**Федеральное государственное автономное образовательное учреждение высшего образования**

**«Национальный исследовательский университет «Высшая школа экономики» Факультет компьютерных наук**

**О Т Ч Е Т**

**по дисциплине**

**«Теория Баз Данных»**

**Выполнил студент гр.**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**(ФИО)**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

***(подпись)***

**Проверил:**

*(должность, ФИО руководителя практики)*

*00.00.2020*

*(дата)*

**2020 год**

Содержание

Содержание 2

1. Описание предметной области 3
   1. Цель 3
   2. Внешние данные 3
   3. Основные сценарии использования 3
2. Концептуальная модель 3
   1. Диаграмма «Сущность-связь» 3
   2. Описание сущностей им связей 4
3. Инфологическая модель 5
   1. Диаграмма «Таблица-связь» 5
   2. Словарь данных 5
4. Даталогическая модель 5
   1. Используемая СУБД и диалект SQL 5
   2. Нормализация данных и организация представлений 6
   3. Хранимые процедуры и триггеры 6
   4. Описание механизмов обеспечения целостности данных 6
   5. DDL-скрипты 6
5. Клиентское приложения 6
   1. Архитектура 6
   2. Сценарии использования 6
   3. Организация доступа к данным 6
   4. Интерфейс с пользователем 6
   5. Отчёты 6
6. Заключение 6
   1. Объёмные характеристики разработки 6
   2. Авторский вклад и комментарии по выполнению проекта 6
7. Источники 6

# Описание предметной области

## Цель

Оптимизация продаж в сети магазинов на Amazon

## Внешние данные

Датасет с товарами с сайта Amazon. Источник: <https://data.world/promptcloud/fashion-products-on-amazon-com>

Содержит данные о товарах, их названии, производители, детальную информацию об отзывах (подробнее напишу). В этом основная суть!

## Основные сценарии использования

На основе отзывов можно определить, какие товары данного производителя популярнее, это полезно производителю. Ему же полезно узнать, на какие отзывы он еще не ответил, это тоже можно сделать. Amazon будет полезно узнать средний рейтинг товаров производителя, чтобы знать, к примеру, кого лучше прорекламировать

# Концептуальная модель

## Диаграмма «Сущность-связь»

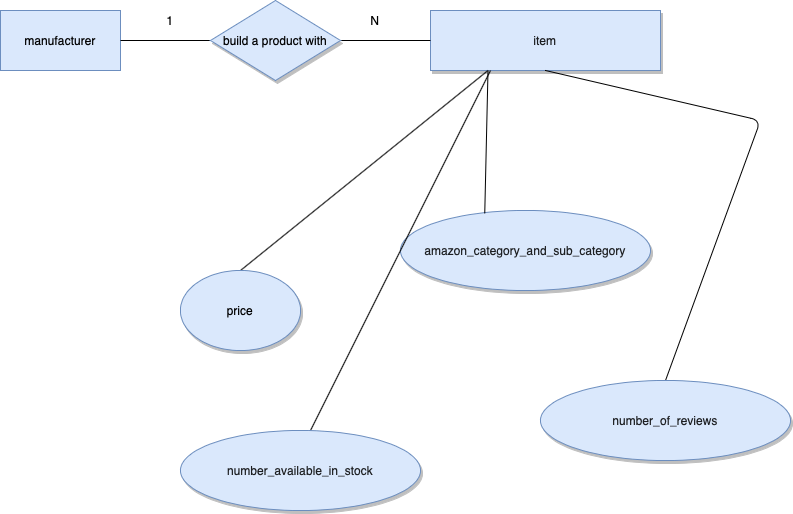
Нотация Ченя, [draw.io](http://draw.io)

Проблемы: при таком описании предметной области (есть товары, есть их производители, есть атрибуты у товаров) у производителя нет никаких атрибутов, есть только вычислимые с помощью атрибутов товаров (к примеру, количество неосвещенных отзывов по всем товарам производителя, средний рейтинг итд). Что с этим делать? Если у какого-то товара 0 отзывов, производитель может понять, что его надо разрекламировать.

Рисунок 2.1. Диаграмма в нотации !!!

## Описание сущностей им связей

Товары и производители. Предметная область: продавцы на Amazon продают товары, потребители пишут отзывы и ставят рейтинги. Словарь сущностей: manufacturer, amazon\_category\_and\_sub\_category. Добавить еще? Описания потом, ибо очевидно.

ВОПРОС: сделать нормальный список items вместо ссылок в customers also bought и сделать связь собственно с тем же названием с покупателями? Также: покупатели – отзывы, сущность покупатели может быть полезной но я не знаю насколько.

ЕЩЕ: замысел следующий. Про производителей ничего неизвестно, про продукты и покупателей — да. Суть в том, чтобы на основании последних двух составлять аналитику для первых, то есть, функционал заключается в вычислении разных метрик для производителей. Нужно ли эти метрики добавлять в список атрибутов тоже заранее?

Нужно еще 2–3 сущности какие-то добавить? Может, категорию и субкатегорию тоже сделать сущностью со связями: товар имеет субкатегорию которая является частью категории?

В предметной области должно быть выделено не менее **5 сущностей и 6 связей**.

1. Хотя бы одна связь должна быть типа «**многие-ко-многим**». При коллективной работе желательно увеличить число сущностей до **7-9**.

# Инфологическая модель

## Диаграмма «Таблица-связь»

С указанием нотации!

A picture containing table, sitting, glass, front

Description automatically generated

Рисунок 3.1. Диаграмма в нотации !!!

## Словарь данных

# Даталогическая модель

## Используемая СУБД и диалект *SQL*

!!!

Использован диалект *Transact-SQL* (Microsoft).

!!!

## Нормализация данных и организация представлений

## Хранимые процедуры и триггеры

## Описание механизмов обеспечения целостности данных

## *DDL*-скрипты

# Клиентское приложения

## Архитектура

## Сценарии использования

## Организация доступа к данным

## Интерфейс с пользователем

## Отчёты

# Заключение

Код проекта находится в (2021)

## Объёмные характеристики разработки

## Авторский вклад и комментарии по выполнению проекта

# Источники

**Database Design and Relational Theory. Normal Forms and All That Jazz** [Book] / auth. Date C.J.. - [б.м.] : O’Reilly Media, 2012. - p. 260.

**SQL Server technical documentation** [Online] / auth. Microsoft. - 10 10, 2020. - https://docs.microsoft.com/en-us/sql/sql-server/.

**Код проекта** [В Интернете]. - 20 09 2021 г.. - http://github.com.