

Practice Assignment

2020-06-09

Prompt

You are working for a local retail firm. They have many, many clients and are interested in some basic information about them. The firm has supplied you with the excel sheet *Masterdata.csv*. They have also asked you to do some basic analysis on the dataset. You are asked to do the following:

Basic

- 1) How many missing values are there in each column? How many missing values are there total? Put your answers in a tibble and name it **missing**.
- 2) How many unique states does the company operate in? Are there any typos in this column? If so, correct the typos. Save the corrected dataframe as **data.cleaned**. Unless otherwise told, work with **data.cleaned**.
- 3) Find the average invoice and the percent of invoices paid. Save them as **avg_inv** and **avg_inv_paid**.

Advanced

- 4) Create a new dataframe where the “Names” column is split into 2 columns: *First Name* and *Last Name*. Nothing else needs to be changed. Save it as **data.cleaned.split**. For part 5, use **data.cleaned.split**.
- 5) Write a function named **client_status** that inputs *last name*, *first name* and outputs:
 - i) state they live in, ii) number of invoices that need to be paid, iii) outstanding debt.Here, outstanding debt means the sum of the value of the invoices not yet paid. Have your code return an error if someone enters a name not in **data.cleaned.split**. Provide a line of code that runs your function for *Jordyn Kang*.