Oklahoma Tax Commission Data Quality Dashboard

Stakeholder Report  
Prepared by: Will Thompson

# Executive Summary

This dashboard project was developed to support the Oklahoma Tax Commission’s mission of promoting tax compliance with transparency and fairness. It identifies potential anomalies and filing errors in tax return data using Python preprocessing, SQLite database integration, and Power BI visualization. The goal is to streamline data quality assurance efforts and deliver clear insights into filing issues across regions and years.

# Objectives

- Detect tax return anomalies using automated scripts  
- Summarize filing statistics such as refund ranges and average return amounts  
- Provide filtering options by region, year, and audit flags  
- Visualize patterns in taxpayer errors to support compliance decisions

# Data Sources and Pipeline

The raw dataset (`tax\_returns.csv`) simulates Oklahoma taxpayer filings, including return amounts, payment dates, and audit flags. The data is loaded into an SQLite database (`tax\_returns.db`) via a Python script (`create\_database.py`). Quality checks are conducted using SQL (`data\_quality\_queries.sql`) to flag issues like premature payments, excessive deductions, and missing values. The processed data is visualized with Power BI.

# Data Quality Rules Applied

1. Refund\_Amount < 0 (Negative Refund)  
2. Payment\_Date < Filing\_Date (Premature Payment)  
3. Paid\_Amount or Return\_Amount is NULL (Missing Info)  
4. Deductions > 50% of Paid\_Amount (High Deductions)

# Power BI Dashboard Overview

The dashboard includes KPIs, filters, error type breakdowns, and taxpayer summaries.

A screenshot of a data sheet

AI-generated content may be incorrect.

# Conclusion & Recommendations

This dashboard demonstrates a scalable approach to data quality analysis using open-source tools and visual reporting. It could be extended with real-time data pipelines, regional drilldowns, and integration with OTC’s existing compliance systems.