

Instructions

- Use any tools that you are comfortable with to perform the tasks required (for example Tableau, Power BI, MS Excel, Python, R).
- Write your solutions in the workspace provided from your certification page.
- Include all of the visualizations you create to complete the tasks.
- Visualizations must be visible in the published version of the workspace. Links to external visualizations will not be accepted.
- You do not need to include any code.
- You must pass all criteria to pass this exam. The full criteria can be found [here](#).

Background

Vivendo is a fast food chain in Brazil with over 200 outlets.

Customers often claim compensation from the company for food poisoning.

The legal team processes these claims. The legal team has offices in four locations.

The legal team wants to improve how long it takes to reply to customers and close claims.

The head of the legal department wants a report on how each location differs in the time it takes to close claims.

Data

The dataset contains one row for each claim.

The dataset can be downloaded from [here](#).

Column Name	Criteria
claim_id	Nominal. The unique identifier of the claim. Missing values are not possible due to the database structure.
time_to_close	Discrete. The number of days to close the claim. Any positive value. Replace missing values with the overall median time to close.
claim_amount	Continuous. The initial claim requested in the currency of Brazil, rounded to 2 decimal places. Replace missing values with the overall median claim amount.
amount_paid	Continuous. Final amount paid. In the currency of Brazil. Rounded to 2 decimal places. Replace missing values with the overall median amount paid.
location	Nominal. Location of the claim, one of "RECIFE", "SAO LUIS", "FORTALEZA", or "NATAL". Remove missing values.
individuals_on_claim	Discrete. Number of individuals on this claim. Minimum 1 person. Replace missing value with 0.
linked_cases	Nominal. Whether this claim is linked to other cases. Either TRUE or FALSE. Replace missing values with FALSE.
cause	Nominal. Cause of the food poisoning. One of "vegetable", "meat" or "unknown". Replace missing values with 'unknown'.

Tasks

Submit your answers directly in the workspace provided.

1. For every column in the data:
 - a. State whether the values match the description given in the table above.
 - b. State the number of missing values in the column.
 - c. Describe what you did to make values match the description if they did not match.
2. Create a visualization that shows the number of claims in each location. Use the visualization to:
 - a. State which category of the variable location has the most observations
 - b. Explain whether the observations are balanced across categories of the variable location
3. Describe the distribution of time to close for all claims. Your answer must include a visualization that shows the distribution.
4. Describe the relationship between time to close and location. Your answer must include a visualization to demonstrate the relationship.