3 января:

Загрузка программы Pentaho DI

Для установки на Mac необходимо через terminal запустить команду

sudo xattr -dr com.apple.quarantine /Applications/data-integration/Data\ Integration.app

После запускаем Data Integration

Загрузка Tableau Desktop и Tableau Pre, активация по студенческой лицензии

Загрузка и установка MySQL

4 января:

Представление схемы BI architecture solutions (схема pdf)

Представление нормализованной модели данных (схема pdf)

Описание таблиц, полей и их взаимосвязей (Excel)

9 января:

Обсуждение открытых вопросов с Ринатом по технической реализации решения:

рассмотрен общий процесс реализации решения лаб. работы (от стадии подготовки данных к загрузки в Tableau)

разобран процесс загрузки и выгрузки данных при помощи Pentaho в БД

выбрана GitHub-платформа для хостинга лабораторной работы

11 января:

Установка MySql Workbench

Создание БД Lab1 с таблицами stg-слоя:

`stg\_customers`

`stg\_purchase\_prices`

`stg\_salary`

`stg\_sales`

(подробное описание таблиц в файле Tables)

Установка соединения БД с Pentaho. Возникла ошибка в подключении, необходимо было скачать .jar (коммутатор) и добавить в папку lib Pehtaho (библиотека MySql), но новая версия MySQL не завязывается с Pentaho, пришлось MySQL сервер скачивать более поздней версией и для нее соответсвующий коммутатор. Соединение было установлено.

Создана тест-версия загрузки данных в БД через Pentaho.

a)      Создание общей базы данных на основе модели данных, включающая необходимые поля, таблицы и их взаимосвязь.

13 января:

В Pentaho были заведены jobs:

Job\_Lab1.kjb - основной, включающий stg\_job.kjb -стеджинг слой и dds\_job.kjb -ддс слой (нормализованная модель)

В stg\_job.kjb были заведены следующие трансформации:

stg\_prices.ktr - загрузка данных листа закупочные цены

stg\_sales.ktr - загрузка данных листа продажи

stg\_cust.ktr - загрузка данных листа покупатели

stg\_salary.ktr - загрузка данных листа зарплаты

15 января:

Переход на БД PostgreSQL и установка DBeaver

Создание БД с таблицами, раннее созданными и загрузка через Pentaho stg слоя в БД (кодировка верная)

18-20 января:

Создание jobs в Pentaho с dds-слоями

Вопросы Ринату по подготовке dds-слоев

21 января:

Таблица Tax не требуется, из БД удалена. Вместо этого при рассчете затрат налоговая ставка = 6%

Созданы dds-слои

Обновлена архитектура данных

b)      Описание процесса загрузки данных в Табло

Подключение БД к Tableau:

Connect to a server -> PostgreSQL ввод данных

server localhost

Database postgres

Port 5432

Username postgres

Password root

На вкладке Data Source настраиваем соединение с БД через PostgreSQL.

У меня три Data Source.

1.Purchase prices

Устанавливаю связь для таблицы dds\_purchase prices left join с dds\_measure, dds\_product, dds\_product\_group

2.Sales

Устанавливаю связь для таблицы dds\_sales left join с dds\_people, dds\_product, dds\_product\_group

3.Salary

Устанавливаю связь для таблицы dds\_salary left join с dds\_people

В итоге получаю три источника данных по разным категориям. Переходя на Sheet их можно увидеть.

следующий функционал по дэшбордам:

a)      общую сумму продаж +

b)      продажи по каждому товару +

c)      продажи по группам товаров +

d)     топовых клиентов (частые клиенты + и клиенты с высоким средним чеком+)

e)      Затраты на покупку товаров  (определение себестоимости товаров)+

f)       Затраты на фонд оплаты труда (зарплаты) по каждому работнику +

g)      Затраты на выплату налогов (НДФЛ для зп+, НДС+, налоги на ведение предпринимательской деятельности) ставка везде 10%

Доп. затраты, не описанные выше, не учитываются.

Отчет формируется за неделю в течение 10 с.

Наценка = 50% от себестоимости товара.

Прибыль рассчитывается программно.