**Insurance Claims Analysis & Fraud Detection Project**

**1. Project Title**

**Insurance Claims Analysis & Fraud Detection Dashboard (MySQL, Power BI)**

**2. Short Summary**

*An end-to-end business intelligence solution to analyze insurance claim patterns, detect potential fraud, and generate actionable insights for underwriters and claims managers.*

**3. Objective / Problem Statement**

The project aims to identify claim trends, highlight high-risk policies, and detect fraudulent claims early. This helps insurance companies reduce losses, improve underwriting accuracy, and speed up claim settlement.

**4. Tools & Technologies Used**

* **Database:** MySQL
* **Visualization**: Power BI
* **Data Cleaning:** MS Excel (CSV UTF-8)
* **Other:** SQL Queries for analysis

**5. Dataset Details**

* **Source:** Kaggle – <https://www.kaggle.com/datasets/shivamb/vehicle-claim-fraud-detection>
* **Size:** ~15421 rows, 39 columns
* **Key Columns**: Month, AccidentArea, PolicyType, VehiclePrice, FraudFound\_P, Age, AgentType
* **Target Column**: FraudFound\_P (Fraud flag)

**6. Methodology / Approach**

**Step 1:** Data Cleaning (Excel)

* Removed empty rows & standardized headers
* Checked data types for numeric/categorical fields

**Step 2: Database Setup (MySQL)**

* Designed schema with correct data types
* Imported CSV using MySQL Workbench
* Verified import with exploratory queries

**Step 3: Data Exploration (SQL)**

* Fraud percentage calculation
* Top policy types in fraud cases
* Average deductible per policy type
* Claim trends by month and vehicle category
* Average claim settlement time
* Claims frequency per customer
* Top 5 most claimed policy types
* High-value or suspicious claims (Fraud detection flags)

**Step 4: Dashboard Creation (Power BI)**

* Connected to MySQL database
* Created KPIs, charts, and filters for interactive insights

**7. Key Insights**

* Fraud rate: 5.95% of total claims
* Sedan - All Peril policy type had the highest fraud cases
* Highest Average deductible is on sports – all perils policytype
* 99.42% of fraud claims had no witnesses present

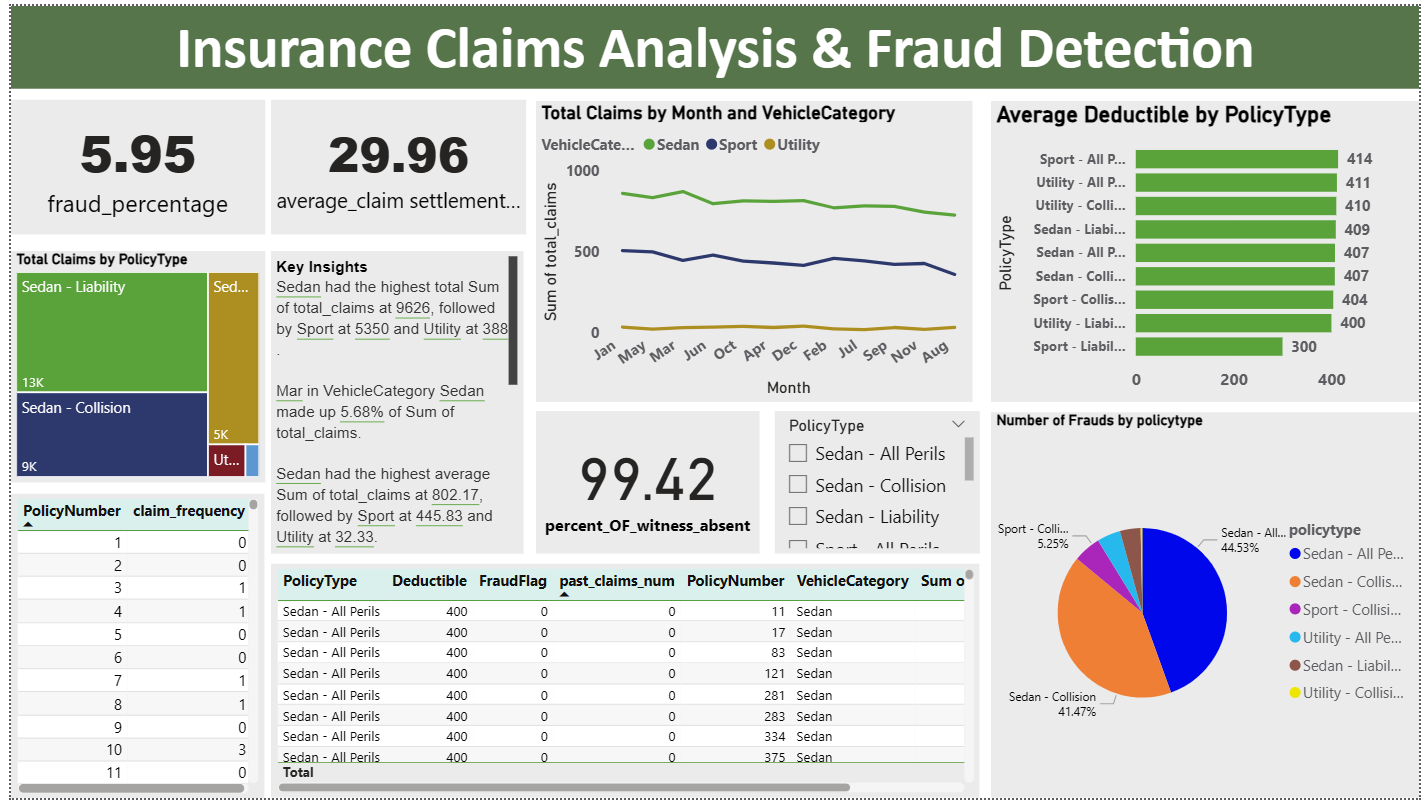
**8. Recommendations**

* **Implement automated fraud detection systems** to flag high-risk claims for review.
* **Investigate "Sedan - All Peril" policies** to understand why they have the highest fraud rates and adjust underwriting or premiums as needed.
* **Mandate stricter verification for claims with no witnesses**, as this is a key indicator of fraud.
* **Provide specialized training** to claims adjusters on recognizing fraud signals.
* **Increase scrutiny on high-deductible policies** (e.g., Sports – All Perils) due to their high value.

**9. Final Deliverables**

* Cleaned dataset (CSV)
* MySQL database schema & queries (SQL file)
* Power BI dashboard (.pbix)
* Final PDF report with insights & recommendations

**10. Screenshots**

****

**11. Conclusion**

This project demonstrates the use of SQL and Power BI to solve a real-world insurance problem. It showcases data cleaning, database management, query writing, and dashboard building skills, making it a valuable portfolio piece for a Business Analyst role.

**12. GitHub / Portfolio Link**

**https://github.com/neelima-collab/Insurance-Claims-Analysis-Fraud-Detection.git**