**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**

***(Insert Power BI dashboard images here)***

**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**

***(If you host files online)***

**📋 Step 0: Plan the Project**

* **Define your goal**:  
  *"Build an end-to-end Insurance Claims Analysis & Fraud Detection Dashboard using MySQL + Power BI."*
* **Identify stakeholders** (hypothetical):  
  *Claims managers, underwriters, fraud investigation team.*
* **List key questions to answer**:
  + What is the fraud rate?
  + Which policy types are most risky?
  + Which regions/vehicle categories have more claims?

📄 *Deliverable:* Write a short **Project Scope Document** (1 page) with:

* Objective
* Tools
* Data source
* Expected outcomes

**🗂 Step 1: Data Preparation**

* **Clean the CSV** in Excel:
  + Remove blanks
  + Standardize formats
  + Save as CSV UTF-8
* **Document what you cleaned** in a separate sheet or Word file:
  + Example: “Replaced ‘?’ with NULL in Age column”, “Converted Year to integer”.

📄 *Deliverable:* Data cleaning log

**🛢 Step 2: Database Creation (MySQL)**

* **Design the table schema** (we already drafted one).
* **Import the CSV** into MySQL.
* **Verify data types** are correct.

📄 *Deliverable:* SQL file (create\_table.sql + import.sql) with comments explaining each step.

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

Run queries to:

* Count total claims
* Fraud claims %
* Avg deductible per policy type
* Top 5 most common vehicle makes in fraud cases
* Monthly trend of claims

📄 *Deliverable:* SQL script with **comments** explaining business meaning of each query.

**📊 Step 4: Power BI Dashboard**

* **Connect Power BI to MySQL**
* **Create visuals**:
  + KPI cards (Total Claims, Fraud %, Avg Claim Amount)
  + Bar chart (Fraud by Policy Type)
  + Heatmap (Claims by Accident Area & Vehicle Price)
  + Time series (Claims over months)
* Add slicers for Policy Type, Year, Accident Area.

📄 *Deliverable:* Power BI .pbix file + a screenshot of dashboard.

**📝 Step 5: Documentation & Storytelling**

This is **critical** for making your resume/portfolio strong:

* Make a **Project Report** (2–4 pages) covering:
  1. **Introduction** (Problem statement, business relevance)
  2. **Tools Used**
  3. **Data Overview**
  4. **Methodology** (Cleaning → SQL → Dashboard)
  5. **Findings & Insights**
  6. **Recommendations**
* Keep visuals and SQL snippets in it.
* End with a **Conclusion**: “This analysis helps underwriters detect fraud early…”

📄 *Deliverable:* PDF report (can be shared with recruiters)

**📌 Step 6: Resume & LinkedIn Integration**

* Add this under **Projects**:

*Developed an end-to-end insurance claims analysis solution using MySQL and Power BI. Designed database schema, wrote SQL queries to detect fraud patterns, and built interactive dashboards for claims managers. Delivered insights on fraud risk by policy type, vehicle category, and claim patterns.*

* Post a LinkedIn update with screenshots & short write-up.

💡 **Tip:** Save all your files (Excel, SQL, Power BI, Report) in a **GitHub repository**. Recruiters love seeing clean, documented projects.

If you want, I can prepare **a project execution checklist** for you so you just tick off tasks as you go — that makes sure you won’t miss anything and your documentation is portfolio-ready.  
Do you want me to make that checklist?