

# Отбор данных

1. Выбор данных из нескольких таблиц
2. Группировка данных
3. Отбор по сгруппированным данным
4. База данных «Касса»

# 1 Выбор данных из нескольких таблиц

## Универсальное отношение

Организация	ФИО работника	Пол	Дата рождения
Самараводоканал	Сафонов Александр Васильевич	м	01.05.1976
Самараводоканал	Тихонов Виктор Михайлович	м	21.07.1966
ООО "СК Монолит"	Носик Михаил Гаврилович	м	01.05.1976
ООО "СК Монолит"	Филиппова Анна Владимирович	ж	28.01.1963
ООО "СК Монолит"	Андрияшкин Владимир Анатольевич	м	11.05.1956
РН-Информ	Меркулова Ольга Сергеевна	ж	25.04.1958
РН-Информ	Ханина Любовь Михайловна	ж	03.03.1953
...			

# 1 Выбор данных из нескольких таблиц

## Хранение данных в СУБД

<u>Номер организации</u>	Организация
1	Самараводоканал
2	ООО "СК Монолит"
3	РН-Информ

<u>Номер организации</u>	<u>Номер работника</u>	ФИО работника	Пол	Дата рождения
1	1	Сафонов Александр Васильевич	м	01.05.1976
1	2	Тихонов Виктор Михайлович	м	21.07.1966
2	1	Носик Михаил Гаврилович	м	01.05.1976
2	2	Филиппова Анна Владимирович	ж	28.01.1963
2	3	Андрияшкин Владимир Анатольевич	м	11.05.1956
3	1	Меркулова Ольга Сергеевна	ж	25.04.1958
3	2	Ханина Любовь Михайловна	ж	03.03.1953

# 1 Выбор данных из нескольких таблиц

Таблицы:

- **firm**

**id\_f** – идентификатор организации

**name** – наименование организации

**id\_f** – простой  
ключ

- **worker**

**id\_f** – идентификатор организации

**id\_w** – идентификатор сотрудника

**name** – ФИО

**sex** – пол

**bdate** – дата рождения

**id\_f , id\_w** – составной ключ

# 1 Выбор данных из нескольких таблиц

Выбор данных из двух таблиц: вывести в одну строку *Наименование организации, ФИО, Пол, Дату рождения*, где идентификатор организации *id\_f* таблицы *firm* равен идентификатору организации *id\_f* таблицы *Worker*

```
SELECT
    firm.name, worker.name, worker.sex,
    worker.bdate
FROM
    firm, worker
WHERE
    firm.id_f = worker.id_f
```

# 1 Выбор данных из нескольких таблиц

Удобнее использовать псевдонимы таблиц

```
SELECT
    a.name, b.name, b.sex, b.bdate
FROM
    firm a, worker b
WHERE
    a.id_f=b.id_f
```

a – псевдоним таблицы firm

b – псевдоним таблицы

worker

# 1 Выбор данных из нескольких таблиц

Выведем сотрудников женского пола с указанием наименования организации, ФИО, пола и даты рождения

```
SELECT
    a.name, b.name, b.sex, b.bdate
FROM
    firm a, worker b
WHERE
    a.id_f=b.id_f and b.sex='Ж'
```

# 1 Выбор данных из нескольких таблиц

Выведем всех людей, отсортировав по ФИО

```
SELECT
    a.name, b.name, b.sex, b.bdate
FROM
    firm a, worker b
WHERE
    a.id_f=b.id_f
ORDER BY
    b.name
```



# 1 Выбор данных из нескольких таблиц

Выведем всех людей, отсортировав по наименованию организации и ФИО

```
SELECT
    a.name, b.name, b.sex, b.bdate
FROM
    firm a, worker b
WHERE
    a.id_f=b.id_f
ORDER BY
    a.name, b.name
```

# Отбор данных

1. Выбор данных из нескольких таблиц
2. **Группировка данных**
3. Отбор по сгруппированным данным
4. База данных «Касса»

## 2 Группировка данных

Группировка – объединение данных по заданному критерию для выделения требуемой информации из данных при помощи агрегатных функций

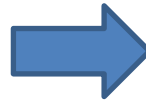
Некоторые агрегатные функции:

- count
- max
- min
- avg
- sum

## 2 Группировка данных

Поле 1	Поле 2	Поле 3
п1_знач1	п2_знач3	п3_знач1
п1_знач2	п2_знач3	п3_знач1
п1_знач2	п2_знач1	п3_знач2
п1_знач1	п2_знач1	п3_знач3
п1_знач3	п2_знач3	п3_знач3
п1_знач3	п2_знач2	п3_знач3
п1_знач2	п2_знач1	п3_знач3
п1_знач1	п2_знач1	п3_знач3

группировка  
по «поле 1»



Поле 1
п1_знач1
п1_знач2
п1_знач3



группировка  
по «поле 1»,  
«поле 3»

Поле 1	Поле 3
п1_знач1	п3_знач1
п1_знач2	п3_знач1
п1_знач2	п3_знач2
п1_знач1	п3_знач3
п1_знач3	п3_знач3
п1_знач2	п3_знач3
п1_знач1	п3_знач3

## 2 Группировка данных

Сгруппировать людей по фамилиям, выяснить сколько людей проживают с одинаковыми фамилиями

```
SELECT
  surname, count(*) 'Кол-во'
FROM
  people
group by surname
```

	surname	кол-во
4	АМАТОВ	1
5	ААСА	2
6	АБ	1
7	АБАБЕКЯН	1
8	АБАБИЙ	19
9	АБАБКОВ	10
10	АБАБКОВА	15
11	АБАГЯН	1
12	АБАДЖАН	1
13	АБАДЖИЕВА	1
14	АБАДЖЯН	17
15	АБАДОВСКАЯ	1
16	АБАДОВСКИЙ	1
17	АБАЕВ	67
18	АБАЕВА	77
19	АБАЗА	12
20	АБАЗАЕВА	1
21	АБАЗАРОВА	2
22	АБАЗДРЕВ	3
23	АБАЗДРЕВА	2
24	АБАЗЕВ	7
25	АБАЗЕВА	7
26	АБАЗИЯН	1
27	АБАЗОВ	3
28	АБАЗОВА	4
29	АБАЗЬЕВ	7
30	АБАЗЬЕВА	12

# Отбор данных

1. Выбор данных из нескольких таблиц
2. Группировка данных
3. **Отбор по сгруппированным данным**
4. База данных «Касса»

### 3 Отбор по сгруппированным данным

**HAVING** определяет условие поиска для группы, обычно используется в предложении GROUP BY. Когда GROUP BY не используется, предложение HAVING работает так же, как и предложение WHERE.

### 3 Отбор по сгруппированным данным

Вывести только те фамилии, под которыми проживают больше 4000 человек (отсортировать по убыванию кол-ва)

```
SELECT
    surname, count(*) 'Кол-во'
FROM
    people
group by
    surname
HAVING
    count(*)>4000
ORDER BY
    count(*) desc
```

	surname	Кол-во
1	ИВАНОВА	11692
2	КУЗНЕЦОВА	10751
3	ИВАНОВ	9443
4	КУЗНЕЦОВ	8571
5	ПОПОВА	6692
6	ПЕТРОВА	6042
7	МАКАРОВА	5423
8	ПОПОВ	5390
9	ВАСИЛЬЕВА	5180
10	ПЕТРОВ	4865
11	ПАВЛОВА	4755
12	ЗАХАРОВА	4690
13	ЕГОРОВА	4484
14	НИКОЛАЕВА	4427
15	СЕМЕНОВА	4345
16	МАКАРОВ	4332
17	КОЗЛОВА	4328
18	ВОЛКОВА	4279
19	СМИРНОВА	4255
20	ВАСИЛЬЕВ	4201
21	АНДРЕЕВА	4177
22	МОРОЗОВА	4174
23	ФРОЛОВА	4097
24	КУЗЬМИНА	4064
25	БОРИСОВА	4032



### 3 Отбор по сгруппированным данным

Вывести только те фамилии, под которыми проживают больше 4000 человек, а так же вывести дату рождения самого старого

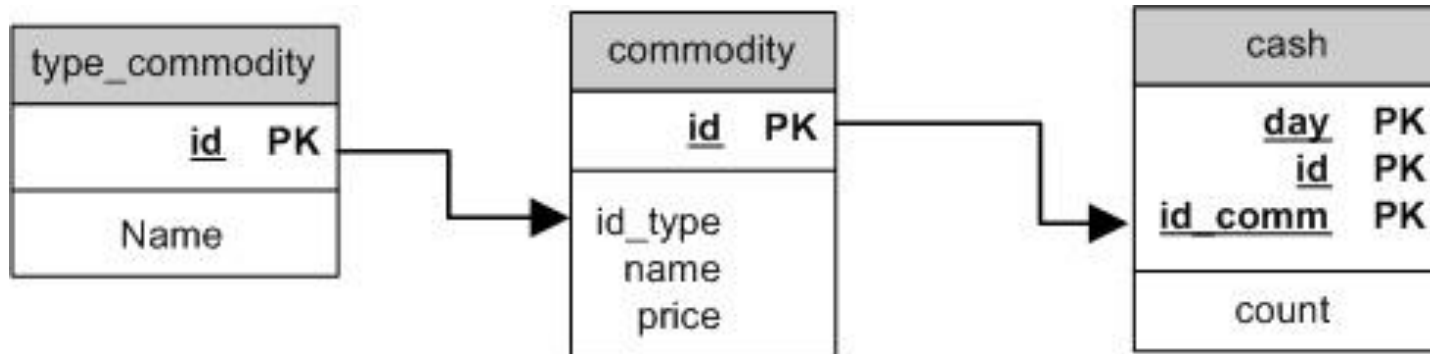
```
SELECT
    surname, count(*) 'Кол-во',
    min(birthday)'др'
FROM
    people
group by
    surname
HAVING
    count(*)>4000
```

	surname	Кол-во	др
1	АНДРЕЕВА	4165	1910-01-01 00:00:00.000
2	БОРИСОВА	4001	1905-11-17 00:00:00.000
3	ВАСИЛЬЕВ	4194	1913-03-16 00:00:00.000
4	ВАСИЛЬЕВА	5158	1912-02-08 00:00:00.000
5	ВОЛКОВА	4263	1911-07-17 00:00:00.000
6	ЕГОРОВА	4475	1910-10-25 00:00:00.000
7	ЗАХАРОВА	4657	1912-08-04 00:00:00.000
8	ИВАНОВ	9367	1916-08-10 00:00:00.000
9	ИВАНОВА	11625	1909-03-14 00:00:00.000
10	КОЗЛОВА	4313	1910-09-28 00:00:00.000
11	КУЗНЕЦОВ	8533	1913-02-23 00:00:00.000
12	КУЗНЕЦОВА	10731	1906-06-24 00:00:00.000
13	КУЗЬМИНА	4053	1910-06-30 00:00:00.000
14	МАКАРОВ	4285	1915-08-10 00:00:00.000
15	МАКАРОВА	5415	1909-01-01 00:00:00.000
16	МОРОЗОВА	4158	1907-10-27 00:00:00.000
17	НИКОЛАЕВА	4406	1912-01-28 00:00:00.000
18	ПАВЛОВА	4745	1911-01-01 00:00:00.000
19	ПЕТРОВ	4827	1912-10-22 00:00:00.000
20	ПЕТРОВА	6015	1911-02-02 00:00:00.000
21	ПОПОВ	5333	1912-07-19 00:00:00.000
22	ПОПОВА	6675	1909-05-12 00:00:00.000
23	СЕМЕНОВА	4311	1909-11-02 00:00:00.000
24	СМИРНОВА	4223	1917-01-26 00:00:00.000
25	ФРОЛОВА	4069	1911-02-01 00:00:00.000

# Отбор данных

1. Выбор данных из нескольких таблиц
2. Группировка данных
3. Отбор по сгруппированным данным
4. **База данных «Касса»**

## 4 База данных «Касса»



**type\_commodity** – тип товара (id – идентификатор типа товара; name – наименование типа товара)

**commodity** – перечень товаров (id – идентификатор товара; id\_type – тип товара; name – наименование товара; price – цена за единицу товара)

**cash** – кассовые чек (day – дата чека; id – номер чека; id\_comm – идентификатор товара; count – кол-во товара)

# 4 База данных «Касса»

type\_commodity

id	name
1	Овощи
2	Фрукты
3	Молочная продукция
4	Колбасы

commodity

id	id_type	name	price
1	1	Томаты на ветке	150.0000
2	1	Томаты черри	250.0000
3	1	Огурцы среднеплодые	100.0000
4	1	Огурцы короткоплод...	120.0000
5	2	Яблоки Сергиана	50.0000
6	4	Сервелат российский	300.0000
7	4	Сервелат коньячный	280.0000
8	1	Картофель рус	30.0000
9	1	Картофель мытый	40.0000
10	2	Груши	90.0000
11	2	Бананы	40.0000
12	2	Лимоны	70.0000
13	4	Сосиски любительский	270.0000
14	2	Киви	120.0000
15	4	Карбонат михайловс...	290.0000
16	3	Сыр маасдам	310.0000
17	3	Сыр гауда	320.0000
19	3	Сыр Король артур	280.0000
20	3	Сыр камамбер	110.0000
21	2	Капуста	50.0000
22	3	Молоко 3.2%	48.0000
23	3	Кефир 3.2 %	44.0000

# 4 База данных «Касса»

cash

day	id	id_comm	count
24.04.2013 ...	1	1	0.5
24.04.2013 ...	1	2	0.3
24.04.2013 ...	1	4	1
24.04.2013 ...	1	12	0.2
24.04.2013 ...	1	22	1
24.04.2013 ...	1	8	3.2
24.04.2013 ...	2	2	0.5
24.04.2013 ...	2	4	2
24.04.2013 ...	2	12	0.3
24.04.2013 ...	2	6	1
24.04.2013 ...	2	22	2
24.04.2013 ...	3	1	1.2
24.04.2013 ...	3	5	1.4
24.04.2013 ...	3	8	2.8
24.04.2013 ...	4	8	3
25.04.2013 ...	5	11	3
25.04.2013 ...	5	5	2
25.04.2013 ...	5	4	0.5
25.04.2013 ...	5	22	2
25.04.2013 ...	6	1	1.6
25.04.2013 ...	6	4	1.3
25.04.2013 ...	6	23	1
25.04.2013 ...	6	6	2
25.04.2013 ...	7	1	1.5
25.04.2013 ...	7	12	0.3

## 4 База данных «Касса»

Запросим данные о продаже каждой единицы товара по каждому чеку с указанием всех известных данных (дата продажи, номер чека, тип товара, наименование товара, цена, кол-во)

## 4 База данных «Касса»

**SELECT**

a.day, a.id, c.name,  
b.name, b.price, a.count

**FROM**

cash a,  
commodity b,  
type\_commodity c

**Where**

a.id\_comm = b.id and  
c.id=b.id\_type

	day	id	name	name	price	count
1	2013-04-24 ...	1	Овощи	Томаты на ветке	150.00	0.5
2	2013-04-24 ...	1	Овощи	Томаты черри	250.00	0.3
3	2013-04-24 ...	1	Овощи	Огурцы короткопло...	120.00	1
4	2013-04-24 ...	1	Фрукты	Лимоны	70.00	0.2
5	2013-04-24 ...	1	Молочная продук...	Молоко 3.2%	48.00	1
6	2013-04-24 ...	1	Овощи	Картофель рус	30.00	3.2
7	2013-04-24 ...	2	Овощи	Томаты черри	250.00	0.5
8	2013-04-24 ...	2	Овощи	Огурцы короткопло...	120.00	2
9	2013-04-24 ...	2	Фрукты	Лимоны	70.00	0.3
10	2013-04-24 ...	2	Колбасы	Сервелат российск...	300.00	1
11	2013-04-24 ...	2	Молочная продук...	Молоко 3.2%	48.00	2
12	2013-04-24 ...	3	Овощи	Томаты на ветке	150.00	1.2
13	2013-04-24 ...	3	Фрукты	Яблоки Сергиана	50.00	1.4
14	2013-04-24 ...	3	Овощи	Картофель рус	30.00	2.8
15	2013-04-24 ...	4	Овощи	Картофель рус	30.00	3
16	2013-04-25 ...	5	Фрукты	Бананы	40.00	3
17	2013-04-25 ...	5	Фрукты	Яблоки Сергиана	50.00	2
18	2013-04-25 ...	5	Овощи	Огурцы короткопло...	120.00	0.5
19	2013-04-25 ...	5	Молочная продук...	Молоко 3.2%	48.00	2
20	2013-04-25 ...	6	Овощи	Томаты на ветке	150.00	1.6
21	2013-04-25 ...	6	Овощи	Огурцы короткопло...	120.00	1.3
22	2013-04-25 ...	6	Молочная продук...	Кефир 3.2 %	44.00	1
23	2013-04-25 ...	6	Колбасы	Сервелат российск...	300.00	2
24	2013-04-25 ...	7	Овощи	Томаты на ветке	150.00	1.5
25	2013-04-25 ...	7	Фрукты	Лимоны	70.00	0.3



## 4 База данных «Касса»

Сколько всего позиций было продано за все дни?

```
SELECT  
count(*)  
FROM  
cash
```

	day	id	name	price	count
1	2013-04-24 ...	1	Томаты на ветке	150.00	0.5
2	2013-04-24 ...	1	Томаты черри	250.00	0.3
3	2013-04-24 ...	1	Огурцы короткоплодые	120.00	1
4	2013-04-24 ...	1	Лимоны	70.00	0.2
5	2013-04-24 ...	1	Молоко 3.2%	48.00	1
6	2013-04-24 ...	1	Картофель рус	30.00	3.2
7	2013-04-24 ...	2	Томаты черри	250.00	0.5
8	2013-04-24 ...	2	Огурцы короткоплодые	120.00	2
9	2013-04-24 ...	2	Лимоны	70.00	0.3
10	2013-04-24 ...	2	Сервелат российский	300.00	1
11	2013-04-24 ...	2	Молоко 3.2%	48.00	2
12	2013-04-24 ...	3	Томаты на ветке	150.00	1.2
13	2013-04-24 ...	3	Яблоки Сергиана	50.00	1.4
14	2013-04-24 ...	3	Картофель рус	30.00	2.8
15	2013-04-24 ...	4	Картофель рус	30.00	3
16	2013-04-25 ...	5	Бананы	40.00	3
17	2013-04-25 ...	5	Яблоки Сергиана	50.00	2
18	2013-04-25 ...	5	Огурцы короткоплодые	120.00	0.5
19	2013-04-25 ...	5	Молоко 3.2%	48.00	2
20	2013-04-25 ...	6	Томаты на ветке	150.00	1.6
21	2013-04-25 ...	6	Огурцы короткоплодые	120.00	1.3
22	2013-04-25 ...	6	Кефир 3.2 %	44.00	1
23	2013-04-25 ...	6	Сервелат российский	300.00	2
24	2013-04-25 ...	7	Томаты на ветке	150.00	1.5
25	2013-04-25 ...	7	Лимоны	70.00	0.3

**Ответ: 25**



## 4 База данных «Касса»

Сколько всего позиций было продано в рамках каждого дня?

```
SELECT  
day, count(*)  
FROM  
cash  
group by  
day
```

	day	(Отсутствует имя столбца)
1	2013-04-24 00:00:00.000	15
2	2013-04-25 00:00:00.000	10

## 4 База данных «Касса»

Рассчитать итоговую сумму по каждому чеку?

1) Рассчитаем расходы по каждой позиции

**SELECT**

a.day, a.id, b.name, b.price,  
a.count, b.price\*a.count

**FROM**

cash a, commodity b

**where**

a.id\_comm = b.id

	day	id	name	price	count	всего
1	2013-04-24 ...	1	Томаты на ветке	150.00	0.5	75
2	2013-04-24 ...	1	Томаты черри	250.00	0.3	75
3	2013-04-24 ...	1	Огурцы короткоплодые	120.00	1	120
4	2013-04-24 ...	1	Лимоны	70.00	0.2	14
5	2013-04-24 ...	1	Молоко 3.2%	48.00	1	48
6	2013-04-24 ...	1	Картофель рус	30.00	3.2	96
7	2013-04-24 ...	2	Томаты черри	250.00	0.5	125
8	2013-04-24 ...	2	Огурцы короткоплодые	120.00	2	240
9	2013-04-24 ...	2	Лимоны	70.00	0.3	21
10	2013-04-24 ...	2	Сервелат российский	300.00	1	300
11	2013-04-24 ...	2	Молоко 3.2%	48.00	2	96
12	2013-04-24 ...	3	Томаты на ветке	150.00	1.2	180
13	2013-04-24 ...	3	Яблоки Сергиана	50.00	1.4	70
14	2013-04-24 ...	3	Картофель рус	30.00	2.8	84
15	2013-04-24 ...	4	Картофель рус	30.00	3	90
16	2013-04-25 ...	5	Бананы	40.00	3	120
17	2013-04-25 ...	5	Яблоки Сергиана	50.00	2	100
18	2013-04-25 ...	5	Огурцы короткоплодые	120.00	0.5	60
19	2013-04-25 ...	5	Молоко 3.2%	48.00	2	96
20	2013-04-25 ...	6	Томаты на ветке	150.00	1.6	240
21	2013-04-25 ...	6	Огурцы короткоплодые	120.00	1.3	156
22	2013-04-25 ...	6	Кефир 3.2 %	44.00	1	44
23	2013-04-25 ...	6	Сервелат российский	300.00	2	600
24	2013-04-25 ...	7	Томаты на ветке	150.00	1.5	225
25	2013-04-25 ...	7	Лимоны	70.00	0.3	21

## 4 База данных «Касса»

2) Сгруппируем данные и применим функцию sum

```
SELECT  
a.day, a.id,  
sum(b.price*a.count) 'Итого'  
FROM  
cash a, commodity b  
where  
a.id_comm = b.id  
group by  
a.day, a.id
```

	day	id	итого
1	2013-04-24 ...	1	428
2	2013-04-24 ...	2	782
3	2013-04-24 ...	3	334
4	2013-04-24 ...	4	90
5	2013-04-25 ...	5	376
6	2013-04-25 ...	6	1040
7	2013-04-25 ...	7	246

## 4 База данных «Касса»

Какой самый популярный товар? Отсортировать по популярности по убыванию

**SELECT**

b.name, count(\*)

**FROM**

cash a, commodity b

**where**

a.id\_comm = b.id

**group by** b.name

**order by** count(\*) desc

	name	(Отсутствует имя столбца)
1	Огурцы короткоплодые	4
2	Томаты на ветке	4
3	Картофель рус	3
4	Лимоны	3
5	Молоко 3.2%	3
6	Томаты черри	2
7	Яблоки Сергиана	2
8	Сервелат российский	2
9	Кефир 3.2 %	1
10	Бананы	1

## 4 База данных «Касса»

Вывести только те товары, популярность которых больше 2. Отсортировать по популярности по убыванию

**SELECT**

b.name, count(\*)

**FROM**

cash a, commodity b

**where**

a.id\_comm = b.id

**group by**

b.name

**having**

count(\*) >=2

**order by**

count(\*) desc

	name	(Отсутствует имя столбца)
1	Огурцы короткоплодые	4
2	Томаты на ветке	4
3	Картофель рус	3
4	Лимоны	3
5	Молоко 3.2%	3
6	Сервелат российский	2
7	Томаты черри	2
8	Яблоки Сергиана	2

## 4 База данных «Касса»

Вывести самый популярный тип товара, отсортировать по популярности по убыванию

**SELECT**

c.name, count(\*)

**FROM**

cash a,

commodity b,

type\_commodity c

**where**

a.id\_comm = b.id and

c.id=b.id\_type

**group by**

c.name

**order by**

count(\*) desc

	name	(Отсутствует имя столбца)
1	Овощи	13
2	Фрукты	6
3	Молочная продукция	4
4	Колбасы	2

# Контрольная точка 1

Состав теста:

- 1) Структура SQL-запроса, 1 шт (низкая сложность);
- 2) Типы данных MS SQL, 1 шт (низкая сложность);
- 3) Запрос с where , 2 (средняя сложность).

Задания низкой сложности – 16.7% рейтинга

Задания средней сложности – 33.3% рейтинга

Время выполнения теста – 15 минут.