUE if the application convert into a loan
  + first\_vintage: The date in which the user had its first loan disbursed
* repayment\_data: dataset including information of the installments of every disbursed loan.
  + loan\_id: loan application unique identifier.
  + installment\_id: installment unique identifier (primary key)
  + installment\_number: Number of the payment to be made.
  + instalment\_duedate: The date in which the payment is due.
  + repaid\_date: The date in which the installment was paid. (Null if the installment has not been paid yet).
  + capital\_due: The amount to be paid by the due date.
  + capital\_paid: The amount that has been paid.

*Note: the data in this case is purely fictitious, and it has no connection to a real case. It has just been designed for applying basic analytics engineering skills.*