

Хранение и представление информации в ГИС

Представление пространственных данных в ГИС

Пространственные атрибуты

- **Атрибут**, реквизит (attribute) – свойство, качественный или количественный признак, характеризующий пространственный объект (но не связанный с его местоуказанием) и ассоциированный с его уникальным номером (идентификатором). Множество атрибутов пространственного объекта образует атрибутивные данные (Геоинформатика, 2008).

Шкалы измерений атрибутивных данных

- ***Номинальная шкала***
(объекты различаются по именам)
- Различны по природе, нельзя сравнивать

Сосна и дуб

- ***Порядковая шкала***
- Качественное сравнение (лучше-хуже, больше-меньше и т.д.)

Сравнение по высоте, по прочности древесины

- ***Интервальная шкала***
- Сравнение на основании измеряемых характеристик

Сравнение по высоте (м)

Типы пространственных атрибутов в ГИС:

Шкалы измерений данных

- *Номинальная шкала*
- *Порядковая шкала*
- *Интервальная шкала*

Типы атрибутов

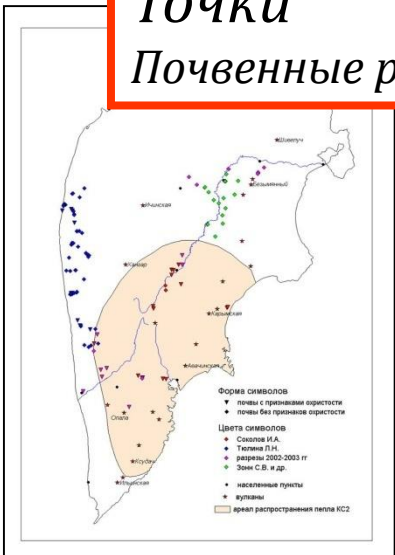
- *Категории*
- *Ранги*
- *Количество*
- *Величина*
- *Отношение*

Категории —

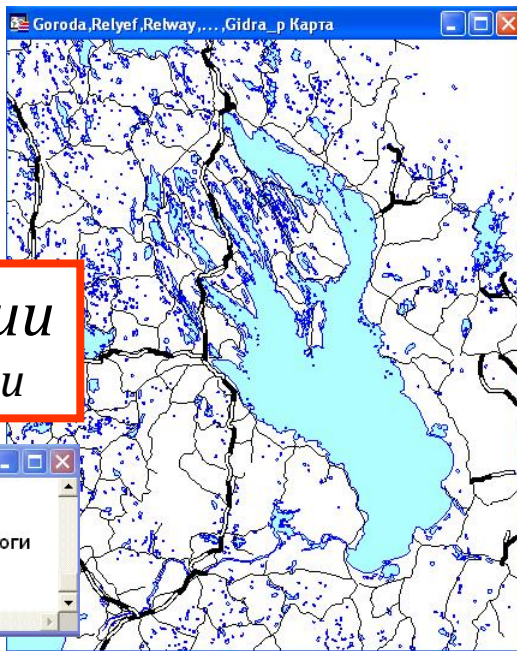
группы подобных объектов

Почвенные разрезы

Почвенные разрезы



Линии Дороги

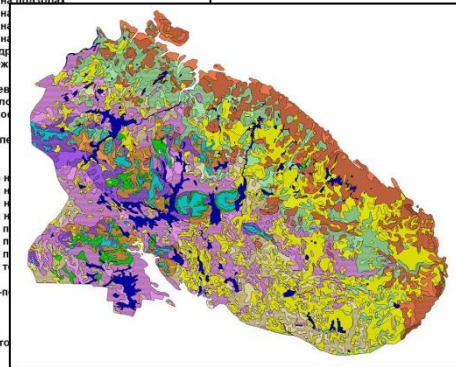


Номинальная шкала

В категории входят объекты одного класса, имеющие характерные отличия от объектов другого класса.

ЛЕГЕНДА

- Тундры на глеезах тундровых
 Тундры на подбурх тундровых
 Тундры на торфяно- и торфянисто-подзолисто-глеевых почвах
 Тундры на подзолах
 Тундры на подзолах глеевых
 Тундры на торфяных болотных почвах
 Тундры на каменистых россыпях
 Лесотундровые березовые редколесья и криволиесья на подбурх тундровых
 Лесотундровые березовые редколесья и криволиесья на подзолах
 Лесотундровые березовые редколесья и криволиесья на подзолах глеевых
 Лесотундровые березовые редколесья и криволиесья на торфяных болотных почвах
 Лесотундровые березовые редколесья и криволиесья на каменистых россыпях
 Северотажные еловые, березово-еловые леса на подбурх тундровых
 Северотажные еловые, березово-еловые леса на подзолах
 Северотажные еловые, березово-еловые леса на
 Северотажные еловые, березово-еловые леса на
 Северотажные еловые, березово-еловые леса на
 Северотажные сосновые леса на подбурх тундл
 Северотажные сосновые леса на глеезах таеж
 Северотажные сосновые леса на подзолах
 Северотажные сосновые леса на подзолах глеев
 Северотажные сосновые леса на торфяных бол
 Северотажные сосновые леса на каменистых ро
 Болота на глеезах тундровых
 Болота на торфяно- и торфянисто-подзолисто-г्ल
 Болота на подзолах глеевых
 Болота на торфяных болотных почвах
 Высокогорная растительность и горные тундры н
 Высокогорная растительность и горные тундры н
 Высокогорная растительность и горные тундры н
 Высокогорная растительность и горные тундры н
 Горные березовые редколесья и криволиесья на л
 Горные березовые редколесья и криволиесья на л
 Горные березовые редколесья и криволиесья на л
 Горные березовые редколесья и криволиесья на т
 Горные еловые леса на подбурх тундровых
 Горные еловые леса на торфяно- и торфянисто-п
 Горные еловые леса на подзолах
 Горные еловые леса на подзолах глеевых
 Горные сосновые леса на подбурх тундровых
 Горные сосновые леса на торфяно- и торфянисто
 Горные сосновые леса на подзолах
 Горные сосновые леса на подзолах глеевых
 Вода



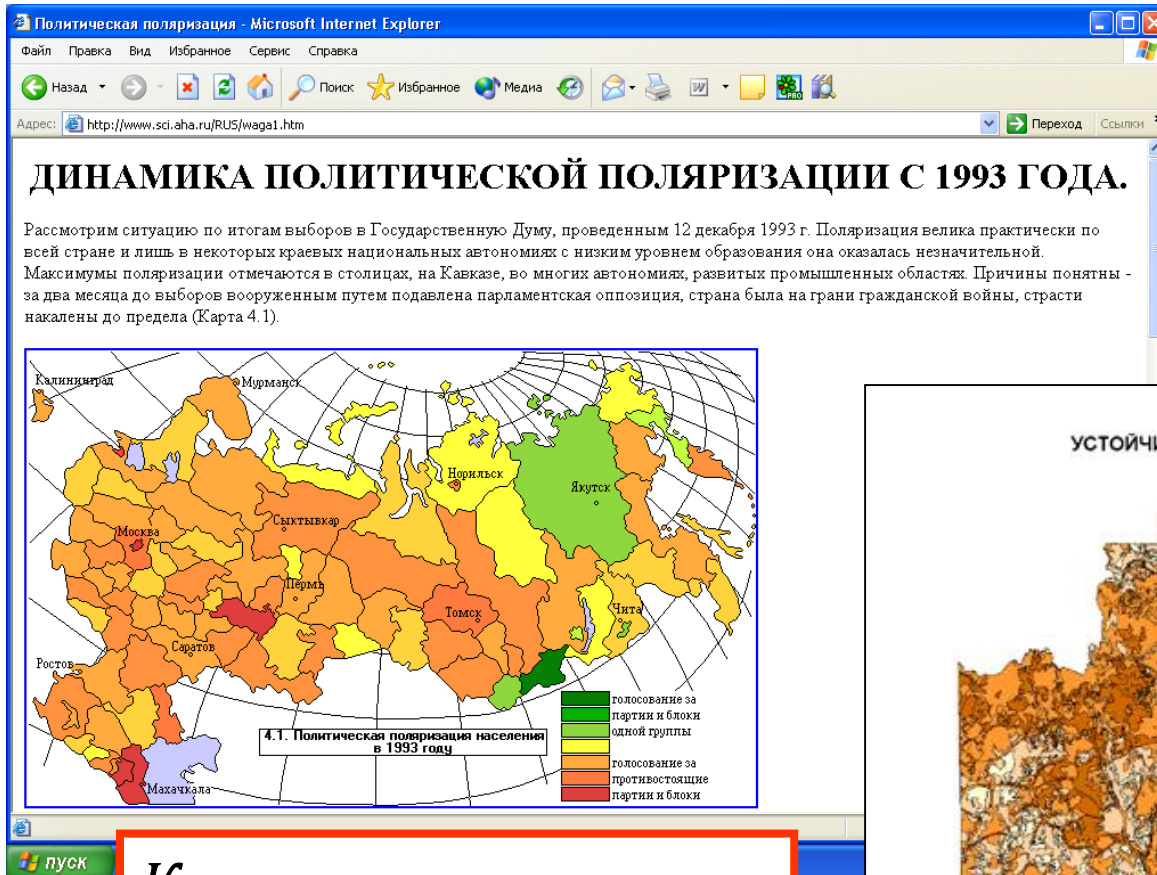
Полигоны

Экосистемы (почвы-растительность)

Ранги –

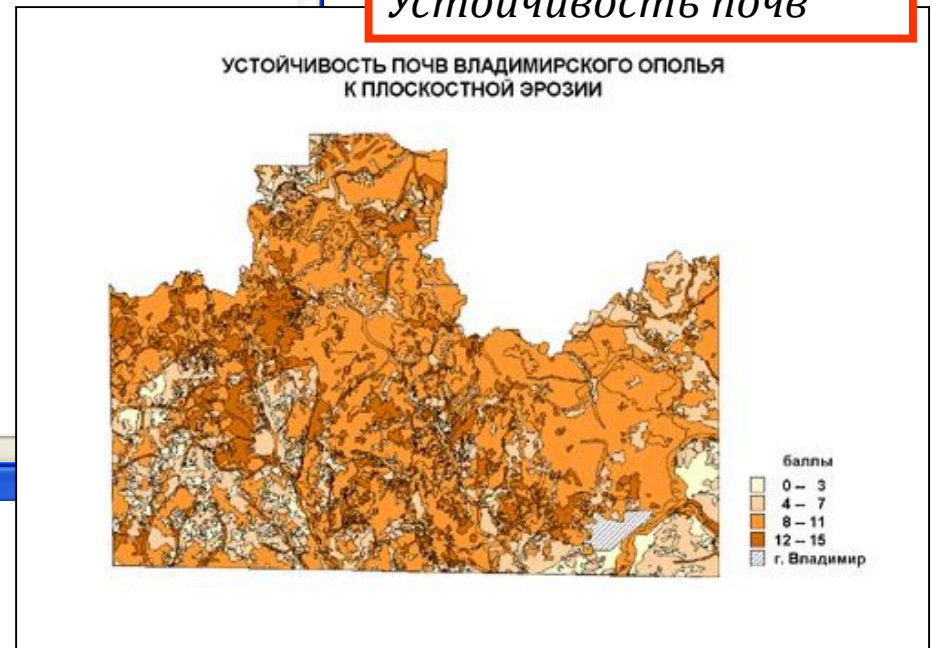
группы объектов, систематизированные в порядке возрастания или убывания величины

Порядковая шкала

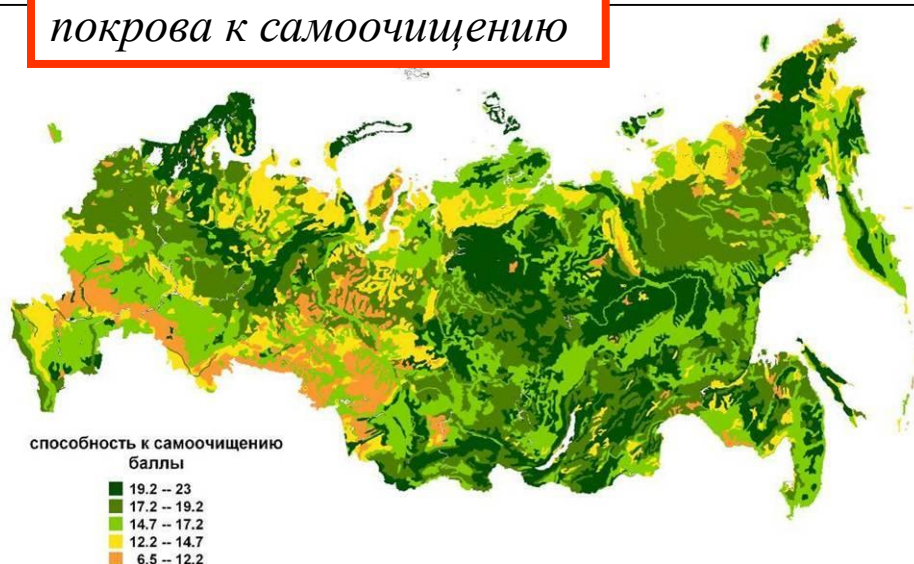


Качественная оценка
Поляризация общества

Балльная оценка
Устойчивость почв



Балльная оценка *Способность почвенного* *покрова к самоочищению*

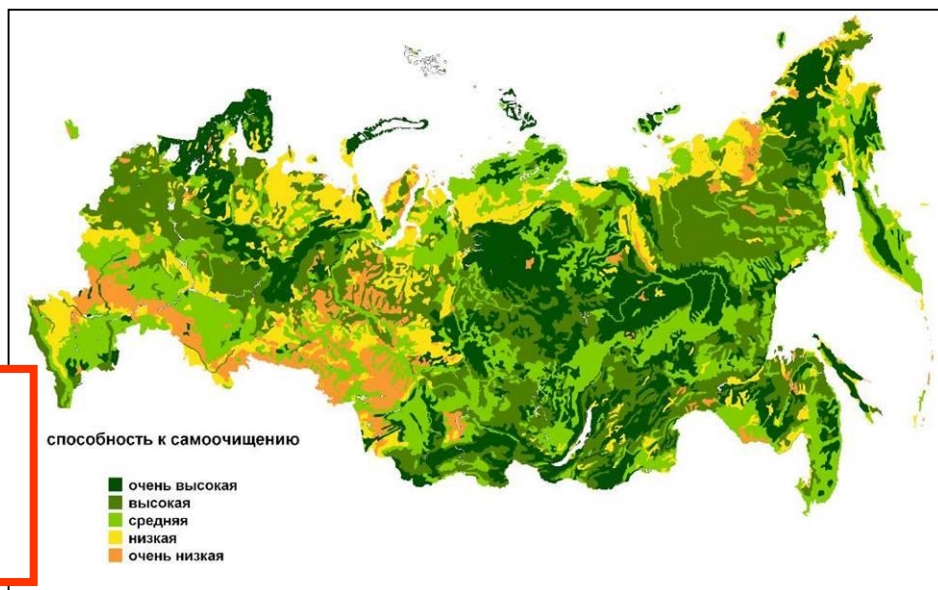


Ранги (баллы) используются,

- когда трудно дать непосредственную оценку величины, характеризующей объект,
- оценка определяется целой комбинацией факторов.

Полуколичественная,
квалиметрическая оценка –
«оценка качества
количеством»

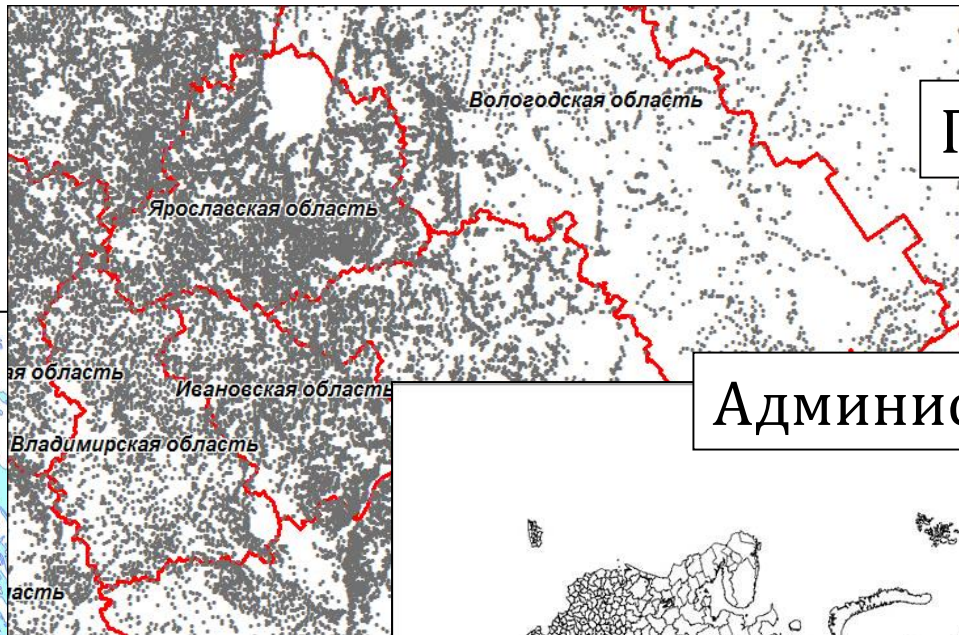
Качественная оценка *Способность почвенного* *покрова к самоочищению*



Количество –

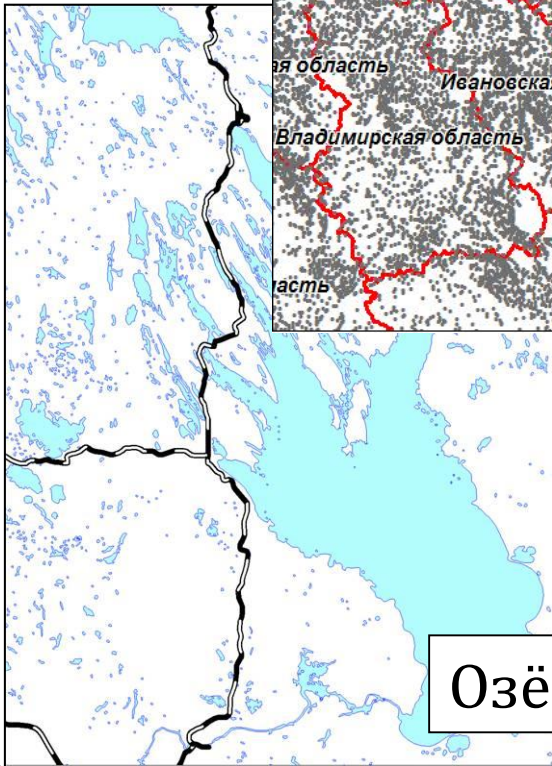
Интервальная шкала

значение, отражающее фактическое число
объектов определённого вида на карте



Города

Административные районы



Озёра

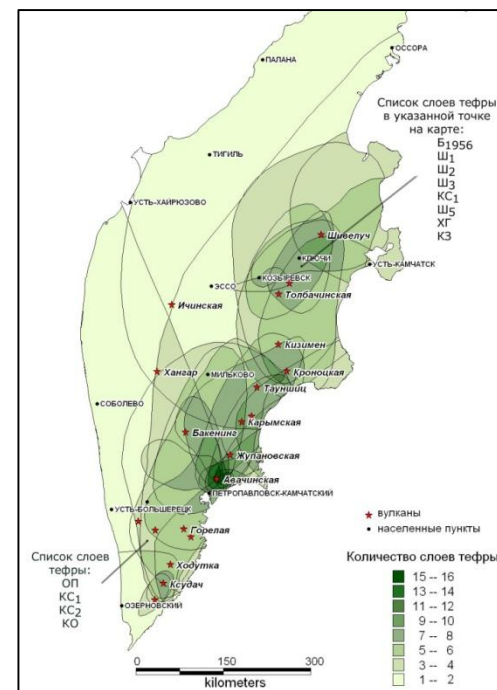


Величина —

Интервальная шкала

числовое значение, связанное с каждым объектом определённого вида на карте

Количественные показатели —
абсолютная оценка.
Разбиение на диапазоны
(классы)

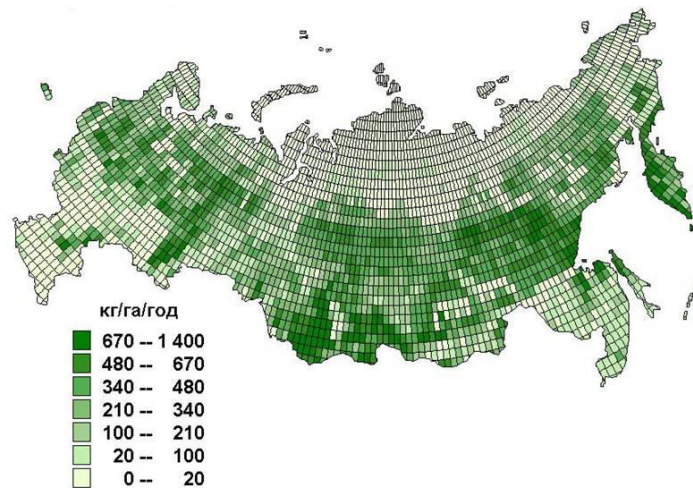


Отношение –

Интервальная шкала

удельное (относящееся к единице измерения)
значение величины, связанное с каждым
объектом определённого вида на карте

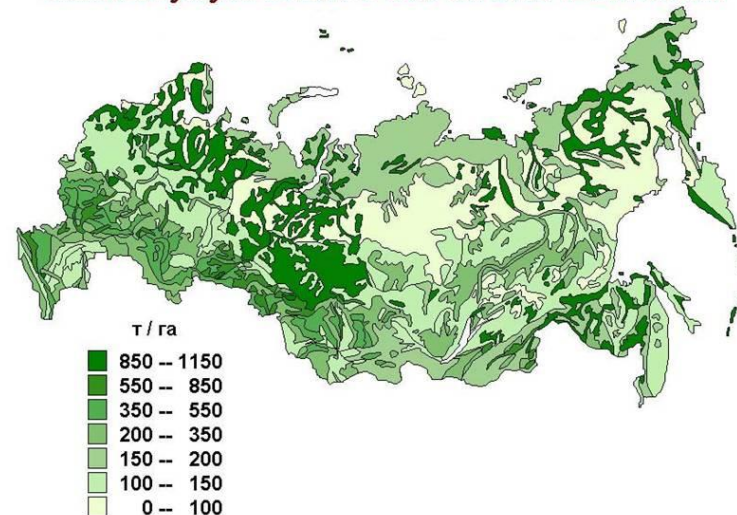
Средние удельные значения ежегодного депонирования
органического углерода в почвенном покрове России



Определяется для
каждого объекта путём
деления одной величины
на другую.

Отражает взаимосвязь
между двумя величинами,
характеризующими объект.

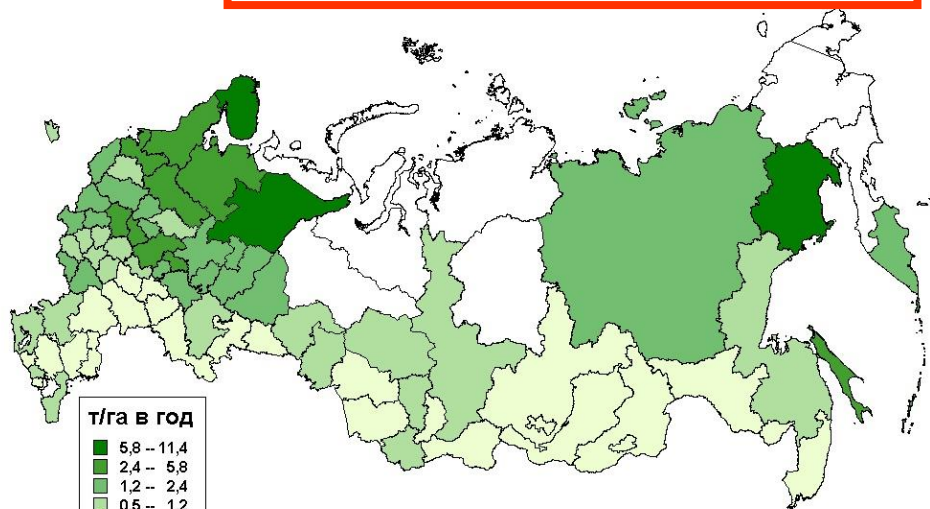
Запасы гумуса в слое 0-100 см в почвах России



ВНЕСЕНИЕ ОРГАНИЧЕ

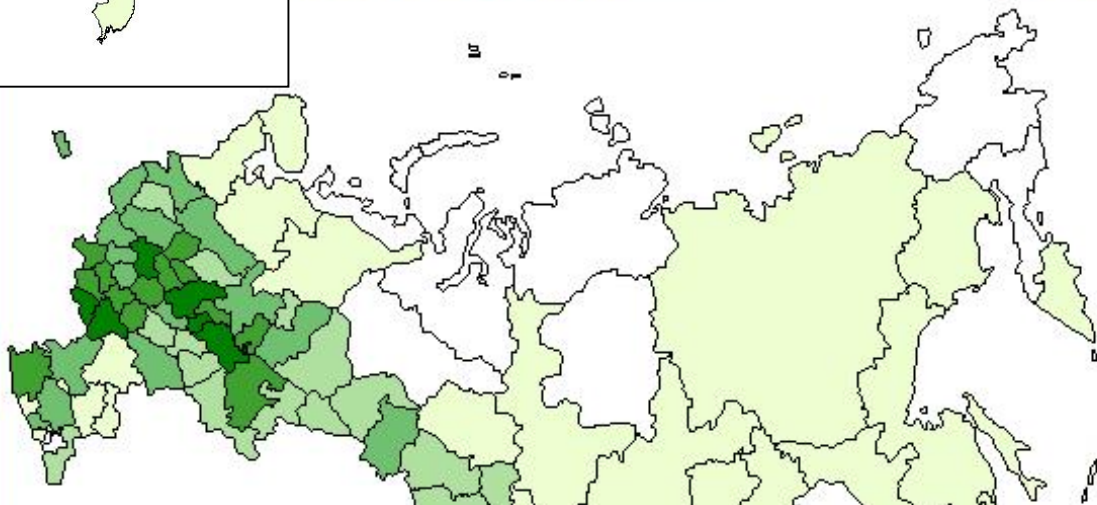
1 га посевной площади

ТУР



Существенное значение имеет *правильное определение территории*, для которой рассчитываются относительные показатели объектов.

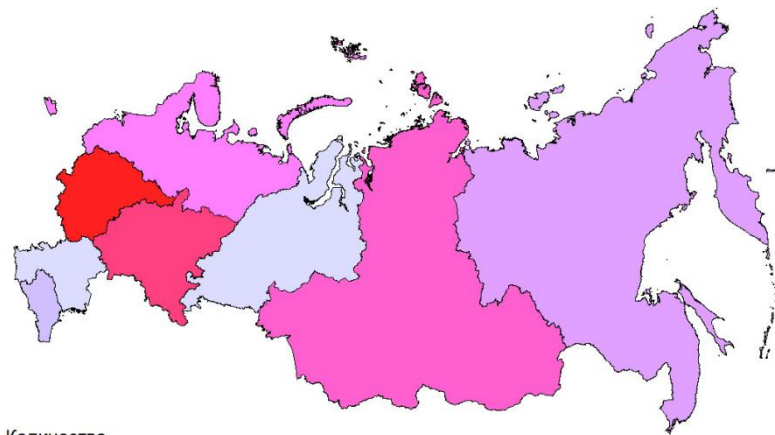
..grid1x1_15M_c_x Карта



1 км² территории субъекта РФ

Различное изображение атрибутов в интервальной шкале

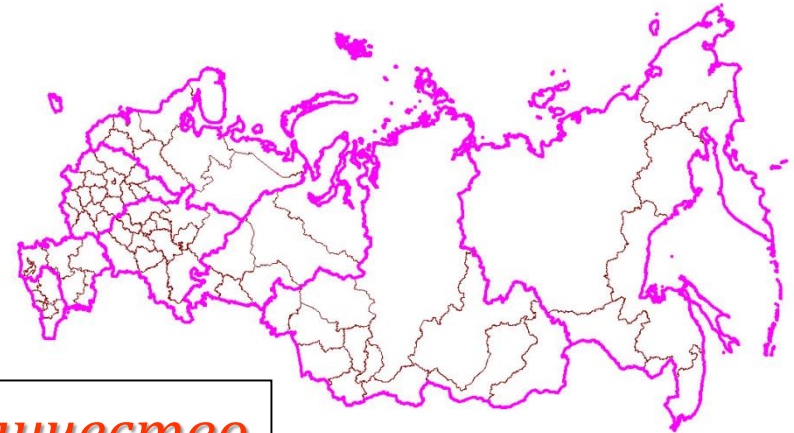
Субъекты РФ в
федеральных округах



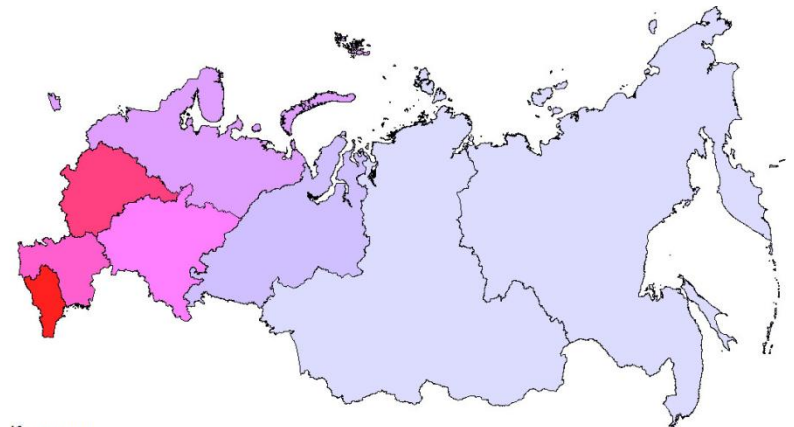
Количество
административных регионов

- 18 (1)
- 14 (1)
- 12 (1)
- 11 (1)
- 9 (1)
- 7 (1)
- 6 (2)

величина



количество



Количество
административных регионов (ед./млн кв.км)

- 40,9 -- 41 (1)
- 27,7 -- 40,9 (1)
- 14,3 -- 27,7 (1)
- 13,5 -- 14,3 (1)
- 6,6 -- 13,5 (1)
- 3,4 -- 6,6 (1)
- 1,4 -- 3,4 (2)

отношение

Хранение и представление информации в ГИС

Структуры компьютерных файлов и модели баз данных

Чего мы хотим от ГИС?

- В ГИС каждый графический объект (его расположение в пространстве и описательные характеристики – атрибуты) должны храниться таким образом, чтобы была возможность работать с данными:

отображать, анализировать, обрабатывать и т.д. любую нужную информацию за приемлемое время.

Операции над данными

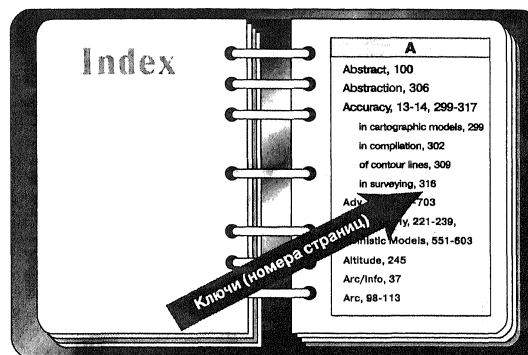
- Идентификация данных и их поиск в БД
- Выборка (чтение) данных из БД
- Добавление (запись) данных в БД
- Удаление данных из БД
- Изменение данных в БД

Информация в компьютере хранится в ***компьютерных файлах (состоят из записей)***

Следующий уровень организации информации – ***базы данных (состоят из файлов)***

Основные структуры компьютерных файлов

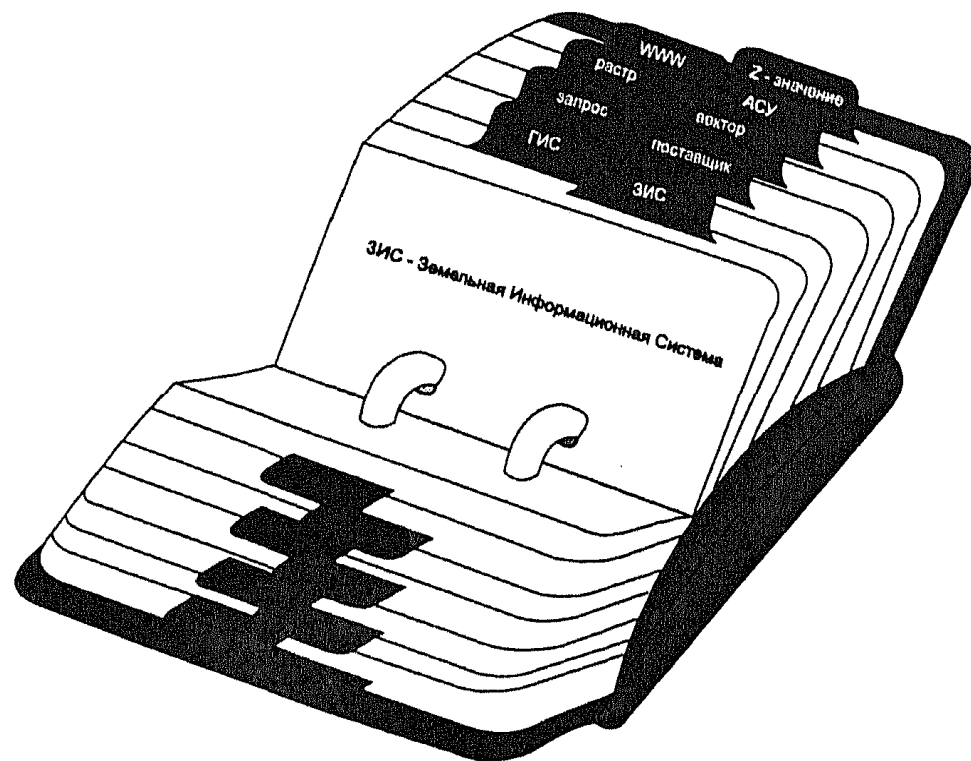
- **Неупорядоченные файлы**
- **Последовательно упорядоченные файлы**
- **Индексированные файлы**



Неупорядоченные файлы

5курс_geomorph_regions	
1	Рын-Пески
2	Верхнетуломский
3	Терский
4	Центральнопонойский
5	Печенгско-Кольский
6	Центральнокольский
7	Западномурманский
8	Кейвинский
9	Восточномурманский
10	Ейско-Сальский
11	Волго-Ахтубинский
12	Ковдозерский
13	Приморско-Нижненеманский
14	Балтийско-Свенцянский
15	Ногайско-Черноземельский
16	Северокарельский
17	Западнокарельский
18	Сарпинский
19	Кубанский
20	Узенско-Чижинский
21	Таманского полуострова
22	Ковдозерско-Топозерский
23	Доно-Оскольский
24	Верхнекамский
25	Вымский
26	Двинско-Мезенский

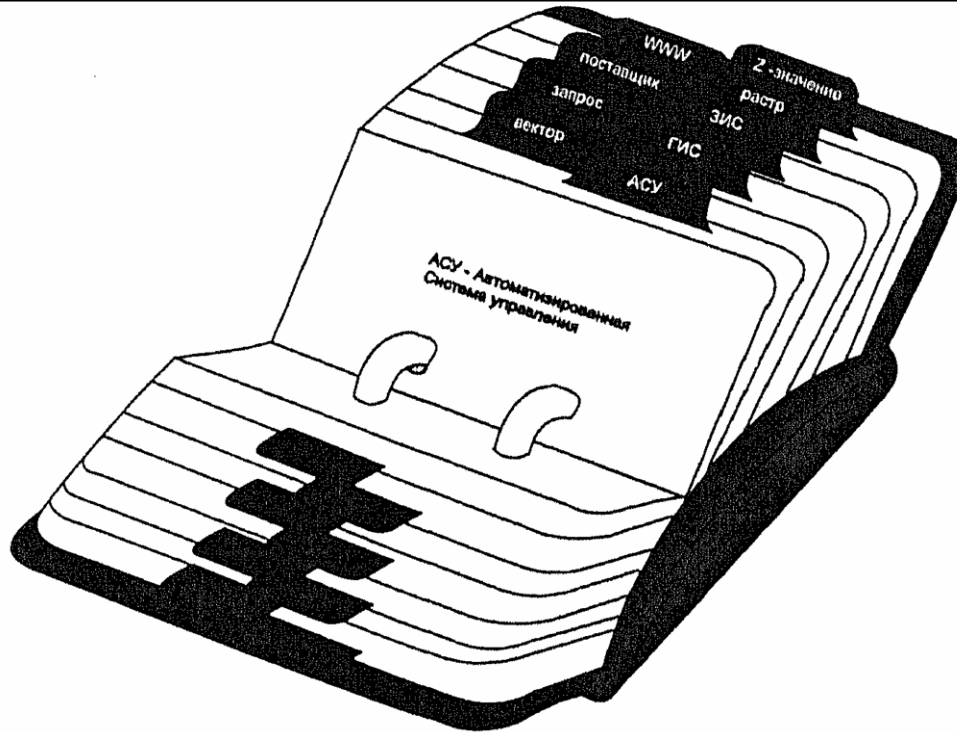
Неэффективный поиск



Данные вносятся «по мере поступления»,
записываются в файл позади всех записей

Последовательно упорядоченные файлы

Поиск делением пополам – дихотомия



5курс_geomorph_regions	
1	Абхазо-Рачинская
2	Абхазо-Сылвинский
3	Абхазо-Свенцанский
4	Белебеевский
5	Белозерский
6	Беломорско-Кулойский
7	Бельско-Камский
8	Березниковский
9	Берсуат-Кумакский
10	Большеземельский
11	Большехобдинский
12	Брянский
13	Бугульминский
14	Бузулукско-Уральский
15	Вайгачский
16	Валдайский
17	Верхневолжский
18	Верхневятский
19	Верхнедонской
20	Верхнекамский
21	Верхнепечорский
22	Верхнесурский

- Записи упорядочены по числам или буквам алфавита
- Каждая новая запись должна вставляться в определенное место последовательности

Индексированные файлы

АТТРИБУТЫ

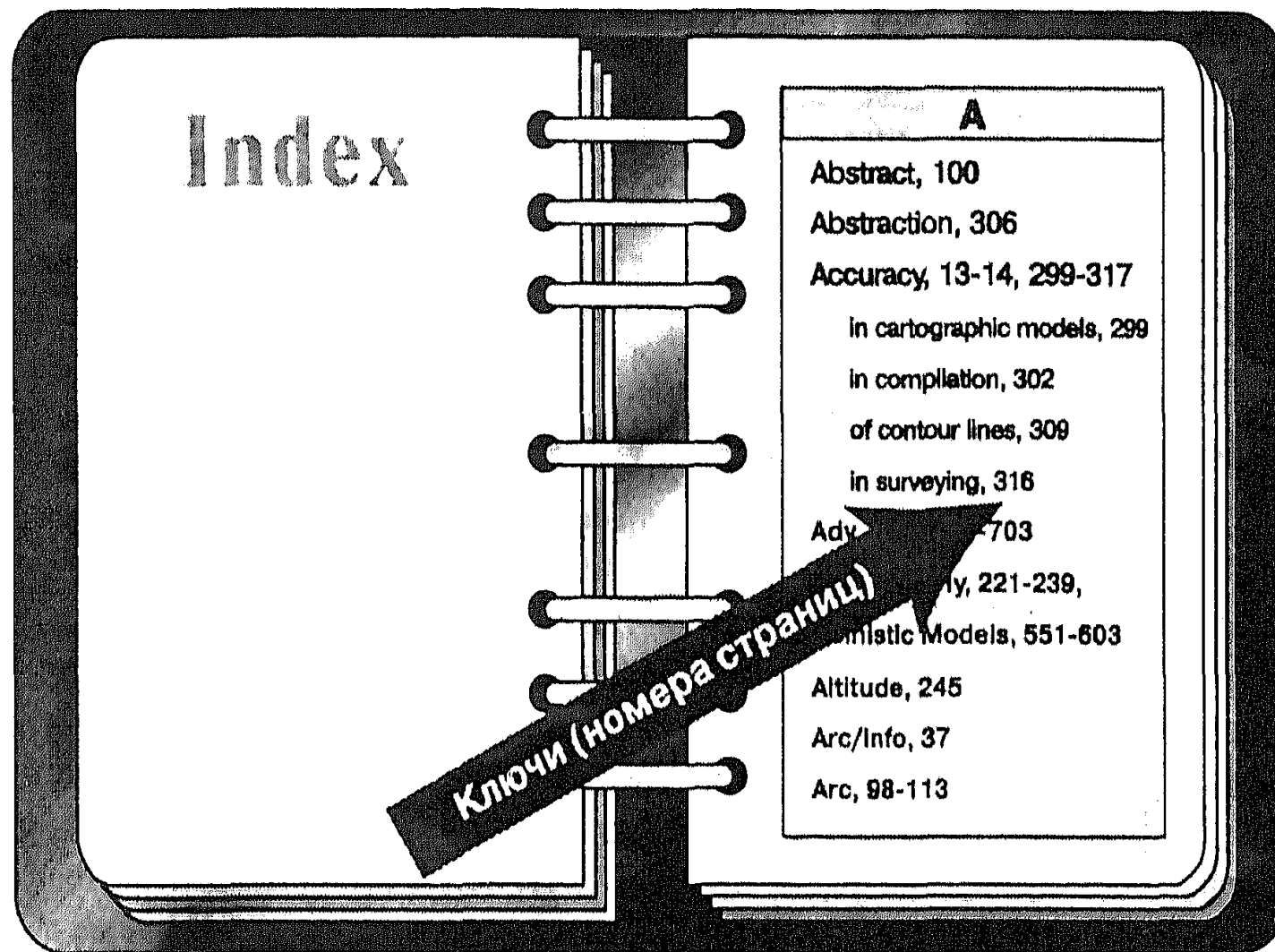
	A	B	C	D
1	РАЙОН	ГЛУБИНА РАСЧЛЕНЕНИЯ,М	АБСОЛЮТНАЯ ВЫСОТА,М	КРУТИЗНА СКЛОНОВ,ГРАД.
2	Рын-Пески	0	0	4.0
3	Верхнетуломский	450	700	11.7
4	Терский	100	225	7.8
5	Центральнопонойский	125	225	8.6
6	Печенгско-Кольский	300	570	10.3
7	Центральнокольский	850	1145	12.4
8	Западномурманский	200	400	7.2
9	Кейвинский	125	335	8.9
10	Восточномурманский	150	270	7.5
11	Ейско-Сальский	45	140	4.0
12	Волго-Ахтубинский			
13	Ковдозерский			
14	Приморско-Нижненеманский			
15	Балтийско-Свенцянский			
16	Ногайско-Черноземельский			
17	Северокарельский			
18	Западнокарельский			
19	Сарпинский			
20	Кубанский			
21	Узенско-Чижинский			
22	Таманского полуострова			
23	Ковдозерско-Топозерский			
24	Доно-Оскольский			
25	Верхнекамский			
26	Вымский			
27	Двинско-Мезенский			

Причины возникновения необходимости индексирования файлов:

- Каждому объекту может быть приписано большое количество атрибутов
- Упорядочить файл по всем атрибутам одновременно невозможно

- Для поиска (сортировки) записей по нескольким атрибутам создается – **внешний индекс**.
- **Внешний индекс** – новый файл (**индексный файл**) в который скопированы значения одного атрибута всех записей вместе с положениями этих записей (ссылками). Записи в новом файле упорядочены по атрибуту – можно осуществлять поиск делением пополам.
- Во внешний индекс можно выносить несколько атрибутов и организовывать поиск по этим атрибутам.
- Исходный файл становится **индексированным**.
- При добавлении записей в исходный файл ссылки на них должны помещаться в соответствующее место индексного файла.

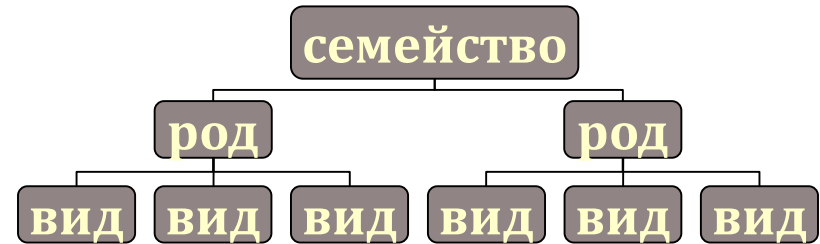
Индексированная структура данных



Район_тип_рельефа		абс_высота_м		
Рын-Пески (Верхнехвалынская песчан		0		
Волгоградское Заволжье (Морская низ		25		
Узенско-Чижинский (Нижнехвалынская		25	фа	склоны
Район_тип_рельефа	глубина_расчл_ср	какая нижнехвалынская	50	0
Рын-				0
Нижн-	Поиск делением пополам – дихотомия			0
Маньчский (Долина Маньча с компл	15	мельский (Верхнечет	50	0
Волховский (Террасированная озер	22.5	зерно-ледниковая ра	65	0.00413888
Двинско-Мезенский (Морская, озерн	25	асированная озерно-	65	2.81955
Северодвинский (Озерно-ледникова	25	(Озерно-ледниковая	75	3.39293
Принева-Эстонский (Предглинтов	25	(Холмисто-моренне	75	3.51327
Псковско-Чудский (Озерно-ледников	25	й (Морская, озерно-л	95	4
Кубанский (Наклонная террасирова	25	Озерно-ледниковая	95	4
Верхневолжский (Озерно-ледникове	27.5	Озерно-ледниковая	100	4.00448
Приильменский (Озерно-ледниковая	27.5	ная дельтовая равн	100	4.05592
Полоцкий (Озерно-ледниковая и зан	27.5	огорный, мелкосопоч	100	4.07067
Верхнепечорский (Ледниковая, озер	30	жский (Аллювиальн	110	4.07395
Горьковско-Заволжский (Аллювиал	30	Озерно-ледниковая,	115	4.1398
Онежско-Двинский (Холмисто-море	30	лжье (Сыртовая ра	115	4.16841
Мещерский (Террасированная озерн	30	кий (Сельговая равн	115	4.19558
Онежско-Сухонский (Ложбина с тер	30	ная террасированне	120	4.25673
Прибеломорский (Озерно-ледникова	30	й (Ложбина с террас	125	4.26317
Окско-Воронежский (Плосковолнист	35	брежно-морская тер	125	4.29906
Приморско-Нижненеманский (Озерн	35	Грядово-холмистая,	130	4.322
Унженский (Моренная и террасиров	40	сированная озерно-с	135	4.32539
Среднепечорский (Террасированная	40	нский (Слабо расчл	135	4.34368
Волжско-Клязьминский (Слабо расч	40	Волгоградское Заволжье (Морская нижне		4.38154
Заветлужский (Плосковолнистая м	40	Большеземельский (Холмисто-грядовая		4.39358
Северный Краинный (Моренно-эро	40	Рославльско-Спас-Деминский (Моренно-с		4.443

Основные типы моделей баз данных

- Иерархическая модель
 - Квадратомическая модель



- Сетевая модель



- Реляционная модель

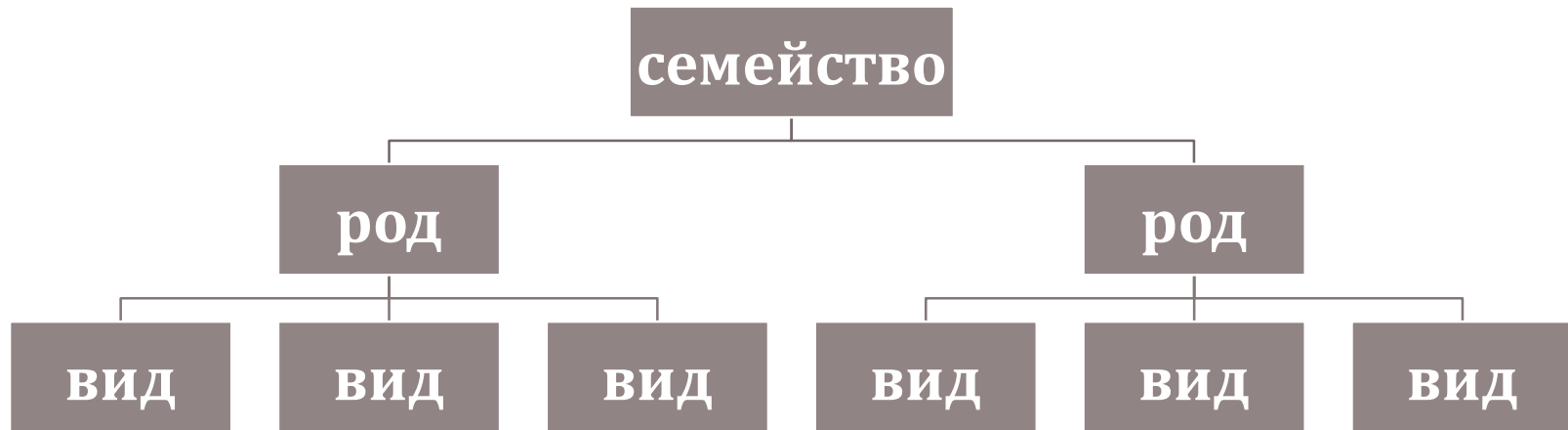
Diagram illustrating a relational model structure using Microsoft Excel. The main table, "почвы_горизонты", contains data for soil horizons. The table structure is as follows:

Но в М	индекс	С	О	О1	О2	О3	А0	А
1	1	Ар						
2	2	Та						
3	3	Тр						
4	88	ТП						
5	5	ПБ						
6	6	Тл						
7	7	ТБ						
8	8	Пл						
9	9	Пл						
10	9	Пл						
11	10	П						
12	11	Пл						
13	90	Пл						
14	14	Пл						
15	92	Пл						
16	15	Пл						
17	16	Пл						
18	19	Пл						
19	17	Пл						
20	19	Пл						

The table is divided into two sections: "почвы_горизонты" and "почвы_2_Список". The "почвы_горизонты" section contains data for soil horizons, and the "почвы_2_Список" section contains data for soil types. The "почвы_2_Список" section is further divided into two sub-sections: "почвы_2_Список" and "почвы_2_Список".

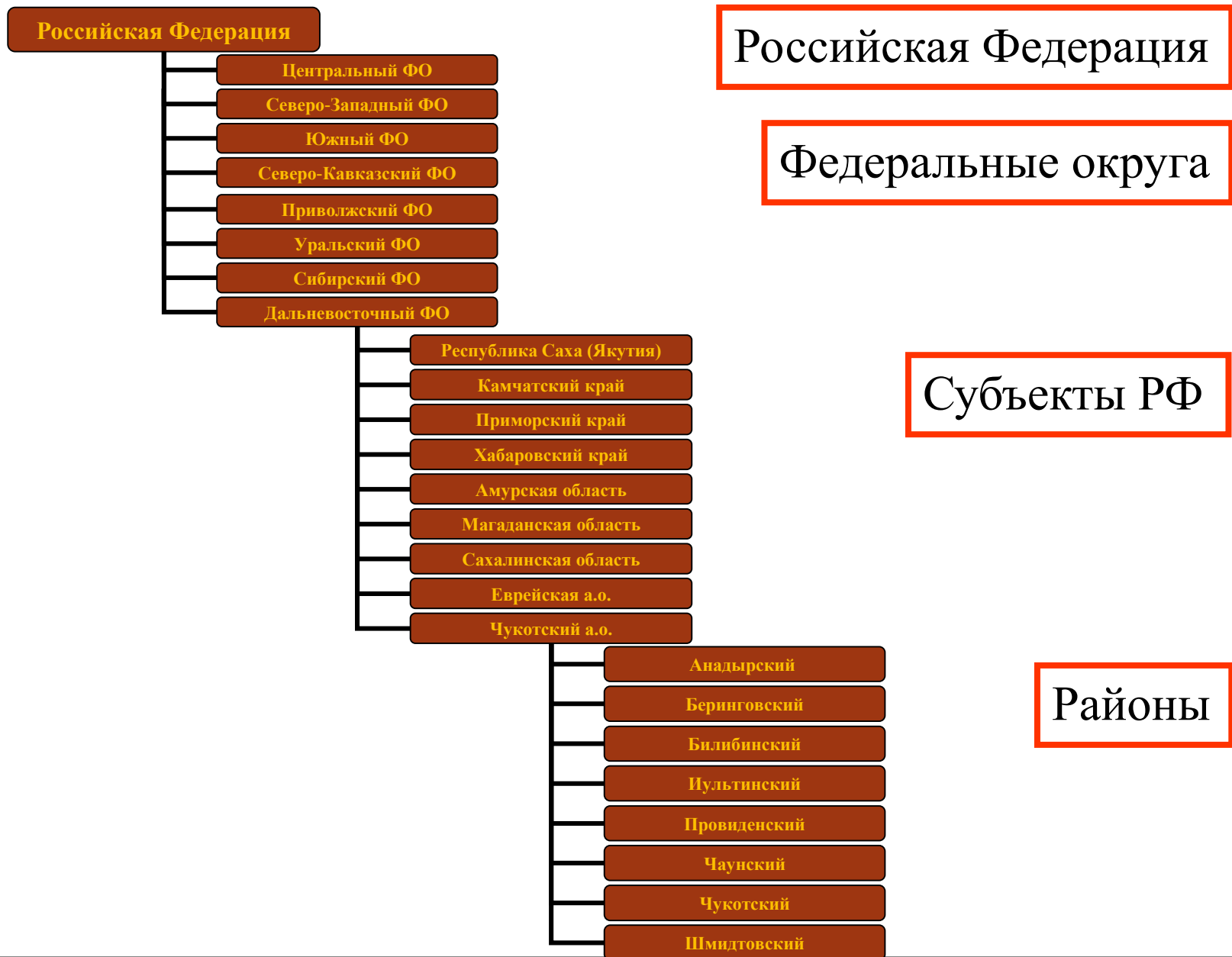
Иерархическая модель

Система классификации
растений и животных



- Записи образуют древовидную структуру и связаны с одной записью, находящейся на более высоком уровне иерархии, т.е. связь один ко многим (1:N).
- Доступ к записям осуществляется путем прохода по строго определенной цепочке узлов дерева.

Схема административного деления России



Квадратомическая модель – иерархическая модель
со связью 1:4

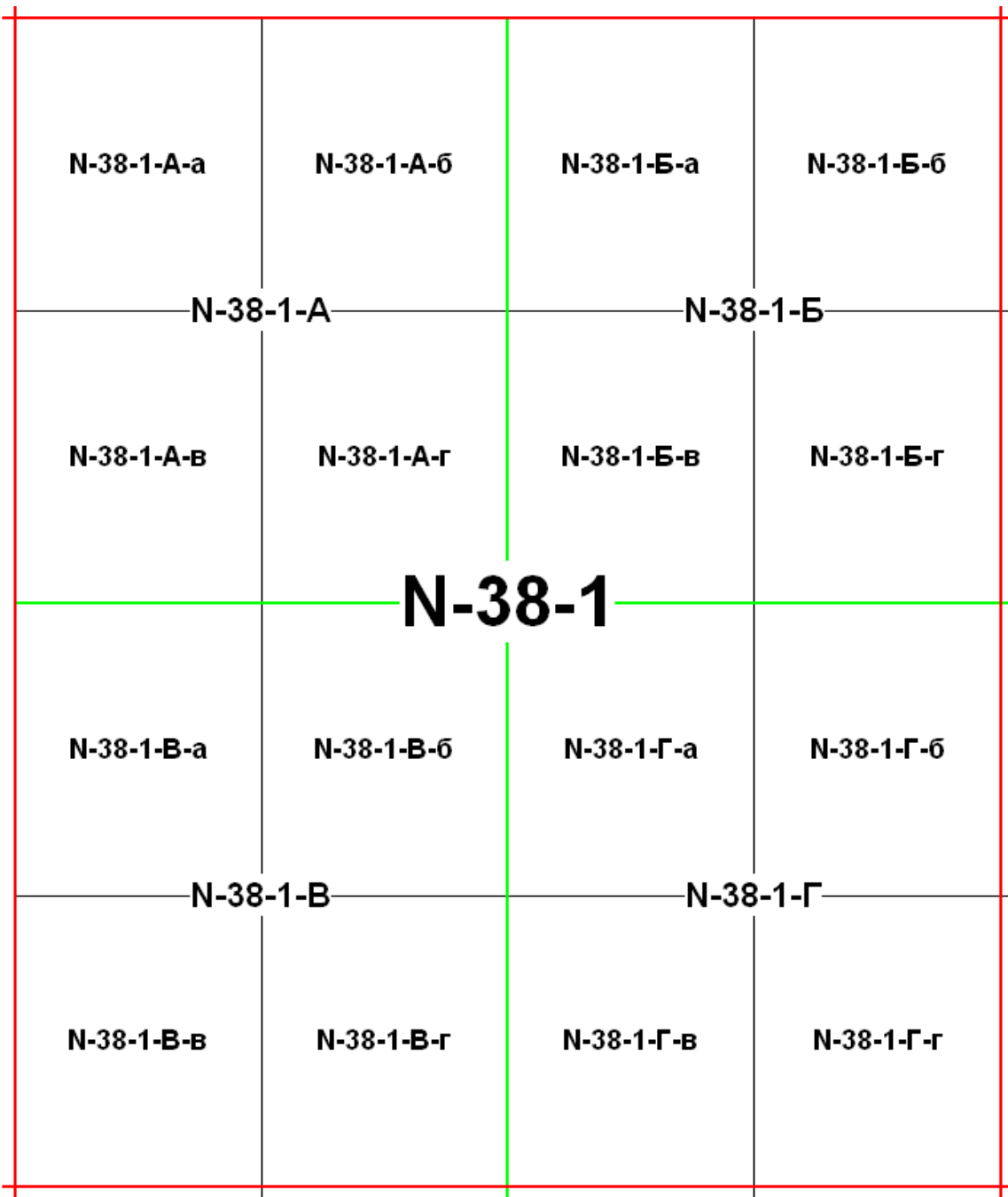


Разграфка топографических карт в России

***Номенклатурные
листы карты***

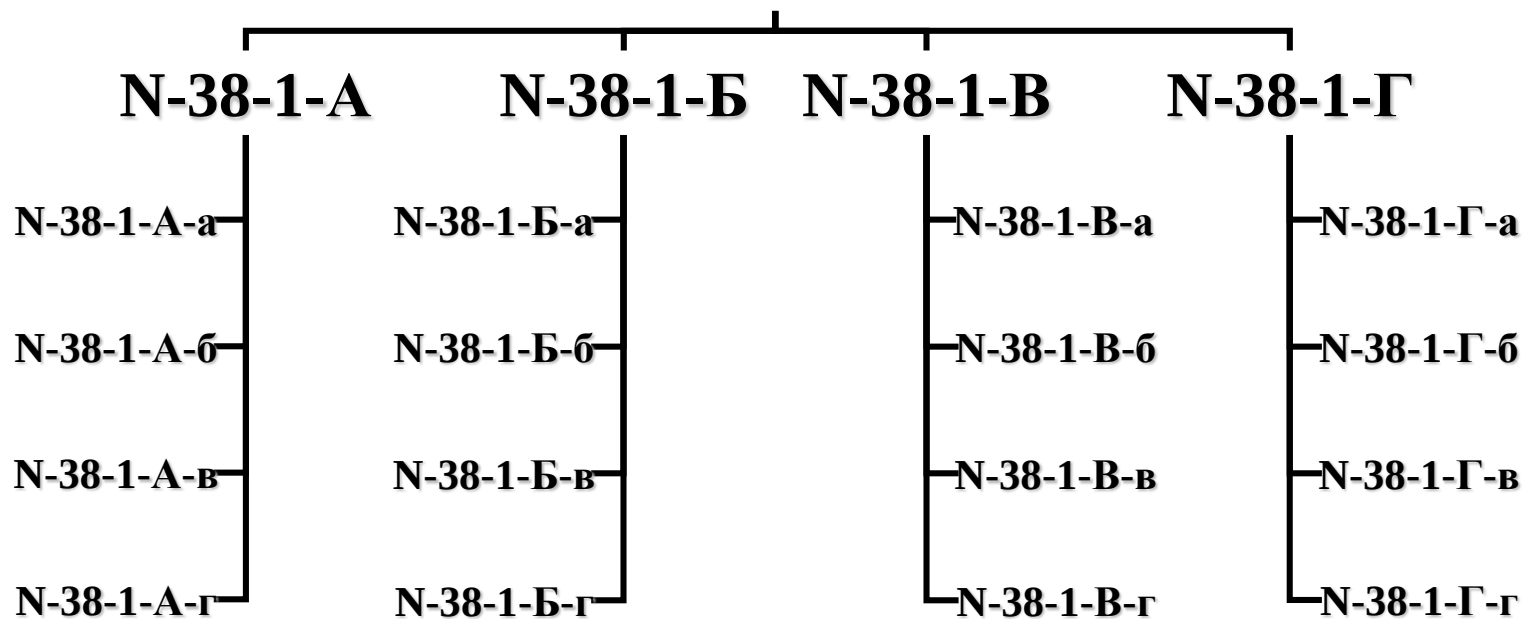
Масштабы:

- N-38-1 – 1:100 000
- N-38-1-A – 1:50 000
- N-38-1-A-a – 1:25 000



Квадратомическое дерево

N-38-1

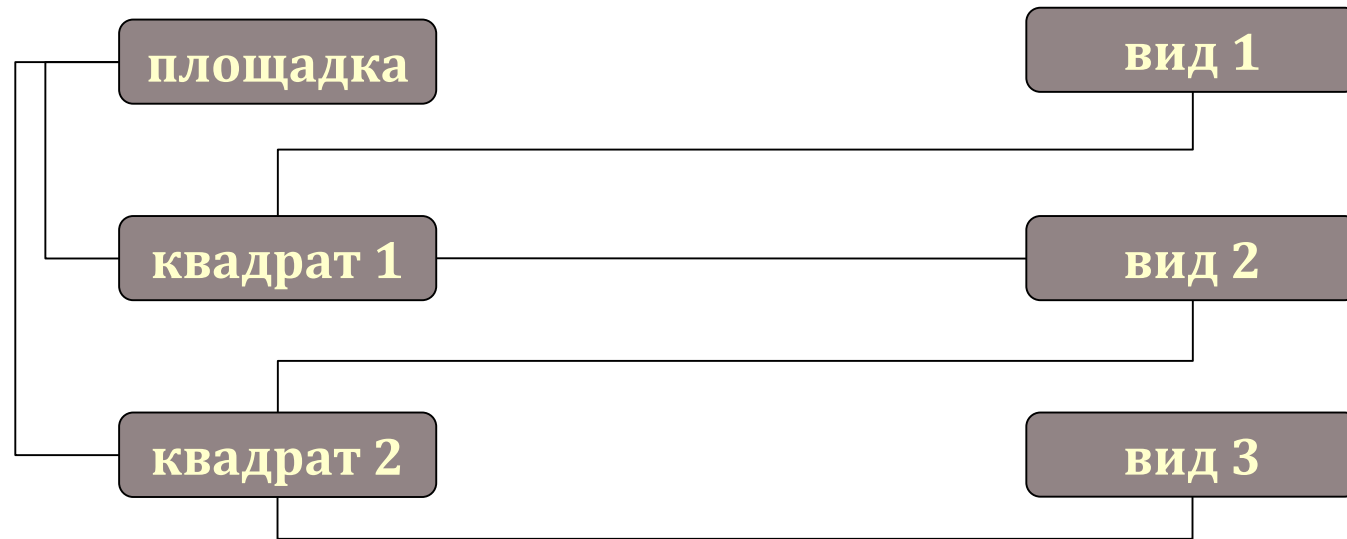


Иерархическая модель

- • Имеет жесткую структуру. Большие индексные файлы.
- Непригодна для использования в сложных системах с оперативной обработкой запросов, не может обеспечить быстроедействие.
- Трудно модифицировать.
- Не используется при большом количестве уровней.
- + • Эффективная организация машинной памяти.
- Применяется для задач с иерархической структурой информации и запросов, составления различного рода классификаторов.

БД 1-го поколения

Сетевая модель



- Усовершенствованная иерархическая структура. Запись в каждом узле может быть связана с несколькими другими узлами.
- Записи содержат указатели, определяющие местоположение других записей, связанных с ними.
- Ускорение доступа к данным.



Сетевая модель

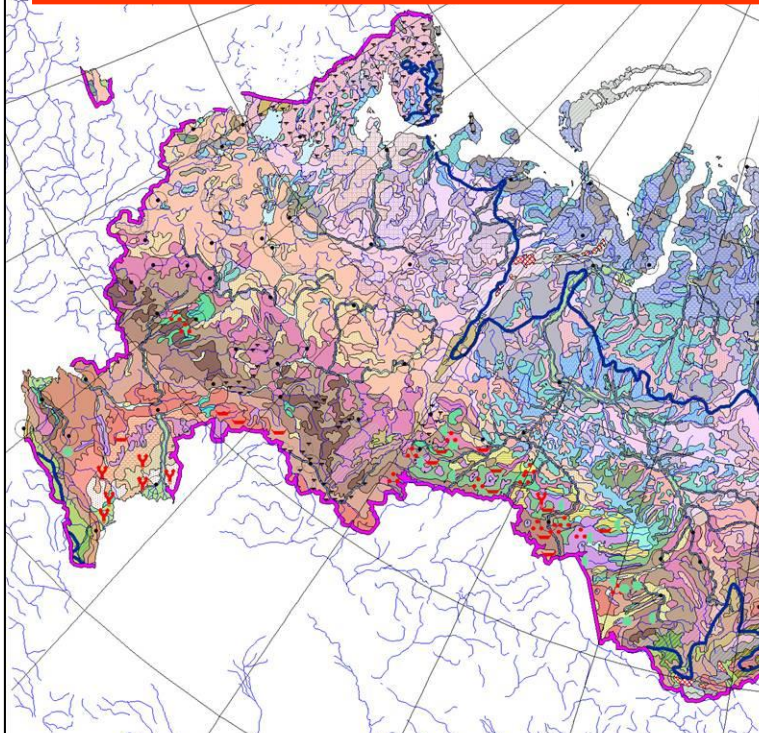
- • Сложна для проектирования и поддержки.
- Трудноизменяемая структура, редактирование записей требует также редактирования указателей.
- + • Эффективна при решении сетевых и коммуникационных задач.

Реляционная модель

- Англ.: relation (отношение; связь)
- **Реляционная модель** – табличная модель данных, основным средством структуризации в которой является отношение (связь). Таблица отражает объекты реального мира и состоит из **строк, или записей** (1 строка – сведения об 1 объекте) и **столбцов, или полей** (1 столбец – 1 характеристика (атрибут) всех объектов).
- В одной базе данных может быть множество таблиц.
- **Первичный ключ** – столбец, значения в котором однозначно идентифицируют каждую строку реляционной модели (каждый объект).
- **Внешний ключ** – столбец в реляционной модели, поддерживающий связь между разными реляционными таблицами.

Количество контуров на карте – 2544

«ID» – *первичный ключ*



ID	Номер_в_легенде	Name	area	площадь_про
2 521	17	Тж	96 491 880 546,18	0,5721
2 522	19	ПБ	9 528 811 289,96	0,0565
2 523	18	Тж\г	5 450 939 052,15	0,0323
2 524	19	ПБ	10 447 205 130,3	0,0619
2 525	48	К	5 636 815 024,15	0,0334
2 526	26	По\дж	12 888 622 539,74	0,0764
2 527	26	По\дж	4 308 683 309,84	0,0255
2 528	92	По	18 010 960 634,2	0,1068
2 529	25	П\дв	315 227 964,96	0,0019
2 530	37	Л	25 147 903 490,26	0,1491
2 531	28	Тж\д	2 785 748 463,17	0,0165
2 532	24	П\д	1 135 011 394,96	0,0067
2 533	37	Л	1 306 568 515,57	0,0077
2 534	41	Ч\о	2 377 360 336,16	0,0141
	17	Тж	54 035 426 150,73	0,3204
	30	Дк	1 354 154 935,94	0,0080
	59	Бп	992 473 381,9	0,0059
	45	Ч\мк	2 710 578 847,19	0,0161
	51	К\сн	20 346 737 919,02	0,1206
	92	По	108 608 835 053,86	0,6440
	26	По\дж	1 697 199 542,84	0,0101
	24	П\д	406 085 542 315,64	9,9999
2 543	26	По\дж	1 433 357 759,39	0,0085
2 544	26	По\дж	2 297 199 333,53	0,0136

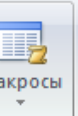
Количество номеров в легенде – 95

(почвы, почвенные комплексы,
непочвенные образования)

«№ в легенде» – *внешний ключ*



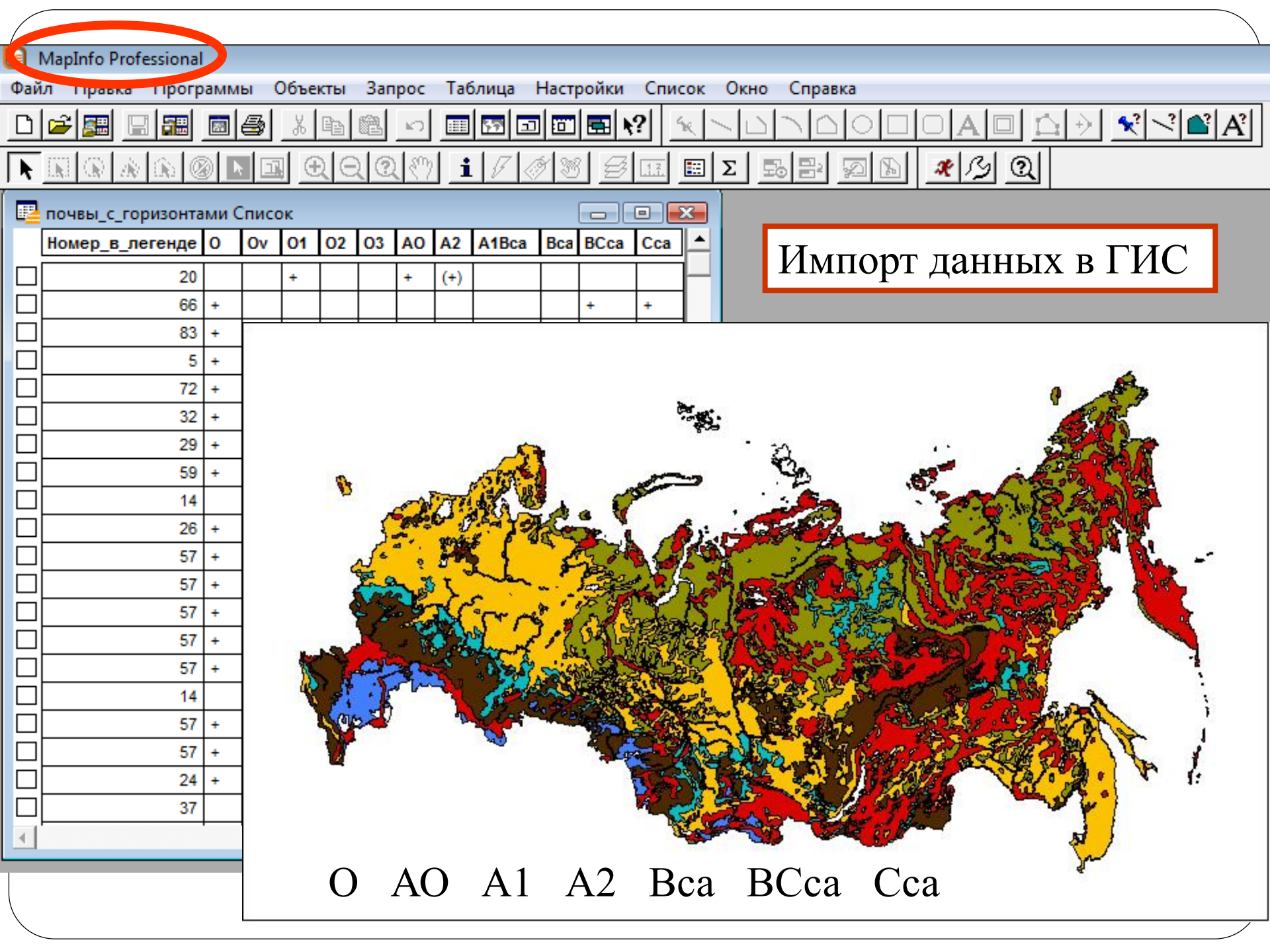
Разработчик



B	Ma
---	----

A2B

Готово



Основные преимущества реляционных СУБД в ГИС

- Возможность манипулирования данными без необходимости знания физической организации БД.
- Удобная работа с атрибутами:
 - можно хранить отдельно от пространственных данных;
 - можно привязывать к пространственным единицам;
 - можно изменять и удалять без изменения пространственных данных.
- Мощное средство аналитической работы с любым количеством атрибутов.
- Освобождает пользователя от всех ограничений, связанных с организацией хранения данных и спецификой аппаратуры.

Реляционная модель

- +**
 - Простота представления и формирования БД.
 - Универсальность и удобство обработки данных.
 - Проблема организации работы решается индексацией записей.
 - Простота модели ввиду отсутствия специальных механизмов навигации (как в иерархической и сетевой моделях данных).
- - Некоторая ограниченность использования при сложных структурах данных – увеличение времени на извлечение данных (как следствие простоты модели).