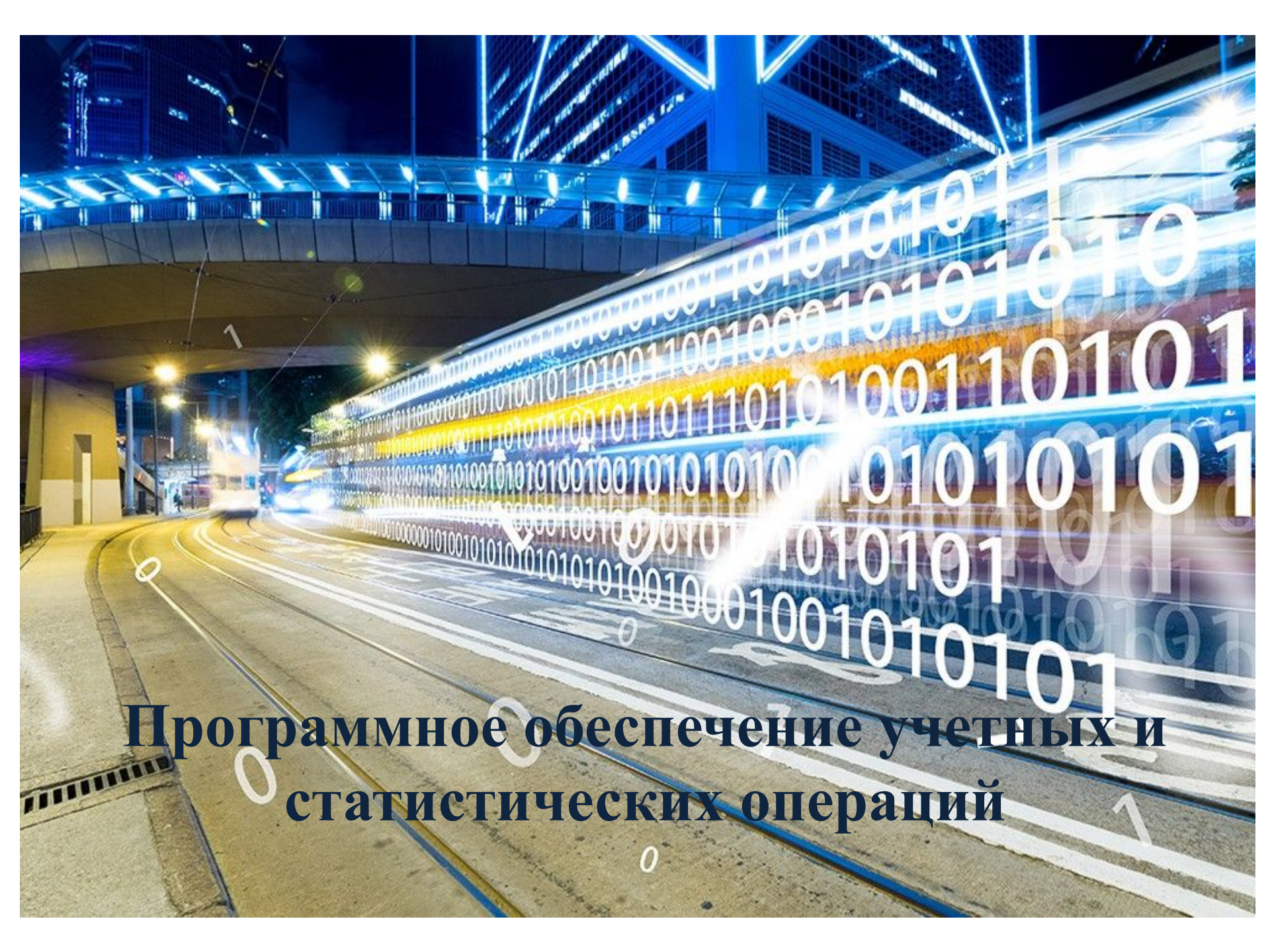




Программное обеспечение учетных и статистических операций

Бизнес-аналитика



Бизнес-аналитика



Бизнес-аналитика = BA (Business Analytics)

Бизнес-аналитика = BI (Business Intelligence)

- Инструменты и приложения управления эффективностью бизнеса PM (performance management)
- Хранилища данных DW (Data Warehouse)
- Методы и инструменты для построения информативных отчётов о текущей ситуации.
- Инструменты и приложения для поиска, анализа, моделирования и доставки информации, необходимой для принятия решений

Бизнес-аналитик

Использует методы бизнес-анализа для аналитики потребностей деятельности организаций -> моделирование деятельности организации, **бизнес-процессов -> системы BPM**

Использует инструментальные, программные средства анализа данных (DM) -> **корпоративные аналитические системы**

Введение в теорию бизнес-анализа

АНАЛИЗ ДАННЫХ — это исследования, связанные с обработкой многомерной системы данных, имеющей множество параметров; средство проверки гипотез и решения задач исследователя.

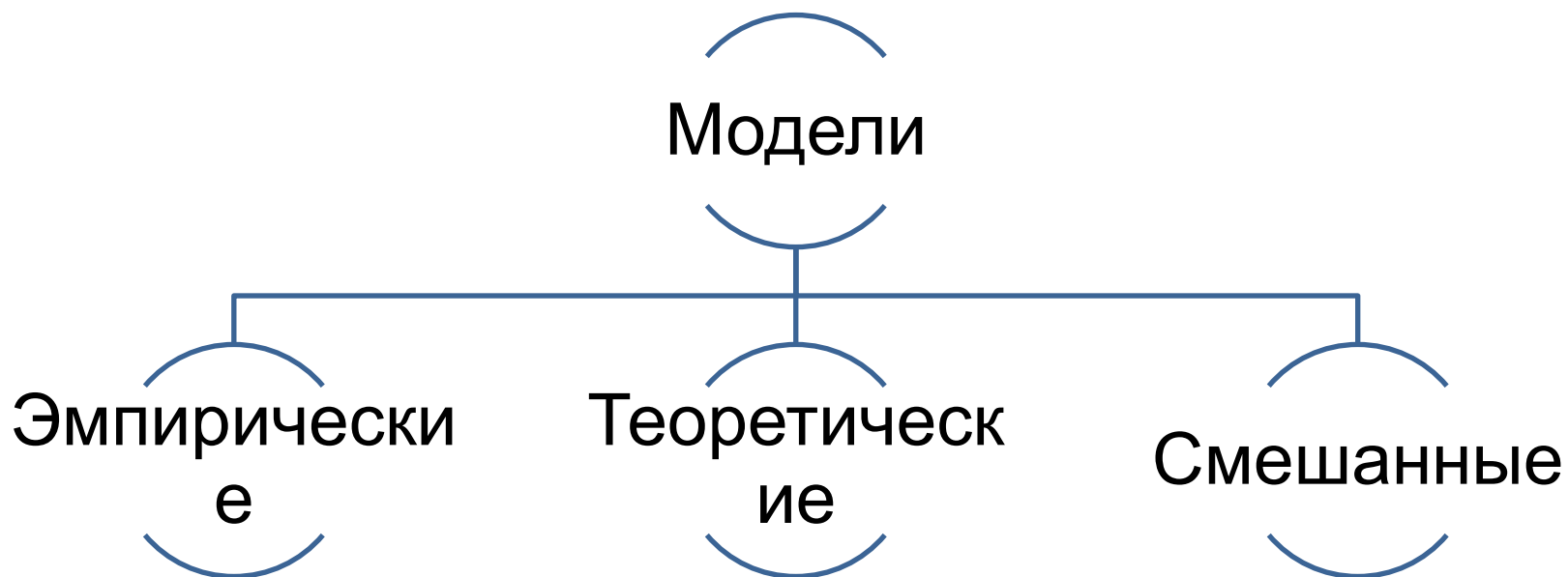
Введение в теорию бизнес-анализа

МОДЕЛЬ («мера», «способ», «сходство с какой-то вещью») — объект или описание объекта, системы для замещения (при определенных условиях, предположениях, гипотезах) одной системы (то есть оригинала) другой системой для лучшего изучения оригинала или воспроизведения каких-либо его свойств.

МОДЕЛИРОВАНИЕ (построение модели)

— универсальный способ изучения окружающего мира, позволяющий обнаруживать зависимости, прогнозировать, разбивать на группы.

Введение в теорию бизнес-анализа



Важные свойства любой модели.

§ Упрощенность. § Конечность. §

Приближенность.

§ Адекватность. § Целостность. §

Замкнутость.

§ Управляемость.

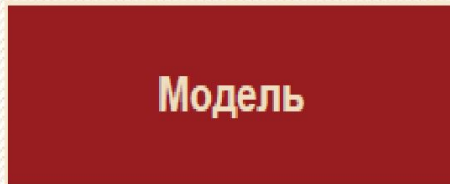
Введение в теорию бизнес-анализа

Аналитический подход к моделированию - модель не «подстраивается» под действительность, а мы пытаемся подобрать существующую аналитическую модель таким образом, чтобы она адекватно отражала реальность.

Исследователь



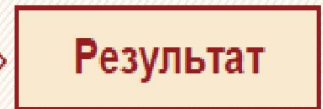
Модель



Данные

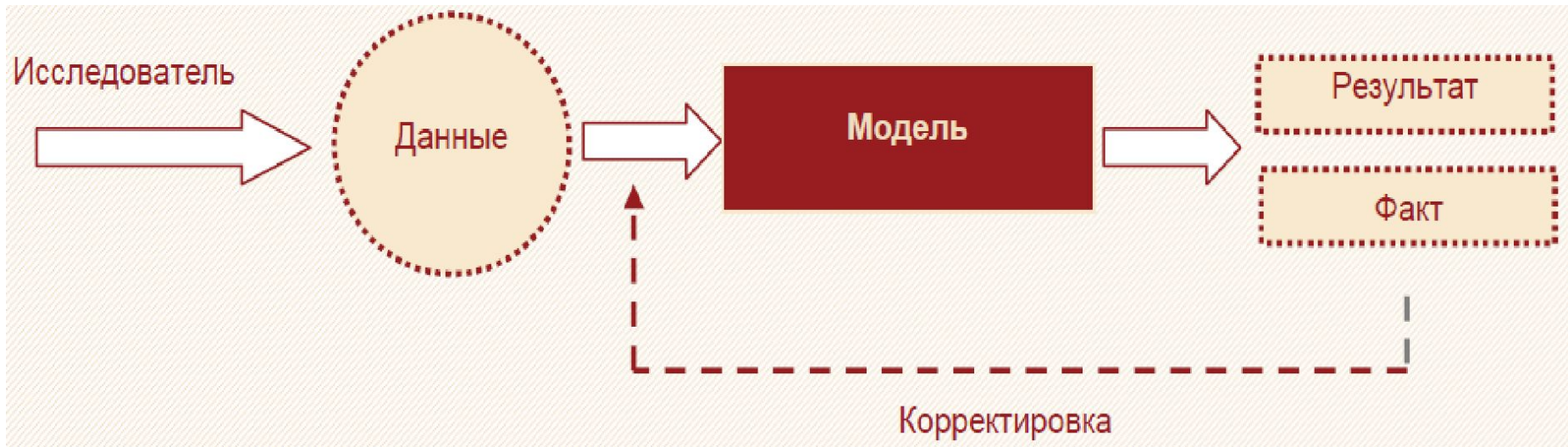


Результат



Введение в теорию бизнес-анализа

Информационный подход к моделированию
— ориентирован на использование данных;
модель «подстраивается» под
действительность.



Введение в теорию бизнес-анализа

Принципы и процесс анализа данных

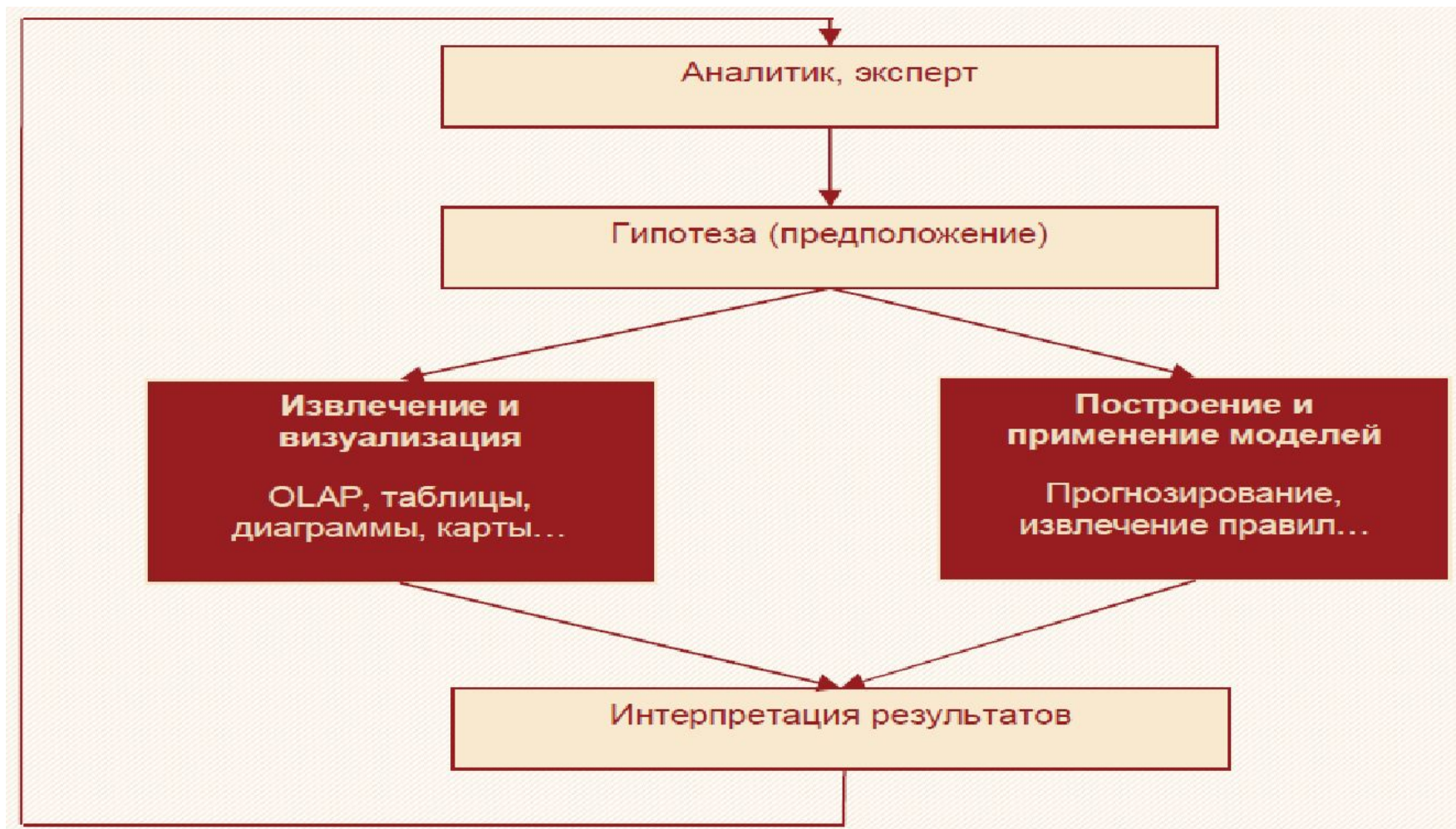
ЭКСПЕРТ — специалист в предметной области, профессионал, эффективно решающий задачи.

АНАЛИТИК — специалист в области анализа и моделирования, на достаточном уровне владеющий инструментальными и программными средствами анализа данных; проводит систематизацию данных, опрос мнений экспертов, координацию действий всех участников проекта по анализу данных.

ГИПОТЕЗА - предположение о влиянии какого-либо фактора или группы факторов на результат.

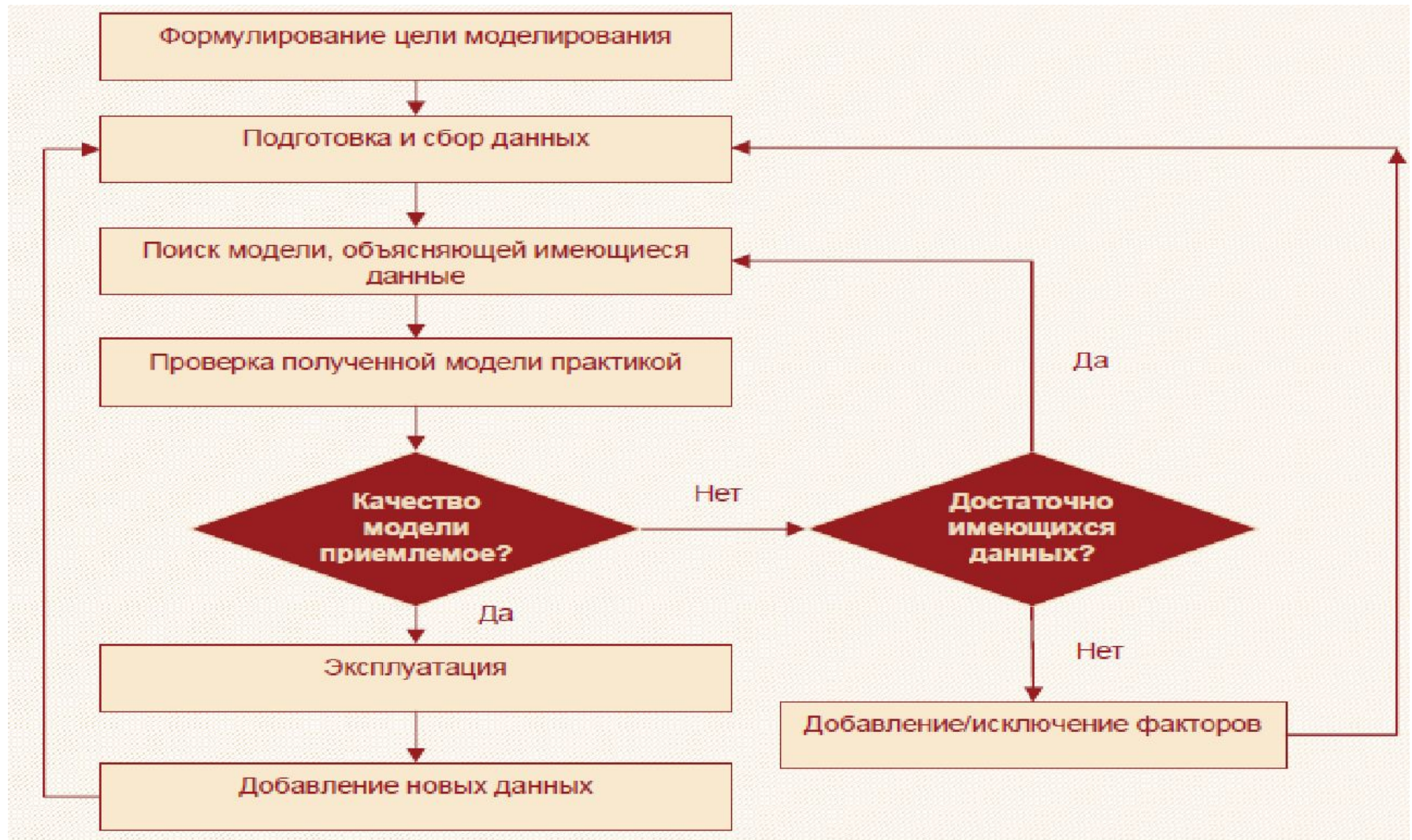
Введение в теорию бизнес-анализа

Принципы и процесс анализа данных



Введение в теорию бизнес-анализа

Процесс построения моделей



Введение в теорию бизнес-анализа

- **Бизнес-аналитика** — это «инструменты и приложения для поиска, анализа, моделирования и доставки информации, необходимой для принятия решений»;
- мультидисциплинарная область на стыке ИТ, БД, алгоритмов интеллектуальной обработки информации, математической статистики и методов визуализации.

Введение в бизнес-анализ

- Формы представления данных
- Типы и виды данных
- Формализация данных
- Методы сбора данных
- Требования, предъявляемые к данным

Бизнес-данные

ДАННЫЕ - сведения, характеризующие систему, явление, процесс или объект, представленные в определенной форме и предназначенные для дальнейшего использования.



Бизнес-данные

Форматы данных

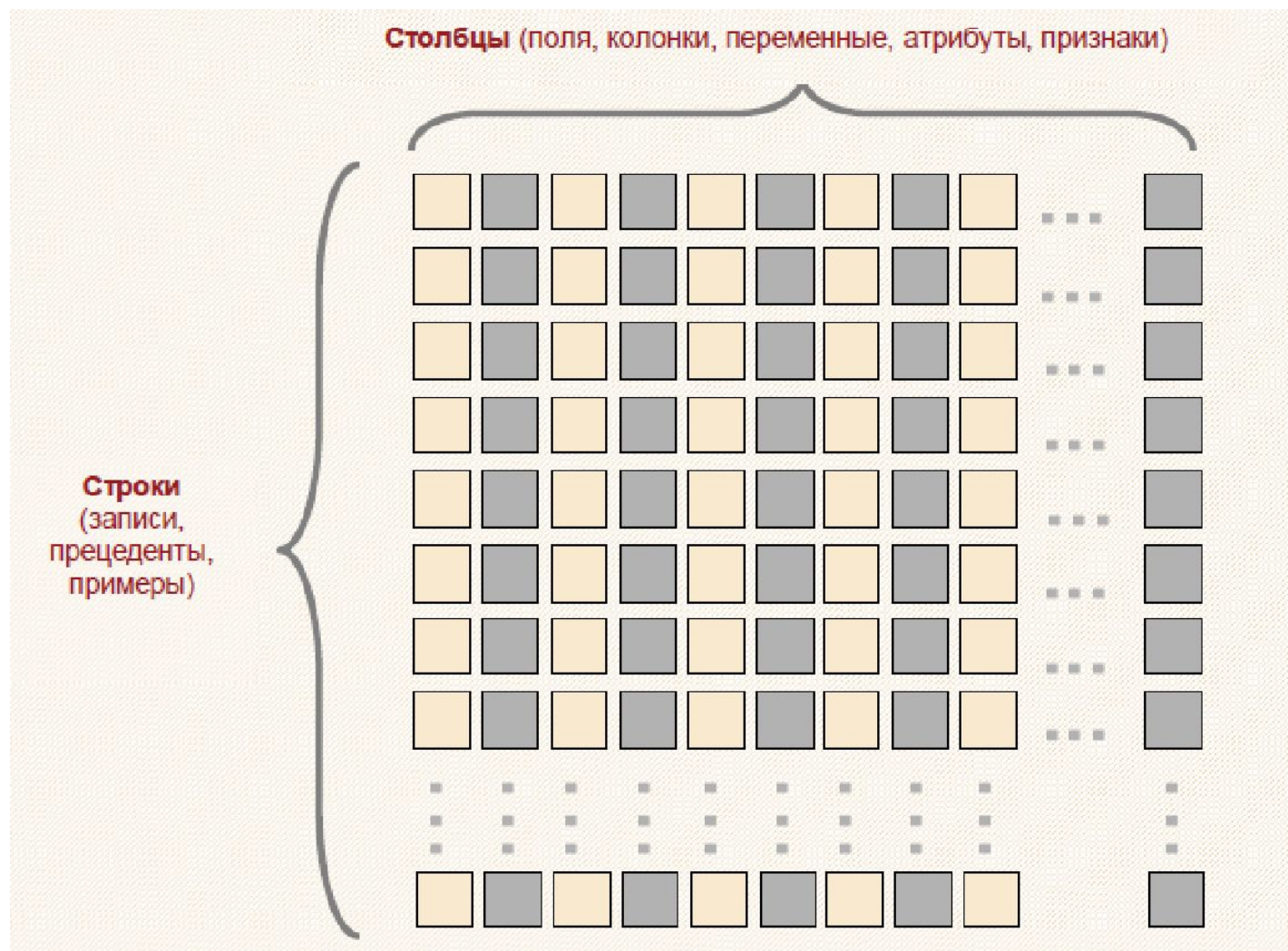
Неструктурированные – произвольная форма: видео, аудио, речь, графика и т.п.

Структурированные – упорядоченные определенным образом с целью применения к ним визуального или машинного анализа: БД и т.п.

Слабоструктурированные – некоторые правила и форматы представления определены в общем виде: строка с адресом, ФИО, дата и т.п.

Бизнес-данные

Таблица-модель хранения структурированных данных



Бизнес-данные

Пример стандартизации слабоструктурированных данных

390045 г. Рязань, ул. Ленина, д. 45 корп. 1



Поле	Значение
Индекс	390045
Город	Рязань
Улица	Ленина
Дом	45
Корпус	1

Бизнес-данные

Типы структурированных данных



Бизнес-данные

Подтипы **строковых** данных

Упорядоченные (ординальные)

Операции «>», «<», «=»

Категориальные

Операции «=», «≠»

Код заявки	Фамилия	Образование	Профессия	Город
Z-01	Иванов	высшее	инженер	Москва
Z-02	Кузнецова	среднее	бухгалтер	Коломна

Бизнес-данные

Вид данных

Виды
данных

Непрерывные – принимают какое угодно значение в любой интервал времени; ариф. операции «имеют смысл»

Дискретные – признак, общее число конечно или бесконечно; ариф. операции «не имеют смысла»

Тип данных	Вид данных	
	Непрерывный	Дискретный
Целый	+	+
Вещественный	+	+
Строковый		+
Логический		+
Дата/время	+	+

Бизнес-данные

Представления наборов данных

Данные

Упорядоченные — каждому столбцу соответствует один фактор, а в каждую строку заносятся упорядоченные по какому-либо признаку события с интервалом периода между строками.

Неупорядоченные — каждому столбцу соответствует фактор, а в каждую строку заносится пример (ситуация, прецедент), упорядоченность строк не требуется.

Бизнес-данные

Пример представления упорядоченных данных

	Дата	Количество	Сумма
	01.01.2004	4	283.31
	01.01.2004	1	72.48
	01.01.2004	1	173.32
	02.01.2004	6	294.84
	02.01.2004	2	405.76
	02.01.2004	12	303.13
	02.01.2004	1	210.5
	03.01.2004	6	521.16
	03.01.2004	3	156.96

	Глубина	ВК	DS
►	887.9	8.85	0.218
	888.1	9.627	0.216
	888.3	14.584	0.217
	888.5	21.647	0.215
	888.7	17.172	0.216
	888.9	6.118	0.215
	889.1	2.886	0.217
	889.3	2.506	0.219

Бизнес-данные

Пример неупорядоченного набора данных

Номер	Банк	Реутеры	Филиалы	Город	Собственные активы
2	Внешторгбанк	.	32	Москва	23236327
3	Газпромбанк	GZPM	27	Москва	9255041
4	ООО "Международный Промышленный банк"	TIBP	4	Москва	26409116
5	Международный Московский Банк	IMBX	1	Москва	1176462
6	ОАО "АЛЬФА-БАНК"	ALFM	17	Москва	12446938
7	ОАО "ПСБ"	ICSP	44	Санкт-Петербург	1275859
8	Банк Москвы	.	34	Москва	3335734
9	АКБ "РОСБАНК" (ОАО)	.	13	Москва	4691449
10	АКБ "ДИБ"	DIBM	0	Москва	2616993

Бизнес-данные

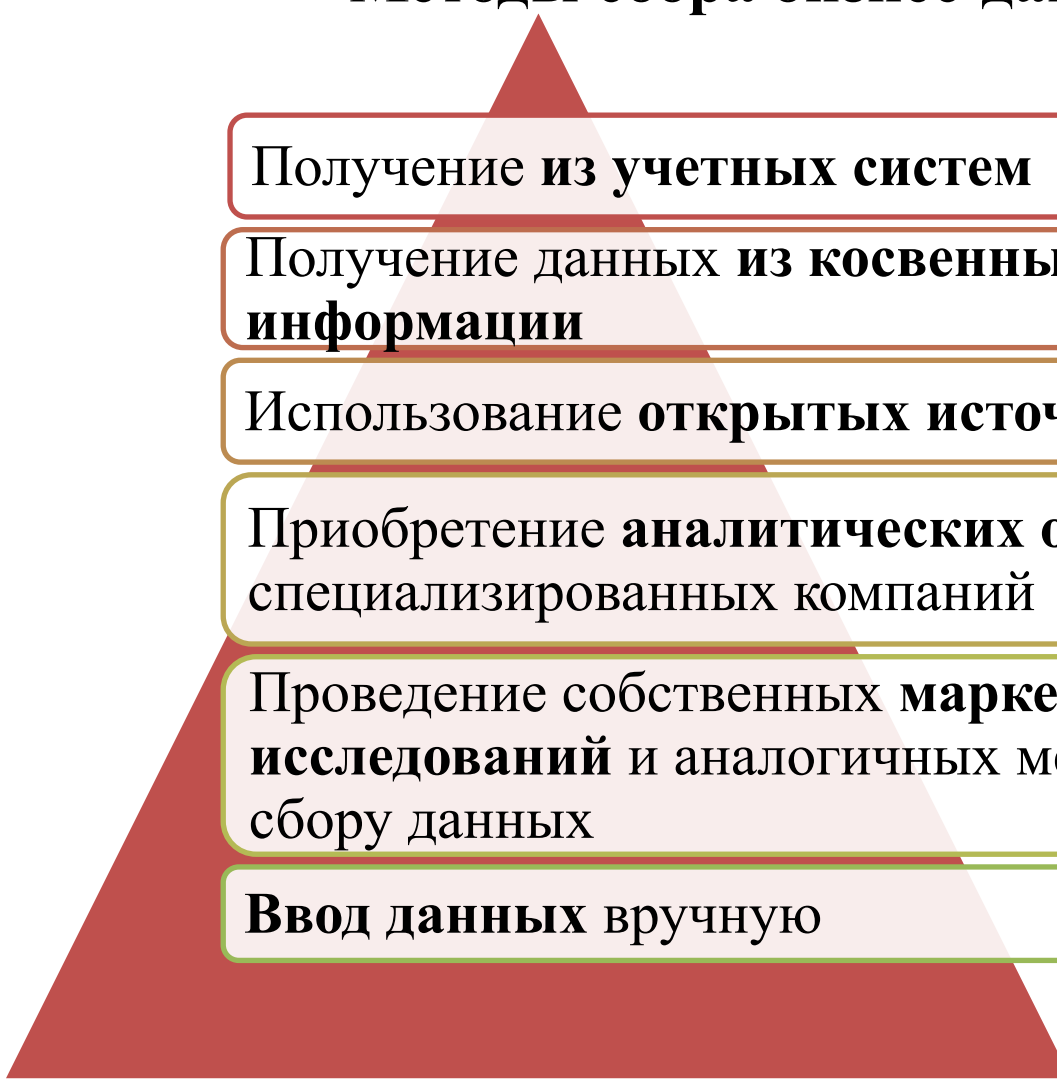
Транзакционные данные

Под *транзакцией* подразумеваются несколько объектов или действий, являющихся логически связанной единицей.



Бизнес-данные

Методы сбора бизнес-данных



Получение из учетных систем

Получение данных из косвенных источников информации

Использование открытых источников

Приобретение аналитических отчетов у специализированных компаний

Проведение собственных маркетинговых исследований и аналогичных мероприятий по сбору данных

Ввод данных вручную

Бизнес-данные

Информативность данных

Неинформативные признаки

содержащие
только одно
значение;

содержащие в
основном одно
значение;

с уникальными
значениями;

между которыми
сильная
корреляция.

Признак
1
1
1
1
1
1
1
1
1
1
1
1

Признак
1
1
1
1
1
0
1
1
1
1
1
1

№ паспорта
0936-866096
8355-512943
8017-098471
2762-945535
0459-997701
6291-817248
0094-883508
6385-082612
9290-732300
7022-736158
3127-709332
4179-171975

Пол	Gender
Жен	0
Жен	0
Жен	0
Муж	1
Муж	1
Жен	0
Жен	0
Муж	1
Муж	1
Жен	0
Жен	0
Муж	1

Бизнес-данные

Требования к бизнес-данным

- Необходимо иметь данные за один полный сезон/цикл с возможностью варьирования интервалов (понедельное, ежемесячное и т. д.)
- Максимальный горизонт прогнозирования зависит от объема данных: данные за 1,5 года — прогноз максимум на 1 месяц; данные за 2–3 года — на 2 месяца.

- Количество примеров (прецедентов) должно быть значительно больше количества факторов.
- Данные должны покрывать как можно больше ситуаций реального процесса.
- Пропорции различных примеров (прецедентов) должны примерно соответствовать реальному процессу

- Целесообразно использовать большой объем данных:
- 300–500 объектов — не менее 10 тыс. транзакций;
- 500–1000 объектов — более 300 тыс. транзакций.

Вопросы - ?