Business/Functional Requirement Document

Problem statement

The business needs a single-page Sales Insights dashboard consolidating performance across years, products, geographies, and sales representatives to support trend analysis and decision-making. Data is dispersed across multiple files and formats, requiring resilient ingestion, standardized modeling, ID cleansing, and robust DAX measures for revenue, cost, profitability, and period-over-period growth.

Objectives

- Deliver an actionable one-page dashboard with intuitive visuals backed by a governed data model.
- Implement resilient ingestion that auto-loads new yearly files and tolerates file removals without refresh failure.
- Standardize keys and data types to enable clean relationships, geo mapping, and time intelligence.

In scope

- Consolidation of yearly Sales files into a single fact table.
- Dimension onboarding: Categories, Geography, Product, SalesRep, SubCategories.
- Transformations for Location split, Date normalization, and GeoKey creation.
- Reusable function to strip the 'ID ' prefix where present in IDs.
- DAX measures: Total Revenue, Total Cost, Gross Profit, Gross Profit MoM Change%, Average Sales per Day, product-level drop/increase, QoQ Growth.
- One-page dashboard with months sorted Jan–Dec on x-axis.

Out of scope

- Changes to source systems or upstream schemas.
- Predictive modeling or ML beyond requested measures.
- Multi-page reports or drill-through pages.

Data sources

- Sales: Folder organized by year files.
- Categories: Excel; Geography: Excel; Product: CSV/Database; SalesRep: Excel; SubCategories: Excel.

Functional requirements

- Resilient ingestion: Removing a year file must not error; adding a new year file must auto-load on refresh into a unified Sales fact.
- Modeling updates: Split Country and City from 'Location'; set geo-friendly data types; normalize Date for formatting/time intelligence.
- Keys: Create unique GeoKey in both Sales and Geography to support relationships.
- Cleansing: Implement a reusable function to remove the 'ID ' prefix in SalesRep and SubCategory IDs.
- Calendar: Use the existing Calendar table in the PBIX for time intelligence.
- Visualization: Assemble a one-page dashboard with diverse visuals; ensure Month sorts Jan–Dec on x-axis.

Non-functional requirements

- Refresh reliability and scalability for new yearly files without manual query edits.
- Correct data categories and types for map visuals and time intelligence.
- Reusable, parameterized cleansing logic to minimize duplication.

Detailed tasks

Task 1.1 – Data ingestion

- Build a query to read and append all files from the Sales yearly folder into a single fact table.
- Configure it to ignore missing files and automatically include new year files at refresh.

Task 2.1 – Transformations

- Split 'Location' into Country and City; set appropriate data categories/types to enable geo maps.
- Normalize Date fields to proper date type/format for sorting and time intelligence.

Task 2.2 – Keys

- Create a unique GeoKey in both Sales and Geography using consistent construction logic (e.g., standardized Country + City).

Task 2.3 – Reusable ID cleansing

- Create a small function that strips the literal prefix 'ID ' from ID columns.
- Invoke this function in SalesRep and SubCategory queries to clean IDs consistently.

Task 3 – DAX calculations

- Total Revenue = Product Retail Price × Units.
- Total Cost = Product Standard Cost × Units.
- Gross Profit = Total Revenue – Total Cost.
- Gross Profit MoM Growth Change%: Month-over-month percentage change of Gross Profit.
- Average Sales per Day: Average of Total Revenue per day, based on days with actual sales.
- Product breakdown: Identify drops/increases by product across periods.
- QoQ Growth: Quarter-over-quarter growth for QBR.

Acceptance criteria

- Ingestion: Refresh succeeds when a year file is removed or a new year file is added; unified Sales fact reflects changes after refresh.
- Modeling & cleansing: Location is split, geo categories set, Date recognized and Month sorting Jan–Dec works via Calendar; reusable ID function applied in SalesRep and SubCategory; GeoKey consistent and unique across joins.
- DAX: Measures for Total Revenue, Total Cost, Gross Profit, MoM %, Avg Sales/Day, Product drop/increase, and QoQ Growth validate on sample checks.
- Reporting: One-page dashboard shows trends, product performance, geography, and time-based growth; Month axis sorted Jan–Dec.

Risks and mitigations

- Inconsistent Location strings may impede GeoKey joins; mitigate with trimming, case normalization, and robust split logic.
- ID formats may vary beyond 'ID '; mitigate with parameterized cleansing and pattern detection.
- Schema drift in new year files; mitigate by aligning schemas and handling optional columns in ingestion.

Deliverables

- Power Query/M ingestion for folder-based Sales consolidation plus reusable ID-cleansing function.
- Star schema model with GeoKey and Calendar relationships.
- DAX measures as specified.
- One-page dashboard PBIX with visuals and correct month sorting.