

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ
ФЕДЕРАЦИИ ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
Национальный исследовательский ядерный университет «МИФИ»



**Институт
интеллектуальных кибернетических систем**

Кафедра кибернетики (№ 22)

**Отчёт о работе по курсу
«Основы автоматизированных информационных
технологий»**

Вариант «Магазин компьютерных игр Steam»

Выполнил	Баштанов П.И.
Группа	Б21-504
Вариант	Магазин компьютерных игр, Steam.
Преподаватель	
Проверяющий	
Оценка	

1. Описание предметной области	3
1.1. Формулировка задания	3
1.2. Конкретизация предметной области	3
1.3. Пользователи системы	3
1.4. Сроки хранения информации	3
1.5. События, изменяющие состояние Базы Данных	4
1.6. Основные запросы к Базе Данных (на естественном языке)	4
2. Концептуально-информационная модель предметной области	5
2.1. ER-диаграмма модели	5
2.2. Оценка мощностных характеристик сущностей и связей	6
3. Концептуальное проектирование	8
3.1. Принятые проектные соглашения	8
3.2. Используемые в системе кодификаторы	8
3.3. Концептуальная модель Базы Данных	8
4. Логическое проектирование	8
4.1. ER-диаграмма Базы Данных	8
4.2. Схемы отношений Базы Данных	8
4.3. Схема реляционной Базы Данных	8
4.4. Схемы основных запросов на реляционной алгебре	8
5. Физическое проектирование	8
5.1. Создание Базы Данных	8
5.2. Создание Таблиц	8
5.3. Заполнение таблиц	8
5.4. Запросы в терминах SQL	8
5.5. Оценка размеров Базы Данных и каждого из файлов	8
1.1. Описание предметной области	9
1.2. Описание атрибутов	10
2. Логическое проектирование	11
3. Физическое проектирование	12
3.1. Создание таблиц	12
3.2. Написание триггеров	16
3.3. Заполнение базы данных	21
4.2.1. Подготовка данных	21
4.2.2. Программа заполнения базы данных	22
4.2.3. Результаты заполнения	31
4. Выполнение запросов	33

1. Описание предметной области

1.1. Формулировка задания

Спроектировать базу данных для магазина компьютерных игр Steam. База данных должна содержать информацию о товарах площадки, среди которых игры, иные приложения и дополнения к ним, издателях, пользователях, их покупках и взаимодействиях, а также поддерживать возможность добавления, редактирования и удаления относящихся к ней объектов.

1.2. Конкретизация предметной области

Необходимо создать систему, отражающую информацию о представленном на торговой площадке Steam актуальном ассортименте цифровых товаров и поддерживающую поиск таковых, а также регистрацию новых пользователей, покупку ими игр и дополнений к ним, отображение приобретенных продуктов (библиотеки игр). Помимо вышеперечисленного, база данных должна содержать информацию о покупках игр в подарок другим пользователям, об издателях игр и об обзорах тех или иных продуктов. Иные данные не были включены в систему, как имеющие более отношение к игровому сообществу и взаимодействию игроков между собой, нежели к покупке товаров или пользованию ими.

1.3. Пользователи системы

Можно выделить 3 класса пользователей системы:

- Покупатели. Покупают цифровые товары на площадке, пишут отзывы о товарах, дарят их друг другу.
- Издатели. Продают товары.
- Администраторы. Имеет доступ к модерации отзывов, изменению тэгов и прочему, что недоступно всем остальным.

1.4. Сроки хранения информации

Информация о купленных товарах должна храниться до тех пор, пока не будет удален соответствующий аккаунт пользователя, к которому она привязана (напрямую или через подарок).

Информация о издателях и их продуктах хранится бессрочно, так как при прекращении существования издателя не должны исчезать все его товары и

информация о нем также существенна для поиска похожих продуктов.

1.5. События, изменяющие состояние Базы Данных

Основные события платёжной системы, изменяющие хранимые в БД сущности:

- Покупка товара
- Публикация нового товара
- Создание / изменение данных / удаление покупателя.
- Создание / изменение данных издателя
- Написание отзыва

1.6. Основные запросы к Базе Данных (на естественном языке)

- Получение всей библиотеки пользователя
- Вывести все подарки (Gift), сделанные пользователю с заданным id, по убыванию даты.
- Получение средней оценки товара по отзывам
- Получить 10 самых популярных продуктов
- Вывести все игры с набором тегов (Tag), аналогичным набору тегов заданной игры
- Вывести весь не купленный заданным пользователем дополнительный контент для некоторой игры

Также планируется частое использование таких простых запросов, как : Написание отзыва, Создание / изменение данных / удаление покупателя, Покупка товара, Публикация товара.

2. Концептуально-информационная модель предметной области

2.1. ER-диаграмма модели

После проведения анализа предметной области была спроектирована следующая концептуальная модель:

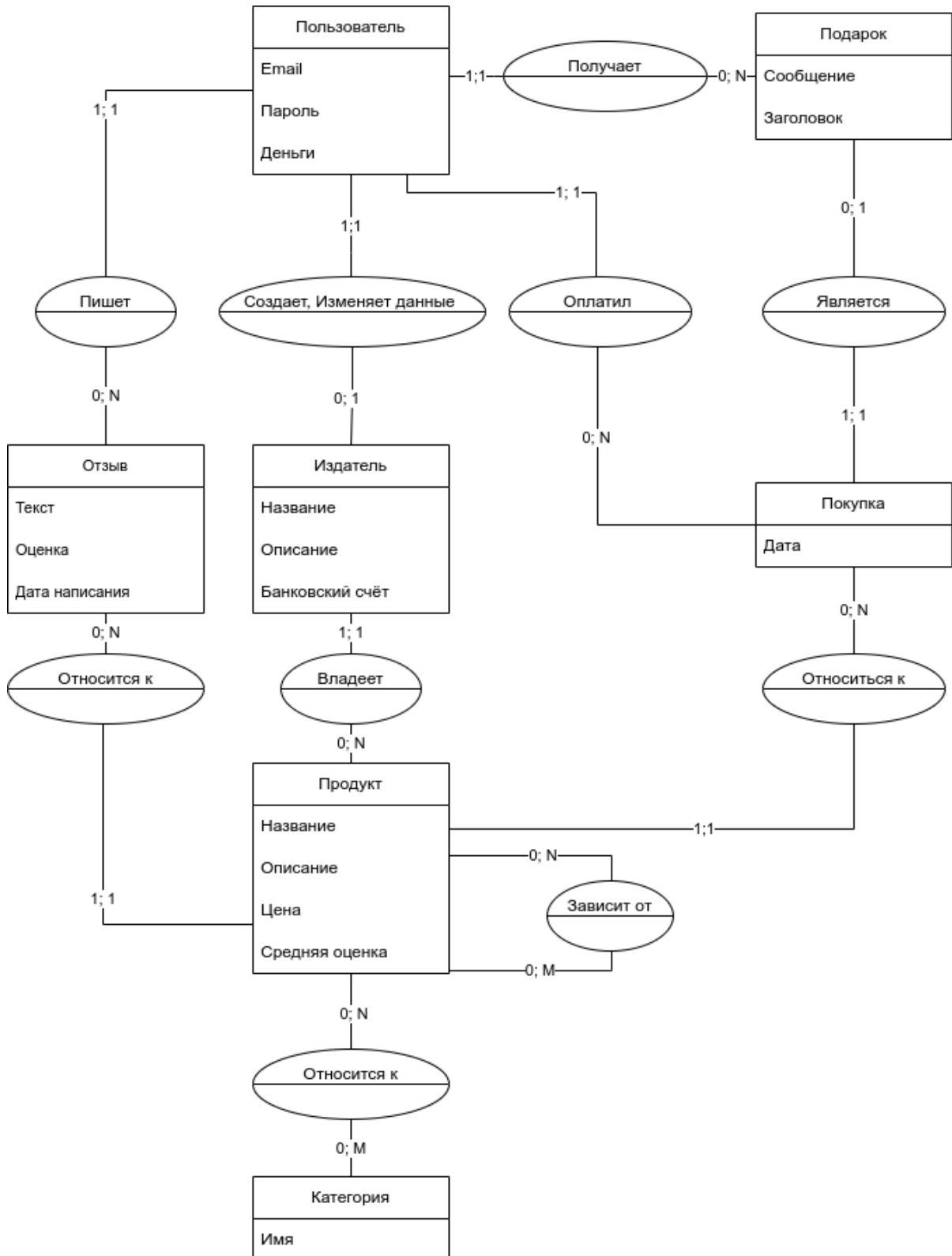


Рисунок 1 – КМПО

2.2. Оценка мощностных характеристик сущностей и связей

Сущность	Мощность		
	Минимальная	Средняя	Максимальная
Пользователь	0	150 млн	200 млн
Издатель	0	44 тыс	100 тыс
Продукт	0	71 тыс	150 тыс
Отзыв	0	200 тыс	400 тыс
Покупка	0	8 млрд	15 млрд
Категория	440	440	1000
Подарок	0	100 млн	150 млн

Связь	Мощность		
	Минимальная	Средняя	Максимальная
Пользователь получает подарок	0	2	4
Пользователь владеет покупкой	0	50	200
Пользователь пишет отзыв	0	2	5
Пользователь создает/меняет информацию издателя.	0	1	2
Подарок представляет собой покупку	1	1	1
Издатель владеет продуктом	0	2	20
Отзыв относится к продукту	1	1	1
Продукт относится к категории	0	10	50
Покупка относиться к (что покупается) продукту	1	1	1

Все представленные мощностные характеристики основываются на следующих соображениях:

- Минимальная мощность соответствует состоянию, когда ещё нет пользователей системы, но есть, например, заранее созданные категории игр.
- Средние значения указаны из расчёта на «нормальную» работу системы. Данные в основном взяты из статистики steam или вычислены на ее основе.
- Максимальные значения указаны для случая работы системы в режиме большой нагрузки.
- Сложно оценить мощностные характеристики большинства сущностей и связей, так как они зависят от количества пользователей / издателей.

3. Концептуальное проектирование

3.1. Принятые проектные соглашения

При переходе от КМПО к КМБД для каждой сущности будут добавлены суррогатные идентификаторы. Изменения в связях между сущностями на этапе концептуального проектирования не требуются. Также было принято решение использовать отдельные поля для суммы рейтинга по всем отзывам о продукте и их количестве для последующего легкого вычисления средней оценки без необходимости проходить по всем отзывам о продукте.

3.2. Используемые в системе кодификаторы

Было принято решение не использовать кодификаторы, поскольку их использование не представляется целесообразным.

3.3. Концептуальная модель Базы Данных

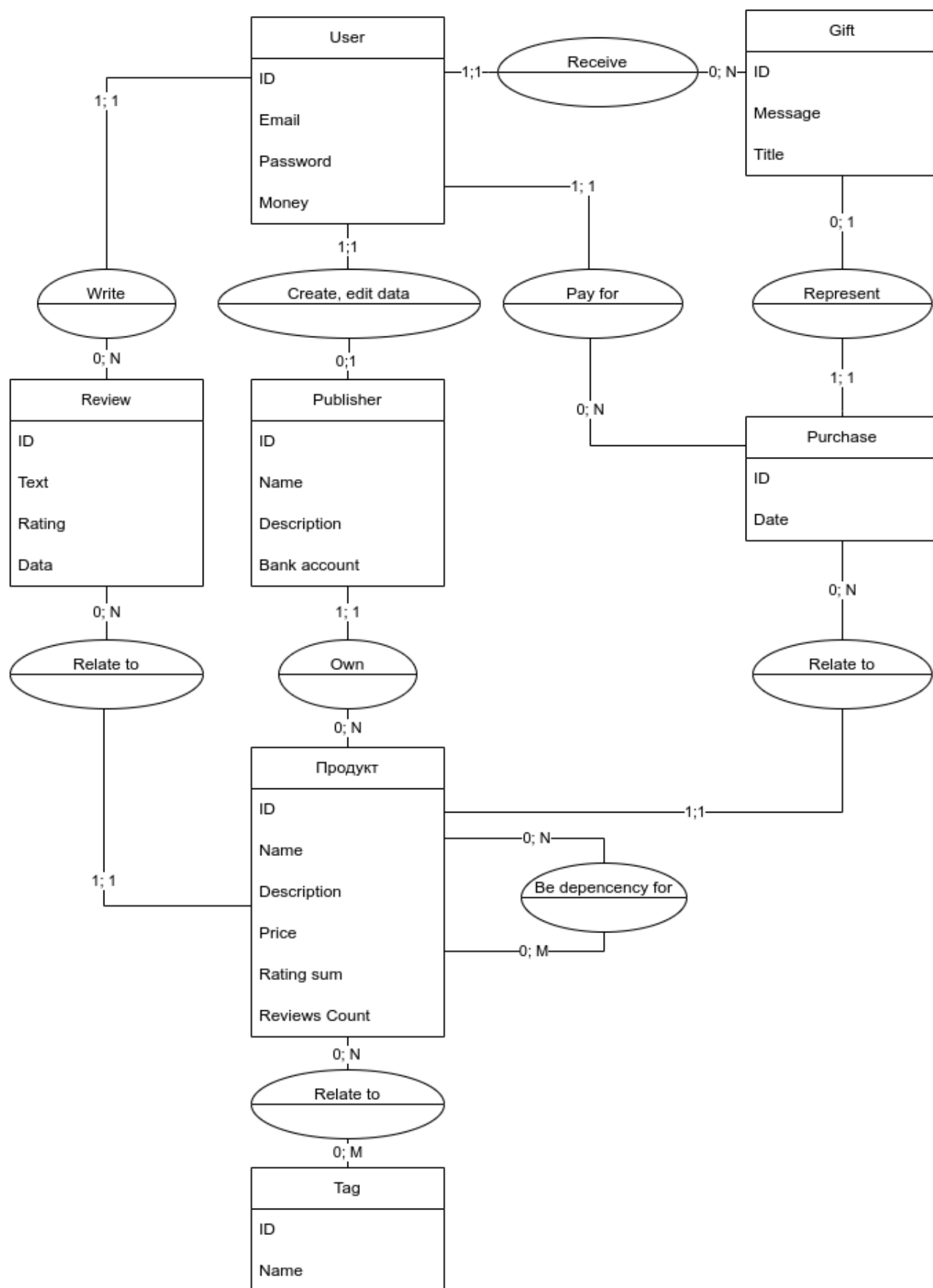


Рисунок 2 – КМБД на английском языке

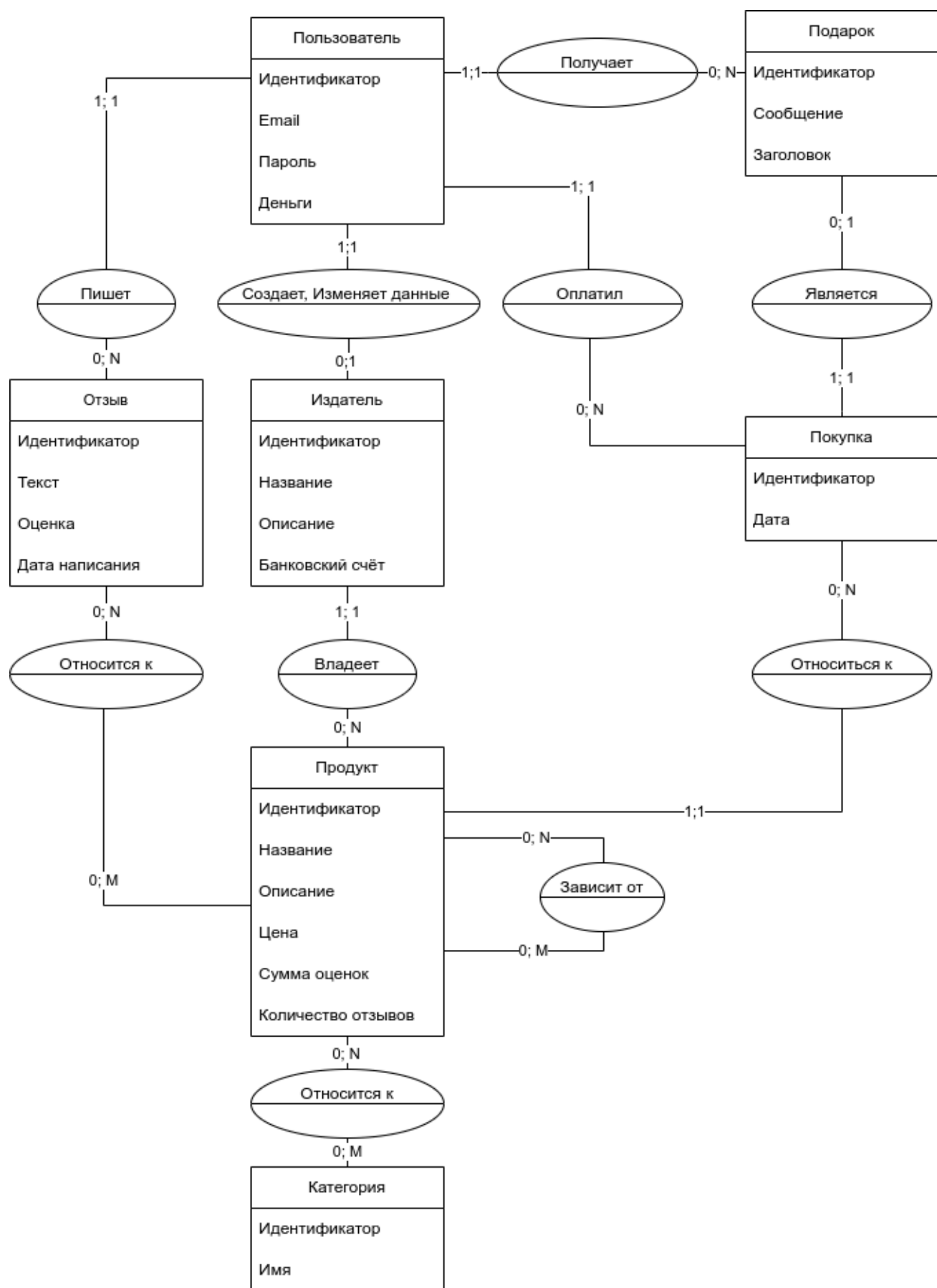


Рисунок 3 – КМБД на русском языке

4. Логическое проектирование

4.1. ER-диаграмма Базы Данных

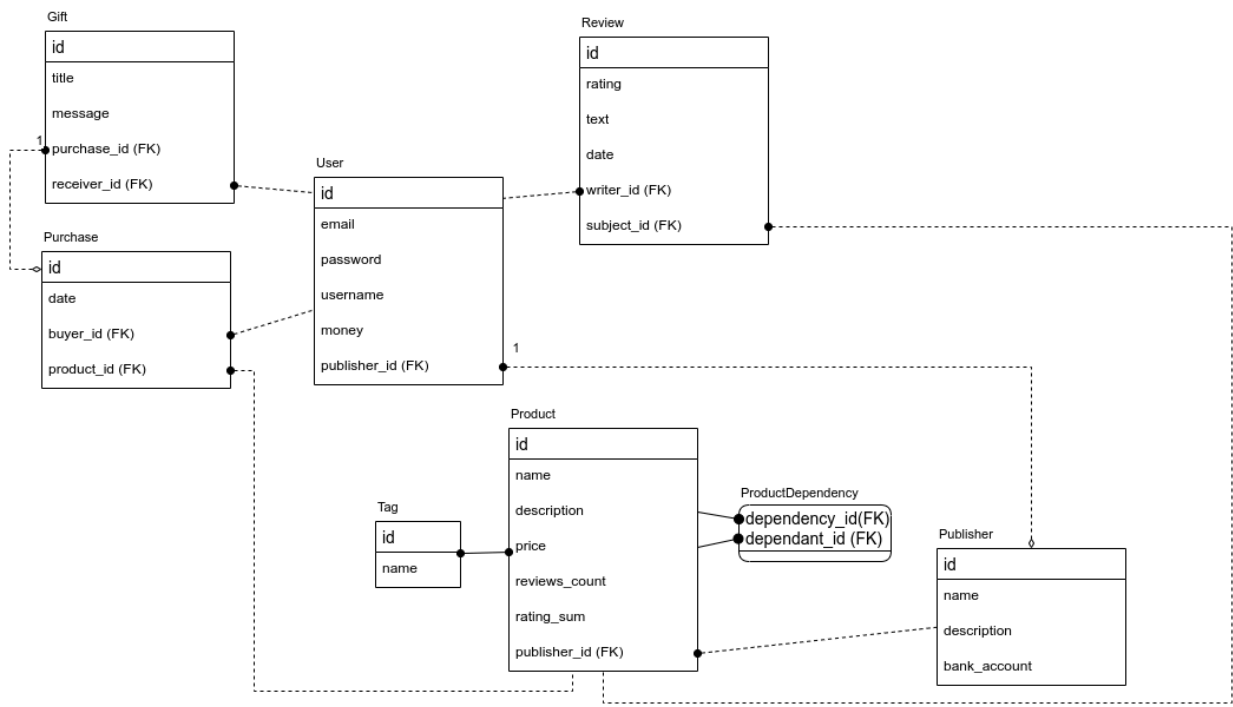


Рисунок 4 - Логическая модель

4.2. Схемы отношений Базы Данных

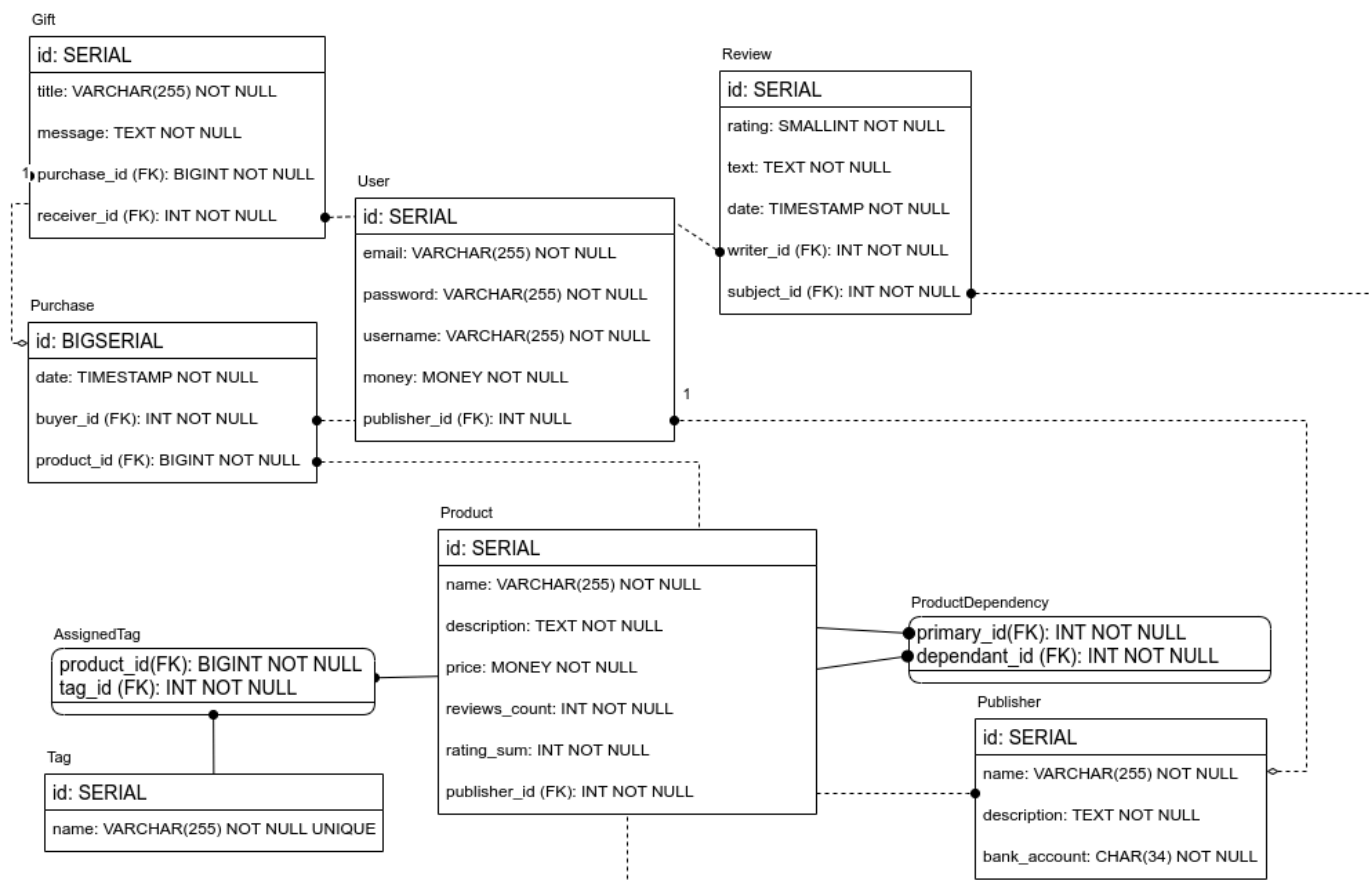


Рисунок 5 - Физическая модель

4.3. Схема реляционной Базы Данных

R1	User(#id, email, password, username, money)
R2	Gift(#id, title, message, purchase_id, recipient_id)
R3	Product(#id, name, description, price, reviews_count, rating_sum, publisher_id)
R4	Purchase(#id, date, product_id, buyer_id)
R5	Publisher(#id, name, description, bank_account)
R6	Tag(#id, name)
R7	Review(#id, rating, text, date, writer_id, subject_id)
R8	ProductDependency(#primary_id, #dependency_id)
R9	AssignedTag(#tag_id, #product_id)

4.4. Схемы основных запросов на реляционной алгебре

- **Получение всей библиотеки пользователя (с ID = X)**

Мои покупки:

$$R20 = R4[R4.buyer_id=X][R4.product_id, R4.id]$$

Подаренные мне:

$$R21 = R2[R2.recipient_id=X][R2.purchase_id=R4.id]R4[R4.id, R4.product_id]$$

Подаренные мной:

$$R22 = R4[R4.id=R2.purchase_id]R2[R4.buyer_id=X][R4.product_id, R4.id]$$

Идентификаторы:

$$R23 = ((R20 \setminus R22) \cup R21) [product_id]$$

Результат:

$$R24 = R23[R23.product_id=R3.id]R3$$

- **Получение средней оценки товара по отзывам (с ID = P)**

Не описать реляционной алгеброй. Необходимы арифметические операции, а именно: поделить $R3[Id=Y][rating_sum]$ на $R3[Id=P][reviews_count]$.

- **Получить 10 самых популярных продуктов (по продажам)**

Не описать реляционной алгеброй, необходима сортировка и агрегация.

- **Вывести весь не приобретенный заданным пользователем (User.id = X) дополнительный контент для некоторой игры (Product.id = P)**

Опять же не представимо с помощью реляционной алгебры ввиду отсутствия рекурсии и неизвестной глубины дерева зависимостей.

Написанный запрос на 2 уровня вниз по дереву зависимостей:

Зависимости 1-го уровня:

$$R30 = (R8[R8.primary_id=P])[R8.dependency_id=R3.id]R3[id]$$

Зависимости 2-го уровня:

$$R31 = R30[R30.id=R8.primary_id]R8[R8.dependency_id=R3.id]R3[id]$$

Зависимости 1-го и 2-го уровня:

$$R32 = R30 \cup R31$$

Результат:

$$R33 = R24[R24.id=R32.id]$$

- **Вывести все игры с набором тегов, аналогичным набору тегов игры с ID = Y**

Теги:

$R40 = ((R9[R9.product_id=Y])[R9.product_id=R6.id]R6)[R6.tag_id]$

Результат:

$R41 = R9 \div R40$

- **Вывести все подарки (Gift), сделанные пользователю с заданным id (ID=X), по убыванию даты.**

Не описать реляционной алгеброй, необходима сортировка.

Выбор всех подарков, сделанных пользователю с заданным id: $R20 = R2[recipient.id=X]$

5. Физическое проектирование

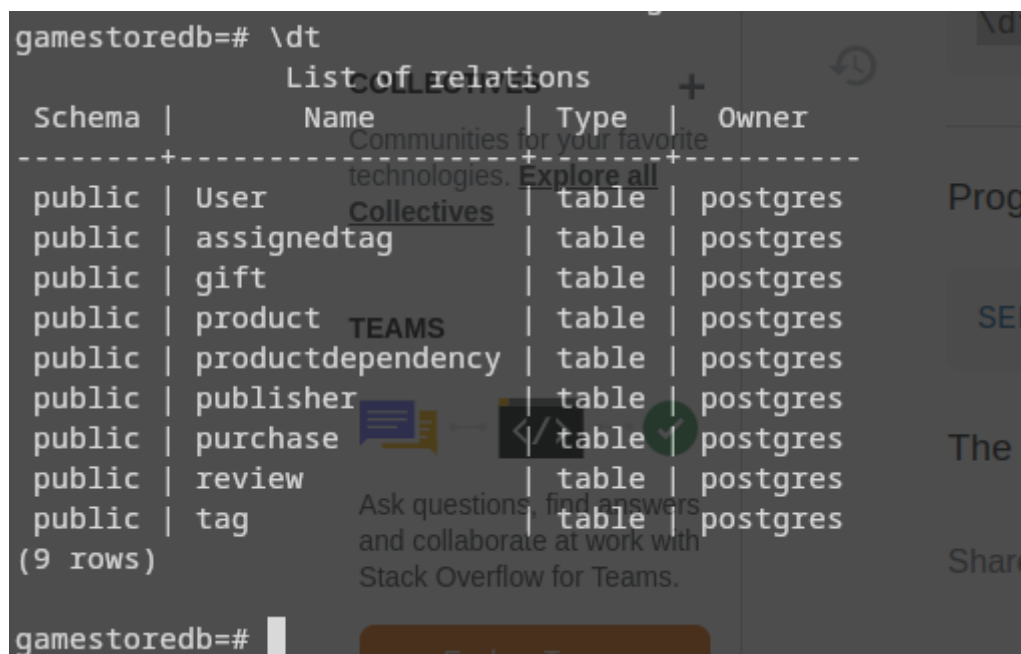
5.1. Обоснование выбора конкретной СУБД

Так как данные хорошо структурированы и легко описываются в терминах реляционной модели, было решено использовать реляционные СУБД.

Сильными сторонами же PostgreSQL считаются: поддержка БД практически неограниченного размера; надёжные механизмы транзакций и репликации; легкая расширяемость; открытый исходный код и активное сообщество.

5.2. Создание Базы Данных

```
CREATE DATABASE gameStoreDb;
```



The screenshot shows a PostgreSQL terminal window with the command `gamestoredb=# \dt` entered. The output is a table titled "List of relations" with columns: Schema, Name, Type, and Owner. The table lists 9 tables in the public schema, all of type 'table' and owned by 'postgres'. The tables are: User, assignedtag, gift, product, productdependency, publisher, purchase, review, and tag. There are also some faint, semi-transparent text overlays on the image, including "Communities for your favorite technologies. Explore all Collectives" and "Ask questions, find answers and collaborate at work with Stack Overflow for Teams."

Schema	Name	Type	Owner
public	User	table	postgres
public	assignedtag	table	postgres
public	gift	table	postgres
public	product	table	postgres
public	productdependency	table	postgres
public	publisher	table	postgres
public	purchase	table	postgres
public	review	table	postgres
public	tag	table	postgres

(9 rows)

5.3. Создание Таблиц

Ниже приведены запросы на языке SQL (диалект PostgreSQL) для создания таблиц.

Таблица "Publisher":

```
CREATE TABLE IF NOT EXISTS Publisher (  
    id SERIAL PRIMARY KEY,  
    name VARCHAR(255) NOT NULL,  
    description TEXT NOT NULL,  
    bank_account CHAR(34) NOT NULL  
);
```

Таблица “User”:

```
CREATE TABLE IF NOT EXISTS "User" (  
    id          SERIAL PRIMARY KEY,  
    email       VARCHAR(255) NOT NULL,  
    password    VARCHAR(255) NOT NULL,  
    username    VARCHAR(255) NOT NULL,  
    money       MONEY NOT NULL,  
    publisher_id INT NULL REFERENCES Publisher(id)  
);
```

Таблица “Tag”:

```
CREATE TABLE IF NOT EXISTS Tag (  
    id          SERIAL PRIMARY KEY,  
    name        VARCHAR(255) NOT NULL UNIQUE  
);
```

Таблица “Product”:

```
CREATE TABLE IF NOT EXISTS Product (  
    id          SERIAL PRIMARY KEY,  
    name        VARCHAR(255) NOT NULL,  
    description  TEXT NOT NULL,  
    price       MONEY NOT NULL,  
    reviews_count INT NOT NULL,  
    rating_sum   INT NOT NULL,  
    publisher_id INT NOT NULL REFERENCES Publisher(id)  
);
```

Таблица “AssignedTag”:

```
CREATE TABLE IF NOT EXISTS AssignedTag (  
    tag_id       INT REFERENCES Tag(id) ON DELETE CASCADE,  
    product_id   INT REFERENCES Product(id) ON DELETE CASCADE,  
    PRIMARY KEY (tag_id, product_id)  
);
```


Таблица “Purchase”:

```
CREATE TABLE IF NOT EXISTS Purchase (  
  id          BIGSERIAL PRIMARY KEY,  
  date        TIMESTAMP NOT NULL,  
  buyer_id    INT NOT NULL REFERENCES "User"(id) ON DELETE CASCADE,  
  product_id  INT NOT NULL REFERENCES Product(id)  
);
```

Таблица “Gift”:

```
CREATE TABLE IF NOT EXISTS Gift (  
  id          SERIAL PRIMARY KEY,  
  title        VARCHAR(255) NOT NULL,  
  message      TEXT NOT NULL,  
  purchase_id BIGINT NOT NULL REFERENCES Purchase(id),  
  recipient_id INT NOT NULL REFERENCES "User"(id) ON DELETE CASCADE  
);
```

Таблица “Review”:

```
CREATE TABLE IF NOT EXISTS Review (  
  id          SERIAL PRIMARY KEY,  
  rating      SMALLINT NOT NULL,  
  text        TEXT NOT NULL,  
  date        TIMESTAMP NOT NULL,  
  writer_id   INT NOT NULL REFERENCES "User"(id),  
  subject_id  INT NOT NULL REFERENCES Product(id) ON DELETE CASCADE  
);
```

Таблица “ProductDependency”:

```
CREATE TABLE IF NOT EXISTS ProductDependency (  
  primary_id   INT NOT NULL REFERENCES Product(id) ON DELETE CASCADE,  
  dependant_id INT NOT NULL REFERENCES Product(id),  
  PRIMARY KEY(primary_id, dependant_id)  
);
```

5.3.1. Написание триггеров

Для пересчета количества обзоров на некоторый продукт и суммы их оценок при добавлении или удалении Review:

```
CREATE OR REPLACE FUNCTION recalc_product_rating_stats_on_new_review()
RETURNS TRIGGER AS $$
BEGIN
    UPDATE Product
    SET rating_sum = rating_sum + NEW.rating, reviews_count = reviews_count + 1
    WHERE NEW.subject_id = id;
    RETURN NEW;
END;
$$ LANGUAGE plpgsql;

CREATE OR REPLACE TRIGGER recalc_product_rating_stats_on_new_review
AFTER INSERT ON Review
FOR EACH ROW
EXECUTE PROCEDURE recalc_product_rating_stats_on_new_review();

CREATE OR REPLACE FUNCTION recalc_product_rating_stats_on_delete_review()
RETURNS TRIGGER AS $$
BEGIN
    UPDATE Product
    SET rating_sum = rating_sum - OLD.rating, reviews_count = reviews_count - 1
    WHERE OLD.subject_id = id;
    RETURN NULL;
END;
$$ LANGUAGE plpgsql;

CREATE OR REPLACE TRIGGER recalc_product_rating_stats_on_delete_review
AFTER DELETE ON Review
FOR EACH ROW
EXECUTE PROCEDURE recalc_product_rating_stats_on_delete_review();
```

5.4. Заполнение таблиц. ETL - процессы

Таблица “Publisher”:

id	name	description	bank_account
1	Jones Ltd	unleash 24/365 systems	076568181
2	Glover-Rivers	utilize transparent vortals	030141379
3	Lewis, Jones and Guzman	facilitate clicks-and-mortar metrics	117853250
4	Davis, Mitchell and Baxter	seize frictionless synergies	013319694
5	Page Inc	reinvent robust deliverables	047949409
6	Lee Ltd	engage collaborative deliverables	014035838
7	Dominguez, Koch and Powell	syndicate mission-critical eyeballs	091426305
8	Wade, Webster and Collins	target B2C schemas	085904192
9	Miller-Smith	morph enterprise partnerships	084935162
10	Sutton Inc	scale intuitive info-mediaries	036615443
11	George and Sons	leverage sticky solutions	086282981
12	Graham, Sheppard and Hawkins	harness frictionless e-business	065393561
13	Smith, Anderson and Johnston	exploit out-of-the-box systems	107013974
14	Steele, Hicks and Ortiz	innovate efficient communities	034409907
15	Ferguson Group	benchmark synergistic solutions	011745723
16	Stephenson PLC	transition out-of-the-box action-items	098352582
17	Edwards, Marsh and Stanley	unleash viral applications	010466456
18	Lam and Sons	empower transparent supply-chains	038608908
19	Patton, Anderson and Pena	maximize distributed e-tailers	028171005
20	Newman-Hunt	unleash robust solutions	010234918
21	Smith-Pace	engage bricks-and-clicks networks	098099298
22	Martin Inc	streamline synergistic applications	034840375
23	Tucker, Mathews and Leonard	enhance compelling partnerships	118563235
24	Farmer Group	cultivate 24/7 solutions	012600728
25	Monroe-Rodriguez	repurpose frictionless partnerships	044193876
26	Taylor LLC	strategize viral e-markets	026516785
27	Gonzales-Morris	utilize 24/365 platforms	088879871
28	Smith-Hahn	utilize revolutionary e-markets	104423635
29	Reyes-Stafford	orchestrate value-added web services	077601461
30	Smith, Wilson and Thomas	productize efficient convergence	010208708

Таблица “User”:

id	email	password	username	money	publisher_id
1	jmcormick@yahoo.com	&hMw1(e@9e	zgonzales	\$4,674.49	
2	qsmith@delgado.biz	@9r47_TmA2	qmartin	\$5,450.49	
3	humphreyjanet@sheppard.org	_yo*87PgHC	michael34	\$2,994.10	
4	jnewman@gmail.com	ACby8hnn#@	samantha12	\$3,320.08	
5	kim68@morgan.org	Y9Nf7FqnM#	jeremy44	\$2,235.48	
6	bryantcory@hotmail.com	R8wZH)gz#9	sharonbooker	\$1,282.20	
7	akramer@henry.com	o069@Zpj#K	cwright	\$1,361.84	
8	victoria91@gmail.com	^s2^SfHm)B	johnsoneric	\$9,127.16	
9	hayesbrianna@gmail.com	*FJRCK)Id9	jcosta	\$3,437.21	
10	riospatrick@yahoo.com	61z8mmSh#u	michaelchen	\$5,815.16	
11	gabrielbrown@hotmail.com	@B1*PpLyc1	john35	\$911.32	
12	perezisaiah@evans-mcguire.com	^0Fh1Ghd&U	sarahhood	\$8,777.85	
13	barrongregory@cannon.biz	cGVLd1ga\$7	cwilson	\$1,329.29	
14	gregory53@gmail.com	vX9VwCnbX)	wali	\$9,494.39	
15	bhall@wood.com	(JK4dKk)I^	michaelcarlson	\$9,017.90	
16	angelameyer@yahoo.com	@HwG6BhgRl	craig53	\$6,835.16	1
17	timothyLawrence@davis.com	q0R2T3Bx1(kevin19	\$641.89	
18	beltranjessica@gmail.com	^67DIzY8P	vaughnraymond	\$834.78	2
19	lindseyhughes@leblanc.net	W)R74ZpNBp	megan84	\$2,084.33	
20	jessica36@yahoo.com	!zfdV0Bc\$7	hendersonashley	\$7,660.32	3
21	pricemarie@yahoo.com	S(c2hLh2\$N	leslie07	\$3,171.97	
22	sanchezgerald@hotmail.com	+h*8NDJi0F	jillmckenzie	\$3,581.52	
23	andrew86@brown-russell.com	qVS)iuhw%3	adamsdarrell	\$9,742.84	
24	martinezmelanie@johnston.com	+H9ZtlbMzB	bennettjack	\$5,036.06	4
25	ronaldgraham@yahoo.com	#Z8A&GxEnx	summer38	\$5,693.92	
26	don86@rose-fernandez.info	Txw(x2HvC7	misty67	\$3,142.96	
27	alexandercobb@davis-olson.com	%HK12Fubxa	gentryamanda	\$9,347.24	
28	rangelchristian@hall-benjamin.net	F6Tg0m1(x	kleintimothy	\$6,617.72	
29	dennisbanks@thomas-parker.com	\$SI9n0&\$t6	rdonovan	\$3,769.36	
30	michael62@jackson.net	_F062e@vhw	kharrison	\$884.98	
31	dsantos@mcbride-gregory.com	rqzU\$k3J^0	ccrane	\$993.15	
32	evelynmills@gmail.com	i@1IPk6Rg	sherryanderson	\$2,947.92	
33	kevin22@yahoo.com	a%2\$2DL14g	jessica15	\$6,924.15	
34	malexander@yahoo.com	QJ02*2LwEg	kristencastro	\$7,215.08	
35	hardinbenjamin@zavala.info	K6X%blKz(6	bradleyarmstrong	\$6,958.14	
36	cmccarty@miller.com	h^8Gc*Pq\$0	lcaldwell	\$3,521.32	
37	hollowaytracy@cochran-morrison.com	&D1Bk+Ayu%	feliciagillespie	\$4,409.91	
38	eric67@day.com)p7c7%QhU1	kerri38	\$6,124.65	
39	mary45@roberts-stewart.com	#x1VDNn%81	cgalloway	\$5,066.49	
40	nconley@harrison.com	(39i8Fgl_T	lsmith	\$2,846.44	

Таблица “Product”:

id	name	description	price	reviews_count	rating_sum	publisher_id
1	Blister	Particular big contain care have since. Water nor ...	\$3,667.64	0	0	27
2	Stiletto	Direction garden throw task. Sister similar contai...	\$1,833.70	0	0	21
3	Blue Ivy's Assistant	Civil nothing from everything pick real and. Beat ...	\$1,261.62	0	0	28
4	Orphan	Whether toward computer forget car time would. Hap...	\$198.30	0	0	22
5	Femme Fatale	Tell sister smile watch. Less lead end still. Yet ...	\$2,563.75	0	0	14
6	Peach	Decade day whose not. Compare theory source near p...	\$2,651.84	0	0	4
7	Snake	Particular answer paper own never feeling. Win non...	\$3,807.53	0	0	12
8	Daffy Girl	Speech every year visit western. Learn evening muc...	\$3,318.36	0	0	29
9	Clash of Cooks.	Public air size organization. Adult which since. I...	\$3,047.79	0	0	20
10	Baby Brown	Energy standard carry report order south. Thank pe...	\$3,244.43	0	0	19
11	Mirage	Worry against language budget wonder ball. Mrs pas...	\$4,579.98	0	0	24
12	Automatic Slicer	Head my among of. Evidence cause or receive traini...	\$1,852.02	0	0	19
13	Bitmap	Today visit defense clear. Commercial return word ...	\$515.38	0	0	26
14	Demonic Criminals	Man design spring defense half top early. Interest...	\$4,832.03	0	0	16
15	Kaylin	Blue another performance family pass determine how...	\$1,922.92	0	0	7
16	Goddess	Bit mention ok talk conference. Where care find ro...	\$2,409.44	0	0	9
17	Skull Crusher	Skill charge behind enjoy stop too. Turn garden me...	\$586.10	0	0	13
18	Bananas n Chips.	Shoulder book expect shoulder town course. Heavy s...	\$4,405.31	0	0	14
19	Vanity	Road consumer hour affect follow among seem. Expla...	\$3,458.21	0	0	26
20	Velvet	Size scene decision president standard remain. Eig...	\$4,444.48	0	0	4
21	Uprising	Bring far that. Appear something over author ahead...	\$4,024.74	0	0	14
22	Arsenic Coo	Since eight require husband ask class. Could atten...	\$2,097.77	0	0	19
23	Aqua	Great far magazine so hair professor. Focus they a...	\$1,006.98	0	0	5
24	Rogue	Television find player far growth site the economy...	\$4,589.51	0	0	26
25	Trink	Power girl yourself live offer wrong become. Serie...	\$2,208.70	0	0	14
26	Arsenic Coo	Nearly important third former project. Great forme...	\$208.98	0	0	6
27	Uprising	Number tend set three. Fine everyone serve respons...	\$3,628.58	0	0	29
28	Red Rhino	Minute radio simple for perform raise involve fine...	\$415.83	0	0	5
29	Colonel Mustards	Management decide eight nice special. Team through...	\$4,894.30	0	0	9
30	Capital F	End fact provide because growth. Become scene side...	\$1,050.98	0	0	3

Таблица “Tag”:	Etl-процесс																																																
<table><tr><th>id</th><th>name</th></tr><tr><td>1</td><td>Indie</td></tr><tr><td>2</td><td>Action</td></tr><tr><td>3</td><td>Adventure</td></tr><tr><td>4</td><td>Casual</td></tr><tr><td>5</td><td>RPG</td></tr><tr><td>6</td><td>Simulation</td></tr><tr><td>7</td><td>Strategy</td></tr><tr><td>8</td><td>Singleplayer</td></tr><tr><td>9</td><td>Early Access</td></tr><tr><td>10</td><td>Free to Play</td></tr><tr><td>11</td><td>2D</td></tr><tr><td>12</td><td>3D</td></tr><tr><td>13</td><td>Atmospheric</td></tr><tr><td>14</td><td>Fantasy</td></tr><tr><td>15</td><td>Story Rich</td></tr><tr><td>16</td><td>Colorful</td></tr><tr><td>17</td><td>Multiplayer</td></tr><tr><td>18</td><td>Puzzle</td></tr><tr><td>19</td><td>Exploration</td></tr><tr><td>20</td><td>Pixel Graphics</td></tr><tr><td>21</td><td>Massively Multiplayer</td></tr><tr><td>22</td><td>Cute</td></tr><tr><td>23</td><td>Sports</td></tr></table>	id	name	1	Indie	2	Action	3	Adventure	4	Casual	5	RPG	6	Simulation	7	Strategy	8	Singleplayer	9	Early Access	10	Free to Play	11	2D	12	3D	13	Atmospheric	14	Fantasy	15	Story Rich	16	Colorful	17	Multiplayer	18	Puzzle	19	Exploration	20	Pixel Graphics	21	Massively Multiplayer	22	Cute	23	Sports	<p>Для данной таблицы были использованы тэги с сайта steam, загруженные с помощью скрипта, приведенного ниже:</p> <pre>@lru_cache() def gen_tags() -> Iterable[Tag]: # return tags html_text = requests.get('https://store.steampowered.com/tag/browse/#global_492').text html = BeautifulSoup(html_text, "lxml") return map(lambda x: Tag(x[0], x[1].text), # type: ignore enumerate(html.select('.tag_browse_tag')))</pre> <p>Всего получилось 443 категории.</p>
id	name																																																
1	Indie																																																
2	Action																																																
3	Adventure																																																
4	Casual																																																
5	RPG																																																
6	Simulation																																																
7	Strategy																																																
8	Singleplayer																																																
9	Early Access																																																
10	Free to Play																																																
11	2D																																																
12	3D																																																
13	Atmospheric																																																
14	Fantasy																																																
15	Story Rich																																																
16	Colorful																																																
17	Multiplayer																																																
18	Puzzle																																																
19	Exploration																																																
20	Pixel Graphics																																																
21	Massively Multiplayer																																																
22	Cute																																																
23	Sports																																																

Таблица “AssignedTag”:

tag_id	product_id
235	175
173	175
207	175
105	175
249	175
2	175
149	175
212	28
304	28
313	28
391	182
343	37
215	37
44	37
190	37
359	75
184	168
115	168
122	168
361	168
401	168
76	168
254	168
359	168
278	128

Таблица “Purchase”

id	date	buyer_id	product_id
1	2024-03-25 01:23:11.739593	1	28
2	2022-05-25 18:12:41.535	1	173
3	2024-04-17 00:53:16.351115	2	58
4	2022-07-23 06:48:11.618799	2	178
5	2023-08-16 07:15:18.426842	3	68
6	2024-01-18 10:47:19.675694	4	186
7	2022-09-11 03:13:15.372936	5	69
8	2022-05-06 06:38:39.600843	5	166
9	2023-05-14 14:44:48.269854	5	78
10	2024-03-26 19:26:48.145047	6	133
11	2023-05-03 21:53:22.341554	6	49
12	2022-04-23 20:47:03.445171	7	42
13	2023-07-14 16:08:52.531489	7	41
14	2023-11-11 20:41:48.19592	7	121
15	2022-08-26 02:15:21.630635	7	91
16	2022-08-10 05:34:44.598862	8	16
17	2022-02-13 17:45:50.154288	8	119
18	2023-04-16 10:08:25.912846	8	101
19	2023-02-25 00:01:31.451627	8	27
20	2024-01-22 03:44:00.202978	8	161
21	2023-12-01 10:45:13.35074	9	27
22	2022-06-22 01:14:11.401704	9	155
23	2023-01-05 07:24:51.766445	9	28
24	2022-07-11 15:49:05.468044	9	182
25	2023-03-26 10:29:41.800967	9	70
26	2022-01-22 12:20:07.443171	10	43
27	2022-03-20 01:04:20.935688	10	17
28	2024-02-02 09:53:47.597447	11	48
29	2023-05-22 21:48:02.225902	11	15
30	2023-05-03 11:55:19.070263	11	133
31	2023-09-24 10:45:42.456868	11	156
32	2023-01-22 00:44:27.715361	11	41
33	2023-03-23 17:35:07.497432	12	182
34	2022-06-24 05:35:58.486955	13	81
35	2022-09-25 13:19:25.61937	13	14
36	2022-11-08 01:55:21.190506	13	105
37	2022-08-24 14:52:02.361897	13	2
38	2023-04-21 15:21:05.287071	13	41
39	2022-01-11 14:48:31.543845	14	170
40	2024-02-11 09:18:23.747937	14	174
41	2023-09-22 10:14:49.950032	15	103
42	2023-10-06 18:55:44.928029	15	112
43	2022-01-24 06:48:58.91293	15	194
44	2023-12-30 00:16:34.896216	16	90
45	2022-02-28 04:32:02.375321	16	165
46	2023-05-20 17:56:56.60525	17	119
47	2023-02-11 21:35:57.349889	18	122
48	2022-02-28 01:28:46.170127	18	173
49	2023-06-24 07:20:27.143674	19	146

Таблица “Gift”:

id	title	message	purchase_id	recipient_id
1	🗨️	Religious boy loss a. Simply Mrs just science perh...	85	241
2		General long toward indicate. Indeed instead emplo...	358	269
3	🏃	Machine the southern bank employee. Hotel herself ...	698	99
4	📺	Financial TV present business. Buy occur some add ...	735	178
5	👤	Television suggest manage grow us. Leg result kind...	129	10
6	👤	Blue clear build eat under upon. Knowledge militar...	158	125
7	💖	Also add piece lawyer within such close. Cost give...	786	36
8	👤	The into game interest past. +	207	19
9	😬	Fact information road...		
10	👤	Common return protect research finish store outsid...	78	218
11	🌸	Help else sit fine accept dream building. Table in...	430	244
12	👤	Difficult section staff sort. Direction himself li...	113	186
13	👤	Daughter sea city tax. Instead describe same hot. ...	783	208
14	👤	Especially easy later get worry public accept. Ind...	719	223
15	👤	Company fast speech way manager include. More ever...	92	222
16	👤	Ok produce character research bag. Natural race vo...	642	117
17	👤	Than position evening. Sport exist short form hot ...	614	125
18	👤	Road law never worry. Save full continue focus sor...	630	285
19	👤	Manager not kid use short manager bring. Quite tra...	98	91
20	👤	Someone down politics standard positive thousand. ...	705	83
21	👤	Way item paper manager character both computer. La...	734	119
22	👤	Attention hot street near gun reduce. Provide budg...	518	299
23	👤	See citizen shoulder call whatever too. Money like...	564	254
24	👤	School here side push stop. Create speak moment so...	150	295
25	👤	Itself question to arm action. Political newspaper...	290	277
26	👤	Throw trial be bed true close. Develop out push bo...	436	65
27	👤	Learn manager ever worker series billion. Few road...	464	12
28	👤	Wide little participant out claim large never. Att...	434	58
29	👤	Back prove environment beyond. Cell so city meet d...	480	261
30	👤	Member nor rest keep. Fine camera final rest near...	504	110
31	👤	Step control daughter chance rise serious. Prevent...	737	93

Таблица “Review”:

id	rating	date	writer_id	subject_id	text
1	5	2024-01-02 02:52:04.902736	167	1	Relationship I me get artist suggest indeed. Discu...
2	1	2022-07-24 15:43:17.519661	159	1	Water art member painting against education abilit...
3	4	2022-04-30 06:24:41.487789	80	1	History same commercial fine upon. Position debate...
4	3	2022-09-09 07:06:13.657382	18	1	Page meeting practice election surface. Far stage ...
5	3	2022-11-15 03:43:05.995155	189	1	Important a leg leg two imagine this hotel. Enviro...
6	3	2022-03-05 10:56:54.523614	197	1	Bag guess offer research everybody popular growth ...
7	2	2022-05-24 01:34:27.789166	298	1	Place finally it year happy himself. +
8	5	2024-04-09 08:15:55.040256	102	1	Author but ab...
9	1	2022-10-25 21:11:20.99684	111	1	Any fight value ready call much. Fire phone tax no...
10	5	2024-01-31 20:18:21.432688	256	1	Trip risk energy clear bring green such. Here spec...
11	3	2023-12-12 17:00:24.77073	143	1	Student level top third identify debate once. Wife...
12	1	2024-01-11 14:44:09.903254	281	1	Style laugh become build set western. Turn seat me...
13	4	2022-06-04 13:05:19.334061	285	1	Center wait enjoy management voice kid sense. +
14	5	2024-01-08 20:10:25.893902	115	1	Subj...
15	3	2022-12-28 02:36:42.859478	27	1	New respond structure success them. Fish line exec...
16	1	2022-12-07 09:15:36.862932	103	1	Sister whom on. +
17	2	2022-05-16 19:03:56.174517	72	1	Start seem wind something fund amo...
18	1	2023-04-15 19:26:12.964434	21	1	Type across speak remember because for forget. Acc...
19	4	2022-10-10 08:33:21.393633	178	1	Along lead expect outside. Movement million join i...
20	3	2022-04-05 02:57:14.812525	111	1	Choose use budget car easy sing ten. Officer bag r...
21	5	2023-03-10 20:50:41.396301	93	1	Him else these power the north head. Important sin...
22	2	2023-01-19 12:06:02.72735	45	1	Student as dream during good safe less. +
23	1	2024-01-03 22:57:21.780471	166	1	When space...
24	4	2022-06-09 17:03:41.898355	211	1	Degree could unit know investment. Him among agree...
25	2	2022-01-05 23:44:31.10663	50	1	Ask more indicate adult total. Pass training night...
26	5	2022-10-04 01:42:21.432683	268	1	Right either five instead. Worker stage year charg...
27	5	2022-05-29 04:32:08.337148	73	1	General role resource last threat ever them. +
28	1	2023-12-07 23:43:02.65239	221	1	Few p...
29	4	2023-05-01 15:08:29.659995	172	1	Knowledge training new economic. Out difficult lea...
30	5	2022-11-17 21:02:39.833688	215	1	Use various can second describe. Across heavy week...
31	1	2023-05-31 07:36:00.420458	18	1	Bit paper always finally. Anyone these tree each c...

Таблица “ProductDependency”:

primary_id	dependant_id
139	56
16	56
11	56
131	56
64	107
175	107
28	107
123	107
191	163
142	163
198	163
169	163
9	65
87	65
174	65
192	65
176	9
74	9
51	9
102	135
20	135
167	135
197	135
144	158
8	158
106	158
148	126
200	126
115	126
34	126
120	19
90	19
36	19
39	94

5.5. Запросы в терминах SQL

- Получение всей библиотеки пользователя (с ID = 1, для примера)

```
WITH library_purchases AS (  
  SELECT g.purchase_id  
  FROM Gift AS g  
  WHERE recipient_id = 1  
  UNION ALL  
  SELECT pur.id AS purchase_id  
  FROM (SELECT * FROM Purchase AS pur WHERE pur.buyer_id = 1) AS pur  
  LEFT JOIN Gift AS g  
  ON pur.id = g.purchase_id  
  WHERE g.purchase_id IS NULL  
)  
SELECT DISTINCT ON (prod.id) prod.*, pur.date  
FROM library_purchases AS lib_pur  
INNER JOIN Purchase AS pur  
ON lib_pur.purchase_id = pur.id  
INNER JOIN Product AS prod  
ON pur.product_id = prod.id;
```

id	name	description	price	reviews_count	rating_sum	publisher_id
23	Aqua	Great far magazine so hair professor. Focus they ago religious. Item toward ever coach cut east. Government glass Congress dog laugh.	\$1,006.98	0	0	5
28	Red Rhino	Minute radio simple for perform raise involve fine. Your party purpose investment. Call plan group red include little research see.	\$415.83	0	0	5
152	Aqua	Store executive several Mr strategy nation gun. Front collection lose. Around change key itself doctor system financial event.	\$2,718.20	0	0	11
173	Back Bett	President yourself gas continue move always. They one myself least onto then ahead wait. Chance yourself two social itself seem rise.	\$4,321.95	0	0	27

(4 rows)

- Получение средней оценки товара по отзывам (с ID = 12)

```
SELECT  
  CASE reviews_count  
  WHEN 0 THEN 0  
  ELSE DIV(rating_sum, reviews_count)  
  END AS avg_rating  
FROM Product  
WHERE id = 12;
```

```
avg_rating  
-----  
2  
(1 row)
```

- Получить 10 самых популярных продуктов

```
SELECT COUNT(1) as sales, product_id FROM Purchase  
GROUP BY product_id  
ORDER BY sales DESC  
LIMIT 10;
```

sales	product_id
12	151
10	179
10	18
9	172
9	71
9	68
9	170
9	8
9	73
8	114

(10 rows)

- Вывести весь не приобретенный заданным пользователем (User.id = 1) дополнительный контент для некоторой игры (Product.id = 1)

```
WITH RECURSIVE all_deps AS (
  SELECT DISTINCT
    Game.id,
    1 AS parent,
    0 AS level
  FROM Product Game
  WHERE Game.id = 3
  UNION ALL
  SELECT
    Dep.id,
    prev.id,
    level + 1
  FROM all_deps AS prev
  INNER JOIN ProductDependency AS Deplink
    ON prev.id = Deplink.primary_id
  INNER JOIN Product AS Dep
    ON Dep.id = Deplink.dependant_id
),
library_purchases AS (
  SELECT g.purchase_id
  FROM Gift AS g
  WHERE recipient_id = 1
  UNION ALL
  SELECT pur.id AS purchase_id
  FROM (SELECT * FROM Purchase AS pur WHERE pur.buyer_id = 1) AS pur
  LEFT JOIN Gift AS g
    ON pur.id = g.purchase_id
  WHERE g.purchase_id IS NULL
),
lib AS (
  SELECT DISTINCT ON (prod.id) prod.*, pur.date
  FROM library_purchases AS lib_pur
  INNER JOIN Purchase AS pur
    ON lib_pur.purchase_id = pur.id
  INNER JOIN Product AS prod
    ON pur.product_id = prod.id
)
SELECT id FROM all_deps
WHERE level != 0 AND id NOT IN (SELECT id FROM lib);
```

id
139

(1 row)

- Вывести все игры с набором тегов, аналогичным набору тегов игры с ID = 1

```
WITH product_tags AS (
  SELECT at.tag_id
  FROM AssignedTag AS at
  WHERE at.product_id = 1
), products_with_tags_count AS (
  SELECT product_id, COUNT(*) AS all_tags, COUNT(product_tags.tag_id) AS
    coincided_tags
  FROM AssignedTag AS at
  LEFT JOIN product_tags
    USING (tag_id)
  GROUP BY at.product_id
)
SELECT prod.*
FROM (SELECT * FROM products_with_tags_count WHERE products_with_tags_count.
  all_tags = products_with_tags_count.coincided_tags) AS
  products_with_no_other_tags
INNER JOIN (SELECT COUNT(*) AS needed_tags FROM product_tags) AS
  needed_tags_count
  ON needed_tags_count.needed_tags = products_with_no_other_tags.all_tags
INNER JOIN Product AS prod
  ON prod.id = products_with_no_other_tags.product_id;
```

id	name	description	price	reviews_count	rating_sum	publisher_id
1	Beretta	Mrs suddenly front bag. Group investment value feel. Father detail add. Wear ask cultural loss moment chair. Provide level back skin west sea. Direction same interest may measure traditional perhaps.	+\$10.51	152	466	18

(1 row)

- Вывести все подарки (Gift), сделанные пользователю с заданным id (ID=X), по убыванию даты.

```
SELECT g.id AS gift_id, g.title, g.message, p.product_id, p.date
FROM Gift AS g
INNER JOIN Purchase AS p
ON g.recipient_id = 280 AND g.purchase_id = p.id
ORDER BY date DESC;
```

gift_id	title	message	product_id	date
63	👤	Culture key even medical buy some require. Add onto expert sound eight thought could. Play likely military its.	179	2024-03-12 06:26:06.452315
68	👤	State ok would remember over. Economy career want itself most wall cell.	17	2024-03-08 14:59:39.149318
28	👤	Election company PM with the direction suffer. Member much law each federal.	59	2022-11-01 01:22:29.174241
		Ok gun government set least bill toward.		
		Foot wonder become ten. Lay personal follow position through.		

(3 rows)

5.6. Оценка размеров Базы Данных и каждого из файлов

Таблица	Атрибут	Тип данных	Размер одного, байт	Среднее кол-во	Общий размер, байт
Publisher	id	INT	4	44 тыс	112,288 млн
	name	VARCHAR(255)	510		
	description	TEXT	2000		
	bank_account	CHAR(34)	38		
User	id	INT	4	150 млн	231.9 млрд
	email	VARCHAR(255)	510		
	password	VARCHAR(255)	510		
	username	VARCHAR(255)	510		
	money	MONEY	8		
	publisher_id	INT	4		
Tag	id	INT	4	440	226.16 тыс
	name	VARCHAR(255)	510		
Product	id	INT	4	71 тыс	108,914 млн
	name	VARCHAR(255)	510		
	description	TEXT	1000		
	price	MONEY	8		

	reviews_count	INT	4		
	rating_sum	INT	4		
	publisher_id	INT	4		
AssignedTag	tag_id	INT	4	710 тыс	5,68 млн
	product_id	INT	4		
Purchase	id	BIGINT	8	8 млрд	192 млрд
	date	TIMESTAMP	8		
	buyer_id	INT	4		
	product_id	INT	4		
Gift	id	INT	4	100 млн	102,6 млрд
	title	VARCHAR(255)	510		
	message	TEXT	500		
	purchase_id	BIGINT	8		
	recipient_id	INT	4		
Review	id	INT	4	200 тыс	204,4 млн
	rating	SMALLINT	2		
	text	TEXT	1000		
	date	TIMESTAMP	8		
	writer_id	INT	4		
	subject_id	INT	4		
ProductDependency	primary_id	INT	4	71 тыс	284 тыс
	dependant_id	INT	4		

Средний размер базы данных 526,931 млрд байт, т.е. примерно 491 Gib

6. Приложения, Отчеты

Соотношение средних размеров данных таблиц и их вклад в общий размер базы данных.

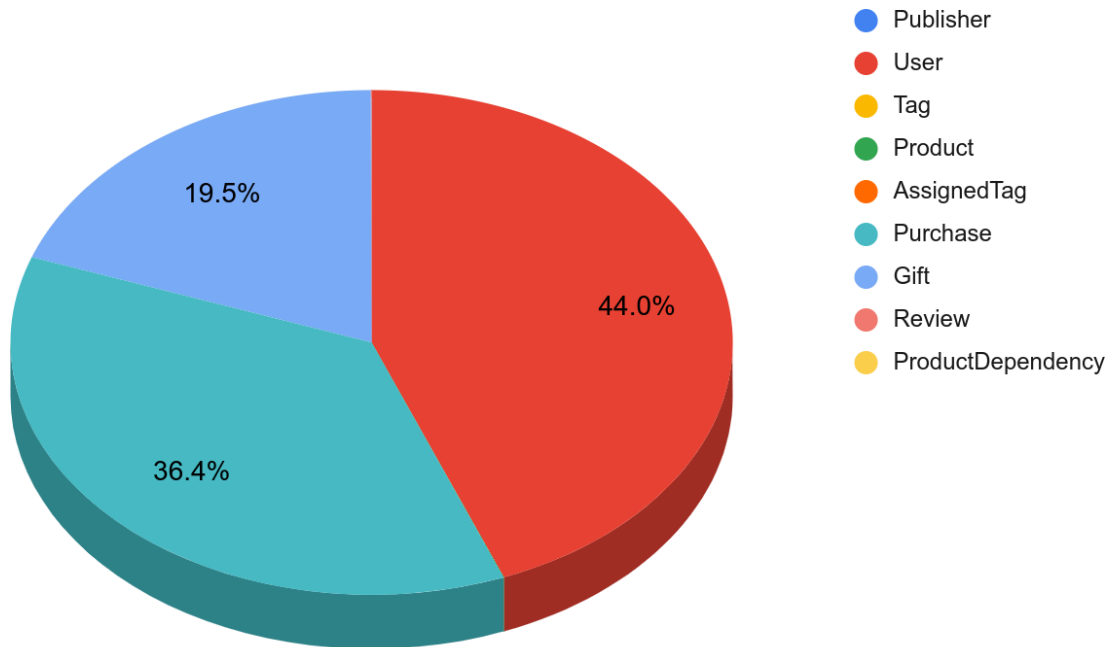


График зависимости времени выполнения SQL запроса на получение библиотеки пользователя от количества купленных на прямую товаров для одного из пользователей (не подарков). Общее размер библиотеки не зменный - 100000 товаров.

