

Background

The insurance industry in Africa, despite the continent being home to 17% of the global population, plays a relatively small role in the worldwide insurance market. In 2017, the total value of insurance premiums in Africa was estimated at \$45 billion. However, the industry accounts for just over 1% of insured catastrophe losses globally, highlighting a significant disparity in insurance coverage and impact compared to other regions. To tackle these issues, we need to better understand the market's challenges, define a clear problem statement, and create effective tools that can improve the industry.

Table 1: The table below shows the main observations that will help us understand and aid in decision-making.

Topics	Key Components
Affordability	<ul style="list-style-type: none">• The daily expense noted is \$25, equivalent to the cost of a simple cheeseburger at McDonald's.• This amount is 10 times lower than the equivalent cost in other countries.• South Africa has the highest average salary among the compared regions, amounting to \$2,088 per month, or about \$70 per day.
Insurance Knowledge	<ul style="list-style-type: none">• Why should I spend my money now for a risk that might not happen?• Who can I target and where
Target Marketing	<ul style="list-style-type: none">• Africa experiences a high unemployment rate of 30%.• Less than 50% of the available market falls below a certain threshold, though the exact metric is not specified.• Which client pool can be optimised?

Data

The insurance_claims dataset is in .xlsx format and can be accessed through the GitHub repository [linked here](#). This dataset will be used in a series of walkthroughs designed to explore various fundamental aspects of data analysis. These tutorials will guide users through techniques to thoroughly examine the dataset and extract meaningful insights. Below are the descriptions of the features in the dataset:

Table 2: The table below highlights the key features and descriptions of the insurance dataset.

Column name	Feature Description
policy_bind_date	Starting date of the insurance policy.
policy_csl	Combined Single Limits - This is the maximum value the insurer will pay out per incident.
policy_annual_premium	The total dollar amount for the yearly premium.
umbrella_limit	Extra insurance that provides protection beyond existing limits and coverages of other policies.
auto_make	Vehicle brand.
auto_model	Vehicle model.
insured_education_level	Highest qualification of the insurer.
policy_deductable	Excess payment before a payout or service is conducted.
insured_occupation	The profession in which the insurer works.
fraud_reported	Y - a fraudulent or false claim N - a legit and valid claim.