

Модели и задачи Data Mining

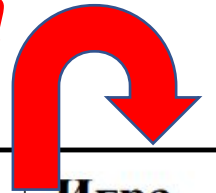
- Data Mining – совокупность большого числа различных методов обнаружения знаний.
- В современной бизнес-аналитике принято выделять в Data Mining описательные (дескриптивные) и предсказательные (предикативные) модели.



Классификация и регрессия

Постановка задачи

В задачах классификации и регрессии требуется определить значение *зависимой переменной объекта* на основании значений других переменных, характеризующих объект.



Наблюдение	Температура	Влажность	Ветер	Игра
Солнце	Жарко	Высокая	Нет	Нет
Солнце	Жарко	Высокая	Есть	Нет
Облачность	Жарко	Высокая	Нет	Да
Дождь	Норма	Высокая	Нет	Да
Дождь	Холодно	Норма	Нет	Да
Дождь	Холодно	Норма	Есть	Нет
Облачность	Холодно	Норма	Есть	Да
Солнце	Норма	Высокая	Нет	Нет

Формально задачу **классификации** и **регрессии** можно описать следующим образом.

Имеется множество объектов:

$$I = \{i_1, i_2, \dots, i_j, \dots, i_n\},$$

где i_j — исследуемый объект.

Каждый объект характеризуется набором переменных:

$$I_j = \{x_1, x_2, \dots, x_h, \dots, x_m, y\},$$

где x_h — независимые переменные, значения которых известны и на основании которых определяется значение зависимой переменной y .

В Data Mining часто **набор независимых переменных** обозначают в виде вектора:

$$X = \{x_1, x_2, \dots, x_j, \dots, x_n\},$$

Каждая переменная x_j может принимать значения из некоторого множества:

$$C_h = \{c_{h1}, c_{h2}, \dots\},$$

Если значениями переменной являются элементы конечного множества, то

говорят, что она имеет **категориальный тип**.

Например, переменная наблюдение принимает значения на множестве значений {солнце, облачность, дождь}.

Если множество значений $C = \{c_1, c_2, \dots, c_r, \dots, c_k\}$ переменной y конечно, то задача называется задачей **классификации**.

Если переменная v принимает значение на **множестве действительных**
классификация: $X \rightarrow \{0, 1, \dots, K\}$, где $0, \dots, K$ – номера классов,

регрессия: $X \rightarrow \mathbb{R}$,

Представление результатов

Правила классификации

В задачах **классификации** и **регрессии** обнаруженная функциональная зависимость между переменными может быть представлена одним из следующих способов:

- классификационные правила;
- деревья решений;
- математические функции.

1. Классификационные правила состоят из двух частей: **условия** и **заключения**:

если (условие) то (заключение).

Условием является проверка одной или нескольких независимых переменных с использованием операций И, ИЛИ, НЕ.

Заключением является значение зависимой переменной или распределение ее вероятности по классам, например:

если (наблюдение = солнце И температура = жарко) то (игра = нет);
если (наблюдение = облачность И температура = холодно) то (игра = да).

Основными достоинствами правил являются легкость их восприятия и запись на естественном языке.

Еще одно преимущество — их относительная независимость.

В набор правил легко добавить новое правило без необходимости изменять уже существующие

Относительная независимость связана с возможной их противоречивостью друг другу.

Если переменные, характеризующие некоторый объект, удовлетворяют условным частям правил с разными заключениями, то возникает неопределенность со значением его зависимой переменной. Например, пусть имеются правила:

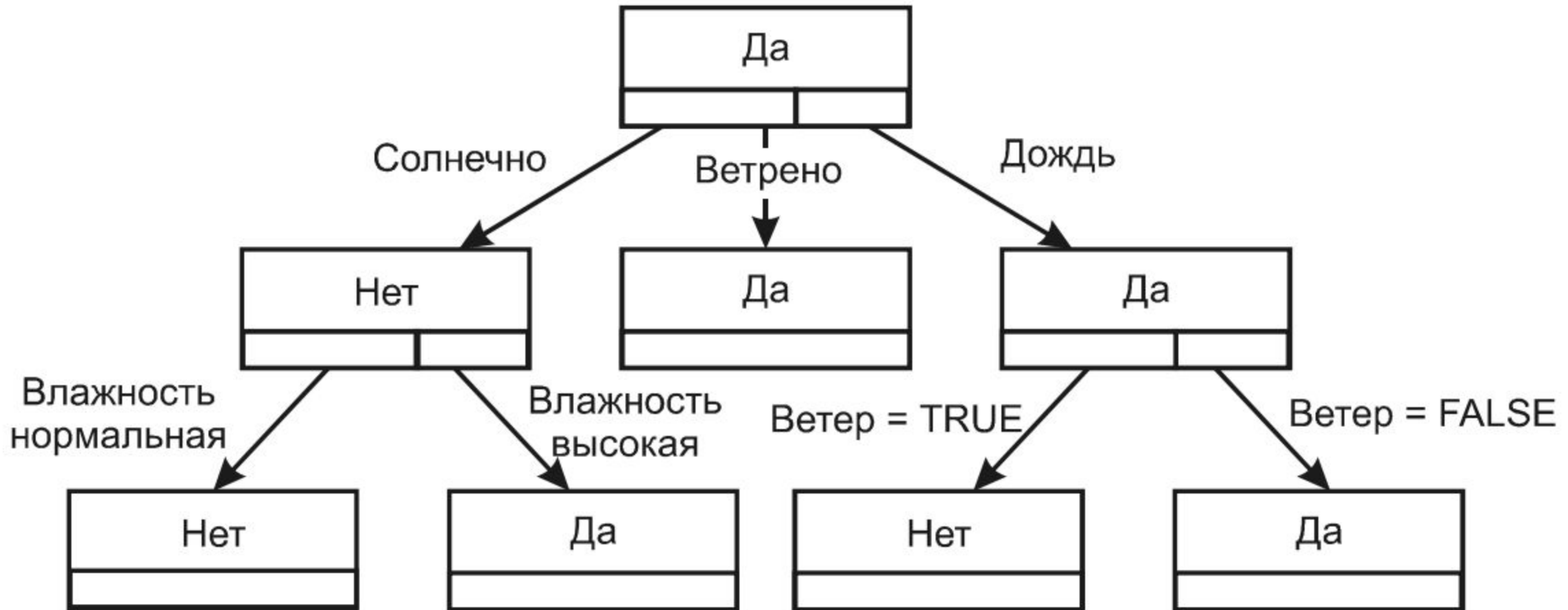
если (наблюдение = солнце) то (игра = нет);

если (наблюдение = облачность И температура = холодно) то (игра = да)

В них объекты, удовлетворяющие условиям второго правила, удовлетворяют и условиям первого правила. Однако вывод делается разный.

Другими словами, в соответствии с этими правилами при одинаковых обстоятельствах будут получены противоречивые указания, что неприемлемо.

2. Деревья решений — это способ представления правил в иерархической, последовательной структуре.



3. Математическая функция выражает отношение зависимой переменной от независимых переменных.

В этом случае анализируемые объекты рассматриваются как точки в $(m + 1)$ -мерном пространстве.

Тогда переменные объекта $i_j = \{x_1, x_2, \dots, x_m, y\}$, рассматриваются как координаты, а функция имеет следующий вид:

$$y_j = \{\omega_0 + \omega x_1 + \omega x_2 + \dots + \omega x_m\}$$

где $\omega_0, \omega_1, \omega_m$ — веса независимых переменных, в поиске которых и состоит задача нахождения классификационной функции.

Очевидно, что все переменные должны быть представлены в виде числовых параметров.

Для преобразования *логических* и *категориальных* переменных

к числовым используют разные способы:

Логические типы, как правило, кодируют цифрами 1 (истина) и 0 (ложь).

Значениями *категориальных* переменных являются имена возможных состояний изучаемого объекта. Имена должны быть перечислены и пронумерованы в списке.

В итоге *категориальная* переменная преобразуется в *числовую* переменную. Например, значение переменной наблюдение = {солнце, облачность, дождь} можно заменить

наблюдение = {0, 1, 2}

Другой способ представления исходно *категориальной* переменной в системе — это замена возможных значений *набором двоичных признаков*.

В наборе столько двоичных признаков, сколько имен содержится в списке возможных состояний объекта.

При анализе объекта значение 1 присваивается тому двоичному признаку, который соответствует состоянию объекта. Остальным присваивается значение 0.

Например, для переменной наблюдения такими значениями будут {001, 010, 100}.

Методы построения правил классификации.

Например, метод *Naive Bayes*

Условная вероятность принадлежности объекта к c_r при равенстве его независимых переменных определенным

$$^3 P(y = c_r | E) = P(E | y = c_r) \cdot P(y = c_r) / P(E)$$

$$P(E | y = c_r) = P(x_1 = c_p^1 | y = c_r) \times P(x_2 = c_d^2 | y = c_r) \times \dots \times$$

$$\times P(x_m = c_b^m | y = c_r).$$

$$P(\text{наблюдение} = \text{солнце} | \text{игра} = \text{да}) = 2/9;$$

$$P(\text{наблюдение} = \text{облачно} | \text{игра} = \text{да}) = 4/9;$$

$$P(\text{наблюдение} = \text{дождь} | \text{игра} = \text{да}) = 3/9;$$

$$P(\text{наблюдение} = \text{солнце} | \text{игра} = \text{нет}) = 3/5;$$

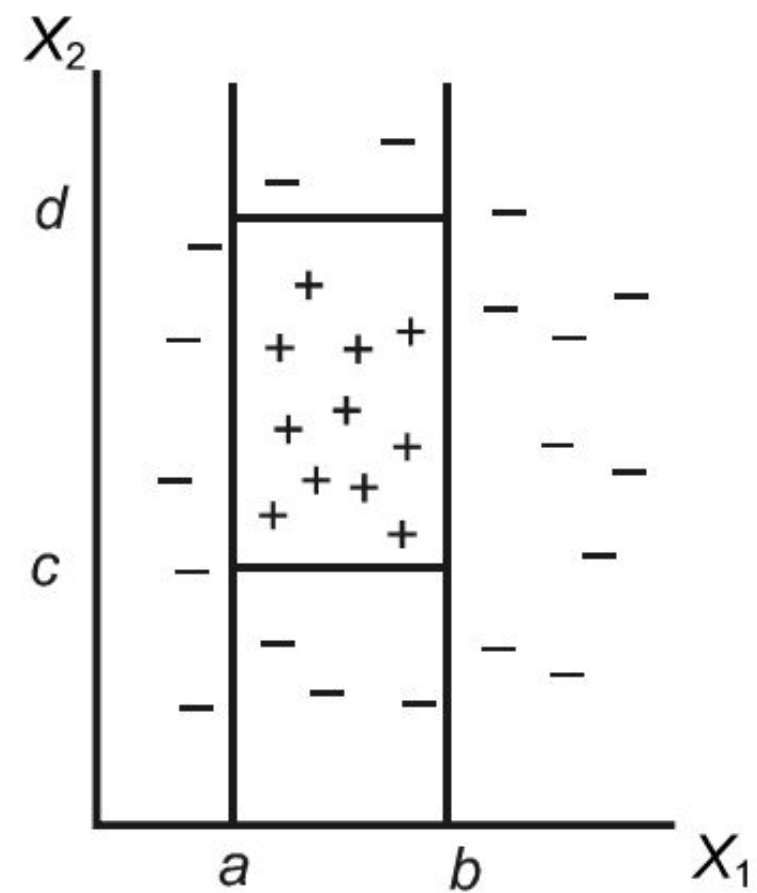
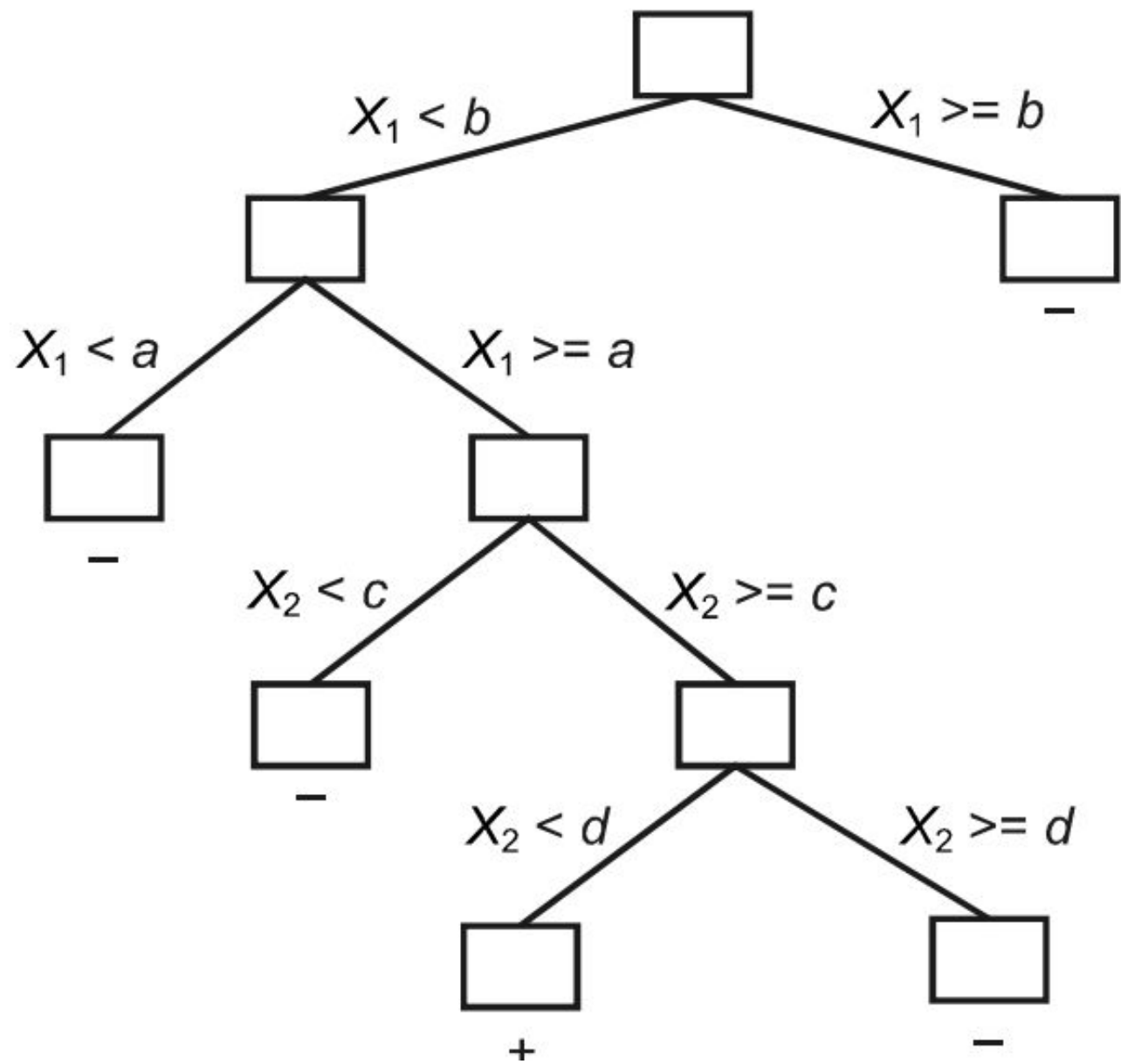
$$P(\text{наблюдение} = \text{облачно} | \text{игра} = \text{нет}) = 0/5;$$

$$P(\text{наблюдение} = \text{дождь} | \text{игра} = \text{нет}) = 2/5.$$

Методы построения деревьев решений. Например, алгоритм *покрытия*

Построение деревьев решений для каждого класса по отдельности.

1. На каждом шаге алгоритма *выбирается значение переменной*, которое разделяет все множество на два подмножества.
2. Разделение должно выполняться так, чтобы все объекты класса, для которого строится дерево, принадлежали *одному подмножеству*.
3. Такое разбиение производится до тех пор, пока не будет построено подмножество, содержащее только объекты *одного класса*.



Возраст	Предписание	Астигматизм	Степень износа	Рекомендации
Юный	Близорукость	Нет	Пониженный	Нет
Юный	Близорукость	Нет	Нормальный	Мягкие
Юный	Близорукость	Да	Пониженный	Нет
Юный	Близорукость	Да	Нормальный	Жесткие
Юный	Дальнозоркость	Нет	Пониженный	Нет
Юный	Дальнозоркость	Нет	Нормальный	Мягкие
Юный	Дальнозоркость	Да	Пониженный	Нет
Юный	Дальнозоркость	Да	Нормальный	Жесткие
Пожилой	Близорукость	Нет	Пониженный	Нет
Пожилой	Близорукость	Нет	Нормальный	Мягкие
Пожилой	Близорукость	Да	Пониженный	Нет
Пожилой	Близорукость	Да	Нормальный	Жесткие
Пожилой	Дальнозоркость	Нет	Пониженный	Нет
Пожилой	Дальнозоркость	Нет	Нормальный	Мягкие
Пожилой	Дальнозоркость	Да	Пониженный	Нет
Пожилой	Дальнозоркость	Да	Нормальный	Нет
Старческий	Близорукость	Нет	Пониженный	Нет
Старческий	Близорукость	Нет	Нормальный	Нет
Старческий	Близорукость	Да	Пониженный	Нет
Старческий	Близорукость	Да	Нормальный	Жесткие
Старческий	Дальнозоркость	Нет	Пониженный	Нет
Старческий	Дальнозоркость	Нет	Нормальный	Мягкие
Старческий	Дальнозоркость	Да	Пониженный	Нет
Старческий	Дальнозоркость	Да	Нормальный	Нет

если (?) то рекомендация = жесткие

возраст = юный - 2/8;
 возраст = пожилой - 1/8;
 возраст = старческий - 1/8;
 предписание = близорукость - 3/12;
 предписание = дальнозоркость - 1/12;
 астигматизм = нет - 0/12;
 астигматизм = да - 4/12;
 степень износа = низкая - 0/12;
 степень износа = нормальная - 4/12.

если (астигматизм = да и ?) то рекомендация = жесткие.

возраст = юный - 2/4;
 возраст = пожилой - 1/4;
 возраст = старческий - 1/4;
 предписание = близорукость - 3/6;
 предписание = дальнозоркость - 1/6;
 степень износа = низкая - 0/6;
 степень износа = нормальная - 4/6.

если (астигматизм = да и степень износа = нормальная) то рекомендация = жесткие.

возраст = юный - 2/2;
 возраст = пожилой - 1/2;
 возраст = старческий - 1/2;
 предписание = близорукость - 3/3;
 предписание = дальнозоркость - 1/3.

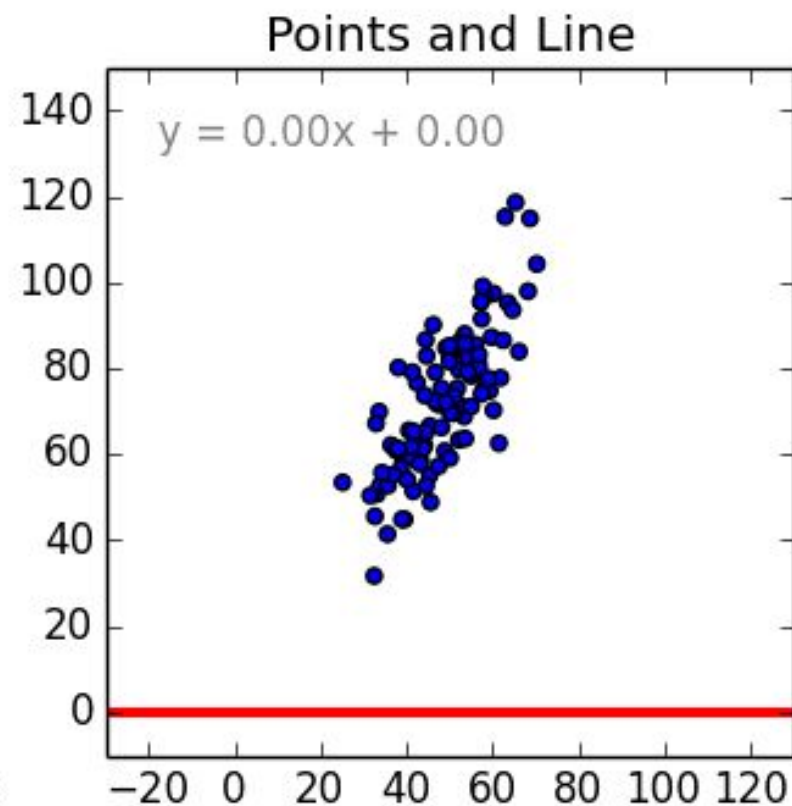
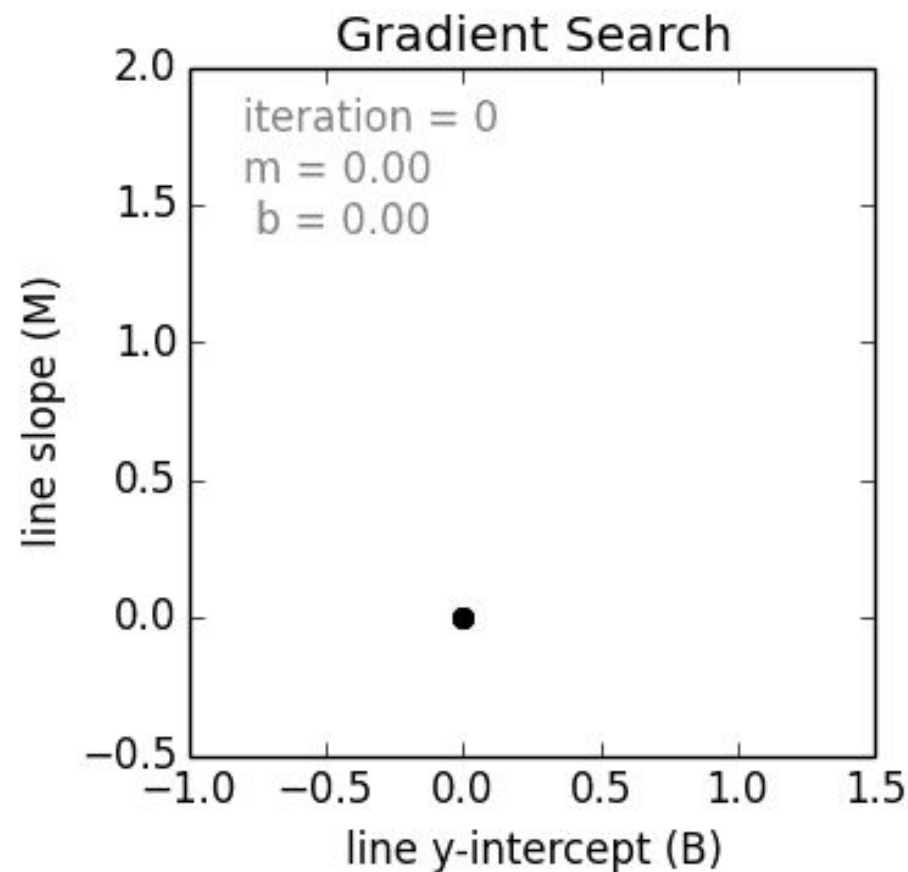
если (астигматизм = да и степень износа = нормальная и предписание = близорукость) то рекомендация = жесткие.

Методы построения математических функций.

Семейство линейных функций

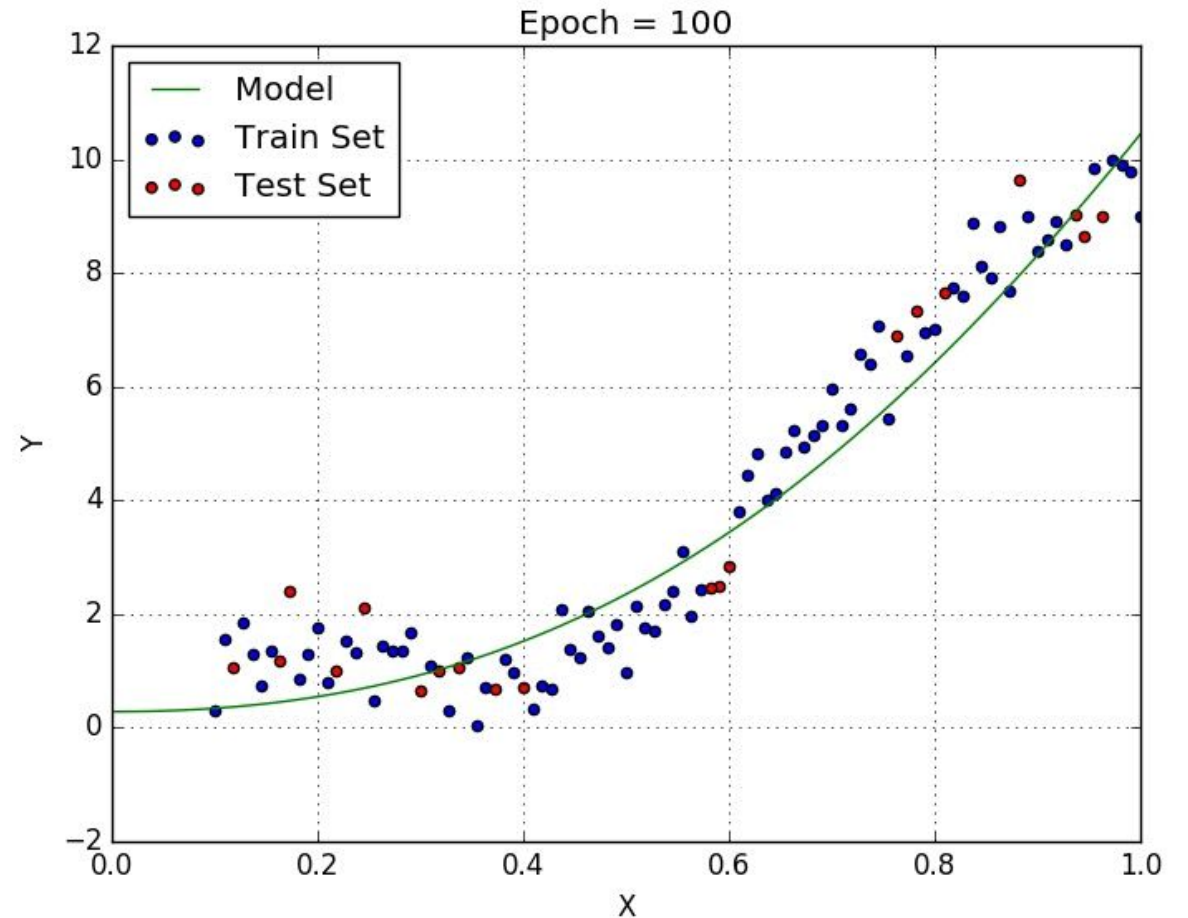
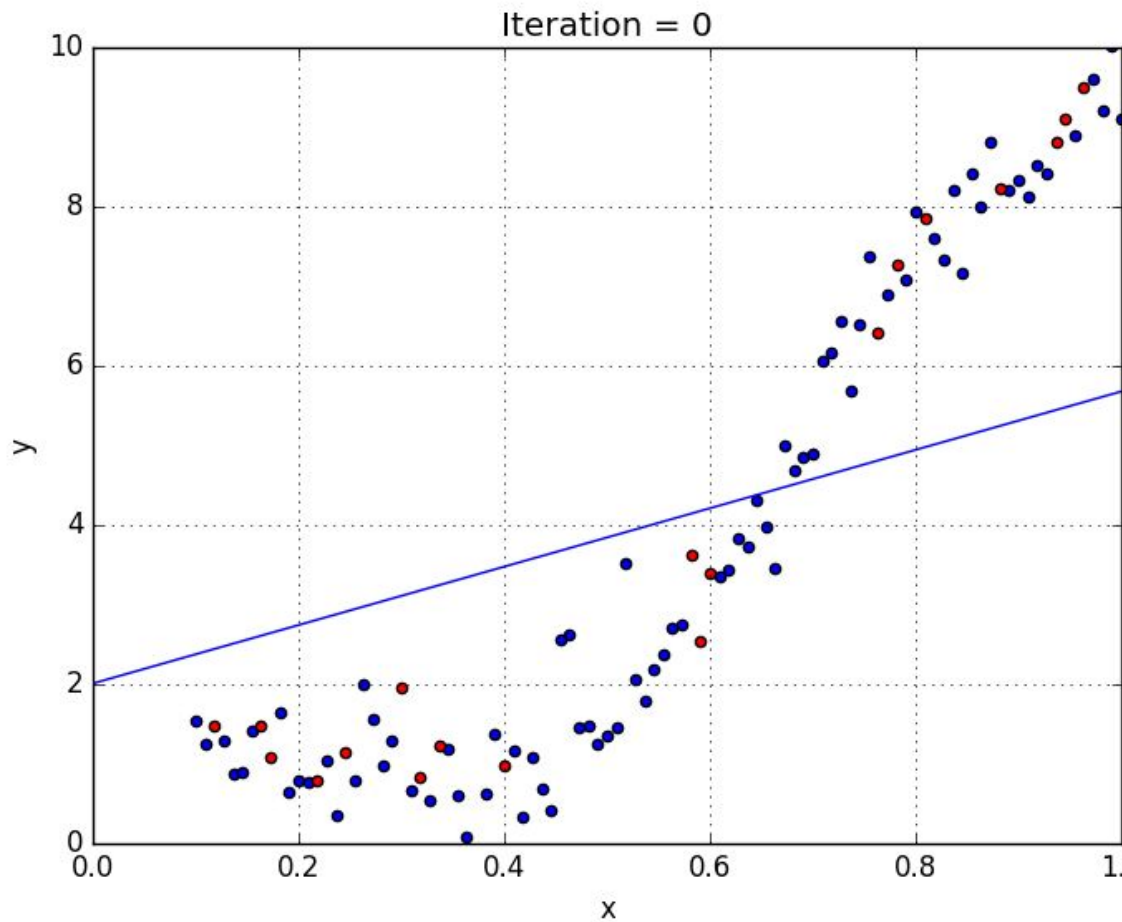
Множественная линейная регрессия: $Y = a_1 * X_1 + a_2 * X_2 + a_3 * X_3 \dots a_n * X_n + b$, где a_n — это коэффициенты, X_n — переменные и b — смещение.

Весовые коэффициенты a_n , а также смещение b вычисляются с применением стохастического градиентного спуска.



Полиномиальная регрессия

В полиномиальной регрессии степень некоторых независимых переменных превышает 1: $Y = a_1 * X_1 + (a_2)^2 * X_2 + (a_3)^4 * X_3 \dots \dots a_n * X_n + b$



Линейная регрессия:

- Легко моделируется, полезна при создании не сложной зависимости, при небольшом количестве данных.
- Обозначения интуитивно-понятны.
- Чувствительна к выбросам.

Полиномиальная регрессия:

- Моделирует нелинейно разделенные данные и сложные взаимосвязи.
- Полный контроль над моделированием переменных объекта (выбор степени).
- Необходимо обладать некоторыми знаниями о данных, для выбора наиболее подходящей степени.
- При неправильном выборе степени модель может быть перенасыщена.

- ***Гребневая (ридж) регрессия***

В случае высокой *коллинеарности* переменных стандартная линейная и полиномиальная регрессии становятся неэффективными.

- ***Регрессия по методу «лассо»***

В регрессии лассо добавляется условие смещения в функцию оптимизации для того, чтобы уменьшить коллинеарность и, следовательно, дисперсию модели.

- ***Регрессия «эластичная сеть»***

Эластичная сеть — это гибрид методов регрессии лассо и гребневой регрессии.

Ансамбли моделей

Разработано множество различных методов и алгоритмов формирования ансамблей.

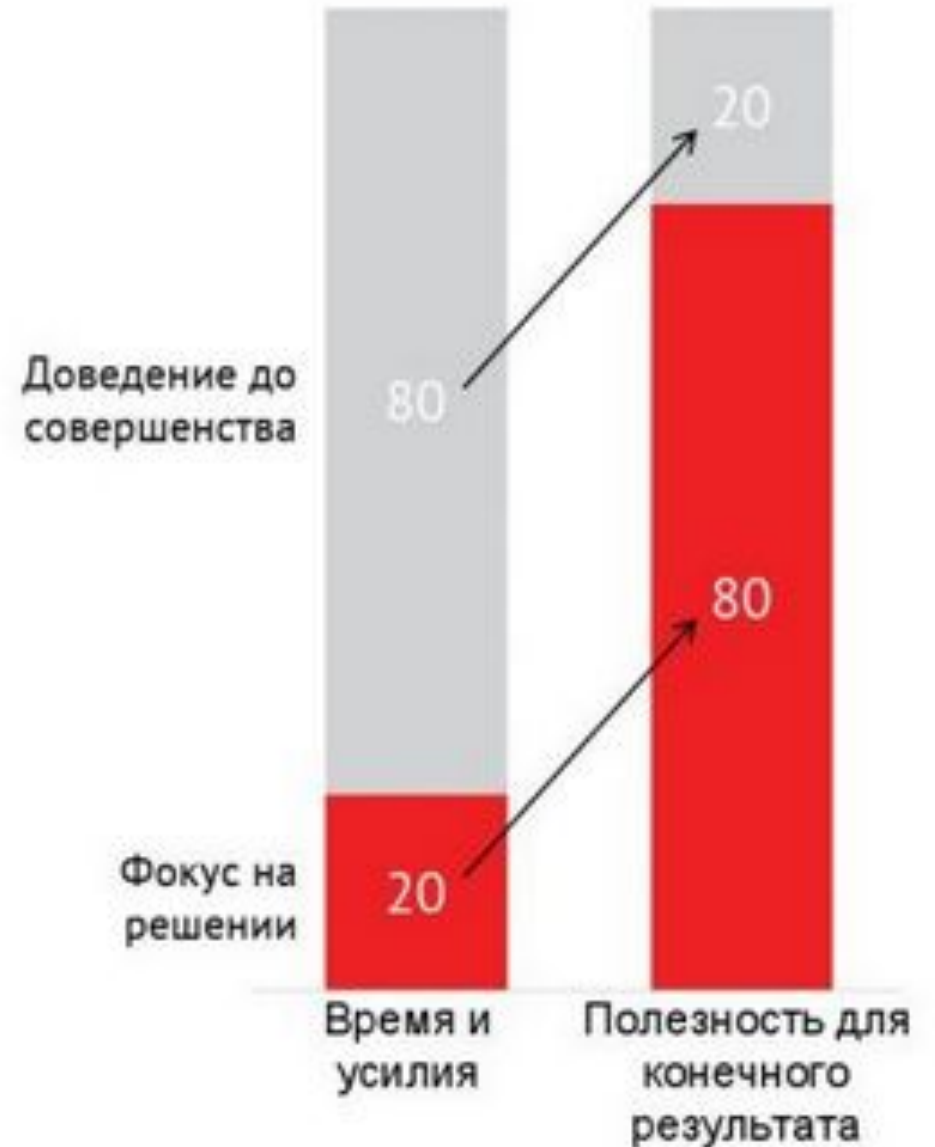
Цель объединения— улучшить (усилить) решение, которое дает отдельная модель.

Использование ансамблей позволяет повысить качество решений, однако такой подход связан с рядом проблем:

- увеличение временных и вычислительных затрат на обучение нескольких моделей;
- сложность интерпретации результатов;
- неоднозначный выбор методов комбинирования результатов, выдаваемых отдельными моделями.

Правило 80/20 (Закон Парето) – эмпирическое правило, названное в честь экономиста и социолога Вильфредо Парето, в наиболее общем виде формулируется так: «**20% усилий дают 80% результата, а остальные 80% усилий - лишь 20% результата**».

Цифры 20 и 80 являются данью заслугам Парето, считать их безусловно точными нельзя.



ABC и XYZ совместный анализ

Популярные методы **классификации** ресурсов:

- по поставщику, по клиенту, расположению зон хранения товара,
- по прибыльности или оборачиваемости...

Товар	Дата	Сумма с учетом скид...
Доска паркетная UPOFLOOR Лос 2085x188x14 дуб красный преми...	01.03.2017, 00:00	13 708,80
Гвоздь оцинкованный 1.4x25, 300 штук, 10 категория, уп	01.03.2017, 00:00	3 311,84
Гипсокартон АБС 9.5x1200x2500, кв.м.	01.03.2017, 00:00	2 271,28
Гипсокартон БЕЛГИПС 12.5x1200x2500, кв.м	01.03.2017, 00:00	2 259,40
Гипсокартон БЕЛГИПС 9.5x1200x2500, влагостойкий, мал.уп., кв.м	01.03.2017, 00:00	2 820,40
Гипсокартон БЕЛГИПС 9.5x1200x2500, кв.м	01.03.2017, 00:00	2 135,76

ab Объект	9.0 Показатель	9.0 Вклад нарастающим итог...	ab Группа
Колеровочная краска SADOLIN Tintorama LSG-FRI для вд красок, ...	5 426 106,62	49,09	A
Грунтовка вд ОПТИМИСТ глубокого проникновения, зеленая эти...	5 398 282,15	49,46	A
Профиль пластиковый HENKE рамочный 7090 вишня 3 м, шт	5 339 053,31	49,82	A
Грунтовка вд полимерная ESCARO AQUASTOP, 10 л, шт	5 193 212,06	50,18	B
Решетка радиаторная КОСТРОМА 600x900мм, бежевая, шт	5 081 600,96	50,53	B

A

B

Существует **классический способ** ранжирования по ABC анализу в части прибыльности где:

A — 80% прибыли

B — 15% прибыли

C — 5% прибыли

Более **детальный способ** деления - добавляется литера **D**.

Ранжирование, в части прибыли, выглядит так:

A — 50% прибыли

B — 30% прибыли

C — 15% прибыли

D — 5% прибыли

Выбор классического или современного способа определяется важностью степени детализации, величиной ассортимента и так далее.

ABC анализ в управлении запасами (1)

Эффективное управление запасами позволяет предприятию удовлетворять ожидания потребителей, создавая товарные запасы, максимизирующие прибыль предприятия.

Ранжирование товара по ABC анализу в части управления запасами, позволяет:

- Расходовать бюджет на закупки более эффективно.
- Эффективнее планировать и распределять страховой запас.
- Лучше удовлетворять спрос покупателей.

Анализ оборачиваемости товара (2)

Оборачиваемость запасов - это показатель обновляемости товара в течение расчетного периода, например, в течение года.

- Категория А - наиболее оборачиваемые товары.
- Категория В - товары со средней оборачиваемостью.
- Категория С - низко оборачиваемые товары.

Оборачиваемость запасов считают двумя способами:

Коэфф. оборачиваемости запасов = Себестоимость продаж / Среднегодовой остаток запасов (1)

Коэфф. оборачиваемость запасов = Выручка / Среднегодовой остаток запасов (2)

Зная коэффициент оборачиваемости в днях, и ранжируя товары по категориям А, В, С по прибыли, менеджер по закупкам планирует закупки.

Расширенный ABC анализ (3)

ABC анализ можно проводить **по частотности заказов**.

Частотность заказов - сколько месяцев в году продается определенный товар.

Например: 1) согласно ABC анализа, товары принадлежащие к категории “В” и “С”, в сумме дают всего 20% от общей прибыли. Но как часто покупают эти товары? 2) Товар дорогой, с высокой рентабельностью и он попадет по прибыли в категорию А, однако, за ним приходят всего пару раз в год. 3) Товар, который не так прибылен, но компания продает его стабильно каждый месяц.

Т.е., имеет смысл дополнительно ранжировать ассортимент по количеству месяцев в году когда товар продавался:

- А - это продажи 10-12 месяцев в году.
- В – 5-9 месяцев в году.
- С- 4 и менее месяцев в году.

Расширенный ABC анализ (4)

Количество обращений означает, сколько отдельных заказов было сделано по каждому товару не зависимо от их количества, стоимости и прибыльности.

Сколько раз покупатель пришел в компанию за месяц. Например, по определенному товару было:

- А – количество обращений от 100 и выше,
- В – 50 – 99 обращений,
- С – менее 50 обращений в месяц.

В итоге, получается **расширенный ABC анализ**:

Расширенный ABC анализ

«AAA» — супер
ТОП

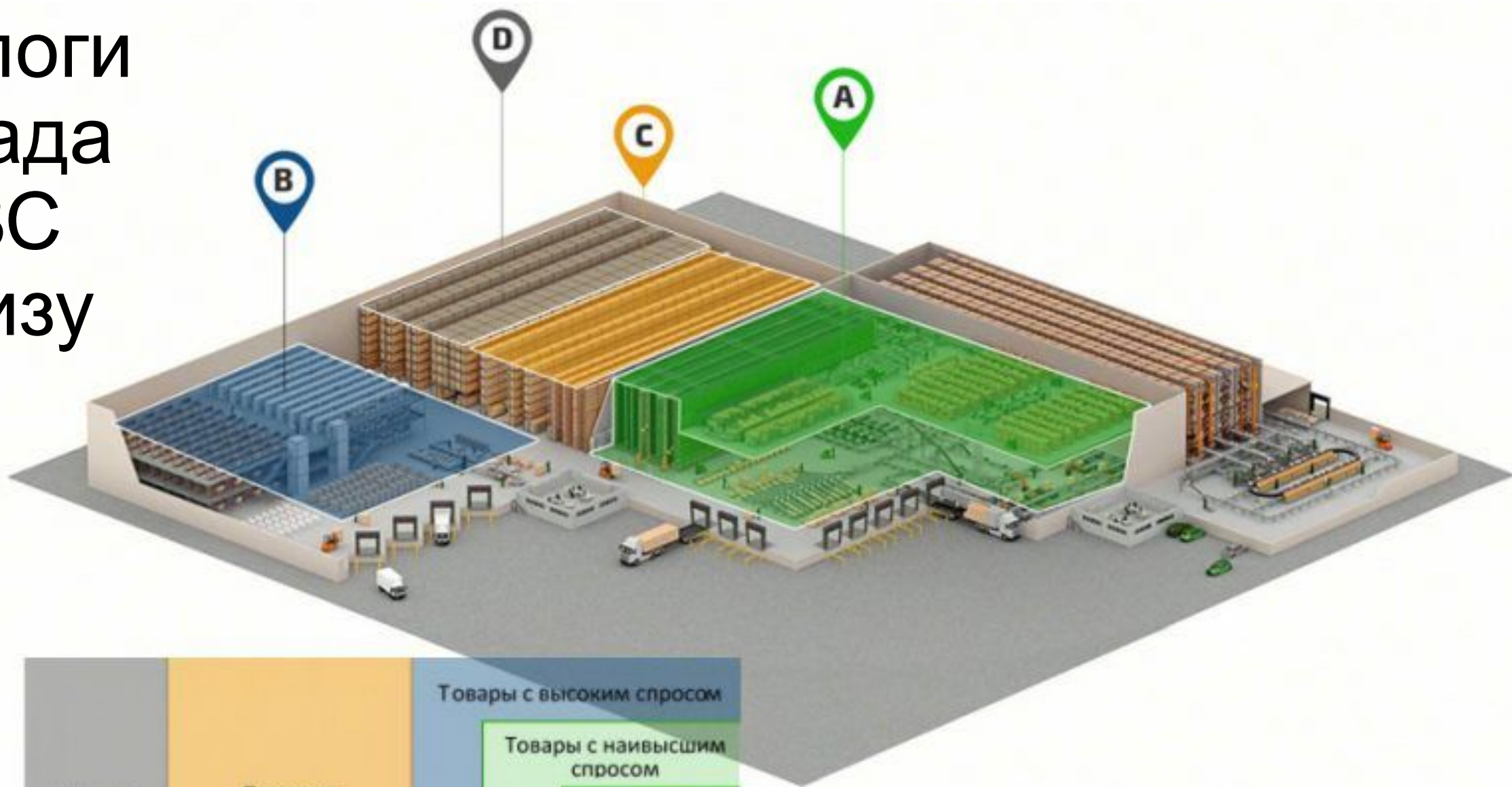
	A	B	C	D	E
1	код товара	наименование товара	ABC прибыль	ABC частотность	ABC обращения
2	1009292202	Магнитола Беларусь	A	A	A
3	1009292203	Ноутбук Sang Tank SUS	A	B	B
4	1009287112	Утюг Киевская Русь	A	B	C
5	1009300201	Батарейки Ямайка	A	A	A
6	1009200456	Микроволновка Волна	B	C	A
7	1009200457	Утюг Шок	B	A	A
8	1009200458	Фен Холодная голова	B	A	A
9	1009200459	Телевизор Samsung 13	B	B	B
10	1009200460	Обогреватель Жара у Жоры	B	C	C
11	1009200461	Полотенцесушитель Бегемот	C	C	C
12	1009200462	Холодильник Вечная память	C	C	C
13	1009200463	Морозильная камера Свежая рыба	C	C	B
14	1009200464	Набор салфеток Чистюля	C	A	A

«CCC» — нужны ли затраты на такой товар?

Выводы по расширенному ABC анализу:

- Товар по прибыльности относится к категории А, но в расширенном ABC анализе имеет категорию ACC - может не иметь большого веса.
- На товар категории ВАА могли не обращать большого внимания, отодвигая на второй план, когда этот товар оказывался в дефиците. Однако, в таком расширенном спектре, этот товар оказывается очень важным.
- Пересмотреть все товары категории САА и так далее.

Топологи я склада по ABC анализу



Простые методы определения границ групп ABC-анализа: Эмпирический метод, метод сумм

Границы определяются по значению суммы двух показателей:

доля по объему **нарастающим** итогом и доля по количеству **нарастающим** итогом.

- Группа А. Нижняя граница (Доля по объему + Доля по количеству) = 100%
- Группа В. Нижняя граница (Доля по объему + Доля по количеству) = 145%
- Группа С. Все оставшееся.

Рекомендуемые границы 80%-15%-5% по объему и 20%-30%-50% по количеству не являются законом природы.

Самая распространенная эмпирическая рекомендация для групп:

- Группа А. 80% по объему, 20% по количеству.
- Группа В. 15% по объему, 30% по количеству.
- Группа С. 5% по объему, 50% по количеству.

Получаем то же самое:

- Группа А. $80\% + 20\% = 100\%$
- Группа В. $80\% + 15\% + 20\% + 30\% = 145\%$
- Группа С. Все оставшееся

N	Товары	Выручка	Выручка Доля	Выручка Доля нарастающим	Группа ABC	Доля по кол-ву	Доля по кол-ву нарастающим	Выручка + кол-во нарастающим
		тыс. руб.	%	%		%	%	%
1	Товар 1	4 400.00	17.2%	17.2%	A	3.33%	3.33%	20.55%
2	Товар 2	4 000.00	15.7%	32.9%	A	3.33%	6.67%	39.54%
3	Товар 3	3 400.00	13.3%	46.2%	A	3.33%	10.00%	56.17%
4	Товар 4	3 000.00	11.7%	57.9%	A	3.33%	13.33%	71.25%
5	Товар 5	2 200.00	8.6%	66.5%	A	3.33%	16.67%	83.19%
6	Товар 6	2 000.00	7.8%	74.3%	A	3.33%	20.00%	94.35%
7	Товар 7	1 100.00	4.3%	78.7%	A	3.33%	23.33%	101.99%
8	Товар 8	900.00	3.5%	82.2%	B	3.33%	26.67%	108.84%
9	Товар 9	700.00	2.7%	84.9%	B	3.33%	30.00%	114.91%
10	Товар 10	600.00	2.3%	87.3%	B	3.33%	33.33%	120.60%
11	Товар 11	500.00	2.0%	89.2%	B	3.33%	36.67%	125.89%
12	Товар 12	400.00	1.6%	90.8%	B	3.33%	40.00%	130.78%
13	Товар 13	300.00	1.2%	92.0%	B	3.33%	43.33%	135.29%
14	Товар 14	200.00	0.8%	92.7%	B	3.33%	46.67%	139.41%
15	Товар 15	200.00	0.8%	93.5%	B	3.33%	50.00%	143.52%
16	Товар 16	150.00	0.6%	94.1%	B	3.33%	53.33%	147.44%
17	Товар 17	140.00	0.5%	94.7%	B	3.33%	56.67%	151.33%
18	Товар 18	130.00	0.5%	95.2%	C	3.33%	60.00%	155.17%
19	Товар 19	120.00	0.5%	95.6%	C	3.33%	63.33%	158.97%
20	Товар 20	110.00	0.4%	96.1%	C	3.33%	66.67%	162.73%
21	Товар 21	105.00	0.4%	96.5%	C	3.33%	70.00%	166.48%
22	Товар 22	100.00	0.4%	96.9%	C	3.33%	73.33%	170.20%
23	Товар 23	100.00	0.4%	97.3%	C	3.33%	76.67%	173.93%
24	Товар 24	100.00	0.4%	97.7%	C	3.33%	80.00%	177.65%
25	Товар 25	100.00	0.4%	98.0%	C	3.33%	83.33%	181.38%
26	Товар 26	100.00	0.4%	98.4%	C	3.33%	86.67%	185.10%
27	Товар 27	100.00	0.4%	98.8%	C	3.33%	90.00%	188.83%
28	Товар 28	100.00	0.4%	99.2%	C	3.33%	93.33%	192.55%
29	Товар 29	100.00	0.4%	99.6%	C	3.33%	96.67%	196.28%
30	Товар 30	100.00	0.4%	100.0%	C	3.33%	100.00%	200.00%
		25 555.00	100.0%					

В этом примере с границами по методу сумм:

- **Группа А – 79%** выручки, **23.3%** наименований
- **Группа В – 15%** выручки, **26.7%** наименований
- **Группа С – 6%** выручки, **50.3%** наименований

XYZ анализ

XYZ анализ, это метод прогноза и анализа стабильности и колебаний спроса продаж по товарам или группам товаров.

Например, спрос на мобильные телефоны марки А составляет:

- Апрель -1000 штук.
- Май — 1100 штук.
- Июнь — 920 штук.

Товар продается стабильно, примерно 1000 штук в месяц. Все колебания спроса в рамках 5- 10%. Этот продукт относится к категории **X**.

Телефон марки В может иметь колебания продаж в рамках 11% – 25% за период. Этот товар отнесем к категории **Y**.

Продажи телефонов марки С мало предсказуемы. Колебания спроса достигают 100%. Такие товары относятся к категории **Z**.

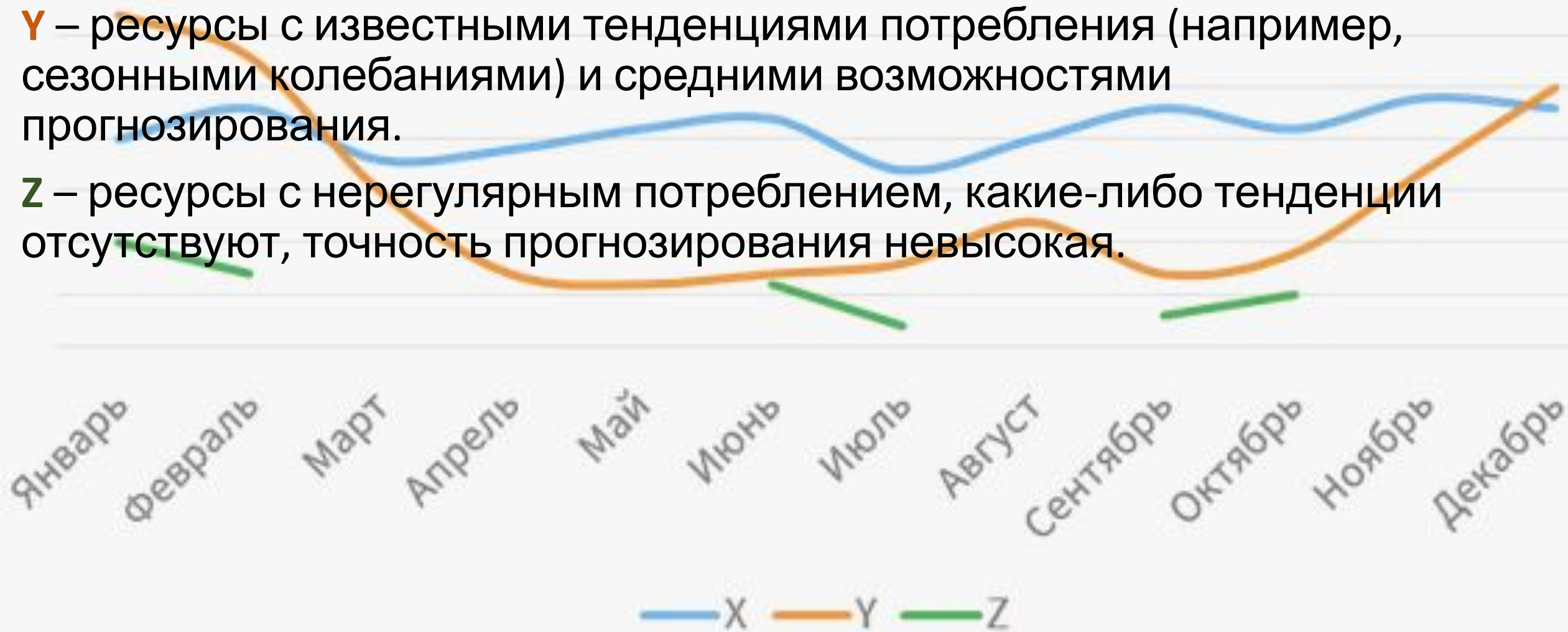
Также к группе **Z** можно отнести сезонные товары.

XYZ-анализ позволяет классифицировать объекты в зависимости от характера потребления и точности прогнозирования его изменения.

X – ресурсы со стабильной величиной потребления и высокой точностью прогноза.

Y – ресурсы с известными тенденциями потребления (например, сезонными колебаниями) и средними возможностями прогнозирования.

Z – ресурсы с нерегулярным потреблением, какие-либо тенденции отсутствуют, точность прогнозирования невысокая.



Формула расчета коэффициента вариации (колебаний) спроса

=СТАНДОТКЛОНП(С2:Е2)/СРЗНАЧ(С2:Е2)

	A	B	C	D	E	F	G
1	код товара	наименование товара	продажи апрель, шт	продажи май, шт	продажи июнь, шт	Итого 2 квартал	коэф. Вариации
2	1009292202	Магнитола Беларусь	45	12	30	87	47%
3	1009292203	Ноутбук Sang Tank SUS	110	90	100	300	8%
4	1009287112	Утюг Киевская Русь	1005	500	1600	3105	43%
5	1009300201	Батарейки Ямайка	10 245	11084	9867	31196	5%

Ноутбук и Батарейки попадают в категорию **X**

Магнитола и Утюг относятся к категории **Z**

XYZ анализ по клиентам

- Клиенты категории **X** — стабильные продажи. По таким клиентам достаточно просто прогнозировать следующие продажи и свои запасы.
- Клиенты категории **Y** — «плавающие» по стабильности продажи.
- Клиенты в категории **Z** — разовые, редкие продажи.

По результатам **XYZ анализа** менеджеры по продажам могут точнее выстраивать свою клиентскую базу.

Аналитика по данному методу может наглядно показать недочеты в работе с клиентами: где 80% усилий дают всего 20%

Матрица ABC-XYZ – единый анализ

Результаты ABC и XYZ анализа можно совместить и получить разделение на 9 групп, которые будут характеризовать ресурсы по обоим критериям.

Таким образом, группа AX будет содержать самые важные ресурсы: наиболее ценные и стабильно потребляемые, а группа CZ – наименее ценные с нерегулярными пот

	<i>X</i>	<i>Y</i>	<i>Z</i>
<i>A</i>	AX <i>высокая ценность, высокая стабильность</i>	AY <i>высокая ценность, средняя стабильность</i>	AZ <i>высокая ценность, низкая стабильность</i>
<i>B</i>	BX <i>средняя ценность, высокая стабильность</i>	BY <i>средняя ценность, средняя стабильность</i>	BZ <i>средняя ценность, низкая стабильность</i>
<i>C</i>	CX <i>низкая ценность, высокая стабильность</i>	CY <i>низкая ценность, средняя стабильность</i>	CZ <i>низкая ценность, низкая стабильность</i>

Набор данных

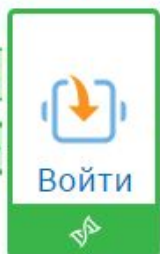
#	31 Дата	9.0 Количество	9.0 Сумма с учетом скид...	ab Группа товара	ab Товар
1	01.03.2017, 00:00	28,00	13 708,80	Напольные покрытия	Доска паркетная UPOFLOOR Лос 2085x188x14 дуб красный преми...
2	01.03.2017, 00:00	112,00	3 311,84	Метизы и крепёж	Гвоздь оцинкованный 1.4x25, 300 штук, 10 категория, уп
3	01.03.2017, 00:00	44,00	2 271,28	Стеновые покрытия	Гипсокартон АБС 9.5x1200x2500, кв.м.
4	01.03.2017, 00:00	44,00	2 259,40	Стеновые покрытия	Гипсокартон БЕЛГИПС 12.5x1200x2500, кв.м
5	01.03.2017, 00:00	44,00	2 820,40	Стеновые покрытия	Гипсокартон БЕЛГИПС 9.5x1200x2500, влагостойкий, мал.уп., кв.м
6	01.03.2017, 00:00	44,00	2 135,76	Стеновые покрытия	Гипсокартон БЕЛГИПС 9.5x1200x2500, кв.м
7	01.03.2017, 00:00	44,00	2 454,76	Стеновые покрытия	Гипсокартон КНАУФ 12.5x1200x2500, кв.м.
8	01.03.2017, 00:00	44,00	2 772,88	Стеновые покрытия	Гипсокартон КНАУФ 12.5x1200x2500, огн., кв.м.
9	01.03.2017, 00:00	44,00	17 442,48	Стеновые покрытия	Лист стекломгнезитовый СМЛ 1220x2440x8, шт
10	01.03.2017, 00:00	70,00	14 422,80	Шпатлёвки	Пигмент блескообразующий TICIANA золото, 0.1 кг, шт
11	01.03.2017, 00:00	60,00	13 557,00	Шпатлёвки	Пигмент блескообразующий TICIANA серебро, 0.1 кг, шт
12	01.03.2017, 00:00	60,00	39 913,80	Шпатлёвки	Покрытие декоративное OPTIMIST ELITE Манна, 15 кг, шт
13	01.03.2017, 00:00	60,00	38 394,60	Шпатлёвки	Покрытие декоративное OPTIMIST ELITE Старый замок, 15 кг, шт
14	01.03.2017, 00:00	50,00	9 001,00	Стеновые покрытия	Решетка радиаторная КОСТРОМА 600x1200мм, бежевая, шт
15	01.03.2017, 00:00	50,00	4 385,50	Стеновые покрытия	Решетка радиаторная КОСТРОМА 600x600мм, бежевая, шт
16	01.03.2017, 00:00	50,00	7 154,50	Стеновые покрытия	Решетка радиаторная КОСТРОМА 600x900мм, бежевая, шт
17	01.03.2017, 00:00	60,00	1 100,40	Шпатлёвки	Шпатлевка ТЕКС РЕ-файн по дереву белая, 0.25 кг, шт
18	01.03.2017, 00:00	60,00	2 675,40	Шпатлёвки	Шпатлевка ТЕКС РЕ-файн по дереву белая, 0.75 кг, шт

Выйти из подмодели

ABC-анализ



XYZ-анализ



Переменные в таблицу



+XYZ



Совместная группа



Статистика по группам



+статистики



ABC-анализ • Быстрый просмотр

#ABC-группы и вклады объектов#

#	ab ГруппаABC	ab Товар	9.0 Вклад нарастающим итог...	9.0 Показатель
57	A	Доска паркетная UPOFLOOR Loc 2085x188x14 бук дымчатый, шт	47,53	6 273 321,64
58	A	Решетка радиаторная КОСТРОМА 600x1500мм, бежевая, шт	47,94	5 898 124,71
59	A	Панель перфорированная LOCATELLI Laccato Quadro 11 HDF, бук,...	48,33	5 783 026,57
60	A	Колеровочная краска SADOLIN Tintorama LSG-FRI для вд красок, ...	48,71	5 545 560,24
61	A	Колеровочная краска SADOLIN Tintorama LSG-FRI для вд красок, ...	49,09	5 426 106,62
62	A	Грунтовка вд ОПТИМИСТ глубокого проникновения, зеленая эти...	49,46	5 398 282,15
63	A	Профиль пластиковый HENKE рамочный 7090 вишня 3 м, шт	49,82	5 339 053,31
64	B	Грунтовка вд полимерная ESCARO AQUASTOP, 10 л, шт	50,18	5 193 212,06
65	B	Решетка радиаторная КОСТРОМА 600x900мм, бежевая, шт	50,53	5 081 600,96
66	B	Краска вд акрилатная SADOLIN Inova Chameleon для стен и пото...	50,87	4 987 256,98
67	B	Доска паркетная TARKETT T-Lock 2272x192x14 дуб натуральный. шт	51,21	4 973 895,80

ABC-анализ • Быстрый просмотр

#ABC-группы и вклады объектов#

#	ab ГруппаABC	ab Товар	9.0 Вклад нарастающим итог...	9.0 Показатель
209	B	покрытие проокное твексокк техническое 4 мм, рулон 15x1 м, ...	79,75	1 681 478,75
210	B	Грунтовка вд BECKERS Skotte Grund Vit 4 л, шт	79,87	1 679 194,84
211	B	Воск бесцветный SENIDECO Effet Matieres Brocart, 0.5 л, шт	79,98	1 655 283,83
212	C	Лента-бордюр 4WALLS для раковин и ванн 30x30x3000, белая, шт	80,09	1 642 336,12
213	C	Плитка напольная KERAMA MARAZZI Ганг кофе 30.2x30.2, кв. м	80,20	1 628 499,67
214	C	Балка декоративная DECOSA Рустик 60x90x4000, темный дуб, шт	80,31	1 602 012,22
215	C	Эмаль алкидная MARSHALL Enamel Parlak для нар/внутр.работ, 6...	80,42	1 593 452,04

XYZ-анализ • Быстрый просмотр

#	ab Товар	ab ГруппаXYZ	9.0 Коэффициент вариации, %	9.0 Среднее значен...
83	доска паркетная АДСКТОР импортный типоразмер 1820x145x14 дуб нату...	Λ	9,10	19 500,45
84	Складной уголок Модерн 28x28x2600 Корень клена, шт	X	9,71	5 413,31
85	Бордюр УРАЛКЕРАМИКА Египет ЕГ707 249x90, шт	X	9,85	1 472,57
86	Колеровочная краска SADOLIN Tintorama LSG-FRI для вд красок, ...	X	10,00	34 104,58
87	Линолеум IVC Wonderwood Jakarta 843 3 м, кв.м	Y	10,32	11 050,89
88	Плитка настенная KERAMA MARAZZI Волна белая 20x25, кв. м	Y	10,65	22 919,78
89	Маты теплоизоляционные URSA Glasswool M-11-2-9000-600-50 (1...	Y	10,67	28 839,12

XYZ-анализ • Быстрый просмотр

#	ab Товар	ab ГруппаXYZ	9.0 Коэффициент вариации, %	9.0 Среднее значен...
253	Плитка настенная KERAMA MARAZZI Неаполь серая 25x40, кв. м	Y	24,94	12 075,15
254	Эмаль алкидная SADOLIN Master-30 WO, 2.5л, шт	Y	24,96	44 351,15
255	Линолеум IVC Tornado Vita 45 3м, кв.м	Y	24,97	9 251,20
256	Цепь короткозвенная 1.5 мм, м	Z	25,06	1 515,70
257	Клей ESKARO Seinaliim для стеклообоев, 5л, шт	Z	25,15	9 237,80
258	Эмаль алкидная MARSHALL Pastel Yarimat 1119 белая п/матовая, ...	Z	25,15	14 049,30

Выходной набор данных					9.0 Вклад нарастаю...	9.0 Показатель	9.0 Коэффициент вариации, %
#	ab	ab	ab	ab Товар			
63	AZ	A	Z	Профиль пластиковый HENKE рамочный 7090 вишня 3 м, шт	49,82	5 339 053,31	42,05
64	BY	B	Y	Грунтовка вд полимерная ESCARO AQUASTOP, 10 л, шт	50,18	5 193 212,06	21,72
65	BZ	B	Z	Решетка радиаторная КОСТРОМА 600x900мм, бежевая, шт	50,53	5 081 600,96	43,23

Выходной набор данных					9.0 Вклад нарастаю...	9.0 Показатель	9.0 Коэффициент вариации, %
#	ab	ab	ab	ab Товар			
207	BZ	B	Z	Эмаль алкидная ALPINA Direkt auf Rost антикоррозийная, по ржа...	79,52	1 695 536,58	26,28
208	BZ	B	Z	Шпатлевка масляно-клеевая ТЕКС Универсал, 30 кг, шт	79,64	1 691 637,05	52,41
209	BZ	B	Z	Покрытие пробковое IBERCORK техническое 4 мм, рулон 15x1 м, ...	79,75	1 681 478,75	37,78
210	BY	B	Y	Грунтовка вд BECKERS Skotte Grund Vit 4 л, шт	79,87	1 679 194,84	16,80
211	BY	B	Y	Воск бесцветный SENIDECO Effet Matieres Brocart, 0.5 л, шт	79,98	1 655 283,83	22,14
212	CZ	C	Z	Лента-бордюр 4WALLS для раковин и ванн 30x30x3000, белая, шт	80,09	1 642 336,12	35,74
213	CZ	C	Z	Плитка напольная KERAMA MARAZZI Ганг кофе 30.2x30.2, кв. м	80,20	1 628 499,67	38,69
214	CZ	C	Z	Балка декоративная DECOSA Рустик 60x90x4000, темный дуб, шт	80,31	1 602 012,22	28,28

Группы	Статистика по группам						
#	ab С..	ab Г	ab Г	ab Товар	9.0 Вклад нарастающим итогом	9.0 Показат...	9.0 Коэффициент вариации, %
60	AZ	A	Z	Колеровочная краска SADOLIN Tintorama ...	48,71	5 545 560,24	43,01
61	AZ	A	Z	Колеровочная краска SADOLIN Tintorama ...	49,09	5 426 106,62	44,86
62	AУ	A	У	Грунтовка вд ОПТИМИСТ глубокого прони...	49,46	5 398 282,15	19,18
63	AZ	A	Z	Профиль пластиковый HENKE рамочный ...	49,82	5 339 053,31	42,05
64	BУ	B	У	Грунтовка вд полимерная ESCARO AQUAS...	50,18	5 193 212,06	21,72
65	BZ	B	Z	Решетка радиаторная КОСТРОМА 600x900...	50,53	5 081 600,96	43,23
66	BZ	B	Z	Краска вд акрилатная SADOLIN Inova Cha...	50,87	4 987 256,98	26,72

Группы	Статистика по группам						
#	ab С..	ab Г	ab Г	ab Товар	9.0 Вклад нарастающим итогом	9.0 Показат...	9.0 Коэффициент вариации, %
208	BZ	B	Z	Шпатлевка масляно-клеевая ЕКС Универ...	79,64	1 691 637,05	52,41
209	BZ	B	Z	Покрытие пробковое IBERCORK техничес...	79,75	1 681 478,75	37,78
210	BУ	B	У	Грунтовка вд BECKERS Skotte Grund Vit 4 л...	79,87	1 679 194,84	16,80
211	BУ	B	У	Воск бесцветный SENIDECO Effet Matieres ...	79,98	1 655 283,83	22,14
212	CZ	C	Z	Лента-бордюр 4WALLS для раковин и ван...	80,09	1 642 336,12	35,74
213	CZ	C	Z	Плитка напольная KERAMA MARAZZI Ганг ...	80,20	1 628 499,67	38,65
214	CZ	C	Z	Балка декоративная DECOSA Рустик 60x90...	80,31	1 602 012,22	28,28

Группы	Статистика по группам	
#	ab Совместная груп...	12 Товар Количество
1	AX	
2	AY	1
3	AZ	62
4	BX	
5	BY	16
6	BZ	132
7	CX	86
8	CY	152
9	CZ	717

1. ABC-XYZ анализ часть товара выводит в супер ТОП (AX).
2. Какой-то товар из категории А обнаруживается в категории AZ (!).
3. Товар с низкой ценностью может иметь высокую стабильность - категория CX.
4. Возможна переоценка значимости товара по категории VX - невысокая маржа, но стабильный доход.
5. Товар из категории CZ? Следует избавляться от таких товаров, если это не компонент другого товара. Сколько такой товар занимает места на складе? Сколько это в «замороженных» деньгах?
6. По товарам группы AX можно увеличить страховой запас.
7. По товарам группы CX и CY можно значительно сократить запас.
8. Расширенный ABC анализ (+ частотность, + количество обращений) и XYZ-анализ совместно дают самую полную картину.
9. Так ли хорош товар, который имеет категорию ACCZ?
10. Так ли плох товар, который имеет категорию СААХ?