

Домашнее задание №1

1. Исходная структура данных

Изначально в базе данных было две таблицы: **customer** и **transaction**, которые содержали ряд избыточных данных и транзитивных зависимостей.

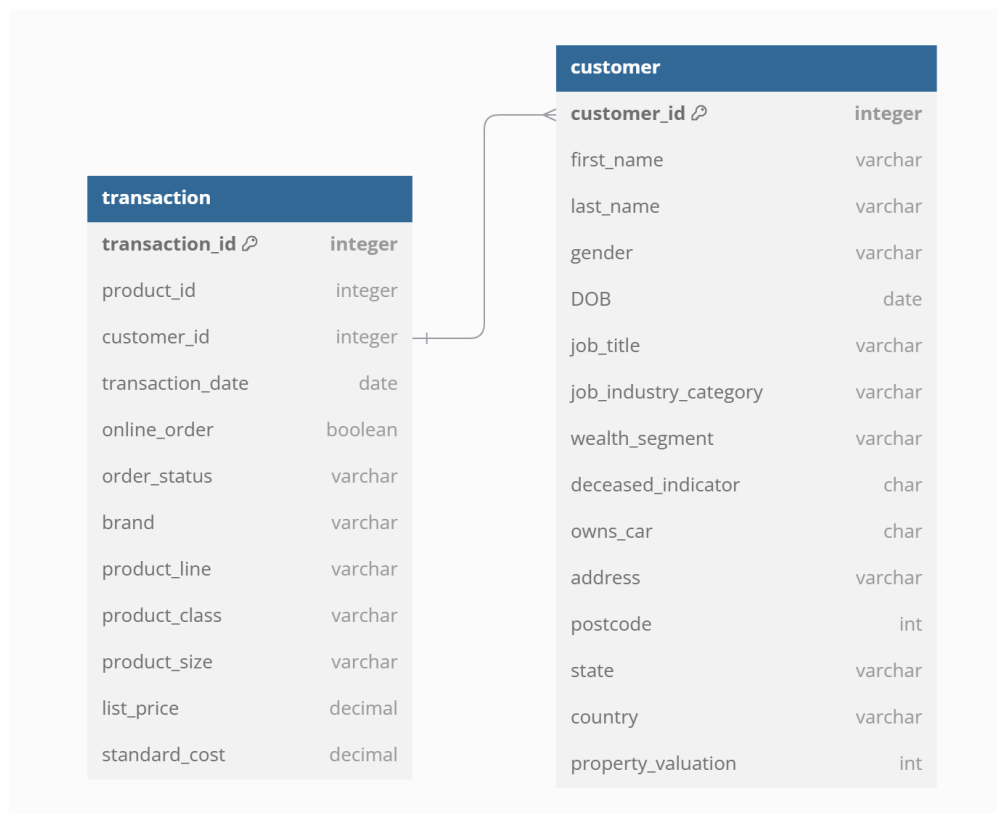


Рис. 1: Исходная структура данных

Отдельно важно отметить, что исходальные данные являлись **неконсистентными**. В рамках данного учебного задания не предпринималось мер по согласованию данных: несогласующиеся данные были просто удалены из базы. В общем случае такой подход, разумеется, некорректен, однако согласование данных на данном этапе потребовало бы приложение усилий, несоразмерных остальному заданию.

2. Почему исходная структура не находилась в нормальной форме?

Нарушение первой нормальной формы (1НФ)

- **Повторяющиеся группы данных:** Атрибуты `brand`, `product_line`, `product_class`, `product_size` дублировались в каждой записи транзакции для одного и того же продукта. Эти данные логически принадлежат таблице продуктов (`product`).

Нарушение второй нормальной формы (2НФ)

- В таблице `transaction` атрибуты `list_price` и `standard_cost` зависят **только** от `product_id`, а не от всего составного ключа `transaction_id`. Это частичная зависимость.

Нарушение третьей нормальной формы (3НФ)

- В таблице `customer` атрибуты `state` и `country` зависят от `postcode`, а не от `customer_id`. Это транзитивная зависимость.

3. Этапы нормализации

Шаг 1: Приведение к первой нормальной форме (1НФ)

Преобразование:

Удалены повторяющиеся группы данных: создана отдельная таблица `product` для хранения информации о продуктах (`brand`, `product_line`, `product_class`, `product_size`).

Результат:

Все данные стали атомарными, без повторяющихся групп.

Шаг 2: Приведение ко второй нормальной форме (2НФ)

Преобразование:

Устранили частичные зависимости: поля `list_price` и `standard_cost` перенесены в таблицу `product`, так как они зависят только от `product_id`.

Результат:

Каждый неключевой атрибут теперь зависит от полного первичного ключа.

Шаг 3: Приведение к третьей нормальной форме (3НФ)

Преобразование:

Устранили транзитивные зависимости: создана отдельная таблица `location`, содержащая уникальные комбинации `postcode`, `state` и `country`. Из таблицы `customer` удалены `state` и `country`, связь с таблицей `location` теперь осуществляется через `postcode`.

Результат:

Каждый неключевой атрибут зависит только от первичного ключа своей таблицы, и транзитивные зависимости устранены.

4. Итоговая структура базы данных

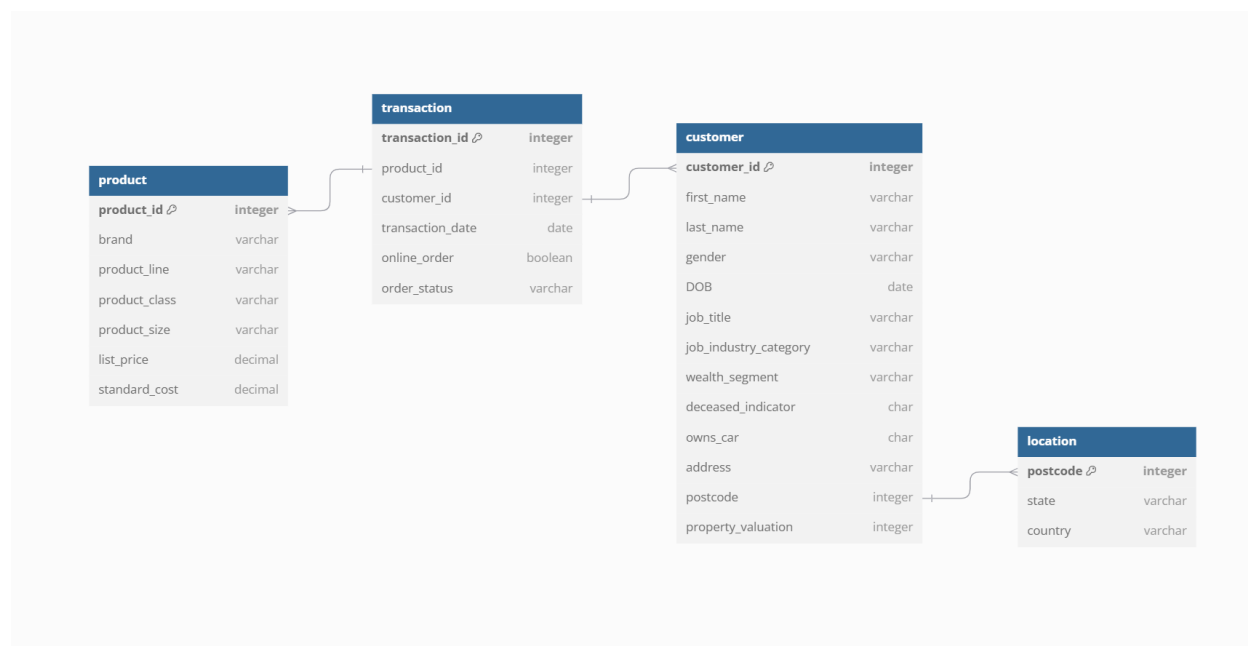


Рис. 2: Итоговая структура базы данных

После нормализации структура базы данных состоит из четырёх таблиц:

1. `customer` — хранит информацию о клиентах и их почтовых индексах.
2. `transaction` — содержит данные о транзакциях, ссылаясь на `customer` и `product`.

3. `product` — содержит информацию о продуктах, включая их характеристики, цену и себестоимость.
4. `location` — содержит уникальные данные о почтовых индексах, регионах и странах.

Почему текущая структура соответствует 3НФ?

- Все таблицы находятся в первой нормальной форме (1НФ) — данные атомарны, без повторяющихся групп.
- Все таблицы находятся во второй нормальной форме (2НФ) — каждый неключевой атрибут зависит от полного первичного ключа.
- Все таблицы находятся в третьей нормальной форме (3НФ) — транзитивные зависимости устранены, данные логически разделены между таблицами.

5. Скриншоты из DBeaver, подтверждающие загрузку данных





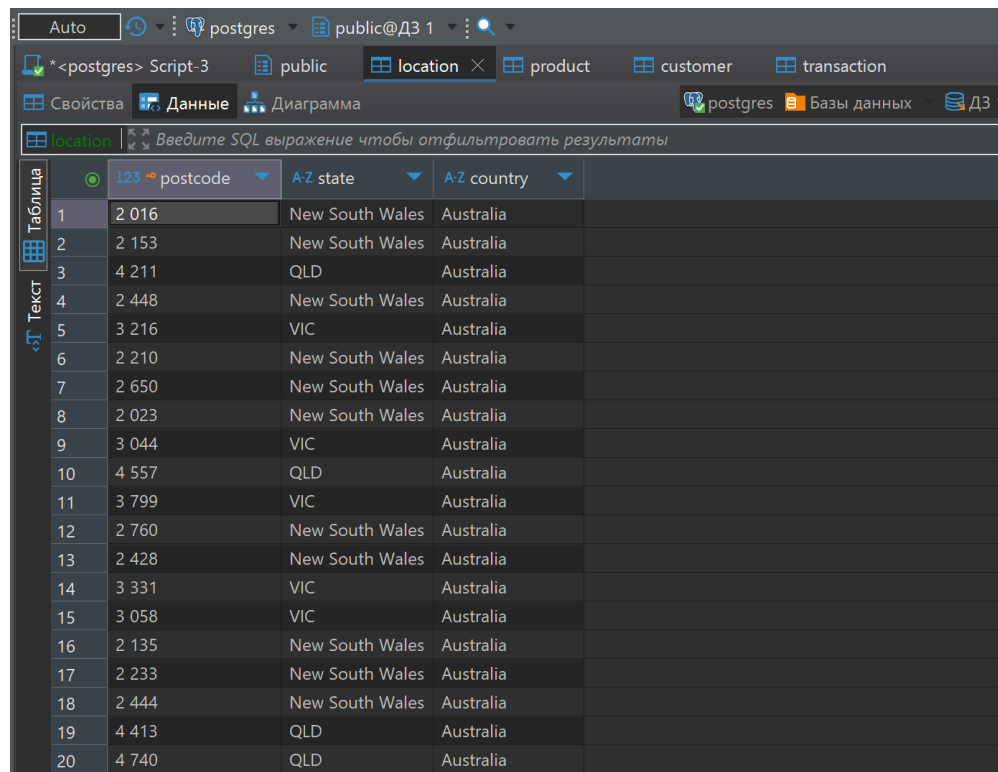
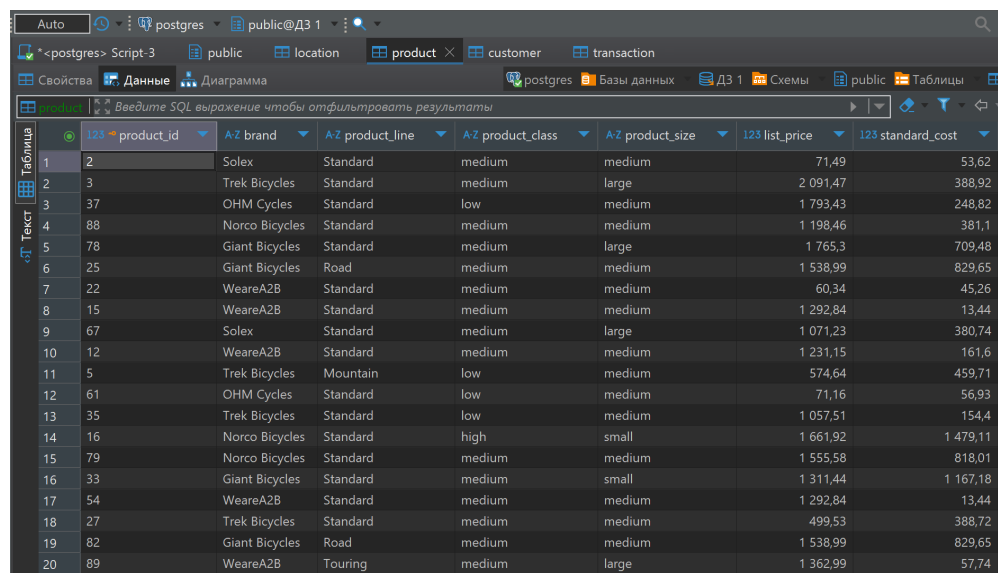
Название	ID объекта	Владелец	Табличное пространство	Примерное число строк
 customer	16 452	postgres		4 000
 location	16 447	postgres		874
 product	16 464	postgres		101
 transaction	16 469	postgres		19 997

Рис. 3: Список таблиц в базе данных



	123 postcode	A-Z state	A-Z country	
1	2 016	New South Wales	Australia	
2	2 153	New South Wales	Australia	
3	4 211	QLD	Australia	
4	2 448	New South Wales	Australia	
5	3 216	VIC	Australia	
6	2 210	New South Wales	Australia	
7	2 650	New South Wales	Australia	
8	2 023	New South Wales	Australia	
9	3 044	VIC	Australia	
10	4 557	QLD	Australia	
11	3 799	VIC	Australia	
12	2 760	New South Wales	Australia	
13	2 428	New South Wales	Australia	
14	3 331	VIC	Australia	
15	3 058	VIC	Australia	
16	2 135	New South Wales	Australia	
17	2 233	New South Wales	Australia	
18	2 444	New South Wales	Australia	
19	4 413	QLD	Australia	
20	4 740	QLD	Australia	

Рис. 4: Фрагмент таблицы location



	123 product_id	A-Z brand	A-Z product_line	A-Z product_class	A-Z product_size	123 list_price	123 standard_cost
1	2	Solex	Standard	medium	medium	71,49	53,62
2	3	Trek Bicycles	Standard	medium	large	2 091,47	388,92
3	37	OHM Cycles	Standard	low	medium	1 793,43	248,82
4	88	Norco Bicycles	Standard	medium	medium	1 198,46	381,1
5	78	Giant Bicycles	Standard	medium	large	1 765,3	709,48
6	25	Giant Bicycles	Road	medium	medium	1 538,99	829,65
7	22	WeareA2B	Standard	medium	medium	60,34	45,26
8	15	WeareA2B	Standard	medium	medium	1 292,84	13,44
9	67	Solex	Standard	medium	large	1 071,23	380,74
10	12	WeareA2B	Standard	medium	medium	1 231,15	161,6
11	5	Trek Bicycles	Mountain	low	medium	574,64	459,71
12	61	OHM Cycles	Standard	low	medium	71,16	56,93
13	35	Trek Bicycles	Standard	low	medium	1 057,51	154,4
14	16	Norco Bicycles	Standard	high	small	1 661,92	1 479,11
15	79	Norco Bicycles	Standard	medium	medium	1 555,58	818,01
16	33	Giant Bicycles	Standard	medium	small	1 311,44	1 167,18
17	54	WeareA2B	Standard	medium	medium	1 292,84	13,44
18	27	Trek Bicycles	Standard	medium	medium	499,53	388,72
19	82	Giant Bicycles	Road	medium	medium	1 538,99	829,65
20	89	WeareA2B	Touring	medium	large	1 362,99	57,74

Рис. 5: Фрагмент таблицы product

customer_id	first_name	last_name	gender	dob	job_title	job_industry_category	wealth_segment	deceased_indicator
1	Laraine	Madendorp	F	1953-10-12	Executive Secretary	Health	Mass Customer	N
2	BI	Bockman	Male	1980-12-16	Administrative Officer	Financial Services	Mass Customer	N
3	Arlin	Dearle	Male	1954-01-20	Recruiting Manager	Property	Mass Customer	N
4	Talbot		Male	1961-10-03		IT	Mass Customer	N
5	Sheila-kathryn	Calton	Female	1977-05-13	Senior Editor	n/a	Affluent Customer	N
6	Curr	Duckhouse	Male	1966-09-16		Retail	High Net Worth	N
7	Fina	Merali	Female	1976-02-23		Financial Services	Affluent Customer	N
8	Rod	Inder	Male	1962-03-30	Media Manager I	n/a	Mass Customer	N
9	Mala	Lind	Female	1973-03-10	Business Systems Development Analyst	Agriculture	Affluent Customer	N
10	Florentze	Birdall	Female	1986-10-11	Senior Quality Engineer	Financial Services	Mass Customer	N
11	Unah	Blatt	Male	1984-04-30		Property	Mass Customer	N
12	Sawyer	Flattman	Male	1984-07-21	Nuclear Power Engineer	Manufacturing	Mass Customer	N
13	Gabriele	Norcross	Male	1955-02-15	Developer I	Financial Services	High Net Worth	N
14	Rayshell	Kitterman	Female	1983-03-25	Account Executive	Financial Services	Affluent Customer	N
15	Enroll	Radage	Male	2000-07-13	Junior Executive	Manufacturing	Mass Customer	N
16	Harlin	Parr	Male	1977-02-27	Media Manager IV	n/a	Mass Customer	N
17	Heath	Faraday	Male	1962-03-19	Sales Associate	n/a	Affluent Customer	N
18	Marje	Neasham	Female	1967-07-06	Professor	n/a	Affluent Customer	N
19	Sorcha	Keyson	Female	2001-04-15	Geological Engineer	Manufacturing	High Net Worth	N
20	Basile	Firth	Male	1980-08-13	Project Manager	Manufacturing	Mass Customer	N

Рис. 6: Фрагмент таблицы customer

product_id	brand	product_line	product_class	product_size	list_price	standard_cost
2	Solex	Standard	medium	medium	71,49	53,62
3	Trek Bicycles	Standard	medium	large	2 091,47	388,92
37	OHM Cycles	Standard	low	medium	1 793,43	248,82
88	Norco Bicycles	Standard	medium	medium	1 198,46	381,1
78	Giant Bicycles	Standard	medium	large	1 765,3	709,48
25	Giant Bicycles	Road	medium	medium	1 538,99	829,65
22	WeareA2B	Standard	medium	medium	60,34	45,26
15	WeareA2B	Standard	medium	medium	1 292,84	13,44
67	Solex	Standard	medium	large	1 071,23	380,74
12	WeareA2B	Standard	medium	medium	1 231,15	161,6
5	Trek Bicycles	Mountain	low	medium	574,64	459,71
61	OHM Cycles	Standard	low	medium	71,16	56,93
35	Trek Bicycles	Standard	low	medium	1 057,51	154,4
16	Norco Bicycles	Standard	high	small	1 661,92	1 479,11
79	Norco Bicycles	Standard	medium	medium	1 555,58	818,01
33	Giant Bicycles	Standard	medium	small	1 311,44	1 167,18
54	WeareA2B	Standard	medium	medium	1 292,84	13,44
27	Trek Bicycles	Standard	medium	medium	499,53	388,72
82	Giant Bicycles	Road	medium	medium	1 538,99	829,65
89	WeareA2B	Touring	medium	large	1 362,99	57,74

Рис. 7: Фрагмент таблицы transition

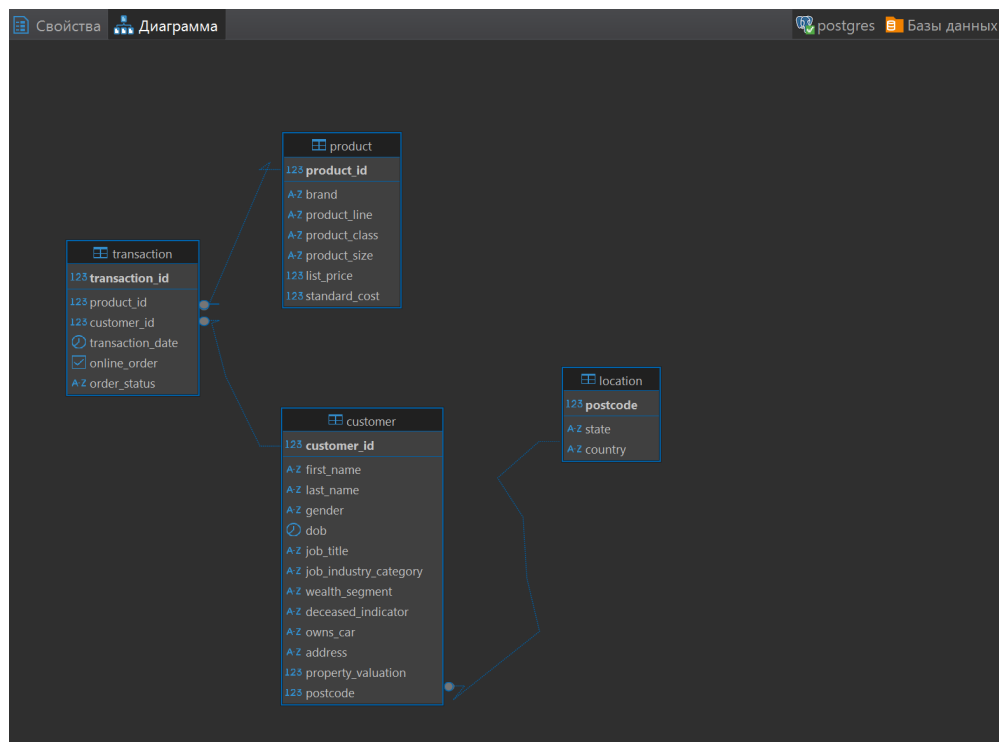


Рис. 8: Диаграмма базы данных