Kurzanleitung zur Erstellung einer ETL Stored Procedure für Fact-Tabellen

1. Anforderungsanalyse

- Zielsetzung klären: Welche Daten sollen in die Fact-Tabelle geladen werden?
- Inkrementelles Laden oder vollständiges Laden: Entscheiden, ob nur neue/aktualisierte Daten (inkrementell) oder alle Daten geladen werden sollen.

2. Design der Stored Procedure

- Parameter und Schritte planen:
 - 1. **Startdatum bestimmen**: Definieren Sie das Startdatum, um nur die relevanten Daten auszuwählen.
 - 2. **Daten extrahieren und transformieren**: Wählen und transformieren Sie die Daten, die nach dem Startdatum geändert wurden.
 - 3. **Daten laden**: Fügen Sie die Daten in die Fact-Tabelle ein.
 - 4. **Letztes Lade-Datum aktualisieren**: Speichern Sie das aktuelle Datum als neues "Last Loading Date".

3. Implementierung

Beispielcode:

```
CREATE PROCEDURE dbo.usp_fact_loading
    @start_date DATETIME

AS

BEGIN

-- Aktuelles Lade-Datum speichern

DECLARE @Loading_date DATETIME = GETDATE();

-- Einfügen der Daten in die Fact-Tabelle

INSERT INTO fact_table (column1, column2, column3, ...)

SELECT column1, column2, column3, ...

FROM source_table

WHERE ISNULL(modify_date, create_date) > @start_date;

-- Aktualisierung des letzten Lade-Datums

UPDATE control_table

SET Last_Loading_Date = @Loading_date;

END
```

4. Testen

• **Verhalten prüfen**: Testen Sie die Prozedur mit verschiedenen Startdaten und Szenarien, um sicherzustellen, dass sie korrekt funktioniert.

5. Bereitstellung und Überwachung

- **Deployment**: Stellen Sie die Prozedur in der Produktionsumgebung bereit.
- **Überwachung**: Überwachen Sie die Ausführung, um sicherzustellen, dass die Daten korrekt und effizient geladen werden.

6. Dokumentation

• **Prozessbeschreibung**: Dokumentieren Sie die Funktion und die wichtigen Details der Stored Procedure.

Zusammenfassung

Diese Anleitung beschreibt, wie Sie eine ETL Stored Procedure für das Laden von Fact-Tabellen entwickeln. Die wesentlichen Schritte umfassen das Festlegen eines Startdatums, das Extrahieren und Laden relevanter Daten, das Aktualisieren des letzten Lade-Datums sowie die Überwachung und Dokumentation des gesamten Prozesses.