

## Домашнее задание 4

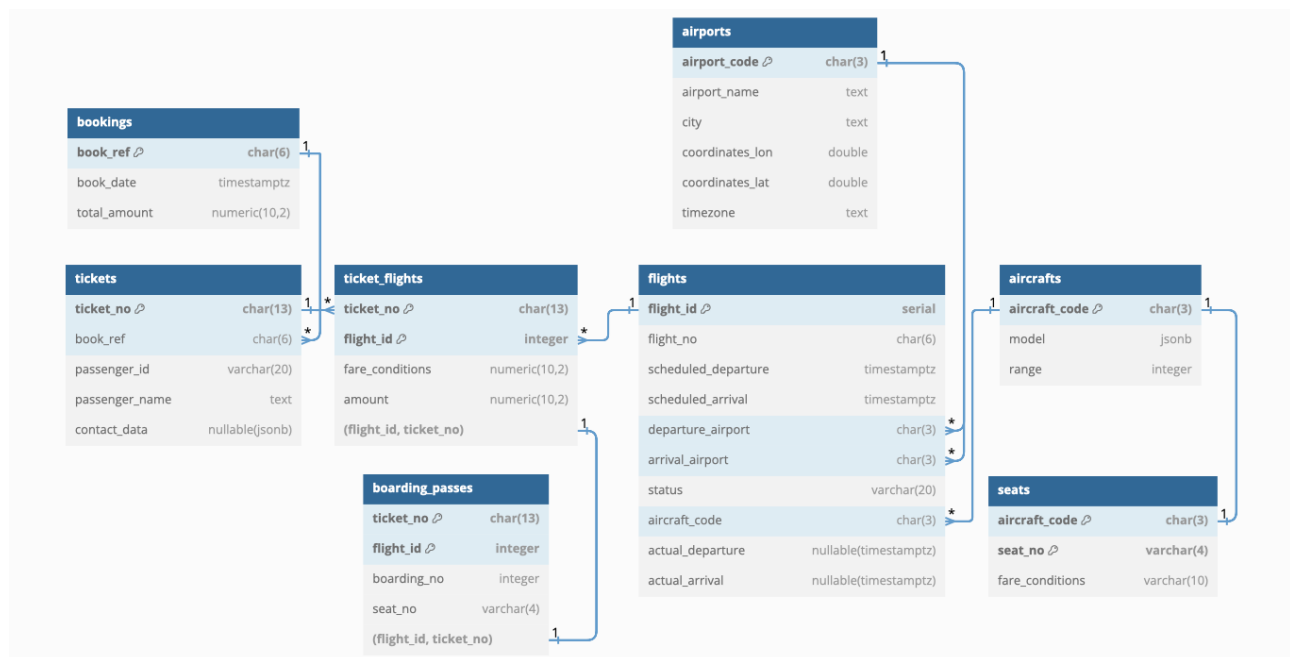
### Important notice

- 1) **К ДЗ №4 можно приступить, не сдав ДЗ №3 полностью.** В таком случае вы должны брать данные напрямую с продовой БД.
- 2) **Это задание выполняется и оценивается индивидуально или в группах до 4 человек!** Если вы используете код из открытых источников (репозитории ваших одноклассников таковыми не считаются) - пожалуйста, указывайте ссылки на них (можно в readme вести лог всех источников, откуда вы берете код). Находить готовые рецепты в интернете - хорошо, списывать - плохо.
- 3) **При обнаружении списывания (одинаковый код в двух репозиториях без указания внешнего источника) оценка будет выставляться студенту (группе), чей коммит с решением был первый.** Остальным - 0 баллов и докладная в УО.
- 4) Чтобы исключить возможность списывания, **рекомендуется сделать ваш репозиторий с домашним заданием приватным.**

### Формулировка

У нас есть поисковик дешевых авиабилетов.

БД системы, которая содержит информацию о забронированных с его помощью билетах, выглядит следующим образом:



Это мы уже с вами видели в ДЗ №1 и №2.

Сейчас мы имеем:

- \* master-хост с БД
- \* async-replica, на которую копируются данные для аналитической нагрузки
- \* пайплайн доставки данных в аналитическую БД
- \* детальный слой DWH в аналитической БД
- \* Витрины с frequent flyer'ами и загруженностью аэропортов

### Задача:

1. **Поднять Open Source BI-инструмент по вашему выбору в docker-compose.** На семинарах мы с вами рассматривали Metabase, Apache Superset и Grafana, но если привычнее поднять и использовать что-то другое - пожалуйста.
2. **Создать аналитический дэшборд по аэропортам**
  1. Цифра - всего активных аэропортов (at least 1 arrival/departure) за последние 30 дней

2. График - всего активных аэропортов (at least 1 arrival/departure) за последние 30 дней в разбивке по дням
3. Pie chart - доля аэропортов в совокупном пассажиропотоке (arrivals+departure)
4. Pie chart - доля аэропортов в общем кол-ве рейсов (arrivals+departure)
3. Создать **аналитический дэшборд по пассажирам**
  1. Цифра - всего уникальных пассажиров за последние 30 дней
  2. Цифра - средний чек за последние 30 дней
  3. Цифра - среднее кол-во перелетов на пассажира за последние 30 дней
  4. График - динамика кол-ва уникальных пассажиров по дням
  5. График - динамика совокупной выручки по дням
  6. Pie chart - доля разных групп пассажиров (см. ДЗ №3) в GMV за последние 30 дней
4. Допускаются отступления в реализации (в зависимости от особенностей вышей BI-системы), но смысл всех перечисленных выше элементов должен быть сохранен.

#### **Бонусные задания:**

1. Бонус за креативность. Красивые визуализации, использование параметризации и т.д.

#### **Сроки**

1. Дедлайн на 100% - 2 недели - **15.12.2024 23:59:59 включительно**
2. Дедлайн на 75% - до конца курса - **18.12.2024 23:59:59 включительно**
3. После **18.12.2024** работы **не принимаются**

#### **Как сдавать ДЗ?**

- Так как не все BI-системы позволяют сохранить дэшборд/подключения, вам нужно **записать скринкаст с демонстрацией работы ваших дэшбордов и загрузить его на любой файлообменник, либо отправить мне его сообщением в Telegram**. На вашей системе, скорее всего, есть нативные инструменты для записи видео с экрана (Windows: Win+Shift+S, Mac: cmd+Shift+5), если таких нет - можно воспользоваться записью экрана во время звонка Teams / Zoom / Телемост / Google Meet.
- Готовое ДЗ (его часть с тем, чтобы поднять BI'ку у себя) загружается на GitHub (приватный репозиторий, для проверки предоставить доступ @mgcrp);
- Домашнее задание №4 можно продолжать делать в том же репозитории, что и домашнее задание №1-3;
- К репозиторию должен быть приложен README с описанием того, что вы сделали и как это запустить;
- Задание сдается в форму: <https://forms.gle/xkWKm6ivbRW1V9pF8>

#### **Критерии оценки**

Балл	Критерий
4	Поднята BI-система
7	Реализован один из двух дэшбордов
10	Реализованы оба дэшборда
До 2-х баллов	Бонус за креативность

Максимальный балл за ДЗ - 12

#### **Как это будет проверяться?**

- 1) **Запуск системы по инструкции** из вашего README (проверка, что поднимается BI)
- 2) **Проверка дэшбордов** по записи экрана