Full **breakdown and explanation** of the work done, including:

* **Why columns were kept or removed**
* **How dimension tables were structured**
* **What each fact table holds**
* Plus a **diagram suggestion** for your ERD tool (e.g., dbdiagram.io, MySQL Workbench)

**✅ OVERVIEW**

We started with 3 source tables:

* SWD\_Collection\_Report
* RME\_ap\_check\_payments\_Report
* RME\_Projects\_Cost\_Dist\_Line\_Report

These tables had overlapping project and task metadata embedded, so we:

1. **Created a new database** called RME\_Clean
2. **Normalized** the schema into:
   * 2 **Dimension Tables**: Projects, Tasks
   * 3 **Fact Tables**: Transactions, Payments, Collections
3. **Removed all redundant columns**
4. **Transferred the cleaned data**

**🧱 TABLE STRUCTURES (Final Design in RME\_Clean)**

**🔹 Dim\_Project – *Project Metadata***

| **Column** | **Source Table** | **Reason** |
| --- | --- | --- |
| PROJECT\_ID | RME\_Projects\_Cost\_Dist\_Line\_Report | Primary key |
| PROJECT\_NUM | RME\_Projects\_Cost\_Dist\_Line\_Report | External ref. |
| PROJECT\_NAME | RME\_Projects\_Cost\_Dist\_Line\_Report | Display name |
| PROJECT\_TYPE | RME\_Projects\_Cost\_Dist\_Line\_Report | Describes cost control type |
| OWNER | All 3 tables | Centralized ownership |

✅ **Deduplicated by PROJECT\_ID**  
❌ Removed from all fact tables to prevent redundancy.

**🔹 Dim\_Task – *Task Info***

| **Column** | **Source Table** | **Reason** |
| --- | --- | --- |
| TASK\_ID | RME\_Projects\_Cost\_Dist\_Line\_Report | Primary key |
| TASK\_NUM | RME\_Projects\_Cost\_Dist\_Line\_Report | External ref. |
| TASK\_NAME | RME\_Projects\_Cost\_Dist\_Line\_Report | Display task title |

✅ One row per task  
❌ Removed from project transaction facts

**🔸 Fact\_ProjectTransactions – *Main Financial Cost Line Data***

| **Column** | **Source Table** | **Notes** |
| --- | --- | --- |
| EXPENDITURE\_ITEM\_ID | Primary Key | Cost line unique ID |
| TRANSACTION\_SOURCE | Source system like Oracle Payables |  |
| PROJECT\_ID | FK to Dim\_Project |  |
| TASK\_ID | FK to Dim\_Task |  |
| EXPENDITURE\_TYPE | Describes type of transaction |  |
| QUANTITY | Actual cost value |  |
| UOM | Unit of measure (e.g. Currency) |  |
| LINE\_NUM | Line ID |  |
| LINE\_DESC | Line description |  |

❌ Removed: PROJECT\_NAME, PROJECT\_TYPE, TASK\_NAME, OWNER

**🔸 Fact\_Payments – *Invoice Payments & Checks***

| **Column** | **Source Table** | **Notes** |
| --- | --- | --- |
| INVOICE\_ID | RME\_ap\_check\_payments\_Report | Often multiple per invoice |
| INVOICE\_NUM | ID string (e.g. OB\_34) |  |
| PAYMENT\_AMOUNT | Positive or negative (refunds) |  |
| CHECK\_DATE | Actual payment date |  |
| VENDOR\_SITE\_CODE | To identify vendor |  |
| STATUS\_LOOKUP\_CODE | Status of invoice |  |
| CATEGORY | e.g. “RME Zero Invoices” |  |

❌ Removed: PROJECT\_NAME, ORG\_NAME, OWNER

**🔸 Fact\_Collections – *Incoming Receipts (OB, PSP, T&D)***

| **Column** | **Source Table** | **Notes** |
| --- | --- | --- |
| CASH\_RECEIPT\_ID | SWD\_Collection\_Report | Unique key |
| RECEIPT\_NUMBER | (e.g. OB\_0026) |  |
| RECEIPT\_DATE | Date collected |  |
| AMOUNT | Gross value |  |
| FUNC\_AMOUNT | Adjusted financial amount |  |
| APPLIED\_AMOUNT | Amount used |  |
| UNAPPLIED\_AMOUNT | Amount left |  |
| RECEIPT\_METHOD\_ID | e.g. Bank, PSP, OB |  |

❌ Removed: PROJECT\_NAME, OWNER, ORG\_NAME, COMMENTS, OLD\_NO

**🔗 Relationships (ER Diagram Star Schema)**

┌────────────────────┐

│ Dim\_Project │◄────────┐

└────────────────────┘ │

▲ │

│ │

┌───────────────────────────┐ │

│ Fact\_ProjectTransactions │ │

└───────────────────────────┘ │

▲ │

│ │

┌────────────────────┐ │

│ Dim\_Task │◄────────┘

┌────────────────────┐

│ Fact\_Payments │

└────────────────────┘

┌────────────────────┐

│ Fact\_Collections │

└────────────────────┘

**✅ Summary of Improvements**

| **Improvement** | **Description** |
| --- | --- |
| 🔄 Removed redundancy | No more repeated PROJECT\_NAME, OWNER, TASK\_NAME |
| 📐 Normalized schema | Clear separation of dimensions and facts |
| 🎯 Faster reporting | Easier joins & efficient queries |
| 🔐 Data integrity | Enforced FK relationships via PROJECT\_ID and TASK\_ID |
| 💼 Ready for Power BI / Looker | Clean structure for BI tools |

Would you like me to generate SQL views to **join** the fact and dimension tables for easier querying?