

Capstone Project

What is a Capstone Project?

- A **Capstone Project** represents the culmination of your learning journey. It brings together all the skills, concepts, and methods you've studied during the AI course and applies them to solve a real-world problem. Think of it as a final test, where you not only demonstrate your understanding of AI but also how to use it effectively in practical scenarios.
- **Hands-On Approach:** Involves actively engaging with the problem by gathering data, building models, and testing solutions. It may include:
 - **Team Discussions:** Collaborating with peers to explore different perspectives and approaches.
 - **Web Search:** Researching online to gather relevant information, datasets, and tools.
 - **Case Studies:** Analyzing previous successful projects to learn from them.

Objectives of the Capstone Project:

1. **Application of Learning:** The goal is to apply theoretical knowledge to practical, real-world issues. This demonstrates your ability to translate academic concepts into actionable solutions.
 - **Example:** If you've learned about neural networks, you should be able to apply them to a project such as image classification.
2. **Communicating Solutions:** It's important to present your findings in a way that non-technical stakeholders can understand. Explaining complex algorithms in simple, clear language is key.

- **Example:** When explaining a model's predictions to a business audience, you would avoid jargon like "backpropagation" and instead focus on how the model benefits the business.
3. **Choosing the Right Algorithm:** You need to analyze the problem carefully to determine the most appropriate algorithm to solve it.
- **Example:** For predicting stock prices (a regression task), you might choose linear regression or a more complex algorithm like a neural network, depending on the dataset and problem complexity.

Key Concepts for Capstone Project:

- **AI Project Cycle:** This is the structured process you follow in any AI project. It includes:
 1. **Problem Definition:** Clearly define the issue you're addressing.
 2. **Data Gathering:** Collect the right data for training your model.
 3. **Feature Definition:** Identify the key factors (features) that influence the outcome.
 4. **Model Construction:** Build and train a suitable AI model.
 5. **Evaluation & Refinement:** Assess the model's performance and make improvements.
 6. **Deployment:** Implement the solution in a real-world setting.
- **Model Validation:** This involves testing your model's performance to ensure it works well. Techniques like RMSE (Root Mean Squared Error), MSE (Mean Squared Error), and MAPE (Mean Absolute Percentage Error) help measure accuracy.

Capstone Project: Policy Bazaar

Project Name :- Policy Bazaar

Team Members Name :- Ahasas Singhal & Chirag Agarwal

Problem Selection :- People face difficulties in selecting the best insurance plan for them.

Introduction

This capstone project focuses on analyzing insurance policies available through Policy Bazaar. The goal is to help users understand the variety of insurance plans available, visualize trends across policy types, and assess claim settlement performances of various insurance providers.

Objective

The objective of this project is to provide an interactive Python-based system for users to browse, compare, and visualize different insurance plans based on their type, premium, coverage, and claim settlement ratio. Users can select specific policy types or view all policies to get a comprehensive understanding.

Dataset Description

Three datasets are used for this project:

- **policies.csv**: Contains details about insurance policies including name, type, premium, coverage, and features.
- **claim_ratios.csv**: Provides the claim settlement ratios of different insurance companies.
- **premium_vs_coverage.csv**: Derived dataset to assist in visualizing premium vs coverage.

Code Implementation

The following Python code was developed for interactive policy browsing and visualization:

```
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np

policies = pd.read_csv(r"C:\\Users\\user23\\Downloads\\policies.csv",
encoding='ISO-8859-1')
claim_ratios = pd.read_csv(r"C:\\Users\\user23\\Downloads\\claim_ratios.csv",
encoding='ISO-8859-1')
premium_vs_coverage =
pd.read_csv(r"C:\\Users\\user23\\Downloads\\premium_vs_coverage.csv",
encoding='ISO-8859-1')

def list_types():
    return policies['Type'].unique()

def plot_policy_type(type_name, df):
    counts = df['Term'].value_counts().sort_index()
    plt.figure(figsize=(8, 4))
    counts.plot(kind='bar', color='skyblue', edgecolor='black')
    plt.title(f"Number of '{type_name}' Plans by Term")
    plt.xlabel('Term (Years)')
    plt.ylabel('Number of Plans')
    plt.tight_layout()
    plt.show()

def plot_claim_ratios():
    plt.figure(figsize=(10, 5))
    plt.bar(claim_ratios['Company'], claim_ratios['CSR_percent'], color='green')
    plt.title("Claim Settlement Ratio by Company")
    plt.xlabel("Company")
    plt.ylabel("CSR (%)")
    plt.xticks(rotation=45)
    plt.tight_layout()
```

```

plt.show()

def plot_premium_vs_coverage():
    df = premium_vs_coverage.dropna(subset=['Coverage'])
    plt.figure(figsize=(10, 5))
    plt.scatter(df['Premium'], df['Coverage'], c='purple', alpha=0.6)
    plt.title("Premium vs Coverage")
    plt.xlabel("Annual Premium (INR)")
    plt.ylabel("Coverage Amount (INR)")
    plt.tight_layout()
    plt.show()

def show_policy_details(df):
    print(df.to_string(index=False))

while True:
    print("Available Policy Types:")
    for i, t in enumerate(list_types(), 1):
        print(f"{i}. {t}")

    print("\nType one or multiple policy types separated by commas (e.g., term, ulip,
whole life) or type 'all' to view everything.")
    x = input("Enter your choice(s): ").lower()

    if x == 'all':
        show_policy_details(policies)
        plot_claim_ratios()
        plot_premium_vs_coverage()
        break

    selected_types = [t.strip() for t in x.split(',')]
    matches = []

    for s in selected_types:
        for t in list_types():
            if t.lower() == s:
                matches.append(t)
                break

    if matches:

```

```

combined_df = policies[policies['Type'].isin(matches)]
show_policy_details(combined_df)

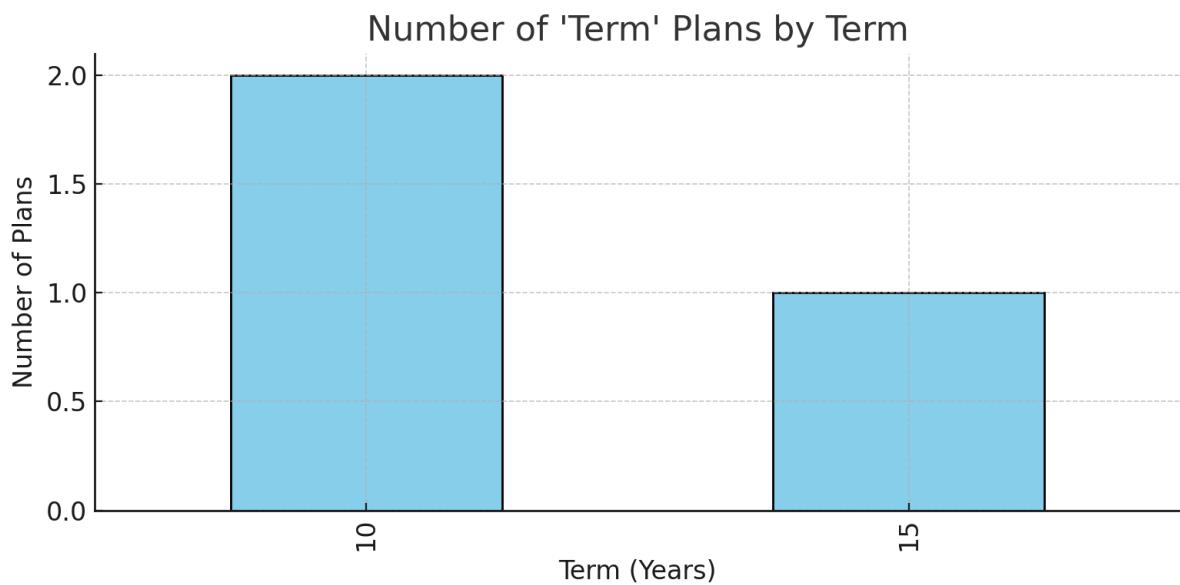
for policy_type in matches:
    df = policies[policies['Type'] == policy_type]
    plot_policy_type(policy_type, df)

plot_claim_ratios()
plot_premium_vs_coverage()
break

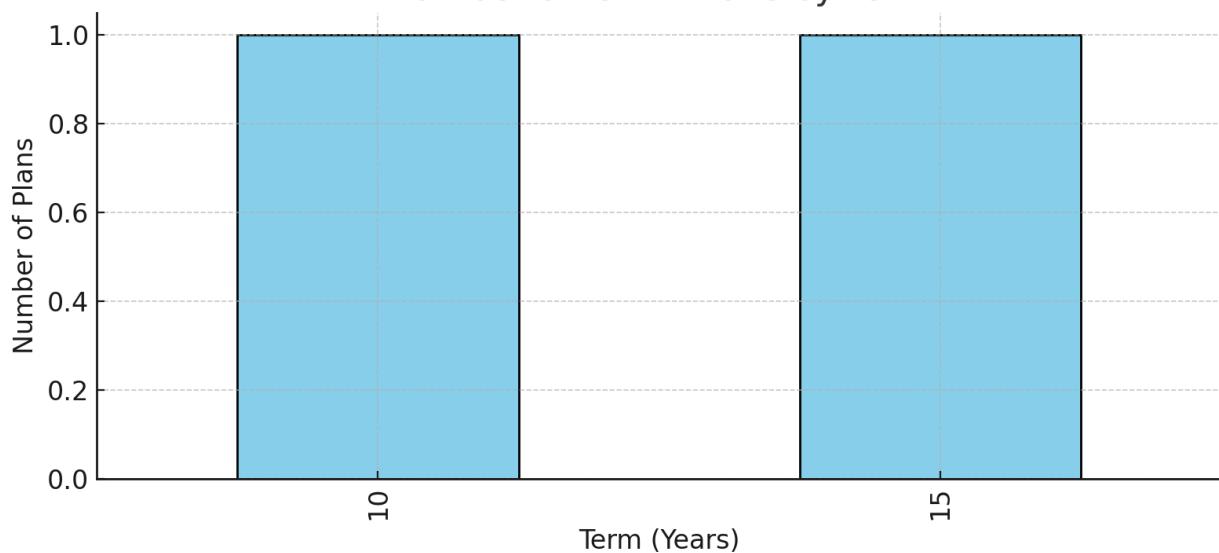
else:
    print("No matching policy types found.")
    print("Available Policy Types:")
    for i, t in enumerate(list_types(), 1):
        print(f"{i}. {t}")
    y = input("Retry? (yes/no): ").lower()
    if 'no' in y:
        break

```

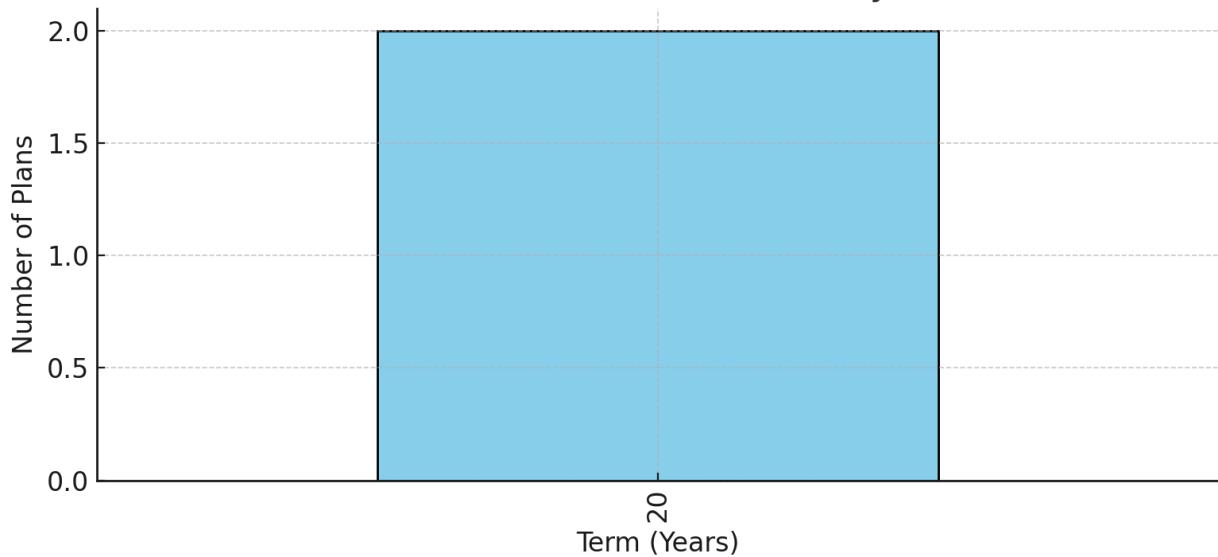
Outcomes -



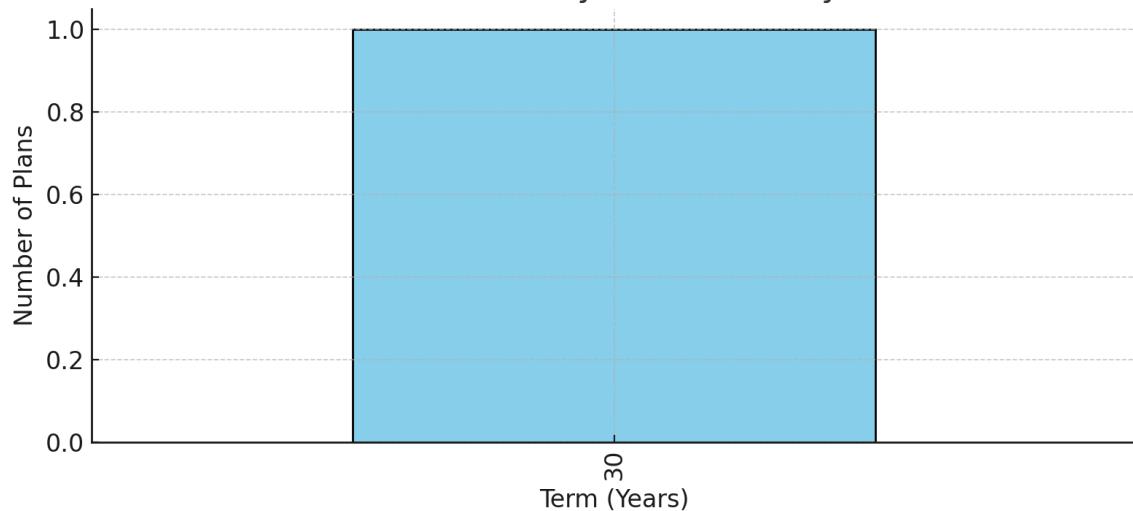
Number of 'ULIP' Plans by Term



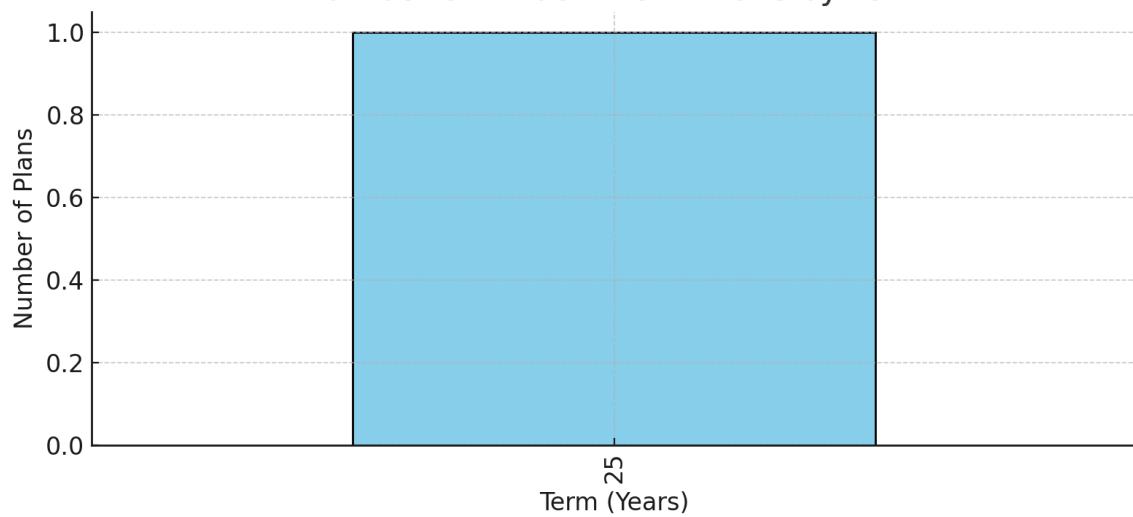
Number of 'Whole Life' Plans by Term

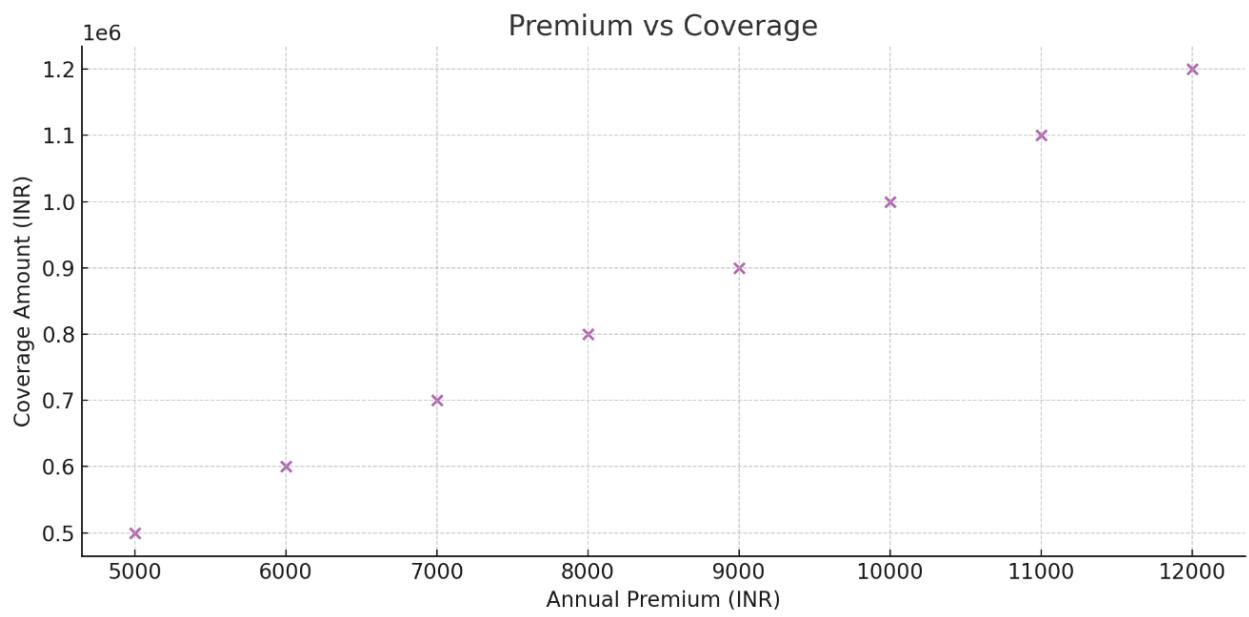


Number of 'Money Back' Plans by Term



Number of 'Endowment' Plans by Term





Data Frames (CSV files) -

POLICIES.CSV

PolicyName	Type	Premium	Coverage	Term	Feature1	Feature2	Claim Settlement Ratio
HDFC Life Click 2 Protect Super	Term	616	10000000	30	Accidental Cover	Tax Benefit	99.96
Max Life Smart Secure Plus Plan	Term	735	10000000	30	Terminal Illness Cover	Family Income Benefit	99.34
ICICI Pru iProtect Smart	Term	710	10000000	30	Waiver of Premium	Rider Options	98.6
LIC Jeevan Umang	Whole Life	1234	5000000	Whole Life	Annual Survival Benefit	Rider Options	98.35
Bajaj Allianz Life Goal Assure	ULIP	2500		20	Wealth Creation	Market Linked Returns	99.04
SBI Life Smart Shield	Term	680	5000000	25	Critical Illness Benefit	Tax Benefit	97.93
Kotak e-Term Plan	Term	599	10000000	30	Critical Illness	Tax Benefit	98.5
Tata AIA Sampoorna Raksha Supreme	Term	670	10000000	30	Accidental Death Benefit	Terminal Illness Cover	99.01
PNB MetLife Mera Term Plan Plus	Term	640	10000000	30	Family Income Benefit	Rider Options	97.8
Bharti AXA Flexi Term Plan	Term	625	10000000	30	Waiver of Premium	Terminal Illness Cover	96.9

CLAIM_RATIOS.CSV

Company	Claim Settled Amount (in Cr.)	CSR (in %)
HDFC Life	2300	99.96
Max Life	1800	99.34
ICICI Prudential	2100	98.6
LIC	4500	98.35
Bajaj Allianz	1300	99.04
SBI Life	1700	97.93
Kotak Life	900	98.5
Tata AIA	1500	99.01

PREMIUIM_VS_COVERAGE.CSV

PolicyName	Premium	Coverage	Term	Type
HDFC Life Click 2 Protect Super	616	10000000	30	Term
Max Life Smart Secure Plus Plan	735	10000000	30	Term
ICICI Pru Protect Smart	710	10000000	30	Term
LIC Jeevan Umang	1234	5000000	Whole Life	Whole Life
Bajaj Allianz Life Goal Assure	2500		20	ULIP
SBI Life Smart Shield	680	5000000	25	Term
Kotak e-Term Plan	599	10000000	30	Term
Tata AIA Sampoorna Raksha Supreme	670	10000000	30	Term
PNB MetLife Mera Term Plan Plus	640	10000000	30	Term
Bharti AXA Flexi Term Plan	625	10000000	30	Term

Conclusion

This project provides an effective platform to explore, filter, and visualize insurance policies, helping users make informed decisions when selecting their insurance coverage.