

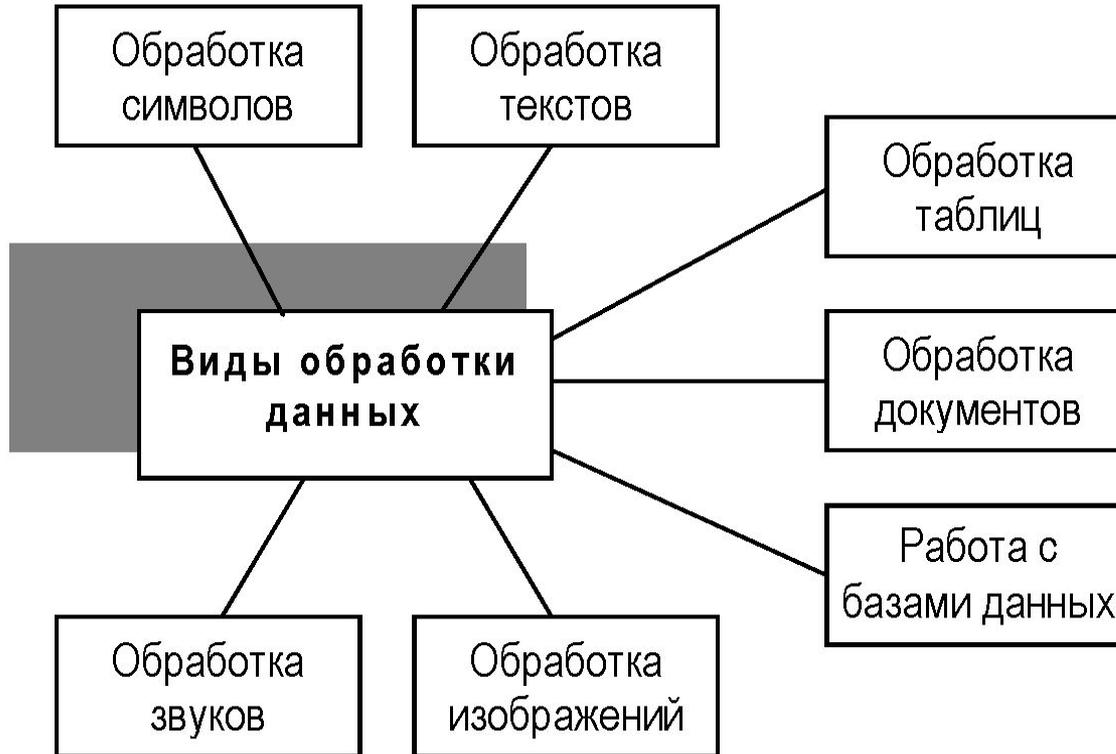
Лекция 9

Алгоритмы обработки информации

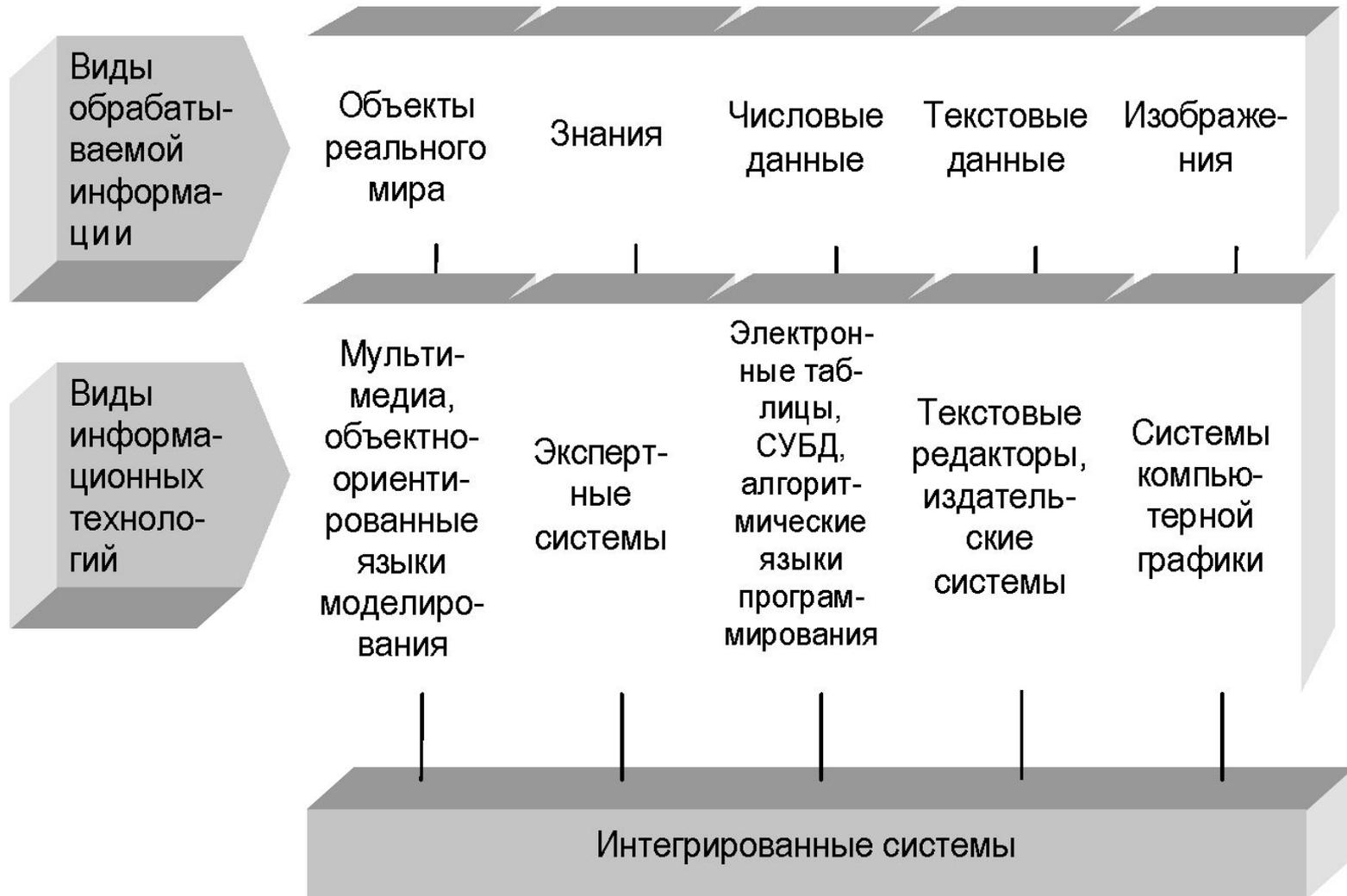
1. Обработка информации: анализ, синтез, преобразование
2. Математический алгоритм
3. Алгоритмизация программы

**ИСХОДНАЯ ИНФОРМАЦИЯ → АЛГОРИТМ ОБРАБОТКИ
→ ИТОГОВАЯ ИНФОРМАЦИЯ (РЕЗУЛЬТАТ)**

Виды обработки данных



Классификация ИТ в зависимости от типа обрабатываемой информации



Технологический процесс обработки данных

ПОДГОТОВИТЕЛЬНЫЙ этап

- подготовка к решению задачи (создание справочников, введение в память компьютера необходимых постоянных данных и др.)

начальный этап

- операциями по сбору, регистрации и размещению документов в базовые массивы (возможна обработка документов, заполненных вручную, однако более эффективным является электронное документирование)

ОСНОВНОЙ

- получение необходимых отчетных форм, когда из компьютерной базы данных извлекаются рабочие массивы, подлежащие группировке по соответствующим ключевым признакам, подсчету по ним итоговых данных с распечаткой в дальнейшем полученных отчетных документов

Типы организации технологических процессов

Предметный

создание параллельно действующих технологических линий, специализирующихся на обработке информации и решении конкретных комплексов задач

Пооперационный (поточный)

последовательное преобразование обрабатываемой информации согласно технологии, представленной в виде непрерывной последовательности сменяющих друг друга операций, выполняемых в автоматическом режиме

Внемашинная технология

- объединяет операции сбора и регистрации данных, запись данных на машинные носители с контролем

Внутримашинная технология

- формирование новых массивов информации;
- упорядочение информационных массивов;
- выборка из массива некоторых частей записи;
- слияние и разделение массивов;
- внесение изменений в массив;
- выполнение арифметических действий над реквизитами в пределах записей, в пределах массивов, над записями нескольких массивов

Первичная обработка

Типы ошибок

- 1. Противоречивость информации;*
- 2. Пропуски в данных;*
- 3. Аномальные значения;*
- 4. Шум;*
- 5. Ошибки ввода данных.*

Проблемы качества данных

Проблемы отдельных источников

Уровень схемы
(Недостаток ограничений целостности, плохой дизайн схемы)

- Уникальность
- Целостность ссылок
- ...

Уровень элемента
(Ошибки при вводе данных)

- Орфографические ошибки
- Избыточность/дубликаты
- Противоречивые значения
- ...

Проблемы множества источников

Уровень схемы
(Неоднородные модели данных и дизайн схем)

- Конфликты наименований
- Структурные конфликты
- ...

Уровень элемента
(Перекрывающиеся, противоречащие и несогласованные данные)

- Несогласованная агрегация
- Несогласованная синхронизация
- ...

Примеры проблем отдельного источника данных на уровне схемы (нарушение ограничения целостности)

Область/ Проблема		Загрязненные данные	Причины/ Примечания
Атрибут	Недопустимые значения	bdate=30.13.70	Значения вне допустимой области
Запись	Нарушенные логических связей	age=22, bdate=12.02.70	Возраст не соответствует году рождения
Тип записи	Нарушение уникальности	emp1=(name="John Smith", SSN="123456") emp2= (name="Peter Miller", SSN="123456")	Нарушена уникальность SSN (номера социального страхования)
Источник	Нарушение целостности ссылок	emp= (name="John Smith", deptno=127)	Не определено подразделение 127

Примеры проблем отдельных источников данных на уровне элемента

Область/ Проблема	Загрязненные данные	Причины/ Примечания	
Атрибут	Утраченные значения	phone= 9999-999999	Неведенные значения (бессмысленные или неопределенные)
	Орфографические ошибки	city="Liipzig"	Обычно опечатки, фонетические ошибки
	Зашифрованные значения и аббревиатуры	experince="B" occupation="DB Prog."	
	Вложенные значения	name="J.Smith 12.02.70 New York"	Множество значений в одном атрибуте (например, в поле свободного формата)
	Значения, не соответствующие своим полям	city="Germany"	
Запись	Нарушенные логических связей	city="Redmond", zip=77777	Индекс должен соответствовать городу
Тип записи	Перестановка слов	name1="J.Smith", name2="Miller P."	Обычно встречается в полях свободного формата
	Дублирующиеся записи	emp1= (name="John Smith",...); emp2= (name="J.Smith",...)	В результате ошибок при вводе данных некое лицо присутствует дважды
	Противоречивые записи	emp1= (name="John Smith", bdate=12.02.70); emp2= (name="J.Smith", bdate=12.12.70)	Один и тот же объект реального окружения описывается различными значениями
Источник	Неверные ссылки	emp= (name="John Smith", deptno=17)	Подразделение 17 определено, но не соответствует объекту

Примеры проблем множества источников на уровне схемы и элемента данных

Потребитель (источник 1)

<i>ID потреб</i>	<i>Имя</i>	<i>Ул</i>	<i>Гор</i>	<i>Пол</i>
11	Kristen Smith	2 Hurley P1	South Fork, MN 48503	0
14	Chrisian Smith	Hurley St 2	S Fork MN	1

Клиент (источник 2)

<i>№ кл</i>	<i>Фам</i>	<i>Имя</i>	<i>Род</i>	<i>Адр</i>	<i>Тел/ Факс</i>
24	Smith	Cristoph	F	23 Harley St, Chicago IL, 60633-2394	333-222-6542/ 333-222-6599
493	Smith	Kris L.	M	2 Hurley Place, South Fork MN, 48503-5998	444-555-6666

Потребители (объединение целевых данных с очищенными)

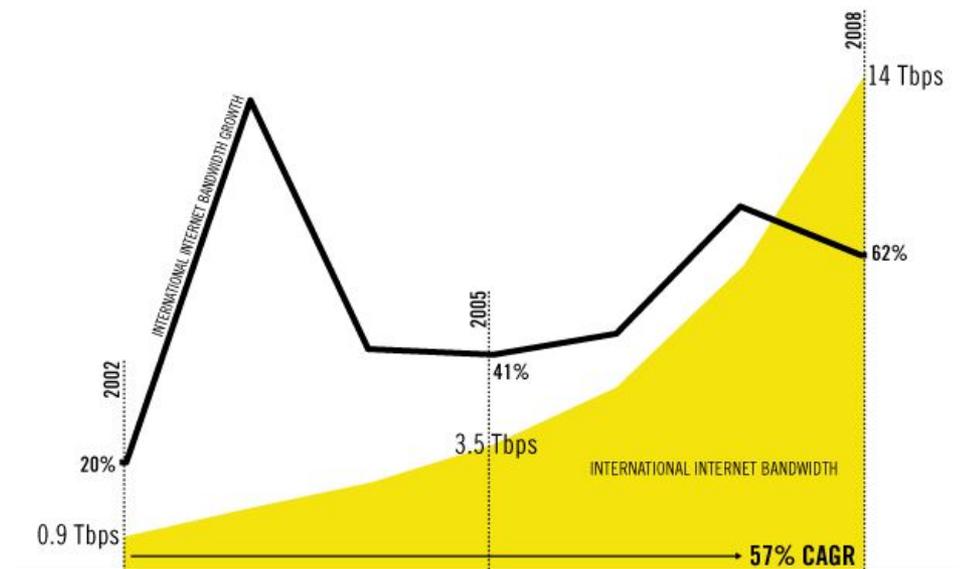
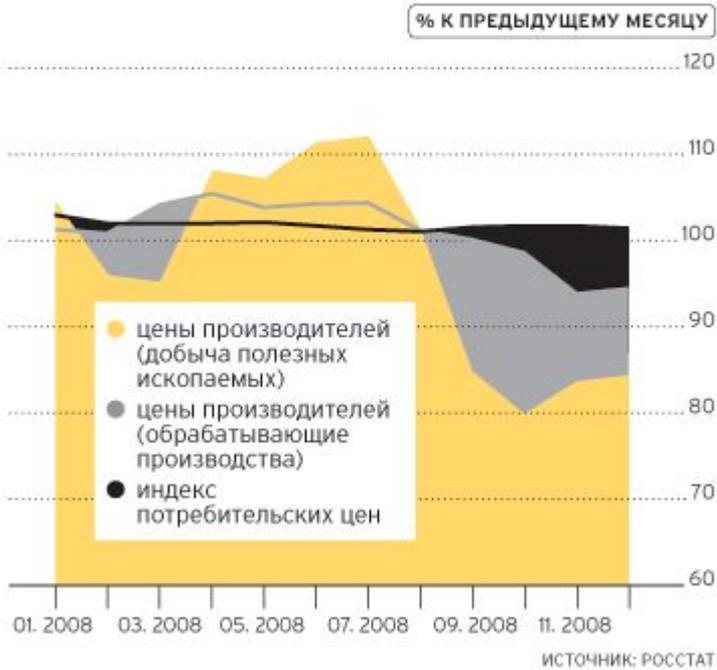
<i>№</i>	<i>Фам</i>	<i>Им</i>	<i>Род</i>	<i>Ул</i>	<i>Гор</i>	<i>Шт</i>	<i>Инд</i>	<i>Тел</i>	<i>Факс</i>	<i>ID потр</i>	<i>№ кл</i>
1	Smith	Kristen L.	F	2 Hurley Place	South Fork	MN	48 503-5998	444-555-666		11	493
2	Smith	Christian	M	2 Hurley Place	South Fork	MN	48 503-5998			24	
3	Smith	Christoph	M	23 Harley Street	Chicago	IL	60 633-2394	333-222-6542	333-222-6599		24

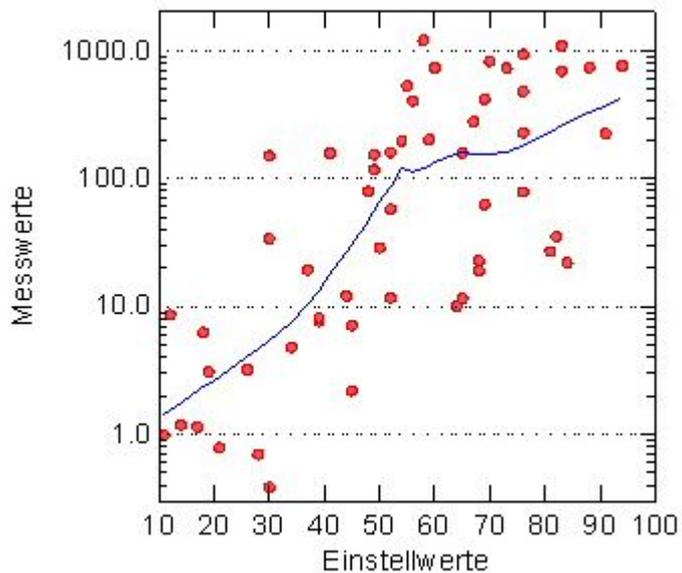
Вторичная обработка информации

Визуализация данных

Не дошло до розницы

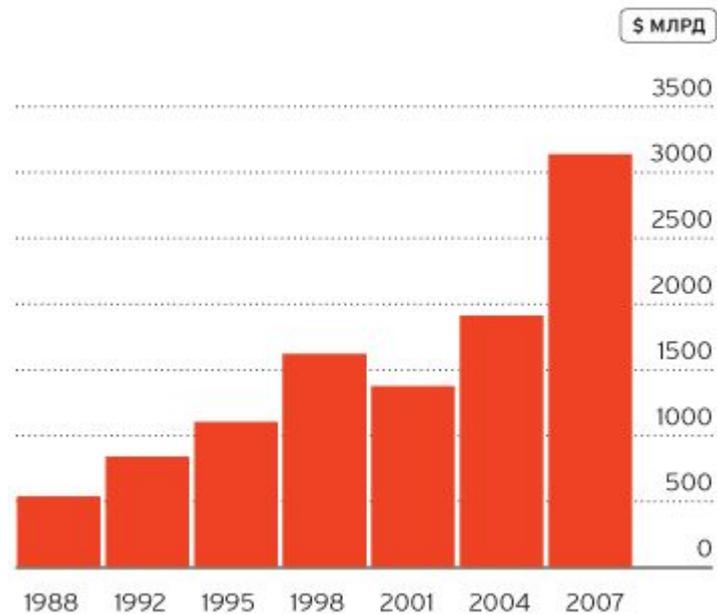
СНИЖЕНИЕ ЦЕН ПРОИЗВОДИТЕЛЕЙ НЕ ПОВЛИЯЛО НА ПОТРЕБИТЕЛЬСКУЮ ИНФЛЯЦИЮ В РОССИИ





От кризиса до кризиса

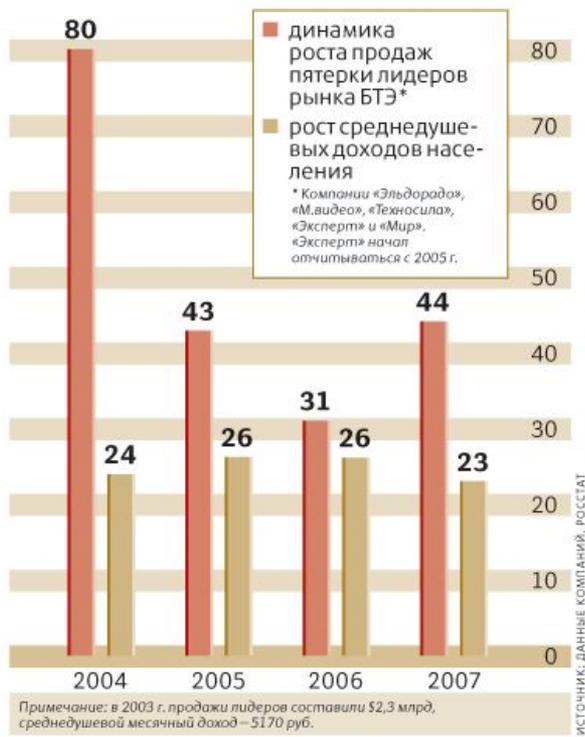
КАК РОС ОБЪЕМ ТОРГОВ НА FOREX



ИСТОЧНИК: TRIENNIAL CENTRAL BANK SURVEY (BIS)

▶ ВПЕРЕДИ ПАРОВОЗА

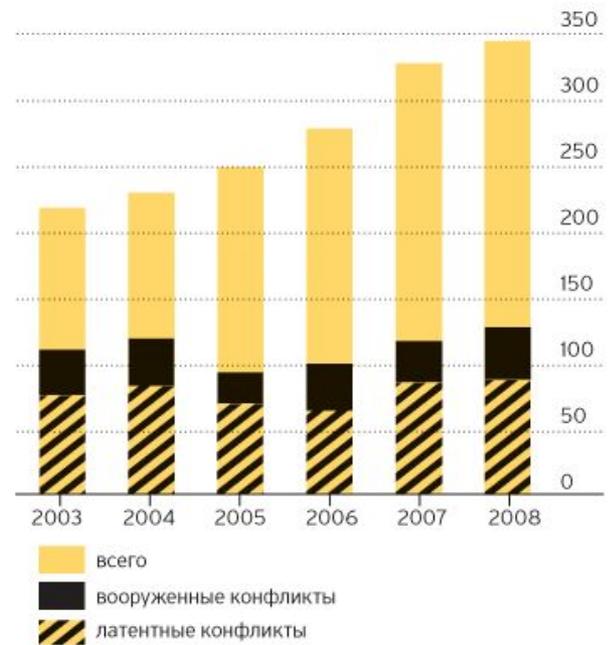
Продажи БТЭ долгое время росли быстрее, чем увеличивались доходы россиян



Примечание: в 2003 г. продажи лидеров составили \$2,3 млрд, среднедушевой месячный доход – \$170 руб.

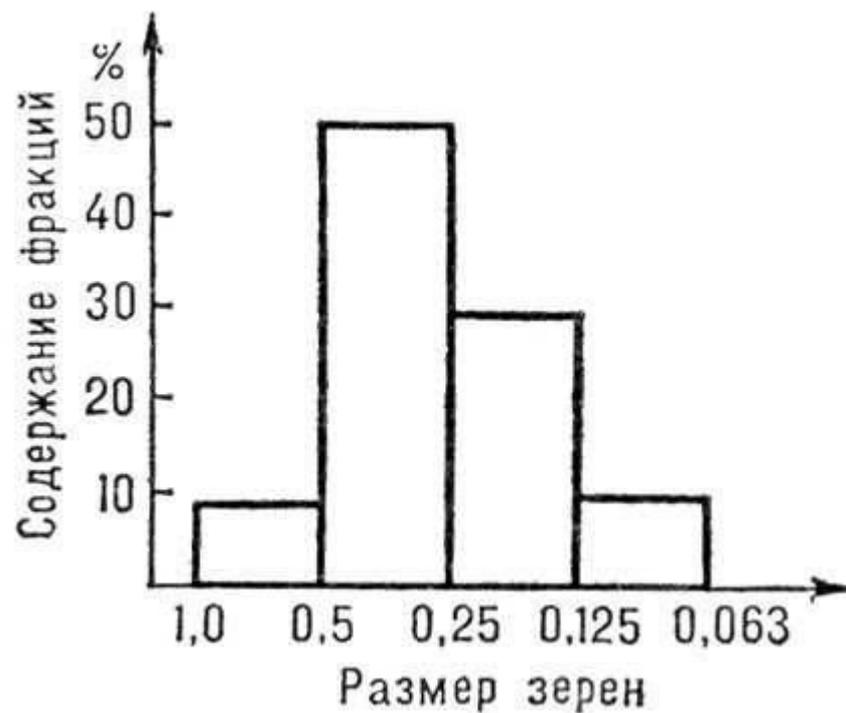
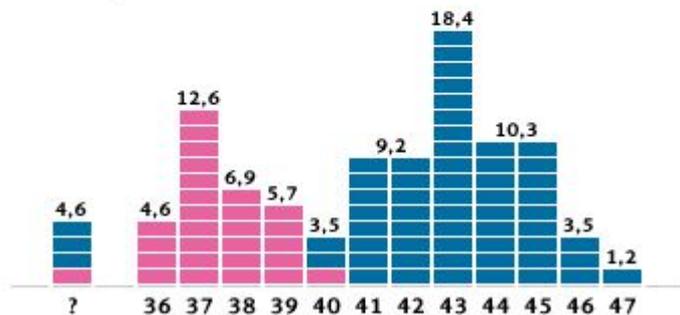
Скрытая угроза

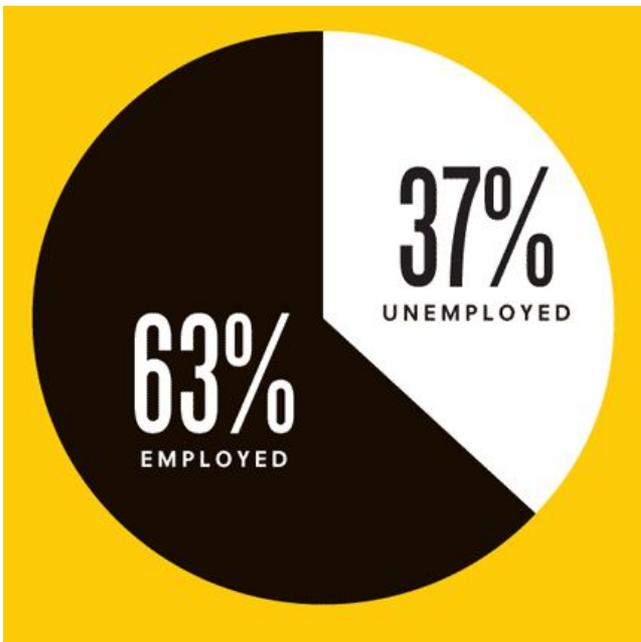
БЫСТРЕЕ ВСЕГО В МИРЕ РАСТЕТ ЧИСЛО КОНФЛИКТОВ, НЕ ПЕРЕШЕДШИХ В СТАДИЮ ВОЙНЫ



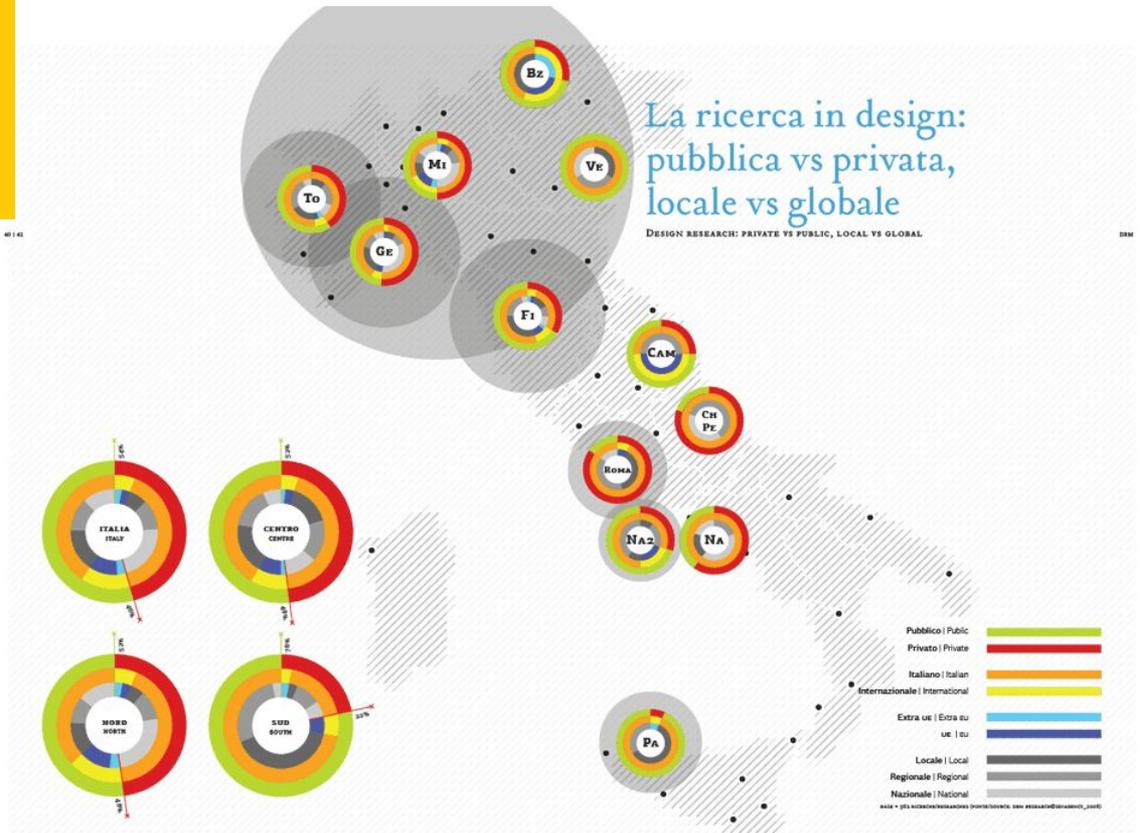
Процентное распределение
размеров **женской** и **мужской** обуви
сотрудников Студии Лебедева

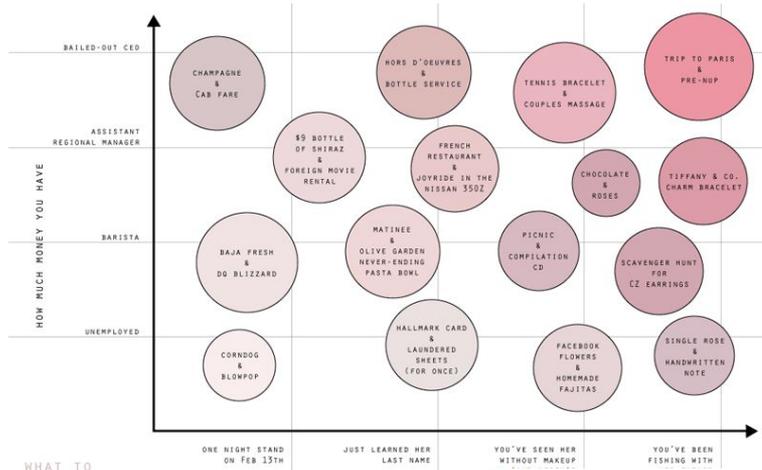
октябрь 2002 г.





01 | 11





WHAT TO DO FOR HER ON VALENTINE'S DAY

HOW MUCH YOU

Главные автодилеры России

Количество дилерских центров

«Инком-авто»

Объем продаж «Инком-авто» автодилеров составляет около половины. В планах компании создание региональной сети из 50 автодилеров

2,3%
Доля рынка в 2007 году

64 176
Объем продаж автомобилей в 2007 году

Марки: Ford, Opel, Chevrolet, Cadillac, Volkswagen, Skoda, Honda, Renault, Peugeot, Fiat, SangYong, Hyundai, KIA, Chevy, Geely, ZK, JMC, Foton, Daewoo, SENS, Lada, GAZ, VAZ

\$940
млн

«Автомир»

Специализируется на низких ценах при покупке, ведет агрессивную маркетинговую политику. Стратегия развития компании — вертикальное развитие, продвижение в регионы

2,9%
Доля рынка в 2007 году

81 200
Объем продаж автомобилей в 2007 году

Марки: Nissan, Ford, Mitsubishi, Opel, Mazda, Chevrolet, Cadillac, Saab, Hummer, Hyundai, Suzuki, Renault, Volkswagen, Peugeot, Fiat, Skoda, Citroen, Kia, Daewoo, SsangYong, Lada, GAZ, Chevy, Lifan

\$1,8
млрд

«Рольф»

Является общенациональным дистрибутором компании Mitsubishi. Самый крупный российский дилер. Активно развивает дополнительные сервисы

4,9%
Доля рынка в 2007 году

137 194
Объем продаж автомобилей в 2007 году

Марки: Audi, Ford, Hyundai, Mazda, Mercedes-Benz, Mitsubishi, Peugeot

\$2,7
млрд

\$1,9
млрд

Major

Собирает больше всего зарплатных брендов, при этом права дилерства по многим из них получены в 2007 году. Построил крупнейший автомобильный комплекс в Магнитогорье

1,8%
Доля рынка в 2007 году

50 000
Объем продаж автомобилей в 2007 году

Марки: Audi, Mercedes-Benz, Toyota, Lexus, Alfa Romeo, Volkswagen, Peugeot, Opel, Hummer, Cadillac, Saab, Chevrolet, Land Rover, Chrysler, Jeep, Dodge, Volvo, Honda, Mazda, Nissan, Ford, Renault, Skoda, MG/Rover

\$896
млн

\$1,3
млрд

«Независимость»

Стратегия выжить в регионах, наращивая темп количества субдилеров. Специализируется на автомобилях среднего сегмента, активно конкурирует с Maza Motors

0,8%
Доля рынка в 2007 году

23 000
Объем продаж автомобилей в 2007 году

Марки: Land Rover, Jaguar, Volvo, Ford, Mazda, Peugeot, Audi, Volkswagen

«Бизнес кар»

Контрольный пакет принадлежит Toyota; дилер сконцентрирован на восточном дистрибуторе, успешные тем самым оборот

1,1%
Доля рынка в 2007 году

29 000
Объем продаж автомобилей в 2007 году

Марки: Toyota, Lexus

\$1
млрд

\$800
млн

\$968
млн

«Авилон — Нью-Йорк моторс Москва»

Обладки всего дилерского центра в Москве, компания является крупным дилером Ford в Европе и крупным полнокомплектным дилером Mercedes-Benz во всех регионах

0,7%
Доля рынка в 2007 году

16 186
Объем продаж автомобилей в 2007 году

Марки: Ford, Mercedes-Benz, Maybach

Genser

Проводит активные политики строительства новых автодилеров в регионах. Продав 50% акций фонду Valga River One Capital Partners

1,1%
Доля рынка в 2007 году

31 379
Объем продаж автомобилей в 2007 году

Марки: Nissan, Infiniti, Ford, Chevrolet, Opel, Cadillac, Hummer, Saab

«Атлант-М»

Самый сильный ин-дистрибутор дилер из «дочки» — работает в Белоруссии и на Украине. Очень быстрый оборот. Делает ставку на регионы и страны СНГ

1%
Доля рынка в 2007 году

29 377
Объем продаж автомобилей в 2007 году

Марки: Saab, Opel, Cadillac, Hummer, Chevrolet, Volkswagen, Audi, Skoda, Ford, Mazda, Toyota, Citroen, Kia, Great Wall, BYD, Jinbei, Brilliance, JMC, Landwind, ShuangHuan

Musa Motors

Специализируется на автомобилях высокого ценового сегмента. Занимается дилером Fiat-Rover. Планирует развиваться в регионах через субдилерство

0,9%
Доля рынка в 2007 году

24 000
Объем продаж автомобилей в 2007 году

Марки: Jaguar, Land Rover, Volvo, Chrysler, Jeep, Dodge, Renault, BMW, Rolls-Royce

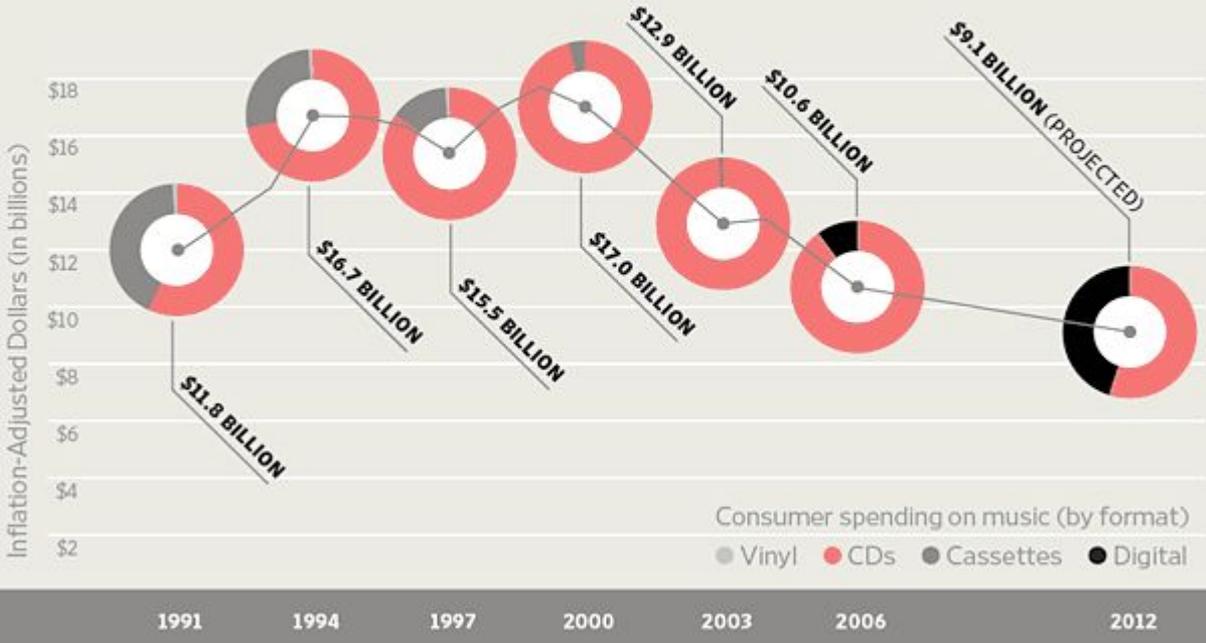
\$936
млн

Кодовые компании по вертикали — количество дилерских центров — «Рейт» продаж в 2007 году

Оборот в 2007 году

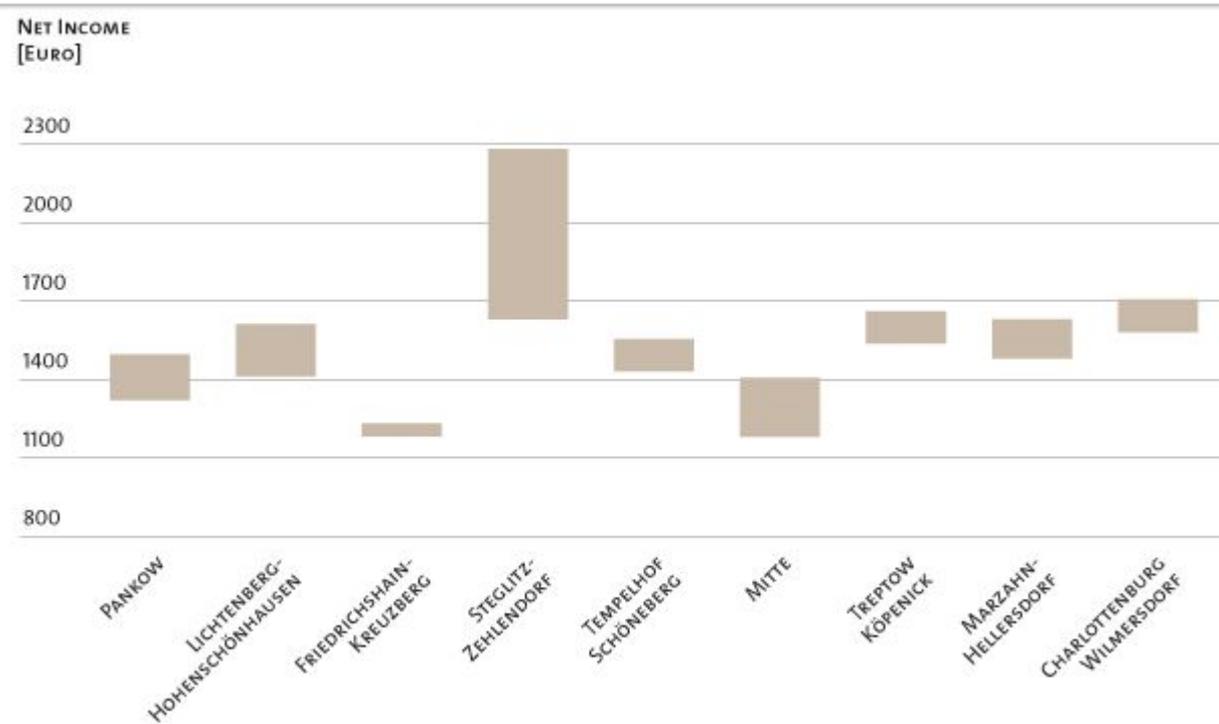
Рост продаж, % (2007 год в сравнении с 2006-м)

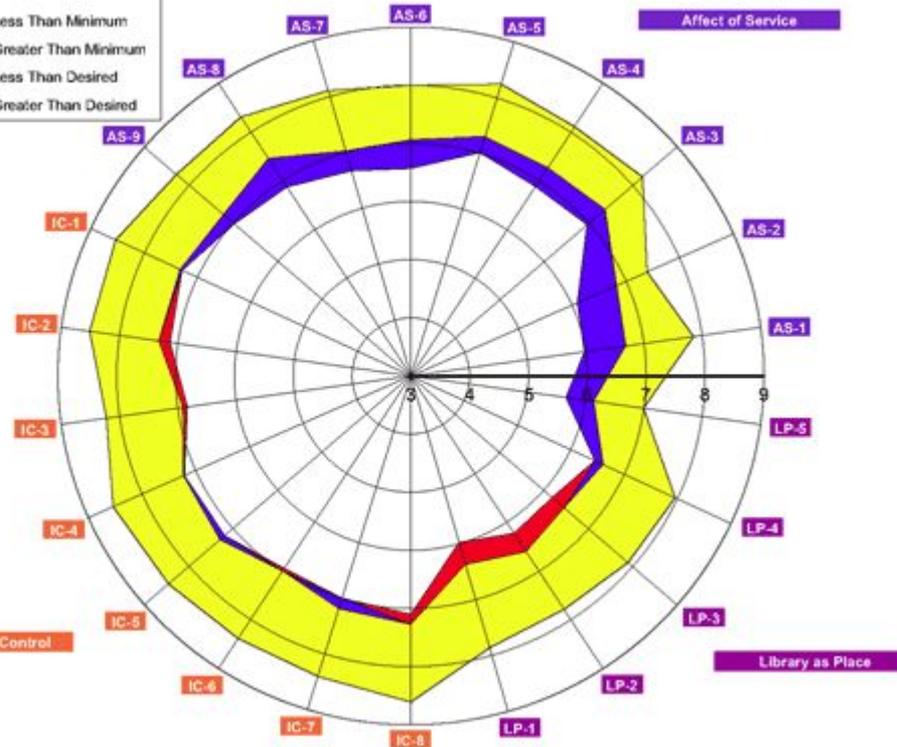
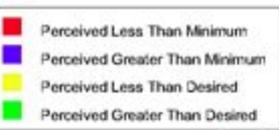
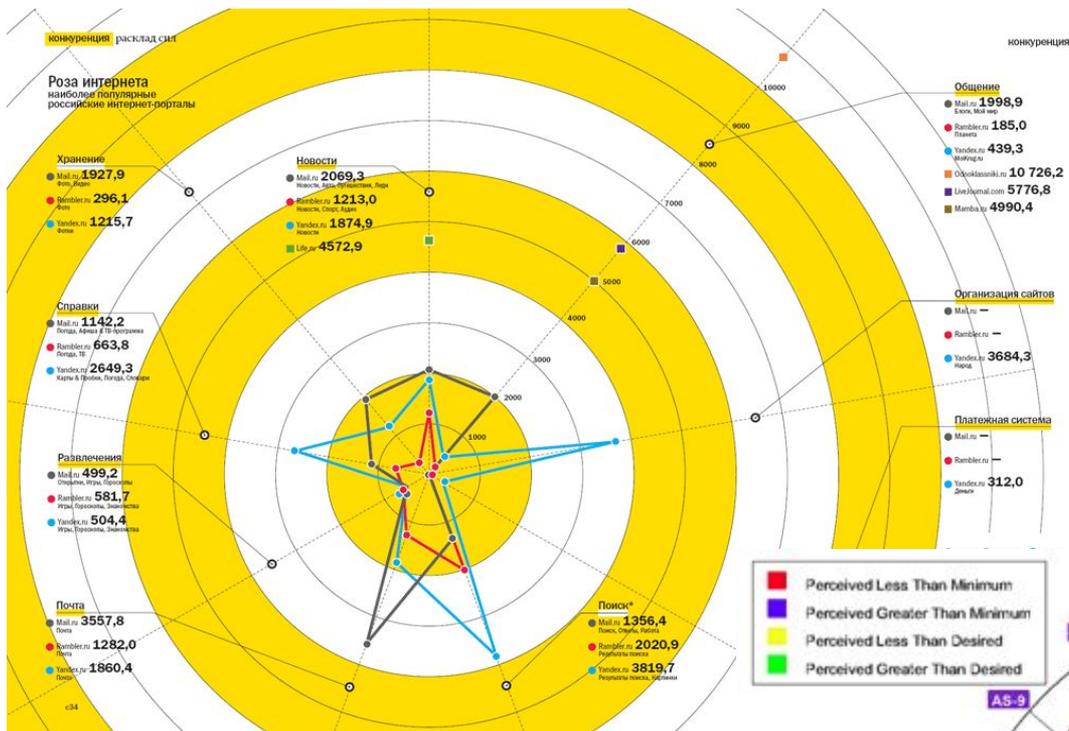
LESS MUSIC IS PURCHASED—AND MORE PURCHASES ARE DIGITAL

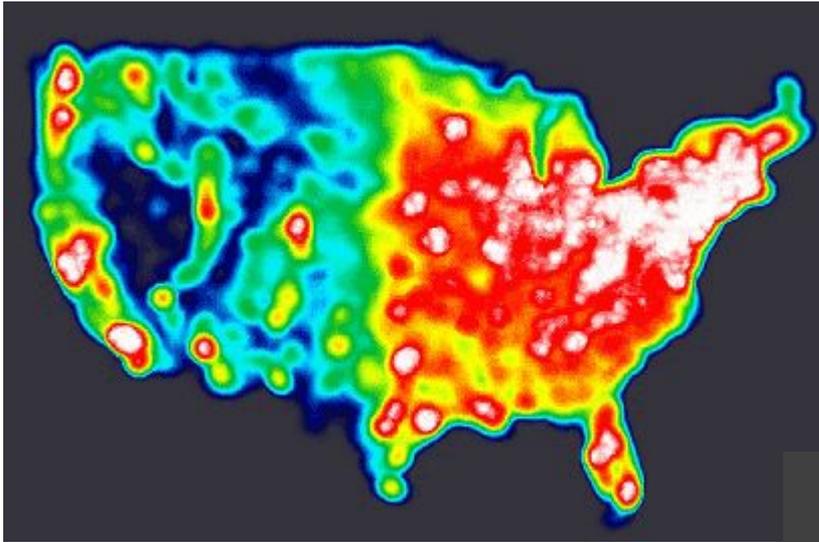


THE INCOME GAP 2005

Range of Average Net Income per Household by Borough







- Connect with an audience of dozens to millions.
- Stop sending mass emails to everyone.
- Archive your thoughts.
- Why the heck not?

Hot Blogs Today

1. [Lorelle on WordPress](#)
2. [churumuri](#)
3. [Scobleizer - Microsoft Geek Blogger](#)
4. [Pomme & Kelly](#)
5. [My Journey to Macintosh](#)
6. [Sam Ideas, Thoughts, Programming](#)
7. [El Blog Oficial de Carlos Sicilia](#)
8. [Qwerty Maniac - The Typo Killer](#)
9. [Football Videos](#)
10. [Nosy Snoop](#)
11. [It's a Definite Maybe](#)

More of the [best blogs on WordPress.com](#).

account?

Username:

Password:

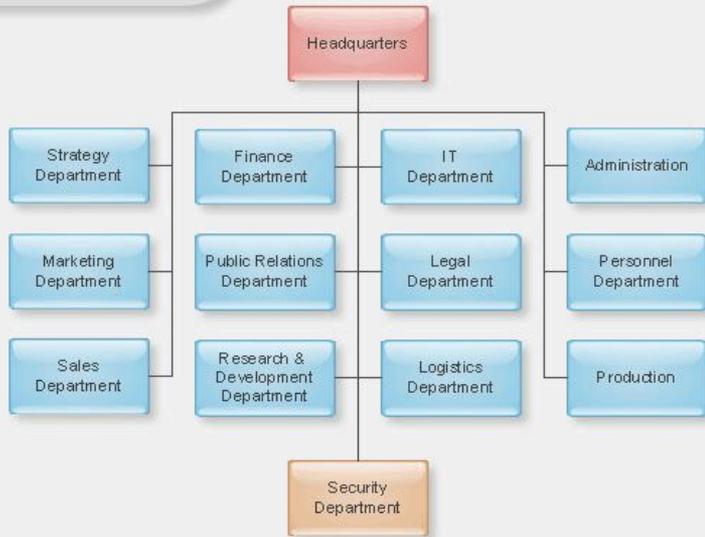
Remember me

Login

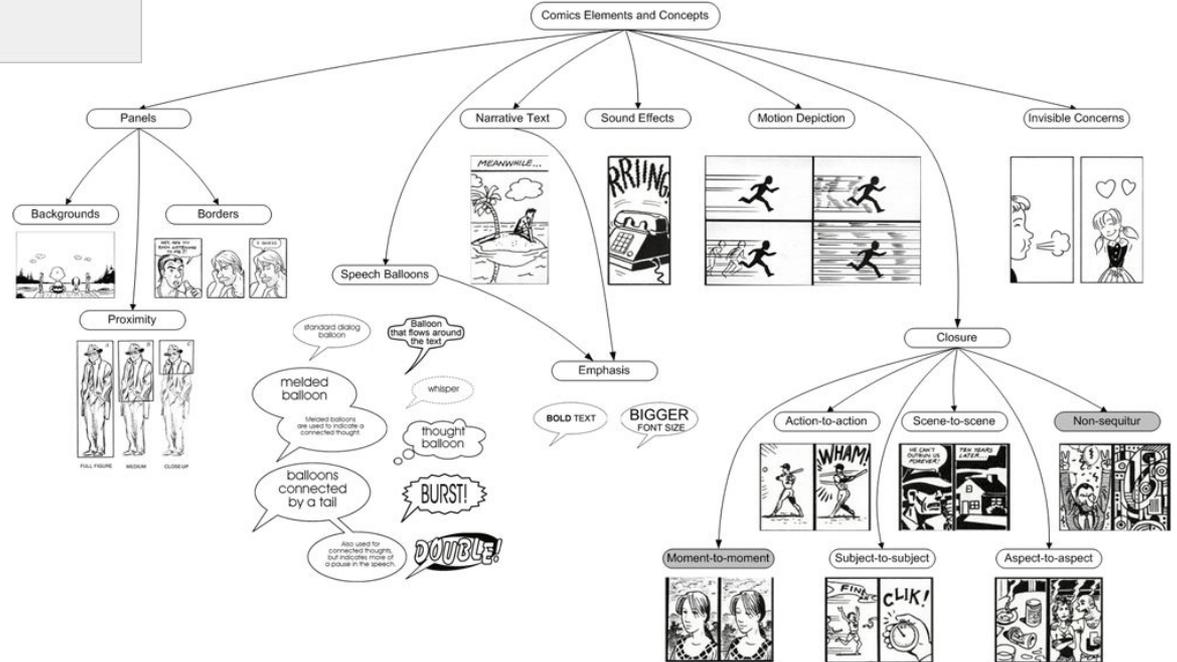
WordPress.com
News Blog

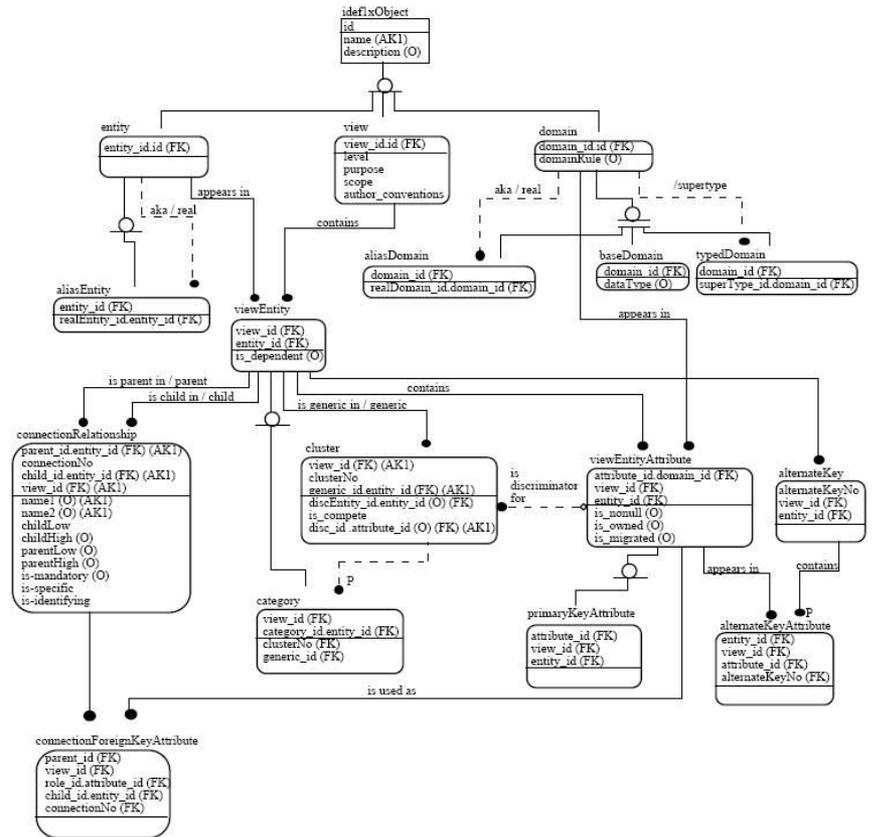
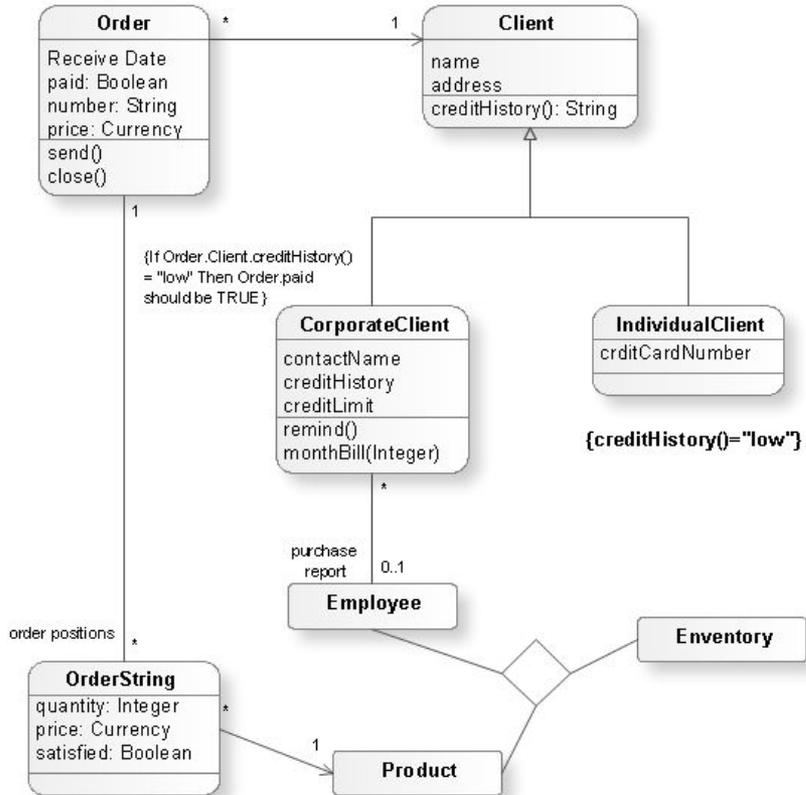
- [Blog Stats Widget](#)
- [Unlimited Blogs](#)
- [Widgets: It's a Plugin!](#)
- [Check Your Spelling](#)
- [Tag Feeds and Paging](#)
- [Andreas04 and Thirteen](#)
- [Login Errors](#)

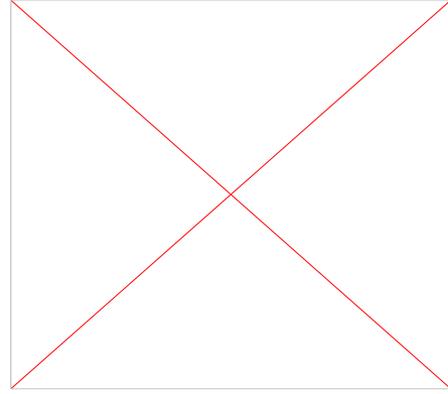
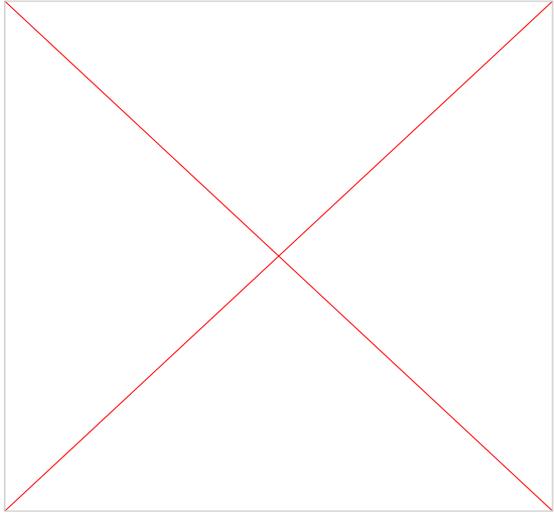
Org Chart

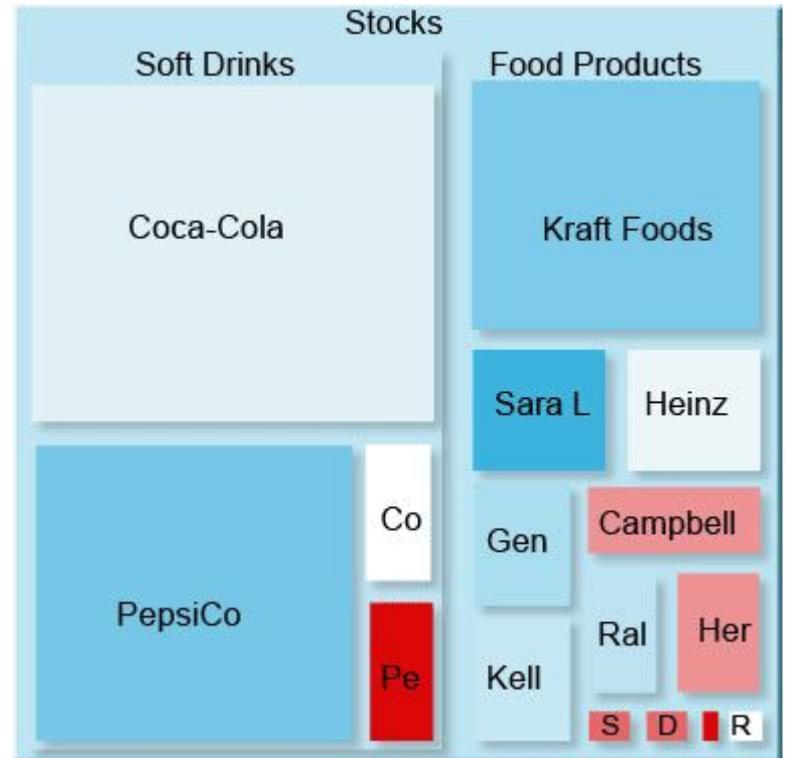
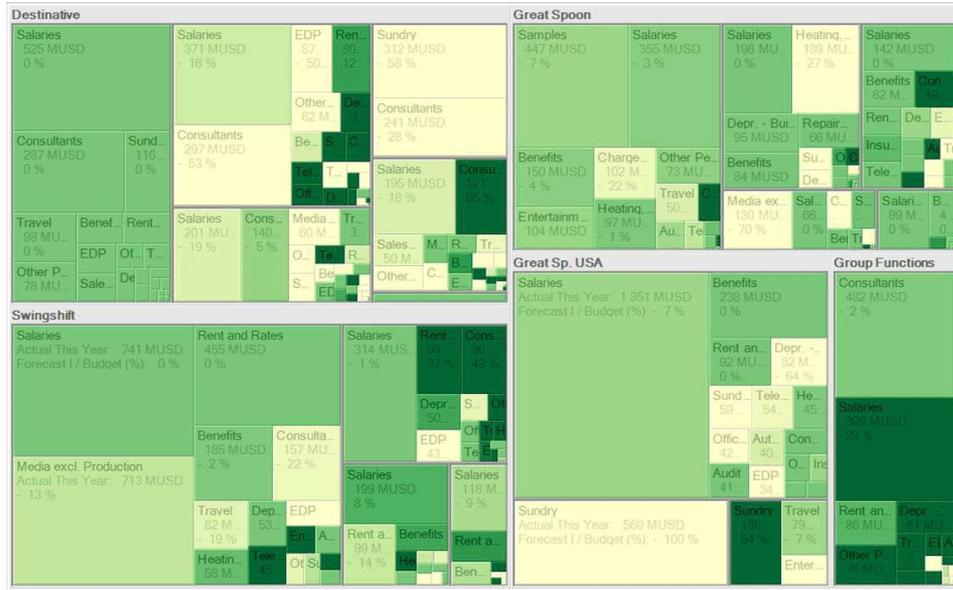


Comics Ontology









Алгоритм линейной регрессии

Шаг 1. Составить матрицы результатов наблюдений и коэффициентов .

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_a \end{bmatrix}, X = \begin{bmatrix} 1 & x_{11} & \dots & x_{1b} \\ 1 & x_{12} & & x_{2b} \\ 1 & x_{13} & & x_{3b} \\ \dots & \dots & & \dots \\ 1 & x_{1b} & & x_{ab} \end{bmatrix}, A = \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \dots \\ \alpha_b \end{bmatrix}, e = \begin{bmatrix} e_1 \\ e_2 \\ \dots \\ e_a \end{bmatrix}.$$

Шаг 2. Определить общий вид уравнения регрессии

$$\hat{y}_i = \alpha_0 + \alpha_1 x_{i1} + \alpha_2 x_{i2} + \dots + \alpha_b x_{ib}$$

Шаг 3. Вычислить вектор-столбец отклонений выборочных значений в матричной форме:

$$e = Y - XA$$

Шаг 4. Оценить e по методу наименьших квадратов:

$$\sum_{i=1}^a e_i^2 = e^T e = (Y - XA)^T (Y - XA) \rightarrow \min,$$

Шаг 5. Рассчитать вектор-столбец коэффициентов:

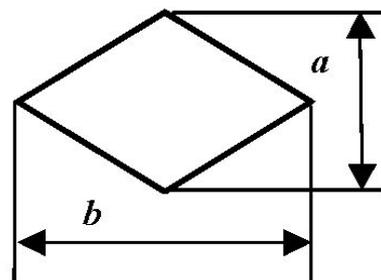
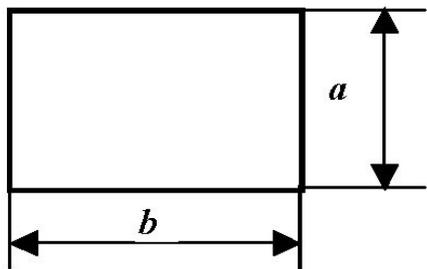
$$A = (X^T Y)^{-1} X^T Y$$

Шаг 6. Получить и интерпретировать уравнение регрессии

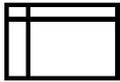
$$S_{op} = 7,28 + 4,6u_i + 4,26t_j + 3,5K_{lm} + 2,8R_q.$$

**ГОСТ 19.701-90. ЕСПД. Схемы алгоритмов,
программ, данных и систем. Условные
обозначения и правила выполнения.**

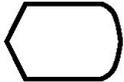
Размеры элементов
схем



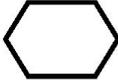
Символы данных

<i>Наименование</i>	<i>Обозначение</i>	<i>Функция</i>
Данные		Отображает данные. Носитель данных не определен.
Запоминаемые данные		Отображает хранимые данные в виде, пригодном для обработки. Носитель данных не определен.
Оперативное запоминающее устройство		Специфический символ данных отображает данные, хранящиеся в оперативном запоминающем устройстве.
Запоминающее устройство с последовательным доступом		Специфический символ данных отображает данные, хранящиеся в запоминающем устройстве с последовательным доступом (магнитная лента, кассета с магнитной лентой, магнитофонная лента).

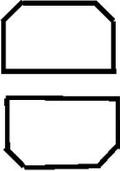
Символы данных

Ручной ввод		<p>Специфический символ данных отображает данные, вводимые вручную во время обработки с устройств любого типа (клавиатура, переключатели, кнопки, световое перо, полосы со штриховым кодом).</p>
Карта		<p>Специфический символ данных отображает данные, представляемые на носителе в виде карты (перфокарты, магнитные карты, карты со считываемыми метками, карты с отрывным ярлыком, карты со сканируемыми метками).</p>
Бумажная лента		<p>Специфический символ данных отображает данные, представляемые на носителе в виде бумажной ленты.</p>
Документ		<p>Специфический символ данных отображает данные, представляемые на носителе в удобочитаемой форме (машинограмма, документ для оптического или магнитного считывания, микрофильм, рулон ленты с итоговыми данными, бланки ввода данных).</p>
Запоминающее устройство с прямым доступом		<p>Специфический символ данных отображает данные, хранящиеся в запоминающем устройстве с прямым доступом (магнитный диск, магнитный барабан, гибкий магнитный диск).</p>
Дисплей		<p>Специфический символ данных отображает данные, представленные в форме для чтения человеком на носителе в виде отображаемого устройства (экран для визуального наблюдения, индикаторы ввода информации).</p>

СИМВОЛЫ

<i>Наименование</i>	<i>Обозначение</i>	<i>Функция</i>
Процесс		Отображает функцию обработки данных любого вида (выполнение определенной операции или группы операций, приводящие к изменению значения, формы или размещения информации или к определению, по которому из нескольких направлений потока следует двигаться).
Решение		Отображает решение или функцию переключательного типа, имеющую один вход и ряд альтернативных выходов, один и только один и только один из которых может быть активен после вычисления условий, определенных внутри этого символа. Соответствующие результаты вычислений могут быть записаны по соседству с линиями, отображающими эти пути.
Подготовка		Символ отображает модификацию команды или группы команд с целью воздействия на некоторую последующую функцию (установка переключателя, модификация индексного регистра или инициализация программы).
Предопределенный процесс		Символ отображает предопределенный процесс, состоящий из одной или нескольких операций или шагов программы, которые определены в другом месте (в программе, модуле).

Символы процесса

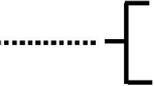
Ручная операция		Символ отображает любой процесс, выполняемый человеком.
Параллельные операции		Символ отображает синхронизацию двух и более параллельных операций.
Граница цикла		Символ, состоящий из двух частей, отображает начало и конец цикла. Обе части символа имеют один и тот же идентификатор. Условия для инициализации, приращения, завершения и т.д. помещаются внутри символа в начале и в конце в зависимости от расположения операции, проверяющей условие.

СИМВОЛЫ ПИНИЙ

<i>Наименование</i>	<i>Обозначение</i>	<i>Функция</i>
Линия		Символ отображает поток данных или управления. При необходимости или для повышения удобочитаемости могут быть добавлены стрелки указатели.
Передача управления		Специфический символ отображает непосредственную передачу управления от одного процесса к другому, иногда с возможностью прямого возвращения к иницирующему процессу после того, как иницированный процесс завершит свои функции. Тип передачи управления должен быть назван внутри символа (например, запрос, вызов, событие).
Канал связи		Специфический символ отображает передачу данных по каналам связи.
Пунктирная линия		Специфический символ отображает альтернативную связь между двумя или более символами. Кроме того, символ используют для обведения аннотируемого участка.

Специальные

СИМВОЛЫ

<i>Наименование</i>	<i>Обозначение</i>	<i>Функция</i>
Соединитель		Символ отображает выход в часть схемы и вход из другой части этой схемы и используется для обрыва линии и продолжения ее в другом месте. Соответствующие символы-соединители должны содержать одно и то же уникальное обозначение.
Терминатор		Символ отображает выход во внешнюю среду и вход из внешней среды (начало или конец схемы программы, внешнее использование и источник или пункт назначения данных).
Канал связи		Символ используют для добавления описательных комментариев или пояснительных записей в целях объяснения или примечаний. Пунктирные линии в символе комментария связаны с соответствующим символом или могут обводить группу символов. Текст комментариев или примечаний должен быть помещен около ограничивающей фигуры.
Пунктирная линия		Символ (три точки) используют в схемах для отображения пропуска символа или группы символов, в которых не определены ни тип, ни число символов.

ГЛАВНОЕ МЕНЮ

