**PAN Card Validation Project – Executive Report**

**Project Overview**

This project focused on ensuring that all PAN (Permanent Account Number) records are accurate, complete, and in line with official standards.  
The work involved cleaning the dataset, checking each PAN against set rules, and producing a clear report showing which records are valid and which are not.

**Problem Statement**

The original dataset of 10,000 PAN records contained incomplete entries, duplicates, and formatting issues.  
Such errors can:

* Create compliance risks.
* Allow fake or incorrect PANs into the system.
* Waste time on manual verification.

The goal was to develop a reliable process to quickly and accurately separate valid PANs from invalid ones.

**Process Overview**

* **Data Review & Cleaning:** Removed incomplete records, eliminated duplicates, and corrected formatting issues using user defined- function in SQL.
* **Validation:** Checked each PAN against official rules for structure and logical consistency using Regular Expression and another user-defined in PostgreSQL.
* **Categorization:** Marked each PAN as either Valid or Invalid.
* **Summary Reporting:** Created a simple report showing overall results.

**Results**

| **Metric** | **Count** |
| --- | --- |
| Total Records | 10,000 |
| Valid PANs | 3,186 |
| Invalid PANs | 5,839 |
| Filtered-out Records | 975 |

**Conclusion & Summary**

* The project delivered a clean, accurate, and compliance-ready set of PAN records.
* The process is fast, repeatable, and scalable for future data.
* The outcome improves data quality, reduces compliance risk, and saves time by removing the need for manual checks.

**Business Value:**

* **Data Retention:** No data was deleted or damaged during this project, correct PANs were displayer with filtering functionality.
* **Higher Accuracy:** Only valid and correctly formatted PANs remain in the system.
* **Risk Reduction:** Early detection of suspicious or non-compliant entries.
* **Efficiency Gains:** Significantly reduced manual validation workload.