

This text contains 3 programming assignments. You can either choose to answer with pseudo coding (for questions 1 and 2 only) or any language of your preference. Any considerations regarding the coding (possible bugs, issues, enhancements, functionalities and etc) can be pointed out if needed.

1. Let  $F_n$ , where  $n \in \mathbb{N}$ , be a sequence given by  $F_n = F_{n-1} + F_{n-2}$ . Also let  $F_0 = 0$  and  $F_1 = 1$ . **Write a function that returns the value of  $F_n$ .**
2. You work in a company that only produces one product. Although, this company only produces one product, it is constantly making changes in this product, specially regarding the weight of the product. The Sales planning team usually analyzes the quantity sold as the number of items, but the current information is only available in Kg. You are asked to write a function that calculates the conversion factor ( $\frac{current\_weight}{old\_weight}$ ) that enables the sales planning team to perform their analysis. The input of your function will always be a data frame (tabular data) as the one presented below

Tabela 1: example of data frame inputs

SKU	Creation Date (yyyy-mm-dd)	SKU weight
1234	2015-08-08	12
1235	2014-12-01	15
111	2018-01-19	10
101	2012-07-20	17

SKU represents the "number version" of the product and it doesn't follow an order. Creation Date, is the date of creation of a specific SKU. SKU weight, is the weight of the SKU in Kg.

The output of your function should also be a data frame, with the same format as the one below.

Tabela 2: example of function output

SKU	SKU conversion factor
9999	1.5
444	1.3
666	0.5
777	1

3. You will be provided a data base to perform some cleaning. The data base contains a sample of the New York city crime data, you have to clean up the data base in your preferred programming language. You will have to create a function to fix the following issues:
  - Incident ZIP: Column 4 "Incident Zip" has mixed types Zip codes that are represented as floats and start with 0 are missing the first digit Some zip codes have the 4 digit extension added. Comparison becomes tough, get rid of the 4 digit extension. Drop zip 0 or bad data. Remove zips less than 10000 and greater than 19999.
  - Closed Date, Latitude and Longitude all have missing values. Get rid of them.
  - Borough Column: Find what's could be wrong or it doesn't make sense in terms of data and get rid of it.
  - Dates and times: Convert dates and times to the same data type so it can analyzed at once.
  - Create a new column that tracks the time it takes to close a complaint call it "Processing Time" and clean it (Drop negative processing time).
  - With the cleansed data create a data base, create a table for Values and at least one Dimension table, **Please write the SQL statement.**