

Технология хранения, поиска и сортировки информации

Базы данных



База данных (БД) -



- файл специального формата, позволяющий в упорядоченном виде хранить данные о группе объектов, обладающих одинаковым набором свойств.
 - БД книжного фонда библиотеки,
 - БД кадрового состава учреждения,
 - БД современных эстрадных песен и др.



Виды БД:



• **Фактографические:**

Содержат краткие сведения об описываемых объектах, представленные в строго определенном формате (библиотека, состав учреждения).

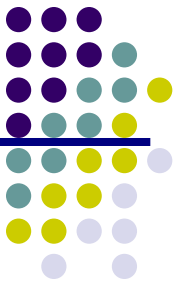


• **Документальные :**

Содержат обширную информацию самого разного типа: текст, графику, звук, мультимедиа.



Определения



База данных (БД) – это хранилище данных о некоторой предметной области, организованное в виде специальной структуры.

Важно:

- данные о некоторой области (не обо всем)
- упорядоченные

Система управления базой данных (СУБД) – это программное обеспечение для работы с БД.

Функции:

- поиск информации в БД
- выполнение несложных расчетов
- вывод отчетов на печать
- редактирование БД

Информационная система – это БД + СУБД.

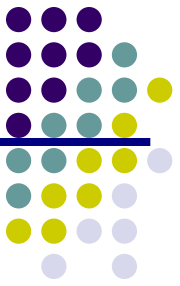


Информационная система -

- это совокупность базы данных и всего комплекса аппаратно-программных средств для ее хранения, изменения и поиска информации, для взаимодействия с пользователем.
- Система продажи билетов на пассажирские поезда и самолеты,
- WWW.



Типы баз данных



- **табличные БД**
данные в виде одной таблицы
- **сетевые БД**
набор узлов, в которых каждый может быть связан с каждым.
- **иерархические БД**
в виде многоуровневой структуры
- **реляционные БД (99,9%)**
набор взаимосвязанных таблиц

Типы БД:



1. Табличные (реляционные) – содержат перечень объектов с одинаковыми свойствами.

№ п/п	Название	Тип процессора	Оперативная память (Мбайт)
1	Compaq	Celeron	64
2	Dell	Pentium III	128
3	IBM	Pentium 4	256

Табличные БД

Модель – картотека

Примеры:

- записная книжка
- каталог в библиотеке

Петров Вася

Суворовский пр., д. 32, кв. 11
275-75-75

ПОЛЯ

записи

Фамилия	Имя	Адрес	Телефон
Петров	Вася	Суворовский пр., д. 32, кв. 11	275-75-75
Иванов	Дима	Кирочная ул., д.25, кв.12	276-76-76



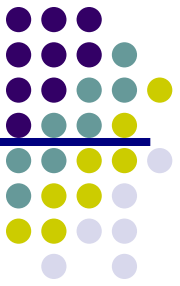
- 1) самая простая структура
- 2) все другие типы БД используют таблицы



во многих случаях – дублирование данных:

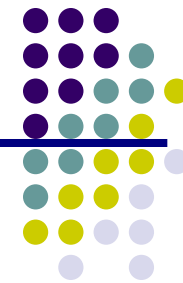
А.С. Пушкин	Сказка о царе Салтане	20 стр.
А.С. Пушкин	Сказка о золотом петушке	12 стр.

Табличные БД



- 1. Количество полей определяется разработчиком и не может изменяться пользователем.**
- 2. Любое поле должно иметь уникальное имя.**
- 3. Поля могут иметь различный тип:**
 - строка символов (длиной до 255 символов)
 - вещественное число (с дробной частью)
 - целое число
 - денежная сумма
 - дата, время, дата и время
 - логическое поле (истина или ложь, да или нет)
 - многострочный текст (MEMO)
 - рисунок, звук или другой объект (объект OLE)
- 4. Поля могут быть обязательными для заполнения или нет.**
- 5. Таблица может содержать сколько угодно записей (это количество ограничено только объемом диска); записи можно добавлять, удалять, редактировать, сортировать, искать.**

Иерархическая БД



Иерархическая БД – это набор данных в виде многоуровневой структуры.

Прайс-лист:

Продавец (уровень 1)

Кей

Товар (уровень 2)

Мониторы

Принтеры

Изготовитель (уровень 3)

Sony

Phillips

Samsung

Модель (уровень 4)

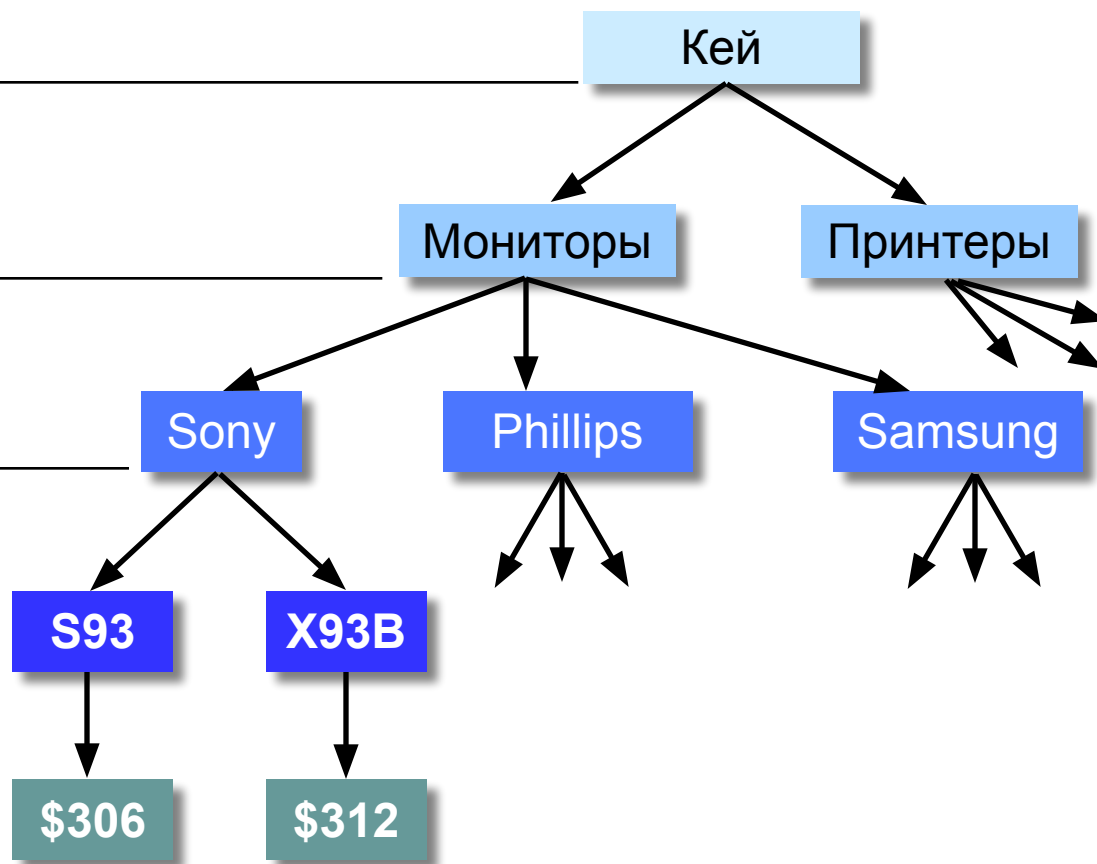
S93

X93B

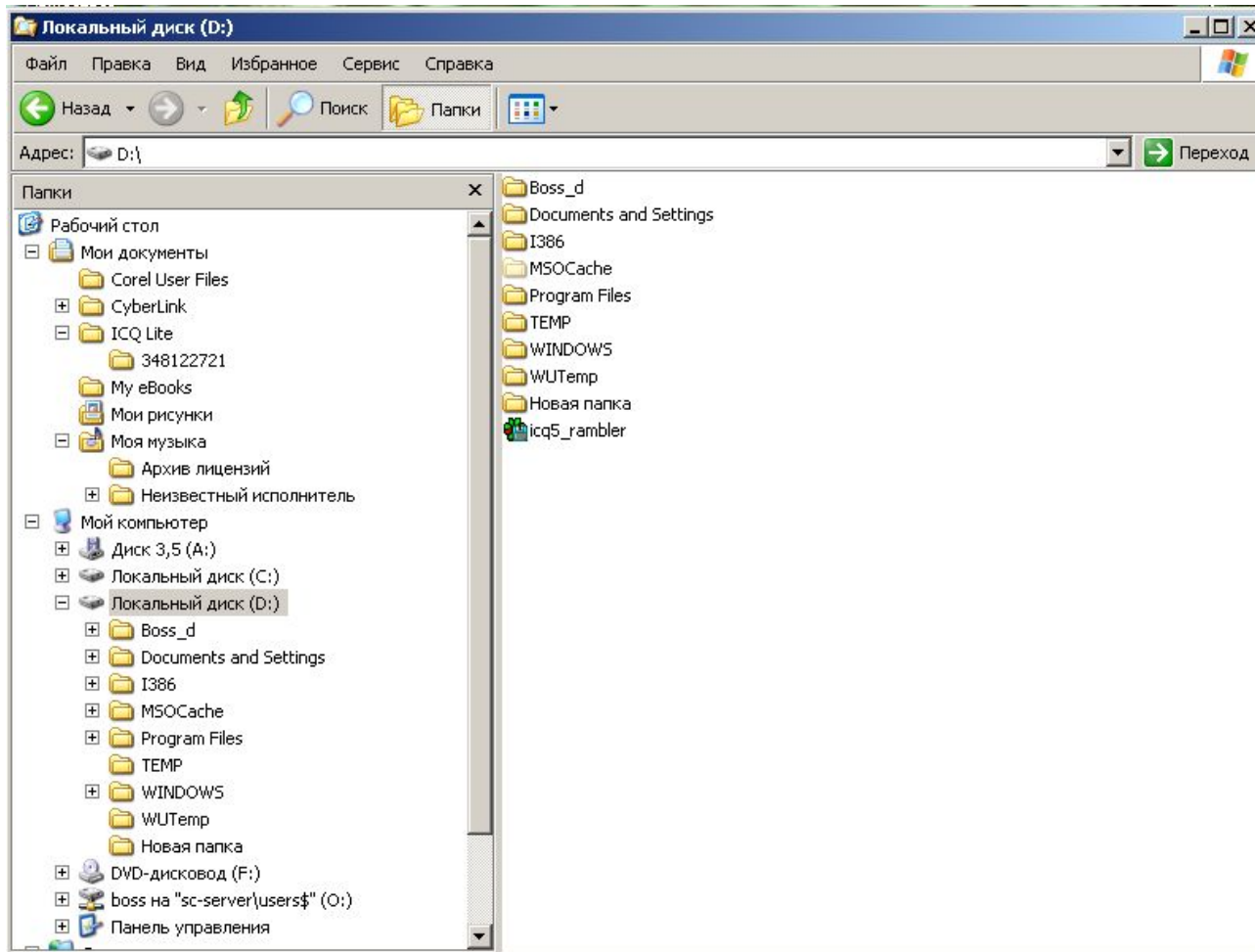
Цена (уровень 5)

\$306

\$312



2. Иерархические – содержат объекты различных уровней



**Иерархическая
БД *Каталог*
папок *Windows***

Иерархическая БД



Приведение к табличной форме:

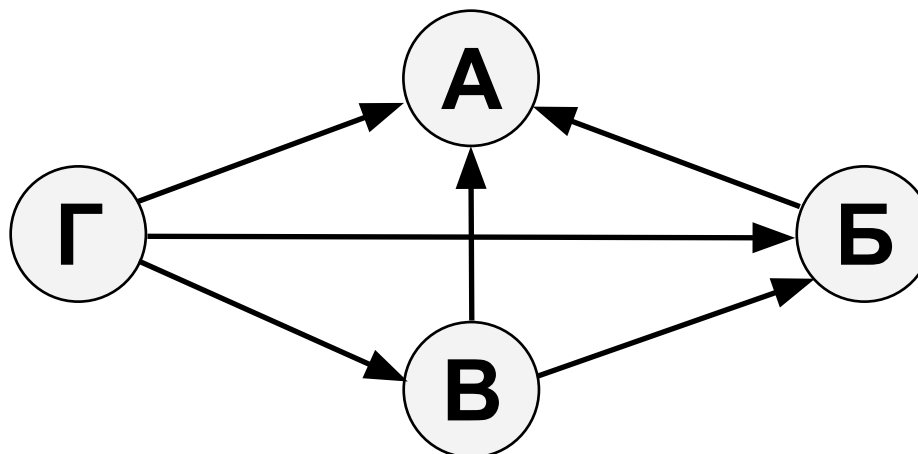
Продавец	Товар	Изготовитель	Модель	Цена
Кей	Монитор	Sony	S93	\$306
Кей	Монитор	Sony	X93B	\$312
Key	Монитор	Phillips	190 B5 CG	\$318
Кей	Монитор	Samsung	SyncMaster 193P	\$452
...				

- 1) дублирование данных
- 2) при изменении адреса фирмы надо менять его во всех строках
- 3) нет защиты от ошибок ввода оператора (*Кей* – *Key*), лучше было бы выбирать из списка

Сетевые БД



Сетевая БД - это набор узлов, в которых каждый может быть связан с каждым.



наиболее полно отражает структуру некоторых задач (например, сетевое планирование в экономике)



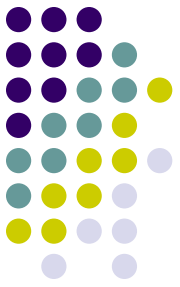
- 1) сложно хранить и искать информацию о всех связях
- 2) запутанность структуры



Можно хранить в виде таблицы, но с дублированием данных!

3. Сетевые – являются обобщением иерархической.

- Всемирная паутина



Система управления базами данных Access



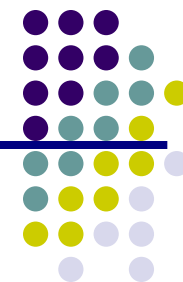
- это программа, позволяющая создавать базы данных, а также обеспечивающая обработку (сортировку) и поиск данных.

	Поле 1	Поле 2	Поле 3	Поле 4	Поле 5
Запись 1
Запись 2

Столбцы – Поля (образуют структуру БД),

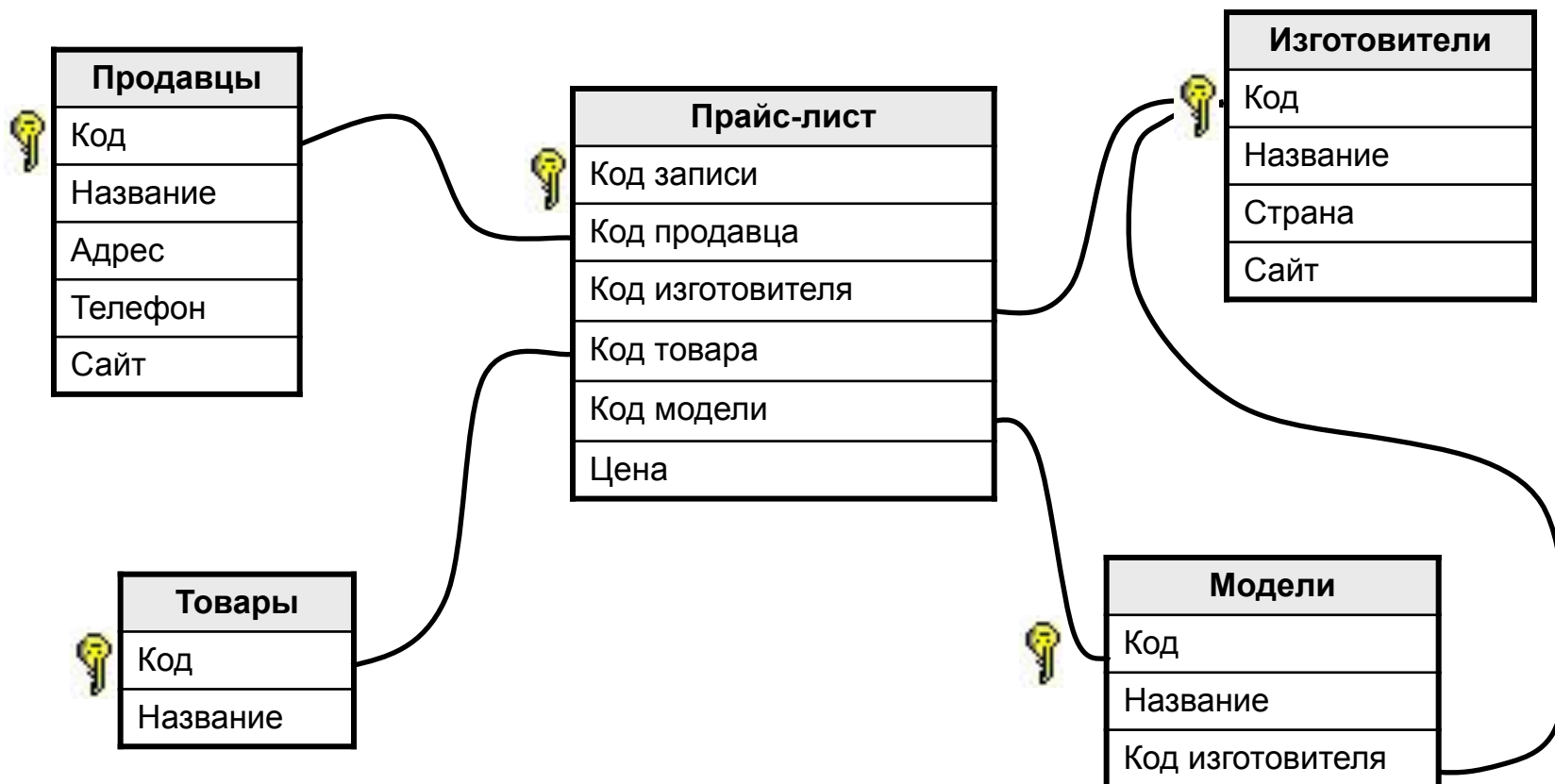
Строки – Записи (составляют информацию БД)

Реляционные БД

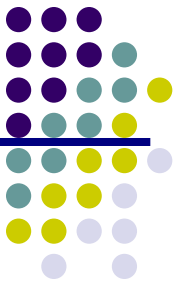




1970-е гг. Э. Кодд, англ. *relation* – отношение.

Реляционная база данных – это набор простых таблиц, между которыми установлены связи (отношения) с помощью числовых кодов.

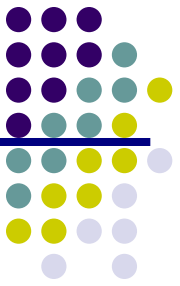


Реляционные БД



-  1) нет дублирования информации;
 - 2) при изменении адреса фирмы, достаточно изменить его в только таблице **Продавцы**;
 - 3) защита от неправильного ввода: можно выбрать только фирму, которая заранее введена в таблицу **Продавцы**;
 - 4) механизм **транзакций**: любые изменения вносятся в базу только тогда, когда они полностью завершены.
-
-  1) сложность структуры (не более 40-50 таблиц);
 - 2) при поиске надо обращаться к нескольким таблицам;
 - 3) нужно поддерживать **целостность**: при удалении фирмы продавца надо удалять все связанные записи из всех таблиц (в СУБД – автоматически, **каскадное удаление**).

Базы данных Access (*Microsoft Office*)

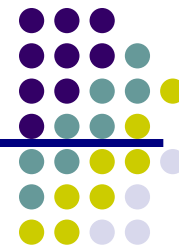


Расширение: *.mdb, один файл

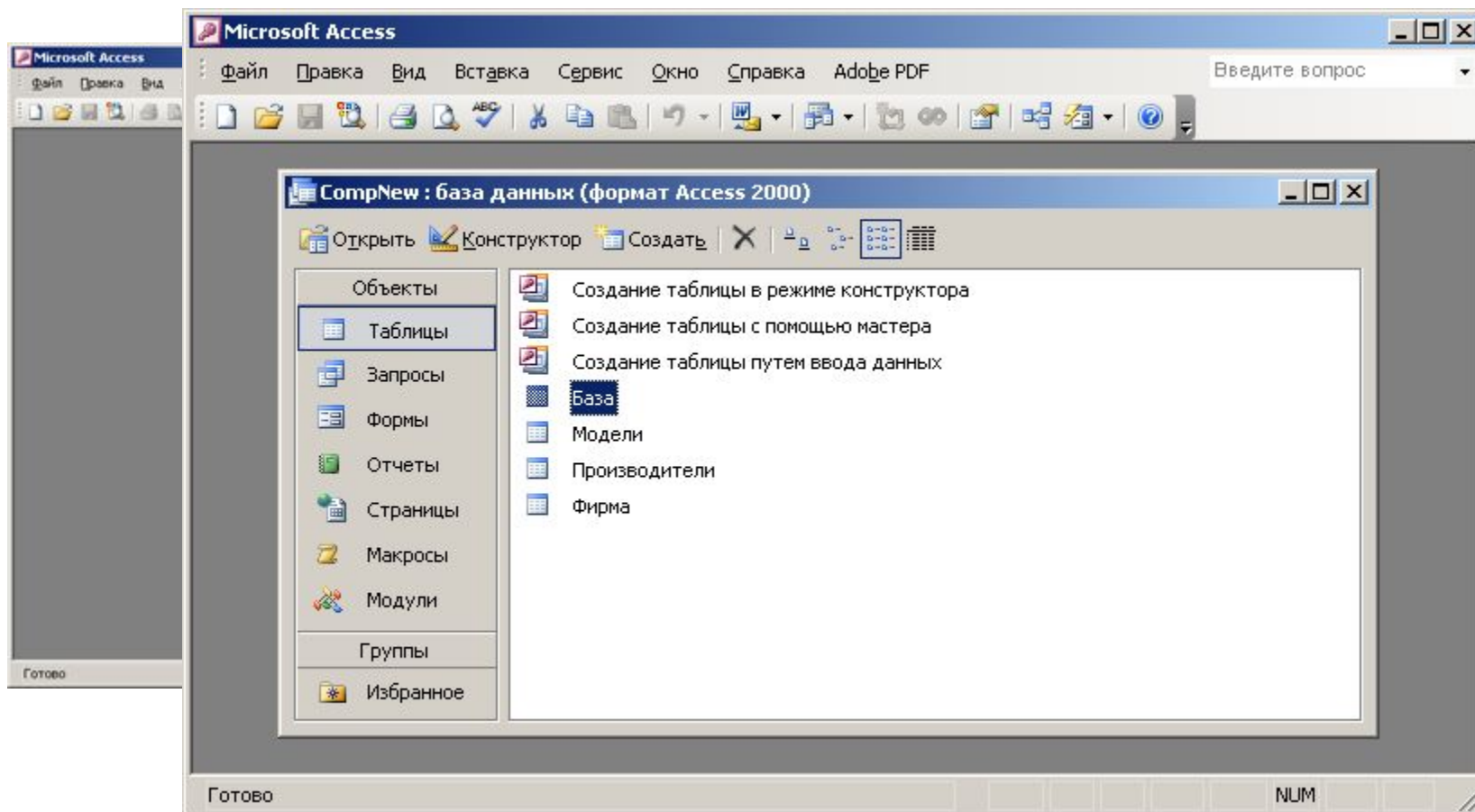
Состав:

- **таблицы;**
- **формы** – диалоговые окна для ввода и редактирования данных;
- **запросы** – обращения к базе данных для выбора нужной информации или изменения базы;
- **отчеты** – документы для вывода на печать;
- **макросы** – средства автоматизации работы;
- **модули** – дополнительные процедура на языке *Visual Basic*.

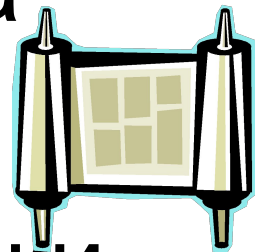
Начало работы



Пуск – Программы – Microsoft Office – Microsoft Access 2003



- **Поле** – основной элемент БД, предназначенный для хранения значений одного параметра реального объекта или процесса (имя, тип, длина, точность).



- **Запись** – совокупность логически связанных полей, характеризующих типичные свойства реального объекта.
- **!В реляционной БД не должно быть совпадающих записей!**

Первичный ключ (уникальное поле) -



- поле (совокупность полей), значение которого не повторяется у разных записей.

Уникальное
поле



№	Название книги	Автор	Год издания	Полка	Инвент. номер
1	Куджо	Стивен Кинг	1996	4	124354
2	Томминокеры	Стивен Кинг	1996	1	562232
3	Темные аллеи	Бунин	1983	4	757744
4	Воющие псы одиночества	Маринина	2005	3	576766

Тип поля -



- определяет множество значений, которые может принимать данное поле в различных записях.

Типы полей:

1. Числовой,
2. Символьный (MEMO),
3. Дата/время,
4. Логический,
5. Счетчик,
6. Денежный,
7. OLE,
8. Текстовый.



№	Название книги	Автор	Год издания	Полка	Инвент. номер
1	Куджо	Стивен Кинг	1996	4	124354
2	Томминокеры	Стивен Кинг	1996	1	562232
3	Темные аллеи	Бунин	1983	4	757744
4	Воющие псы одиночества	Маринина	2005	3	576766

счетчик

ТЕКСТОВЫЙ

дата

ЧИСЛОВОЙ

Условия выбора и простые логические выражения

