**Документация, описывающая процедуру ETL и все особенности его загрузки**

Загрузка осуществляется посредством job под названием avia.kjb. Сначала происходит загрузка таблиц измерений, последней- загрузка данных в таблицу фактов.

**1. Таблица измерений «самолеты»:**

Процесс начинается с подключения к таблице bookings.aircrafts. Перед загрузкой в таблицу измерений происходят проверки качества данных:

-нет пропусков в дальности полета;

-нет пропусков в названии кода;

-нет дублирования одного и того же кода.

**2. Таблица измерений «пассажиры»:**

Процесс начинается с подключения к таблице bookings.tickets. Выбираются уникальные записи и добавляется уникальный id, проводятся проверки качества: заполнение id, наличие имени и фамилии, заполнены адреса и тд.

**3. Таблица измерений «тарифы:**

Процесс начинается с подключения к таблице bookings.ticket\_flights, проверка качества данных: стоимость не должна быть нулевой, тариф должен быть заполнен. Добавляется id и данные загружаются в таблицу dim.tariff.

**4. Таблица измерений «аэропорты»:**

Процесс начинается с подключения к таблице bookings.airports. Добавляется поле, которое считает количество букв в коде аэропорта, проверки: в коде только три буквы, код должен быть заполнен, город должен быть заполнен. После проверки данные выгружаются в таблицу dim.airports.

**5. Таблица измерений «календарь»:** заполняется при помощи запроса sql.

**6. Таблица фактов:**

Процесс начинается с подключения к таблице bookings.flights. Далее присоединяются таблицы tickets\_flights и tickets. Далее фильтры: только проданные билеты и только прилетевшие самолеты.Вычисляется разница между фактическим передвижением самолетов и расписанием. Проверки качества данных: статус полета должен быть заполнен, время отлета должно быть раньше времени прибытия, пассажир должен присутствовать. Далее замена id из таблиц на технический id из словарей. Выбирается перечень полей и осуществляется загрузка.