**Тз**

1.Создать оперативную базу данных в любом из видов баз данных(не важно реляционной или нет)

Варианты:

а.Если делать через MS SQL - можно скачать adventureworks сэмпл данных от microsoft

б.Скачать сэмплы с любых источников, важно что бы было как минимум три таблицы или более (два dimension и один fact)

г.3 таблицы:

1.Таблица продаж - sales : customerId, productId, qty

2.Таблица клиентов - customers: id, name, country

3.Таблица продуктов- products : id, name, groupname

И вставить данные вручную

Это будет оперативной базой данных

2.Создать дополнительные три базы данных на другом инстансе от оперативной (обязательно реляционная mysql, postgresql, etc….):

mrr, stg, dwh - нужно почитать и понять почему мы строим 3 базы и для чего каждая из них

3.Во все таблицы с мерами мы добавляем приставку fact в имени, c измерениями dimension + приставка имени базы данных. Пример: если таблица в базе данных mrr то таблица с продажами будет mrr\_fact\_sales, если таблица с продуктами в stg то stg\_dim\_products. В именах только английский, и все имена одного формата(камел кейс или снейк кейс).Почитать что такое fact и dimension

4.Создать ETL(airflow, nifi, spark, SSIS либо любой другой ETL) с переходами данными из оперативной базы в mrr -> stg -> dwh

Из оперативной базы данных в mrr брать данные с помощью high water mark(дельта).Для этого создать таблицу high\_water\_mark в который будет последний день или апдейт каждой таблицы. В mrr вытягиваем в параметр время в соответствии с таблицей источникам и в dwh обновляем high\_water\_mark последним значением которое есть в таблице.

5.В каждом пакете/даге/процессе сделать систему логов которые будут писаться в созданную для этого таблицу, время исполнения пакета + если есть ошибка(это делаеться в event handler)

6.Создать процедуру и использовать там cursor, try и catch(при ошибки будет писаться лог в созданную для этого таблицу), сделать какую нибудь функцию.

Все процедуры и функции сохранить в базе данных dwh.

7.Сделать дашборд и модель данных на ваше усмотрение в Power BI из данных в dwh.Обязательно с бизнес логикой(что то простое на ваше усмотрение)

8.Создать скрипт который будет делать backup для трех баз данных(mrr, stg, dwh).