



Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Московский государственный технический университет
имени Н.Э. Баумана
(национальный исследовательский университет)»
(МГТУ им. Н.Э. Баумана)

ФАКУЛЬТЕТ _____ ИНФОРМАТИКА И СИСТЕМЫ УПРАВЛЕНИЯ _____

КАФЕДРА _____ СИСТЕМЫ ОБРАБОТКИ ИНФОРМАЦИИ И УПРАВЛЕНИЯ (ИУ5) _____

РАСЧЕТНО-ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

к домашнему заданию

по дисциплине: Оперативный анализ данных/Средства визуализации данных

на тему: Разработка мобильного приложения

Студент _____
(Группа)

ИУ5-53Б
(Подпись, дата)

Назаров М.М.
(И.О.Фамилия)

Руководитель

(Подпись, дата)

Маслеников К.Ю.
(И.О.Фамилия)

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Московский государственный технический университет имени Н.Э. Баумана
(национальный исследовательский университет)»
(МГТУ им. Н.Э. Баумана)

УТВЕРЖДАЮ

Руководитель секции ИУ5
(Индекс)

« » 20 г.
(И.О.Фамилия)

З А Д А Н И Е
на выполнение домашнего задания

по дисциплине оперативный анализ данных
по теме разработка мобильного приложения

Студент группы ИУ5-53Б

Назаров Максим Михайлович
(Фамилия, имя, отчество)

График выполнения домашнего задания: 25% к нед., 50% к нед., 75% к нед., 100% к нед.

Техническое задание _____

Оформление научно-исследовательской работы:

Расчетно-пояснительная записка на листах формата А4.
Перечень графического (иллюстративного) материала (чертежи, плакаты, слайды и т.п.)

Дата выдачи задания « » 20 г.

Руководитель домашнего задания

(Подпись, дата) (И.О.Фамилия)

Студент

(Подпись, дата) (И.О.Фамилия)

Оглавление

Введение.....	4
Описание ПО.....	4
Описание бизнес-процессов в нотации BPMN.....	4
Инфологическая модель БД	5
Хранилище данных	6
Визуализация данных.....	7
Выбор СУБД.....	11
План и схема архивации, моделирование восстановления данных.....	11
Заключение.....	13
Литература	13

Введение

При выполнении домашнего задания были поставлены следующие цели:

- получение навыков составления BPMN диаграмм;
- освоение программы Tableau Desktop;
- получение навыков грамотной и информативной визуализации больших объемов данных.

В результате выполнения домашнего задания должно быть создано приложение с расширением *.twb, которое будет содержать необходимые данные и их визуализацию с помощью инструментов Tableau Sense.

Описание ПО

Предметная область – мобильные приложения Google Play Market. Мобильное приложение — программное обеспечение, предназначенное для работы на смартфонах, планшетах и других мобильных устройствах, разработанное для конкретной платформы (iOS, Android, Windows Phone и т. д.). Многие мобильные приложения предустановлены на самом устройстве или могут быть загружены на него из онлайн-магазинов приложений, таких как App Store, Google Play, и других, бесплатно или за плату. Их мы и рассмотрим.

Приложения, предоставляемые магазинами, разделяются на категории и возрастной рейтинг. Задача, стоящая перед аналитиками компаний — верно улавливать, какая категория наиболее выгодна для разработки в данный момент.

Для этой цели полезно использовать результаты аналитики магазина приложений, который отслеживает все популярные категории и возрастные рейтинги конкретных типов приложений.

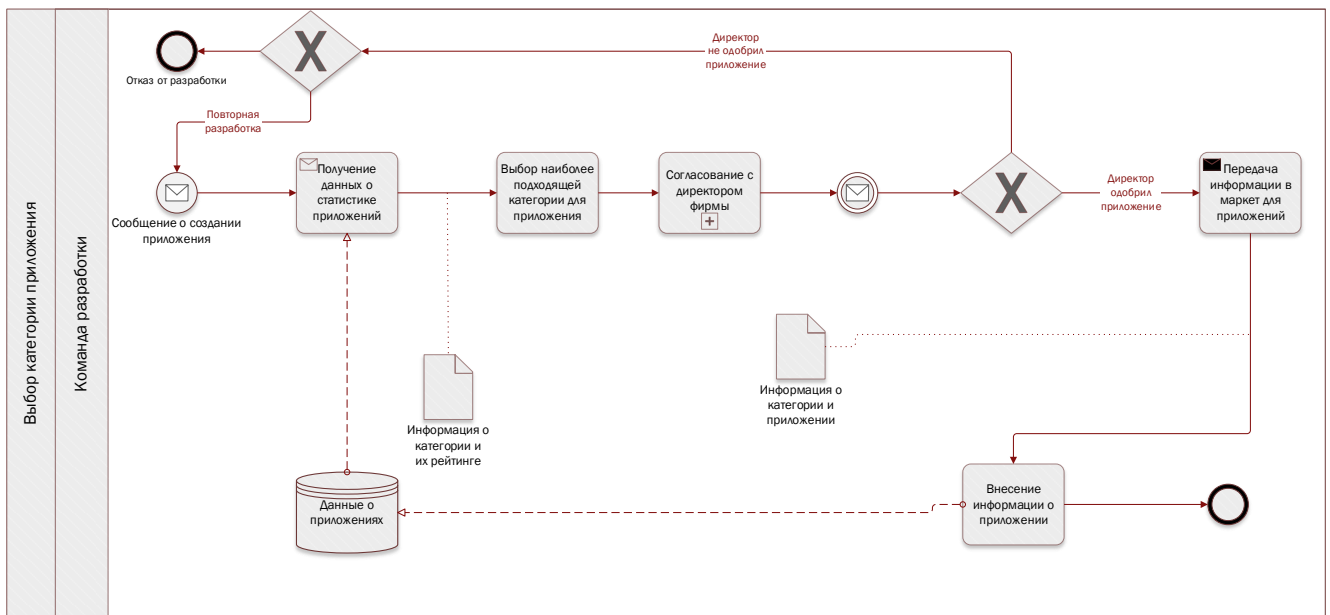
При выборе жанра приложения одним из признаков будет наличие или отсутствие популярности у приложений такой категории. К примеру, если категория Puzzle хорошо сейчас скачивается, то логично начать разработку приложения этой категории.

Описание бизнес-процессов в нотации BPMN

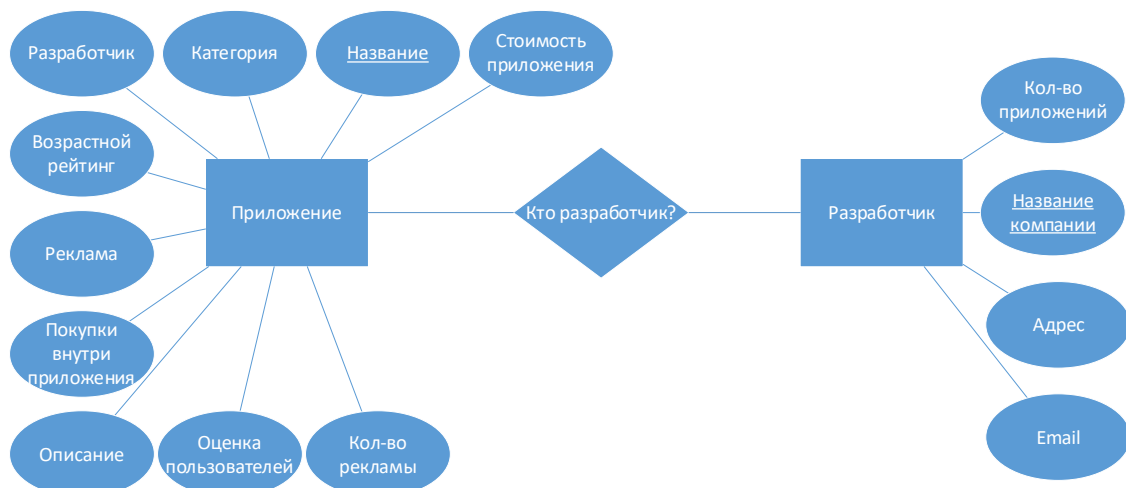
Процесс выбор категории для приложения:

После того как компания решила создать свое приложение они должны выбрать для него категорию. Сначала они запрашивают сведения обо всех приложениях после чего начинают выбирать наиболее подходящий вариант.

Затем аналитики согласуют свое вариант с директором и отправляют информацию о своем приложении в базу данных о мобильных приложениях.



Инфологическая модель БД



Атрибуты:

1. Название – char(30)
2. Название компании - char(30)
3. Категория – char(20)
4. Стоимость приложения - integer
5. Разработчик - char(30)
6. Возрастной рейтинг - integer
7. Реклама - bit
8. Покупки внутри приложения -bit
9. Описание – char(60)
- 10.Оценки пользователей - integer
- 11.Кол-во рекламы - integer
- 12.Кол-во приложений - integer
- 13.Адрес – char(40)
- 14.Email – char(40)

Сущности:

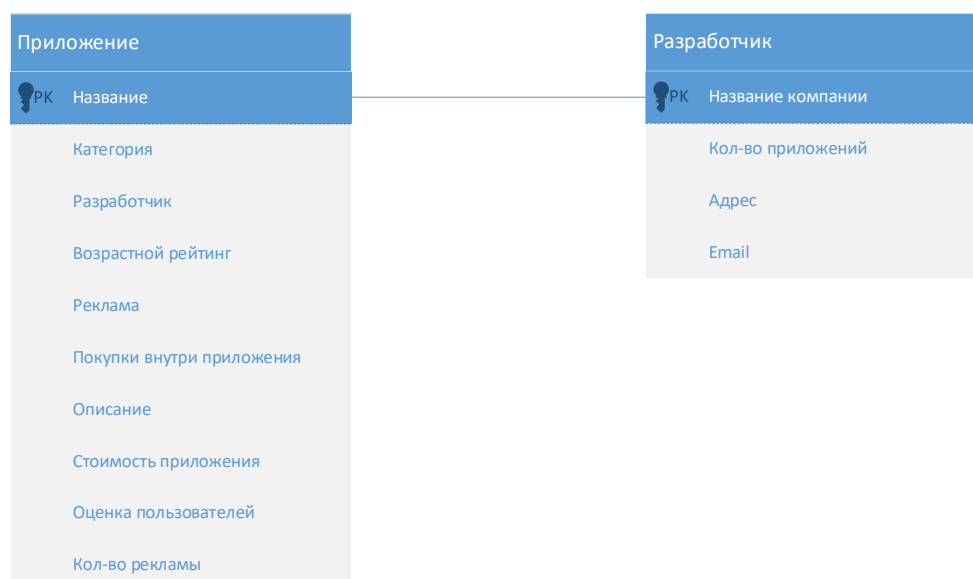
1. Приложение(Название, категория, стоимость приложения, разработчик, возрастной рейтинг, реклама, покупки внутри приложения, описание, оценка пользователей, кол-во рекламы)
2. Разработчик(Название компании, кол-во приложений, адрес, Email)

Связь между сущностями:

Кто разработчик? (Разработчик - Приложение), тип 1-М от Разработчик к
Название компании

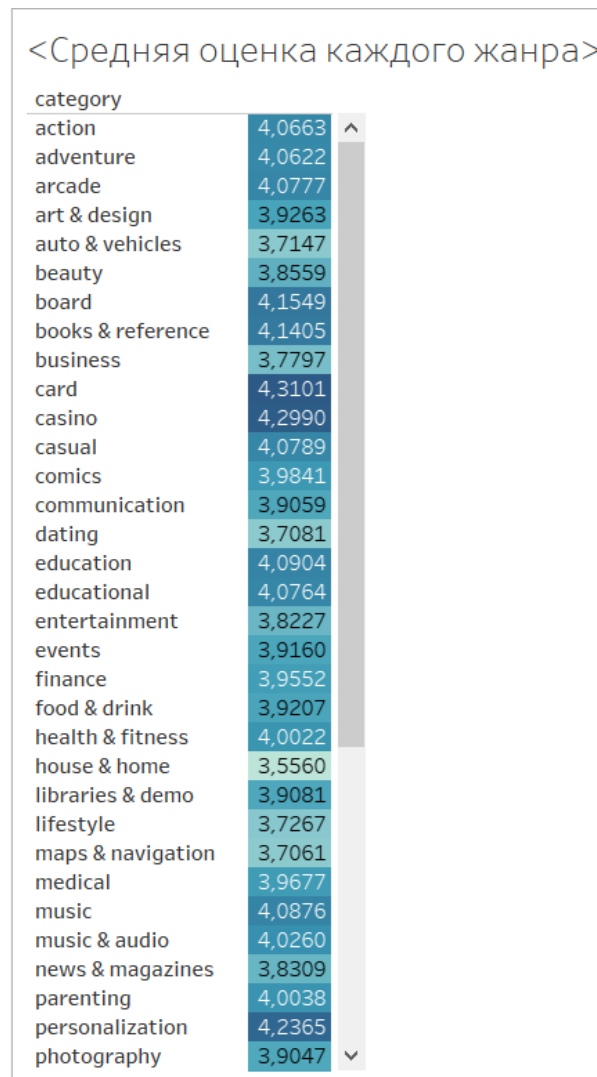
Хранилище данных

Схема хранилища данных



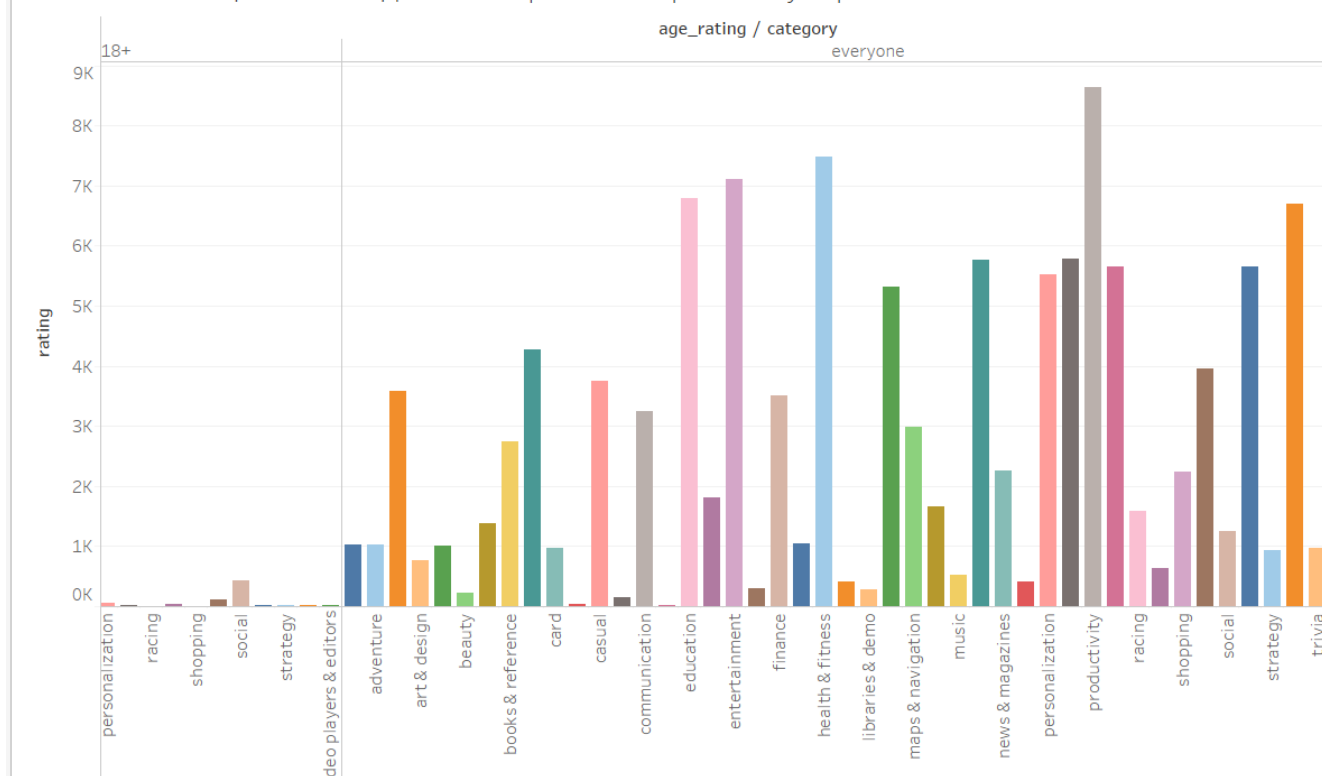
Визуализация данных

Для выбора категории для приложения необходимо знать какие категории наиболее востребованы. Для этого мы рассмотрим график средней оценки пользователя для каждого жанра. Это позволит определить к каким категориям пользователи относятся более лояльно и ставят выше оценки.



Так же важно учитывать количество конкурентов и какие жанры чаще всего делают. Так как спрос рождает предложения то график покажет на что людей больше всего интересует. График показывает возрастной рейтинг и какая категория популярней всего.

<Количество оценок каждой категории по возрастному ограничению>



Круговые диаграммы показывают распределение приложений на возрастные рейтинги. Показывая какую аудиторию большая часть разработчиков пытается охватить.

<Количество загрузок относительно возрастного рейтинга>

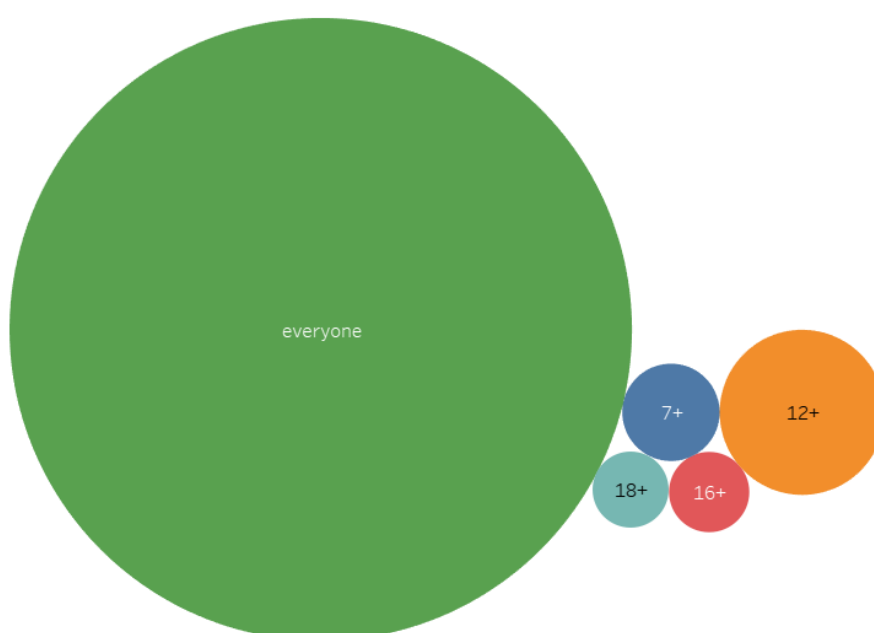


Диаграмма количества рекламы в приложениях к категориям показывает в каких категориях и как сильно монетизируют свои приложения другие разработчики. Это диаграмма позволяет понять сколько рекламы можно внедрить в приложение для его монетизации.

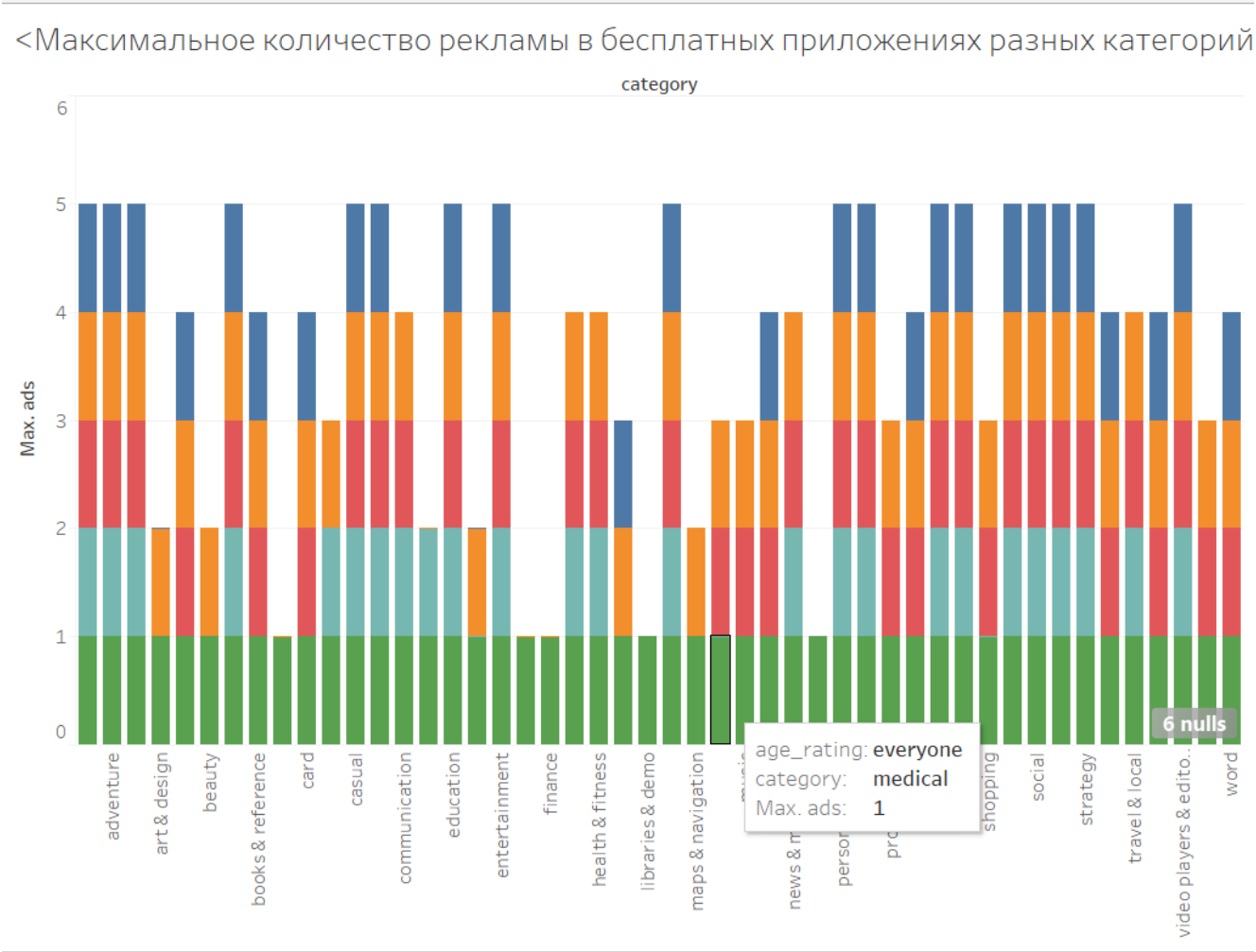
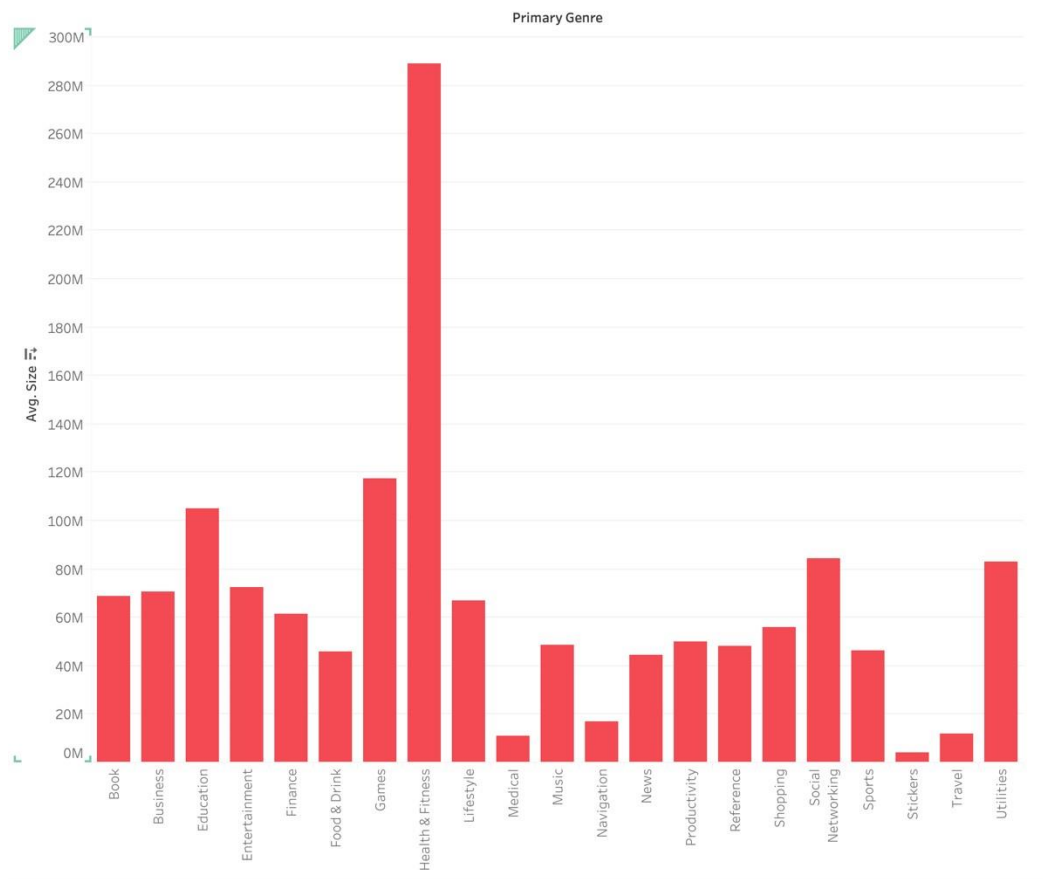
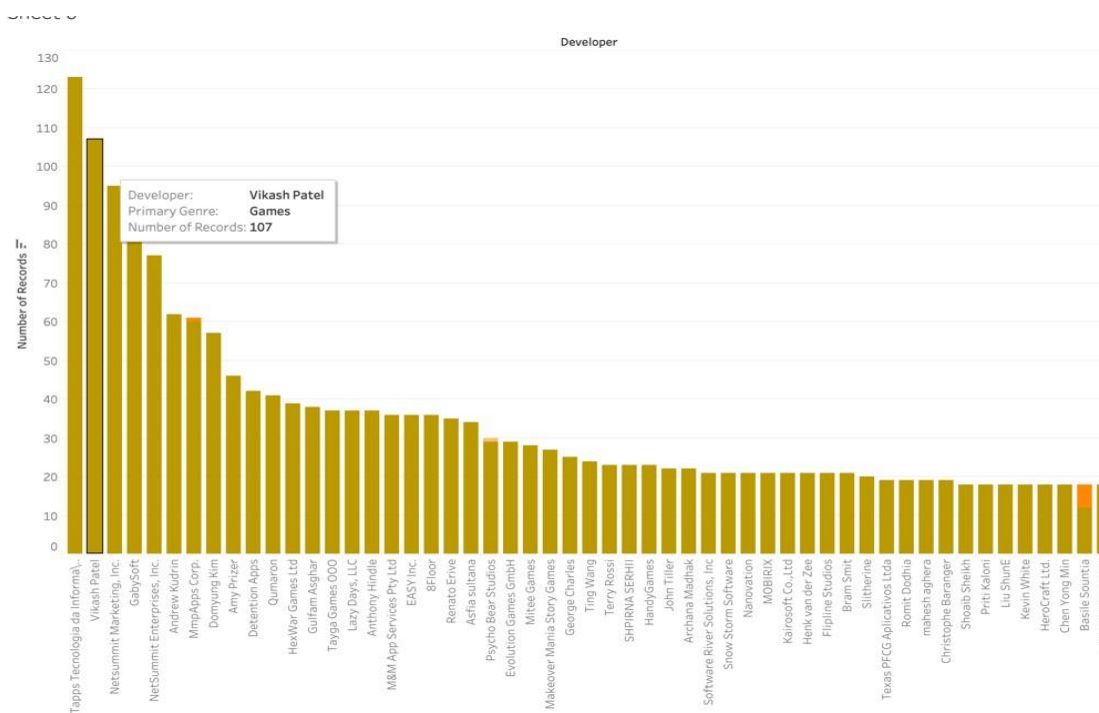


Диаграмма среднего размера приложения к категории показывает для каких жанров требуется больше памяти.



Рассматривая столбчатую диаграмму можно узнать на какие категории больше всего ориентированы разработчики. Это позволяет увидеть что кампании выпускают большое количество приложений одинаково жанра что дает понять что этот жанр является прибыльным.



Выбор СУБД

В качестве СУБД был выбран MS SQL Server по следующим критериям:

- 1) Предусмотренные типы данных. MS SQL Server поддерживает такие категории данных, как точные числа, приближенные числа, дата и время, символьные строки, двоичные данные и проч.
- 2) Реализация языка запросов. Язык SQL является одним из самых распространённых языков запросов.
- 3) Мобильность. MS SQL Server не зависим от среды, в которой он работает.
- 4) Масштабируемость. MS SQL Server может соответствовать росту информационной системы, причем рост может проявляться в увеличении числа пользователей, объема хранимых данных и объеме обрабатываемой информации.
- 5) Многоязыковая поддержка. MS SQL Server поддерживает большое количество национальных языков, что расширяет область применения системы и приложений, построенных на ее основе.
- 6) Документация. MS SQL Server имеет полную и подробную документацию.
- 7) Стоимость. Возможность бесплатного использования MS SQL Server в учебных целях.

План и схема архивации, моделирование восстановления данных

На рисунке 1 представлена схема архивации данных.

Раз в месяц предусматривается полное резервное копирование. Делать это чаще не имеет смысла, т.к. вероятность аварийной ситуации и утери всех данных невелика. Кроме того, раз в сутки предусматривается дифференциальное резервное копирование. Этот период времени оптимален для того, чтобы защитить данные от вирусных программ или незначительных сбоев, не требующих восстановления БД с нуля. При дифференциальном резервном копировании происходит сохранение только тех файлов, которые были подвергнуты изменению с момента последнего полного копирования. Благодаря этому данный вид копирования можно применять относительно часто.

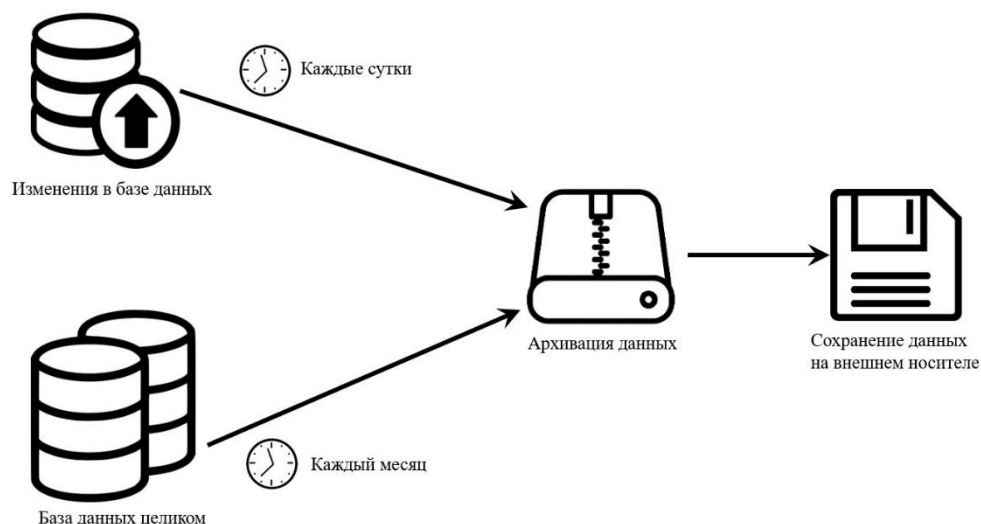


Рисунок 1 – Схема архивации данных

На рисунке 2 представлен план архивации данных за один месяц (ноябрь 2020). Все ежедневные изменения сохраняются на внешнем носителе после завершения рабочего дня.

Ноябрь 2020						
понедельник	вторник	среда	четверг	пятница	суббота	воскресенье
октябрь 26	27	28	29	30	31	ноября 1
Сохранение изменений в базе данных за день	Сохранение изменений в базе данных за день	Сохранение изменений в базе данных за день	Сохранение изменений в базе данных за день	Сохранение изменений в базе данных за день	Полное резервное копирование базы данных	Сохранение изменений в базе данных за день
2	3	4	5	6	7	8
Сохранение изменений в базе данных за день	Сохранение изменений в базе данных за день	Сохранение изменений в базе данных за день	Сохранение изменений в базе данных за день	Сохранение изменений в базе данных за день	Сохранение изменений в базе данных за день	Сохранение изменений в базе данных за день
9	10	11	12	13	14	15
Сохранение изменений в базе данных за день	Сохранение изменений в базе данных за день	Сохранение изменений в базе данных за день	Сохранение изменений в базе данных за день	Сохранение изменений в базе данных за день	Сохранение изменений в базе данных за день	Сохранение изменений в базе данных за день
16	17	18	19	20	21	22
Сохранение изменений в базе данных за день	Сохранение изменений в базе данных за день	Сохранение изменений в базе данных за день	Сохранение изменений в базе данных за день	Сохранение изменений в базе данных за день	Сохранение изменений в базе данных за день	Сохранение изменений в базе данных за день
23	24	25	26	27	28	29
Сохранение изменений в базе данных за день	Сохранение изменений в базе данных за день	Сохранение изменений в базе данных за день	Сохранение изменений в базе данных за день	Сохранение изменений в базе данных за день	Сохранение изменений в базе данных за день	Сохранение изменений в базе данных за день
30						
Полное резервное копирование базы данных						

Рисунок 2 – План архивации данных

На рисунке 3 представлена схема восстановления данных.

При незначительных сбоях или заражении вирусом может быть достаточно восстановить данные с момента последнего дифференциального резервного копирования. Но при утрате всех данных потребуется полное восстановление базы данных с нуля.

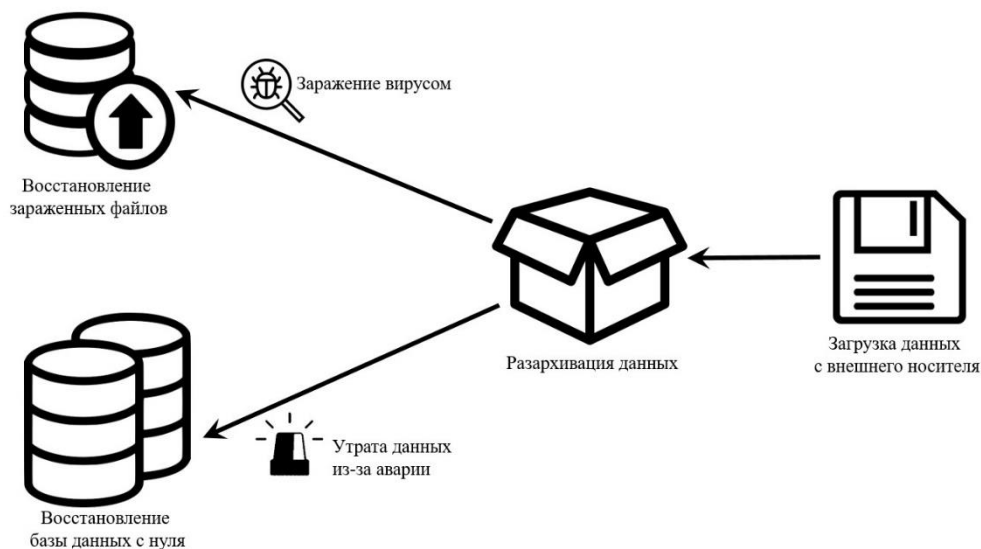


Рисунок 3 – Схема восстановления данных

Заключение

В процессе выполнения домашнего задания были достигнуты поставленные цели:

- были получены навыки составления BPMN диаграмм;
- была освоена программа Tableau Desktop;
- были получены навыки грамотной и информативной визуализации больших объемов данных.

Разработанное приложение содержит необходимые данные и их визуализацию, облегчающую принятие решений. Были использованы такие инструменты визуализации Tableau Desktop, как таблица, сводная таблица, столбчатая диаграмма, круговая диаграмма.

Литература

1. Н. Б. Паклин, В.И. Орешков. Бизнес аналитика, от данных к знаниям СПб.: Питер 2013г.

2. Tutorial: Get Started with Tableau Desktop Version: 2020.3

<https://help.tableau.com/current/guides/get-started-tutorial/en-us/get-started-tutorial-home.htm>