Data Exercise

Data Evangelist

September 2021

Introduction

This assignment is intended to assess skills in data analysis, visualization, interpretation and presentation skills. This assignment can be tackled in multiple ways - and we're mostly excited to see what you can come up with. This assignment is accompanied by two data files - you can choose to use one, two or both in combination.

Description of deliverable

We expect that at the interview, you've prepared 10 minutes worth of material relating to this assignment. It can take the form of a presentation, a set of concrete, informative visualizations or just a summary of interesting findings - or all of the above. You're free to use any tools you like.

How to think about the Data

The two data files in question relates to 1) Customer Subscriptions, 2) Claims.

1. A customer subscription is an agreement that UNDO and a customer go into, where the customer pays a monthly subscription fee in exchange for insurance services. A policy may include multiple insurances. For example, one customer may choose to purchase liability insurance and bike insurance, while another may choose to purchase accident insurance. The price of these services vary across customers and products.
The user_sample.csv dataset includes the total price of the chosen insurances as

well
as the date from which the agreement goes into effect - it also includes a list of the chosen coverages/insurances, and an id on that user.

2. The claims_sample.csv dataset relates to claims. In insurance a claim is when a customer chooses to activate their insurance with the intent of receiving a lump sum back. Say, I have a bike insurance; my bike is worth 5.000DKK; my bike gets stolen - in this case, I'd reach out to UNDO and ask for a reimbursement such that I can go out and purchase a new, equivalent bike. This ask for reimbursement is what's consideredaclaim.
Not all claims are valid though - some customers will attempt to claim things that aren't covered by their selected products - and others will purposely destroy their itemswiththeintentofgettingmoneypaidout.Theclaims_sample.csvdatasetis only concerned with claims; not whether that claim was paid out or deemed fraudulent. The table includes information about what type of claim was made; e.g. Bike Theft, Dropped Smartphone, Smashed Screen on laptop etc, it also includes information about when the claim was made, by which user, the number of items involved in the claim (if a bag gets stolen, many items may need reimbursement), and the original purchase price of the items.

Data File Explanation

user_sample.csv
Contains 2.000 rows and 6 columns.

user_id: Is an id - can be linked with user_id in claims_sample.csv
user_started_date: The date that the user signed up with UNDO
monthly_price: The that the customer pays each month
coverage_types_selected: The coverage enumerations that the users selected coverages pertain to.

coverages_selected: The coverages as described in the app. 1-to-1 with
covrage_types_selected. area: T he part of Denmark that the customer is located.

claims_sample.csv
Contains 229 rows and 9 columns.

purchase_amount: The price that the customer originally paid when they purchased the product items_claimed: The number of items that the customer has claimed

user_id: An id - can be linked with user_id in user_sample.csv

reason: A categorical variable that on occasion will hold information about what happened to the claimed item

claim_filed_date: The date that the customer filed a claim with UNDO

claim_coverage_triggered: The coverage used to file the claim

agreement_type: The type of insurance the triggered coverage falls under

claim_coverage_type_triggered: See user_sample.coverage_type_selected. 1-to-1 with claim_coverage_triggered

platform: Did the user file the claim from an iOS device or an Android device.

note
Some variables contain information in Danish. The contents of these variables primarily pertain to coverages, geography or claim notes.
Misunderstandings due to translation or confusion because of language will not be penalized.