# Documentation for the Property Claims Dataset

This document provides a detailed description of the dataset generated for analyzing property insurance claims. The dataset is designed to simulate realistic claims data, focusing on granular cause-of-loss (CoL) details to support subrogation, underwriting, and loss prevention efforts.

## Directory Structure

output/

└── property\_claims\_MM-DD.csv

## Example Prompts

### 1. Subrogation Analysis

"List all claims with subrogation potential marked as 'Yes' and their estimated recovery amounts."

"Identify the most common loss categories associated with subrogation potential."

"Analyze the recovery amounts for claims with preventable losses."

### 2. Preventive Actions

"List all claims marked as 'Preventable Loss' and their recommended preventive actions."

"Identify the most frequent origins of failure for preventable losses."

"Analyze the financial impact of preventable losses by loss category."

### 3. Financial Insights

"Calculate the total estimated loss amount for all claims."

"Identify the top 5 claims with the highest estimated loss amounts."

"Analyze the average recovery amount for claims with subrogation potential."

## Dataset Description

### Property Claims Data

Description: Comprehensive dataset containing detailed property insurance claims data, including cause-of-loss details, financial metrics, and subrogation potential.

Records: 500 entries

### Columns:

* Claim ID: Unique identifier for each claim.
* Policyholder ID: Unique identifier for the policyholder.
* Date of Loss: Date when the loss occurred.
* Location of Loss: Geographic location of the loss (e.g., city, state).
* Loss Category: Broad category of the loss (e.g., Water Damage, Fire, Theft).
* Loss Subcategory: Specific type of loss within the category (e.g., Pipe Burst, Electrical Fire).
* Estimated Loss Amount ($): Estimated financial impact of the loss.
* Claim Status: Current status of the claim:  
  - Open  
  - Closed  
  - Denied
* Subrogation Potential: Indicates whether the claim has subrogation potential:  
  - Yes  
  - No  
  - Pending Review
* Recovery Amount ($): Amount recovered through subrogation efforts.
* Preventable Loss: Indicates whether the loss was preventable (Yes/No).
* Recommended Preventive Action: Suggested actions to prevent similar losses in the future.
* Component of Failure: Primary component involved in the loss (e.g., Roof, Pipe, HVAC System).
* Subcomponent of Failure: Specific subcomponent involved (e.g., Valve, Wiring).
* Origin of Failure: Cause of the failure (e.g., Faulty Installation, Wear and Tear, Weather Event).
* Manufacturer/Brand: Manufacturer or brand of the failed component.
* Adjuster Notes: Additional notes or observations from the claims adjuster.
* Underwriting Impact: Indicates whether the claim has implications for underwriting (Yes/No).

## Use Cases

1. 1. Subrogation and Recovery Analysis: Identify claims with subrogation potential and analyze recovery amounts to improve financial outcomes.
2. 2. Loss Prevention: Analyze preventable losses and recommend actions to reduce future claims.
3. 3. Underwriting Insights: Use granular cause-of-loss data to refine underwriting guidelines and assess risk more accurately.
4. 4. Financial Impact Analysis: Evaluate the financial impact of claims by category, subcategory, and location to prioritize risk management efforts.
5. 5. Component and Failure Analysis: Identify patterns in component failures to address recurring issues and improve product quality.
6. 6. Temporal and Geographical Trends: Analyze seasonal and regional trends in claims to inform resource allocation and risk mitigation strategies.