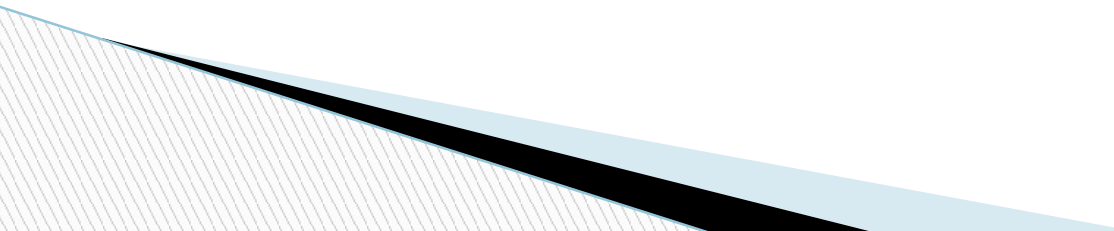


SPSS (Statistical Package for Social Sciences или в новой интерпретации — Superior Performing Software Systems) — система (программный пакет) статистической обработки информации, которая предоставляет пользователю широкие возможности преобразования и анализа данных, а также наглядного представления полученных результатов



- Файл исходной базы данных для проведения статистического анализа в **SPSS** формируется в редакторе данных (**Data Editor**). Редактор данных имеет две вкладки: «**Представление переменные**» (**Variable View**) и «**Представление данные**» (**Data View**). Данные вкладки представляют собой таблицы, содержащие информацию о данных, собранных для проведения анализа.

- Во вкладке редактора данных «Представление переменные» представлена таблица с данными, описывающими значения переменных. Каждый **столбец** отображает **переменную** (вопрос анкеты), каждая **строка** — **отдельное наблюдение** (объект сбора информации). В качестве объектов сбора информации могут выступать люди, предприятия, продукты, бренды и т.д.



	Имя	Тип	Ширина	Знаков ...	Метка	Значения	Пропущенн...	Столбцы	Выравнивание	Мера	Роль
1	Номер	Числовой	8	0	Номер респондента	Нет	Нет	4	По право...	Номинальная	Входная
2	Потребитель	Числовой	8	0	Пьете ли вы пакетированный чай?	{1, да}...	99	8	По право...	Номинальная	Входная
3	Время	Числовой	8	0	В какое время дня вы предпочитаете пить пакетированный чай?	{1, с 6 утра ...	99	8	По право...	Шкалы	Входная
4	Частота	Числовой	8	0	Как часто вы пьете пакетированный чай?	{1, каждый ...	99	8	По право...	Номинальная	Входная
5	Причина	Числовой	8	0	Почему вы пьете пакетированный чай?	{1, это удоб...	99	8	По право...	Номинальная	Входная
6	Причина_дру...	Текстовая	100	0	Другая причина	Нет	Нет	11	По левом...	Номинальная	Входная
7	Lipton	Числовой	8	0	Какие марки пакетированного чая вы предпочитаете? (Lipton)	{1, да}...	99	8	По право...	Номинальная	Входная
8	Greenfield	Числовой	8	0	Какие марки пакетированного чая вы предпочитаете? (Greenfield)	{1, да}...	99	8	По право...	Номинальная	Входная
9	Ahmad	Числовой	8	0	Какие марки пакетированного чая вы предпочитаете? (Ahmad)	{1, да}...	99	8	По право...	Номинальная	Входная
10	Brooke_Bond	Числовой	8	0	Какие марки пакетированного чая вы предпочитаете? (Brooke Bond)	{1, да}...	99	8	По право...	Номинальная	Входная
11	Другие_марки	Текстовая	100	0	Какие марки пакетированного чая вы предпочитаете? (Другие марки)	Нет	Нет	8	По левом...	Номинальная	Входная
12	Состав	Числовой	8	0	Оцените важность характеристик пакетированного чая среди представленных ниже по шкале от 1...	{1, не важн...	99	8	По право...	Порядковая	Входная
13	Вкус	Числовой	8	0	Оцените важность характеристик пакетированного чая среди представленных ниже по шкале от 1...	{1, не важн...	99	8	По право...	Порядковая	Входная
14	Цена	Числовой	8	0	Оцените важность характеристик пакетированного чая среди представленных ниже по шкале от 1...	{1, не важн...	99	8	По право...	Порядковая	Входная
15	Страна_прои...	Числовой	8	0	Оцените важность характеристик пакетированного чая среди представленных ниже по шкале от 1...	{1, не важн...	99	8	По право...	Порядковая	Входная
16	Цвет_при_за...	Числовой	8	0	Оцените важность характеристик пакетированного чая среди представленных ниже по шкале от 1...	{1, не важн...	99	8	По право...	Порядковая	Входная
17	Скорость_за...	Числовой	8	0	Оцените важность характеристик пакетированного чая среди представленных ниже по шкале от 1...	{1, не важн...	99	8	По право...	Порядковая	Входная
18	Упаковка_па...	Числовой	8	0	Оцените важность характеристик пакетированного чая среди представленных ниже по шкале от 1...	{1, не важн...	99	8	По право...	Порядковая	Входная
19	Материалы_...	Числовой	8	0	Оцените важность характеристик пакетированного чая среди представленных ниже по шкале от 1...	{1, не важн...	99	8	По право...	Порядковая	Входная
20	Приспособле...	Числовой	8	0	Оцените важность характеристик пакетированного чая среди представленных ниже по шкале от 1...	{1, не важн...	99	8	По право...	Порядковая	Входная
21	Коп_во_вкус...	Числовой	8	0	Оцените важность характеристик пакетированного чая среди представленных ниже по шкале от 1...	{1, не важн...	99	8	По право...	Порядковая	Входная
22	Другие_хар_к...	Текстовая	100	0	Оцените важность характеристик пакетированного чая среди представленных ниже по шкале от 1...	Нет	Нет	8	По левом...	Номинальная	Входная
23	Выбор	Числовой	8	0	Что в большей степени влияет на Ваш выбор при покупке пакетированного чая?	{1, собстве...	99	8	По право...	Номинальная	Входная
24	Возраст	Числовой	8	0	Возраст респондента	Нет	99	8	По право...	Шкалы	Входная
25	Пол	Числовой	8	0	Пол респондента	{1, жен }...	99	8	По право...	Номинальная	Входная
26	Семья	Числовой	8	0	Семейное положение	{1, не заму...	99	8	По право...	Номинальная	Входная
27	Доход	Числовой	8	0	Доход на одного члена семьи в месяц	{1, менее 1...	99	8	По право...	Шкалы	Входная
28											
29											

Представление Данных Представление Переменных



33 : Страна_производ

Показано 27 переменных из 27

	Номер	Потребитель	Время	Частота	Причина	Причина_другая_причина	Lipton	Greenfield	Ahmad	Brooke_Bond	Другие_марк_и	Состав	Вкус	Цена	Страна_производства	Цвет_при_заварке	Скорость_заварки	Упаковка
1	1	да	с 14 часов	1-2 раза в ...	это быстро		нет	да	да	нет		скорее ва...	очень важно	скорее ва...	скорее ва...	затрудняю...	затрудняю...	очень ва...
2	2	да	с 22 часов...	каждый де...	другие пр...	дешево	да	нет	нет	да	Яшма	не важно	скорее ва...	очень важно	не важно	очень важно	очень важно	не ва...
3	3	да	с 6 утра д...	1-2 раза в ...	это удобно		да	нет	нет	да	Майский	не важно	скорее ва...	очень важно	не важно	очень важно	очень важно	не ва...
4	4	нет	99	99	99		99	99	99	99		99	99	99	99	99	99	99
5	5	да	с 14 часов	3-5 раз в н...	это вкусно		нет	да	да	нет		скорее ва...	очень важно	скорее ва...	скорее ва...	затрудняю...	затрудняю...	очень ва...
6	6	да	с 22 часов...	1-2 раза в ...	другие пр...	под рукой	да	нет	нет	да	Японский	не важно	скорее ва...	очень важно	не важно	очень важно	очень важно	не ва...
7	7	да	с 6 утра д...	3-5 раз в н...	это вкусно		да	нет	нет	да		не важно	скорее ва...	очень важно	не важно	очень важно	очень важно	не ва...
8	8	нет	99	99	99		99	99	99	99		99	99	99	99	99	99	99
9	9	да	с 6 утра д...	1-2 раза в ...	другие пр...	дал сосед	да	нет	нет	да		не важно	скорее ва...	очень важно	не важно	очень важно	очень важно	не ва...
10	10	да	с 22 часов...	1-2 раза в ...	это вкусно		да	нет	нет	да	Лисла	не важно	скорее ва...	очень важно	не важно	очень важно	очень важно	не ва...
11	11	да	с 14 часов...	3-5 раз в н...	это быстро		нет	да	да	нет		скорее ва...	очень важно	скорее ва...	скорее ва...	затрудняю...	затрудняю...	очень ва...
12	12	да	с 22 часов...	реже 1 раз...	другие пр...	взял у родителей	да	нет	нет	да	Принцесса...	не важно	скорее ва...	очень важно	не важно	очень важно	очень важно	не ва...
13	13	да	с 6 утра д...	1-2 раза в ...	это удобно		да	нет	нет	да		не важно	скорее ва...	очень важно	не важно	очень важно	очень важно	не ва...
14	14	нет	99	99	99		99	99	99	99		99	99	99	99	99	99	99
15	15	да	с 14 часов...	реже 1 раз...	это удобно		да	нет	нет	да	Яшма	не важно	скорее ва...	очень важно	не важно	очень важно	очень важно	не ва...
16	16	да	с 6 утра д...	1-2 раза в ...	это удобно		да	нет	нет	да		не важно	скорее ва...	очень важно	не важно	очень важно	очень важно	не ва...
17	17	да	с 14 часов...	реже 1 раз...	это быстро		нет	да	да	нет		скорее ва...	очень важно	скорее ва...	скорее ва...	затрудняю...	затрудняю...	очень ва...
18	18	да	с 14 часов...	1-2 раза в ...	это удобно		да	нет	нет	да		не важно	скорее ва...	очень важно	не важно	очень важно	очень важно	не ва...
19	19	да	с 6 утра д...	реже 1 раз...	это удобно		да	нет	нет	да		не важно	скорее ва...	очень важно	не важно	очень важно	очень важно	не ва...
20	20	да	с 14 часов...	3-5 раз в н...	это быстро		нет	да	да	нет		скорее ва...	очень важно	скорее ва...	скорее ва...	затрудняю...	затрудняю...	очень ва...
21	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
22	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
23	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
24	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
25	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
26	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
27	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Представление Данных Представление Переменных

Применение видов статистического анализа данных в зависимости от типов переменных (зависимых и независимых)

Вид переменных		Независимые переменные	
		Метрическая шкала	Номинальная (порядковая) шкала
Зависимые переменные	Метрическая шкала*	Регрессионный анализ	Дисперсионный анализ
	Номинальная (порядковая) шкала	Дискриминантный анализ	Таблицы сопряженности

***Метрическая шкала** – переменные, измеренные по относительной или интервальной шкале

В случае **кластерного и факторного** видов статистического анализа возможно использование любого типа переменных (метрических и не метрических).

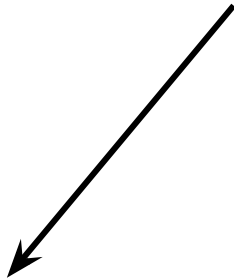
Дисперсионный анализ

**изучает влияние категориальной
переменной (фактора)- номинальной или
порядковой на количественные
переменные (интервальные или
относительные)**

Примеры задач, решаемых с помощью дисперсионного анализа:

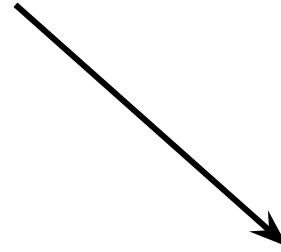
- Как уровень рекламы и уровень цен (высокий, средний, низкий) одновременно влияют на продажи товара данной торговой марки?
- Связан ли выбор потребителей данной торговой марки с уровнем образования (ниже среднего, среднее, колледж, высшее) и возрастом?
- Как осведомленность об универмаге (высокая, средняя, низкая) и представление о нем (позитивное, нейтральное, негативное) влияют на предпочтение потребителем этого магазина?

Дисперсионный анализ



одномерный

- однофакторный
- многofакторный
- многofакторный



многомерный

-

Число зависимых переменных	Число независимых переменных	Применяемы й метод статистическ ого анализа
1	1	Однофакторны й дисперсионный анализ
1	2	Двухфакторный дисперсионный анализ
1	3	Трехфакторный дисперсионный анализ
Минимум 2	1 и более	Многомерный дисперсионный

Независимая переменная №1 Не метрическая!	Независимая переменная №2 Не метрическая!		
	Регион 1	Регион 2	Регион 3
	Показатели объема продаж (тыс. штук) (зависимая переменная) Метрическая!		
Тип упаковки А	3567	5673	6478
Тип упаковки В	4567	2567	3569
Тип упаковки С	7856	4769	4736

Однофакторный дисперсионный анализ

Устанавливает, значимо ли различаются средние значения нескольких независимых выборок.

Нулевая гипотеза (H_0) утверждает, что k генеральных совокупностей имеют одно и то же среднее значение. То есть категориальный фактор не влияет на количественную переменную.

Альтернативная гипотеза (H_1) утверждает, что средние значения не все равны между собой. По крайней мере они различаются у двух совокупностей.

Для проверки нулевой гипотезы обычно используют **тест Ливина** **F-критерий Фишера** и **величину значимости** (значимость полученного результата).

Если величина значимости меньше величины 0,05, то делается вывод о том, что гипотеза о равенстве средних значений отвергнута с вероятностью ошибки 0%, то есть различия в средних значениях для разных групп неслучайны!

- Прежде всего мы проведем **однофакторный одномерный дисперсионный анализ** и
- **установим, насколько значимо различается кратность покупок в различных возрастных**
- **группах респондентов (1 — младше 18 лет; 2 — 19-35 лет; 3 — 36-60 лет; 4 — старше 60**
- **лет).**

- Рассмотрим теперь ситуацию, когда необходимо исследовать сразу две независимые переменные (и взаимодействия между ними), то есть выполнить **двухфакторный одномерный дисперсионный анализ**.
- Исходные данные останутся такими же, как в предыдущем примере, однако теперь мы будем устанавливать **различие в кратности покупок сырков возрастными и половыми группами (переменная q3)**.

ДИСКРИМИНАНТНЫЙ АНАЛИЗ

Определение

Дискриминантный анализ –

раздел многомерного статистического анализа, включающий в себя методы

КЛАССИФИКАЦИИ многомерных

наблюдений по принципу максимально

СХОДСТВА при наличии **ОБУЧАЮЩИХ**

признаков.

Ограничения использования

Дискриминантный анализ используется в тех случаях, когда заранее известно **число** групп, на которые необходимо разбить набор объектов, а также имеется набор объектов, по которым уже известно, к каким группам они принадлежат.

(пример: люди, купившие товар одной из конкурирующих марок)

Ограничения использования

Дискриминантный анализ используется в случае, когда зависимая переменная - **номинальная** или **порядковая** (ее еще называют классифицирующей), а независимые переменные – **количественные** (**допускаются** порядковые)

Задачи, решаемые с помощью применения дискриминантного анализа:

- Определение решающих правил, позволяющих по значению количественных признаков (независимых переменных) отнести каждый объект к одному из известных классов – *составление классифицирующей функции*;
- Классификация неизвестных объектов при наличии данных по известным объектам, то есть прогнозирование, в какой группе будет относиться неизвестный объект по известным о нем характеристикам.

Примеры задач

- Определить, на основании каких характеристик потребитель выбирает товар и предсказать выбор еще не определившихся потребителей (! ограничение: либо берем количественные характеристики, либо опрос по важности характеристик по шкале важности)
- Определить, существует ли зависимость факта покупки («купил»/«не купил») в магазине от таких переменных как время посещения магазина, время пребывания в магазине, количество людей в магазине, время консультации с продавцом, ...

Примеры задач

- - обосновать выбор категории заведений общественного питания (кафе, бар, ресторан, кофейня, кондитерская) в зависимости от степени выраженности различных факторов (чувство голода посетителя, наличие компании, ассортимент заведения, количество людей в заведении, ...);
- - обосновать выбор марок спортивной одежды различными группами потребителей (профессионалы, любители, приверженцы активного отдыха, новички).

Пример

- Определить, существует ли зависимость выбора йогурта от важности для потребителя таких характеристик как состав и вкус
- Опрашивались потребители 2 марок йогурта: Активия и Даниссимо
- Важность вкусовых характеристик и состава была оценена респондентами по 8-балльной шкале, где 1 – совсем не важно, 8 – критически важно

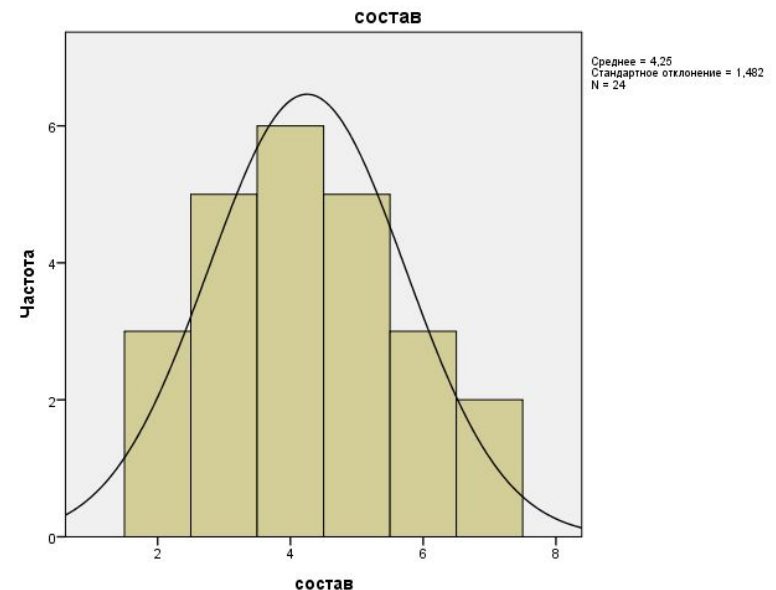
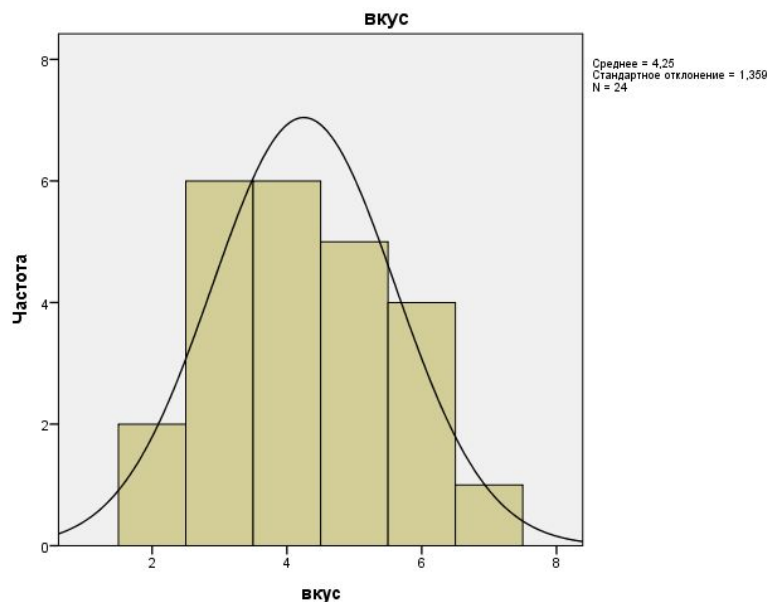


	номер	вкус	состав	выбор	пе
1	1	2	3	Активия	
2	2	3	4	Активия	
3	3	6	5	Активия	
4	4	4	4	Активия	
5	5	3	2	Активия	
6	6	4	7	Активия	
7	7	3	5	Активия	
8	8	2	4	Активия	
9	9	5	6	Активия	
10	10	3	6	Активия	
11	11	3	3	Активия	
12	12	4	5	Активия	
13	13	5	4	Даниссимо	
14	14	4	3	Даниссимо	
15	15	7	5	Даниссимо	
16	16	3	3	Даниссимо	
17	17	4	4	Даниссимо	
18	18	5	2	Даниссимо	

2. Проверка нормальности распределения независимых переменных

* Дискриминантный анализ создавался для анализа нормально распределенных переменных

** Если одна или несколько независимых переменных имеют распределение, отличное от нормального, то использовать дискриминантный анализ **МОЖНО**, но необходимо указать это в **ограничениях исследования**



6. Прогнозирование – классификация последующих объектов по выявленной модели

13	13	5	4	Даниссимо	Даниссимо
14	14	4	3	Даниссимо	Даниссимо
15	15	7	5	Даниссимо	Даниссимо
16	16	3	3	Даниссимо	Активия
17	17	4	4	Даниссимо	Активия
18	18	5	2	Даниссимо	Даниссимо
19	19	4	2	Даниссимо	Даниссимо
20	20	5	5	Даниссимо	Даниссимо
21	21	6	7	Даниссимо	Даниссимо
22	22	5	3	Даниссимо	Даниссимо
23	23	6	4	Даниссимо	Даниссимо
24	24	6	6	Даниссимо	Даниссимо
25	25	5	3	.	Даниссимо
26	26	4	8	.	Активия
27	27	6	4	.	Даниссимо
28					

Задача

Компания занимается продажей подержанных автомобилей и хочет составить модель для прогнозирования факта покупки авто в течение первого месяца после выставления на продажу в зависимости от следующих характеристик:

- Цена авто (в у.е.)
- Техническое состояние (оценено по 10-балльной шкале, где 1 – очень плохое, 10 – отличное)
- Возраст (в годах)
- Пробег (в км)

Корреляционный анализ

выявляет наличие, а также определяет направление и силу линейной связи между несколькими переменными, имеющими интервальный, порядковый или дихотомический тип шкалы.

Основные характеристики переменных

	Зависимые переменные		Независимые переменные	
Метод	Количество	Тип	Количество	Тип
<i>Корреляционный анализ</i>	-	-	Любое	Интервальная Порядковая Дихотомическая
<i>Линейная регрессия</i>	Одна	Интервальная Порядковая	Любое	Интервальная Порядковая Дихотомическая

Коэффициенты корреляции

Коэффициент корреляции	Распределение переменных	Шкала	Формула
Корреляция Пирсона (корреляция моментов произведения)	Нормальное	Интервальная Количественная	$r_{xy} = \frac{\Sigma(x_i - \bar{x}) \times (y_i - \bar{y})}{\sqrt{\Sigma(x_i - \bar{x})^2 \times \Sigma(y_i - \bar{y})^2}}$
Ранговая корреляция Спирмена или Кендала	Не является нормальным (хотя бы у одной переменной)	Порядковая шкала (хотя бы у одной переменной)	Спирмен: $r_s = 1 - \frac{6 \Sigma_i d_i^2}{N(N^2 - 1)}$ Кендала: $\tau = \frac{P(p) - P(q)}{N \frac{(N-1)}{2}}$ (рекомендуется при наличии выбросов)

Пример: Корреляция Пирсона

Каков Ваш среднемесячный доход в расчете на одного члена семьи?

Ответы	Коды в SPSS	
	Порядковая переменная	Интервальная переменная
■ до \$100 (1);	1	50
■ от \$ 100 до \$ 300;	2	200
■ от \$ 300 до \$ 600;	3	450
■ от \$ 600 до \$ 1000;	4	800
■ от \$ 1000 до \$ 1500;	5	1250
■ свыше \$1500.	6	1750

Как часто Вы посещаете рестораны?

Ответы	Коды в SPSS	
	Порядковая переменная	Интервальная переменная
■ более 1 раза в день;	1	60
■ примерно 1 раз в день;	2	30
■ 2-3 раза в неделю;	3	10
■ примерно 1 раз в неделю;	4	4
■ 2-3 раза в месяц;	5	2,5
■ примерно 1 раз в месяц;	6	1
■ реже 1 раза в	7	0,5

Пример: Корреляция Спирмена

Какие факторы для Вас наиболее важны при выборе одежды?

- Высокое качество одежды.
- Доступные цены.
- Широта ассортимента одежды.
- Близость к дому или работе.
- Высокое качество обслуживания.
- Красивый интерьер магазина.

Оцените, пожалуйста, следующие характеристики данного магазина одежды (в котором происходит опрос) по пятибалльной шкале (от 1 — очень плохо до 5 — отлично)

- ■ Высокое качество одежды.
- ■ Доступные цены.
- ■ Широта ассортимента одежды.
- ■ Близость к дому или работе.
- ■ Высокое качество обслуживания.
- ■ Красивый интерьер магазина.
- ■ Ваша общая оценка работы данного магазина.

Регрессионный анализ

- определяет степень детерминированности различий значений (вариаций) зависимой переменной у разных наблюдений независимой(ыми) переменной(ыми);
- предсказывает значения зависимой переменной с помощью независимой(ых);
- определяет вклад отдельных независимых переменных в вариацию зависимой.

Регрессионный анализ в маркетинговых исследованиях

- Какие частные параметры продукта оказывают влияние на общее впечатление потребителей от данного продукта?

Например, требуется установить, как влияет возраст и пол респондента на частоту покупок шоколадок (построение уравнения с целью прогноза).

- Какие частные характеристики продукта в большей степени влияют на цену продукта?

*Например, требуется установить, что влияет в большей степени на **цену**: материал продукта или цвет продукта (установление соотношения между различными частными параметрами по силе и направлению влияния на общее впечатление).*

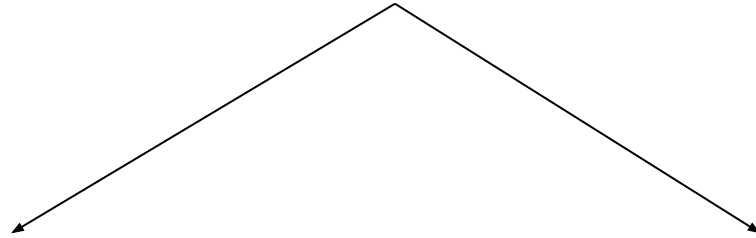
- Как ведет себя одна переменная в зависимости от изменения другой?

Например, необходимо построить график зависимости осведомленности о шоколадках и частоты покупки. Как изменится частота покупок при увеличении осведомленности покупателя на 10%. (графическое прогнозирование – только для двух переменных)

Основные характеристики переменных

	Зависимые переменные		Независимые переменные	
Метод	Количество	Тип	Количество	Тип
<i>Корреляционный анализ</i>	-	-	Любое	Интервальная Порядковая Дихотомическая
<i>Линейная регрессия</i>	Одна	Интервальная Порядковая	Любое	Интервальная Порядковая Дихотомическая

Регрессия

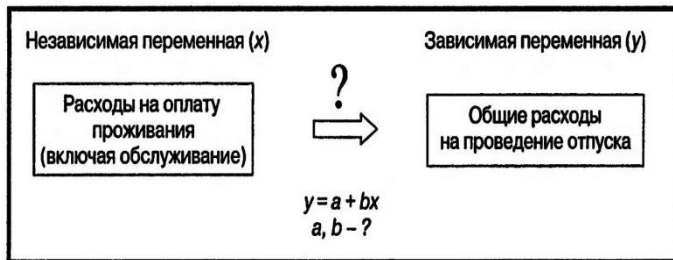


Простая регрессия

$$y = a + bx$$

Множественная
регрессия

$$y = a + b_1x_1 + b_2x_2 + \dots + b_nx_n$$



Пример: множественная линейная регрессия

Респонденты трёх классов:

- Первый класс
- Бизнес-класс
- Эконом-класс

Необходимо:

- 1) Выявить наиболее значимые для респондентов параметры обслуживания на борту.
- 2) Установить, какое влияние оказывают оценки частных параметров обслуживания на борту на общее впечатление авиапассажира от полета.

Оцените по пятибалльной шкале следующие характеристики сервиса на борту авиакомпании X (1 – очень плохо, 5 – отлично):

- комфортабельность салона,
- работа бортпроводников,
- питание во время полета,
- цена билетов,
- спиртные напитки,
- дорожные наборы,
- аудиопрограммы,
- видеопрограммы,
- пресса,
- общая оценка.

Пример: множественная линейная регрессия

$СБ = 0,78 + 0,20К + 0.20Б + 0,08ПП + 0.07С + 0Д0Н + 0,08В + 0Д2П$, где

- СБ — общая оценка сервиса на борту;
- К — комфортабельность салона;
- Б — работа бортпроводников;
- ПП — питание во время полета;
- С — спиртные напитки;
- Н — дорожные наборы;
- В — видеопрограмма;
- П — пресса.



	Имя	Тип	Ширина	Знаков ...	Метка	Значения	Пропущенн...	Ширина ...	Выравнивание	Шкала	Роль
1	price	Числовой	4	0	цена	Нет	Нет	8	☰ По право...	🔧 Шкала	👉 Входная
2	sqft	Числовой	4	0	площадь	Нет	Нет	8	☰ По право...	🔧 Шкала	👉 Входная
3	feats	Числовой	1	0	удобства	Нет	Нет	8	☰ По право...	📊 Порядковый	👉 Входная
4	north_e	Числовой	1	0	район	{0, не северо-восток}...	Нет	8	☰ По право...	📊 Порядковый	👉 Входная
5	cust	Числовой	1	0	тип застройки	{0, не обычная застройка}...	Нет	8	☰ По право...	📊 Порядковый	👉 Входная
6	corner	Числовой	1	0	угол	{0, не угловое здание}...	Нет	8	☰ По право...	📊 Порядковый	👉 Входная
7	age2	Числовой	8	0	возраст	Нет	-9999	8	☰ По право...	🔧 Шкала	👉 Входная
8	tax2	Числовой	8	0	налоги	Нет	-9999	8	☰ По право...	🔧 Шкала	👉 Входная
9											
10											
11											
12											
13											
14											
15											
16											
17											
18											
19											
20											
21											
22											

Представление Данные Представление Переменные

Факторный анализ

позволяет разделить массив переменных на малое число групп (факторов).

Классификация переменных по различным факторам (группам) производится на основе коэффициента корреляции между исследуемыми переменными.

В один фактор объединяются переменные, которые имеют высокий коэффициент корреляции друг с другом и не коррелируют или имеют низкий коэффициент корреляции с другими переменными, входящими в состав других факторов.

Факторный анализ в маркетинговых исследованиях

Сегментирование рынка

Факторный анализ применяется для выявления агрегатных переменных, являющихся основанием для сегментирования потребителей.

Например, потребители плавленых сыров могут характеризоваться различной степенью значимости, которую они видят в исследуемых характеристиках данного продукта (респондентов просят оценить по пятибалльной шкале важность нескольких характеристик плавленых сыров: срок хранения, калорийность, процент жирности и т. д.). Здесь факторный анализ позволит выявить целевые сегменты потребителей на основании значимости для них различных групп факторов:

- покупатели, ориентирующиеся при выборе плавленого сыра преимущественно на ценовые факторы (стоимость, скидки);
- покупатели, ориентирующиеся на качество исследуемого продукта (срок хранения, состав ингредиентов, вкус);
- покупатели, выбирающие сыр в основном по внешнему виду (дизайн упаковки).

Факторный анализ в маркетинговых исследованиях

Изучение продукта и бенчмаркинг продукта

В данном случае факторный анализ помогает выявить агрегатные параметры продукта, влияющие на выбор потребителя.

Например, различные марки шоколадных конфет могут быть оценены по следующим макрокатегориям:

- качество (ингредиенты, вкус),
- полезность для здоровья (наличие сахара, калорийность),
- цена.

Факторный анализ в маркетинговых исследованиях

Рекламные и медиа-исследования

Факторный анализ может использоваться для выявления скрытых мотивов поведения потребителей при восприятии рекламы.

Ценообразование

Факторный анализ используется для выявления особенностей поведения потребителей, чувствительных к цене.

Например, данная категория респондентов может характеризоваться повышенным вниманием к ценовым факторам при выборе продукта, низкими доходами, большой численностью семьи и т. д.

Основные характеристики переменных

	Зависимые переменные		Независимые переменные	
Метод	Количество	Тип	Количество	Тип
<i>Факторный анализ</i>	Нет	-	Любое	Любой



	Имя	Тип	Ширина	Знаков ...	Метка	Значения	Пропущенн...	Ширина ...	Выравнивание	Шкала
1	V1	Числовой	8	2	Я потребляю напиток, чтобы освежить дыхание	Нет	Нет	8	☰ По право...	Шкала
2	V2	Числовой	8	2	Я потребляю напиток, чтобы чувствовать себя увереннее	Нет	Нет	8	☰ По право...	Шкала
3	V3	Числовой	8	2	Я потребляю напиток, чтобы освежить полость рта	Нет	Нет	8	☰ По право...	Шкала
4	V4	Числовой	8	2	Я потребляю напиток, чтобы избавиться от неприятного вкуса во рту	Нет	Нет	8	☰ По право...	Шкала
5	V5	Числовой	8	2	Я потребляю напиток как заменитель других напитков	Нет	Нет	8	☰ По право...	Шкала
6	V6	Числовой	8	2	Я потребляю напиток, когда мне хочется что-нибудь съесть	Нет	Нет	8	☰ По право...	Шкала
7	V7	Числовой	8	2	Напиток помогает мне сконцентрироваться	Нет	Нет	8	☰ По право...	Шкала
8	V8	Числовой	8	2	Я потребляю напиток, чтобы отвлечься на несколько минут и подумать	Нет	Нет	8	☰ По право...	Шкала
9	V9	Числовой	8	2	Я потребляю напиток для того, чтобы сжечь жировые отложения	Нет	Нет	8	☰ По право...	Шкала
10	V10	Числовой	8	2	Я потребляю напиток для того, чтобы избавиться от лишнего веса	Нет	Нет	8	☰ По право...	Шкала
11	V11	Числовой	8	2	Я потребляю напиток для улучшения самочувствия	Нет	Нет	8	☰ По право...	Шкала
12										
13										
14										
15										
16										
17										
18										
19										
20										
21										
22										

Пример: факторный анализ

q1. Авиакомпания X обладает репутацией компании, превосходно обслуживающей пассажиров.

q2. Авиакомпания X может конкурировать с лучшими авиакомпаниями мира.

q3. Я верю, что у авиакомпании X есть перспективное будущее в мировой авиации.

q4. Я знаю, какой будет стратегия развития авиакомпании X в будущем.

q5. Я горжусь тем, что работаю в авиакомпании X.

q6. Внутри авиакомпании X хорошее взаимодействие между подразделениями.

q7. Каждый сотрудник авиакомпании прикладывает все усилия для того, чтобы обеспечить ее успех.

q8. Сейчас авиакомпания X быстро улучшается.

q9. Нам предстоит долгий путь, прежде чем мы сможем претендовать на то, чтобы называться авиакомпанией мирового класса.

q10. Авиакомпания X действительно заботится о пассажирах.

q11. Среди сотрудников авиакомпании имеет место высокая степень удовлетворенности работой.

q12. Я верю, что менеджеры высшего звена прикладывают все усилия для достижения успеха авиакомпании.

q13. Мне нравится, как в настоящее время авиакомпания X представлена визуально широкой общественности (в плане цветовой гаммы и фирменного стиля).

q14. Авиакомпания X — лицо России.

q15. Мы выглядим «вчерашним днем» по сравнению с другими авиакомпаниями.

q16. Обслуживание авиакомпании X является последовательным и узнаваемым во всем мире.

q17. Я бы не хотел, чтобы авиакомпания X менялась.

q18. Авиакомпания X необходимо меняться для того, чтобы использовать в полной мере имеющийся потенциал.

q19. Я думаю, что авиакомпании X необходимо представить себя в визуальном плане более современно.

q20. Изменения в авиакомпании X будут позитивным моментом.

q21. Авиакомпания X — эффективная авиакомпания.

q22. Я бы хотел, чтобы имидж авиакомпании X улучшился с точки зрения иностранных пассажиров.

q23. Авиакомпания X — лучше, чем многие другие.

q24. Я бы хотел, чтобы лидером в отрасли на территории России была российская авиакомпания.

Выявить схожие (то есть тесно коррелирующие между собой) утверждения и разделить их на несколько однородных групп — описывающих различные аспекты (макропараметры) конкурентной позиции авиакомпании X на рынке. Другими словами, выделить группы схожих по значению параметров авиакомпании,

Пример: факторный анализ

q01	Статистика заставляет меня плакать
q02	Мои друзья будут думать, что я глупый, так как я не могу разобраться в SPSS
q03	Стандартные отклонения возбуждают меня
q04	Мне снится, что Пирсон нападает на меня с коэффициентами корреляции
q05	Я не понимаю статистику
q06	У меня маленький опыт работы с компьютером
q07	Все компьютеры ненавидят меня
q08	Я никогда не был хорош в математике
q09	Мои друзья лучше справляются со статистикой
q10	Компьютеры нужны только для игр
q11	У меня было плохо с математикой в школе
q12	Люди пытаются сказать, что SPSS делает статистику проще, но это не так
q13	Я боюсь, что я нанесу непоправимый ущерб из-за моей некомпетентности в компьютерах
q14	Компьютеры имеют свое сознание и специально ломаются, когда я начинаю работать с ними
q15	Компьютеры существуют, чтобы достать меня
q16	Я открыто плачу при упоминании о центральной тенденции
q17	Я впадаю в кому каждый раз, когда вижу уравнение
q18	SPSS всегда ломается, когда я пытаюсь ее использовать
q19	Все смотрят на меня, когда я пользуюсь SPSS
q20	Я не могу спать при мысли о собственных векторах
q21	Я просыпаюсь под моим пуховым одеялом, думая, что я в ловушке под нормальным распределением
q22	Мои друзья лучше справляются с SPSS, чем я
q23	Если я буду хорош в статистике, мои друзья подумают, что я кретин

Необходимо:
Разделить
респондентов на
группы, путем
снижения
размерности.

Иерархический кластерный анализ

Классификационный метод анализа данных

Цель – выделения однородных групп (**кластеров**) из исследуемой совокупности объектов (потребителей, продуктов, брендов и т.д.).

Эти кластеры должны быть однородными внутри и разнородными между собой!

Иерархический кластерный анализ используется для классификации набора объектов, когда заранее не известно число групп, на которые нужно этот набор разбить. Группы, на которые разбита выборка, называются кластерами. Число групп заранее не задается.

Примеры задач, решаемых с помощью кластерного анализа:

- определение групп потребителей – сегментация (выделение существующих/потенциальных);**
- когда на рынке присутствует большой выбор товаров одного назначения под разными торговыми марками. Необходимо разбить товары на группы схожих товаров;**
- определение потенциальных групп потребителей. Результаты классификации используются, чтобы в дальнейшем для разных групп определить оптимальные цены на услуги, оптимальные тарифы.**

Объекты исследования (туристы)	Характеристики объектов (переменные, по которым производится разделение на кластеры)	
	Возраст	Интересы (мотивы поведения)
Турист №1		
Турист №2		
Турист №125....		

Таб.1 Исходные переменные

Объекты исследо- вания (возраст ные группы туристов)	Характеристики объектов (переменные, по которым производится разделение на кластеры)			
	Развлече- ния	Спец. предложен- ия	Спокойны й отдых	Спор- т
17-18 лет				
19-24 года				
..... 65-70 лет				

Tab.2 Конкретизированные переменные

Пример:

- В ходе исследования было опрошено 745 авиапассажиров, летавших одной из 22 российских и зарубежных авиакомпаний. Авиапассажиров просили оценить по пятибалльной шкале — от 1 (очень плохо) до 5 (отлично) — семь параметров работы наземного персонала авиакомпаний в процессе регистрации пассажиров на рейс: вежливость, профессионализм, оперативность, готовность помочь, регулирование очереди, внешний вид, работа персонала в целом.

- **Требуется:**
- Сегментировать исследуемые авиакомпании по **уровню воспринимаемого авиапассажирами качества работы наземного персонала.**
- Итак, у нас есть файл данных, который состоит из семи интервальных переменных, обозначающих оценки качества работы наземного персонала различных авиакомпаний (q13-q19), представленные в единой пятибалльной шкале. Файл данных содержит одновариантную переменную q4, указывающую выбранные респондентами авиакомпании (всего 22 наименования). Проведем кластерный анализ и определим, на какие целевые группы можно разделить данные авиакомпании.

КОНДЖОИНТ-АНАЛИЗ

Конджоинт-анализ - метод для определения самого лучшего набора атрибутов, представляющих продукт или услугу.

Цель конджоинт-анализа - измерение степени предпочтения потребителем одного из конкурирующих продуктов (услуг) в условиях предположения о комплексной оценке всех атрибутов, составляющих продукт.

Задачи, решаемые с помощью применения КОНДЖОИНТ-анализа:

- Определить наилучшую (оптимальную) конфигурацию свойств продукта или услуги
- Произвести сравнение свойств (атрибутов) продуктов с целью выявления тех из них, которые оказывают наибольшее влияние на покупательские решения.
- Достоинством метода является возможность выявить латентные факторы, влияющие на поведение потребителей.
- Сегментация рынка на основе потребительских предпочтений, например, оценка размера сегмента, отдающего предпочтение сервису, либо сегмента, чувствительного в первую очередь к цене, и т.д. Подобная информация позволяет выбрать наиболее привлекательные сегменты рынка и разработать стратегию работы с выбранными сегментами

Примеры

- Выбор оптимальной конфигурации товара: например, поиск “компромисса” между количеством и сложностью встроенных функций комнатного кондиционера и показателями потребления кондиционером электроэнергии.
- Выбор оптимального сочетания цены и размера упаковки для любого товара рынка FMCG
- Определение значимости атрибутов товаров или услуг: изучение приоритетов потребителей по ключевым атрибутам товаров или услуг, например, сервис, цены, имидж, качество, широта ассортимента для розничной сети.

Пример в рамках курсовой работы

У оператора мобильной связи имеется несколько тарифов, отличающихся по таким характеристикам, как стоимость звонков внутри и вне сети, стоимость смс и ммс сообщений, стоимость различных интернет-услуг, наличие включенных пакетов и дополнительных бонусов. По результатам опроса потребителей с помощью применения процедуры конджоинт анализа необходимо:

- Определить, какие отличительные характеристики оказывают наибольшее влияние на выбор потребителя;
- Составить профиль оптимального тарифа;
- Выбрать тариф оператора, наиболее близкий к оптимальному и сформулировать управленческое решение.

Последовательность действий (в теории) - 1

Необходимо определить внешний вид упаковки нового шампуня, которая будет наиболее благосклонно воспринята покупателями.

1. Выделить характеристики, по которым различаются упаковки шампуней: - объем упаковки, - форма упаковки, - основной цвет тюбика, - рисунок / цвет рисунка, - материал, из которого изготовлена упаковка, - размер крышки и т.д.
2. ВЫБРАТЬ НАИБОЛЕЕ ВАЖНЫЕ ХАРАКТЕРИСТИКИ и для каждой их характеристик составить список всех возможных значений фактора:

АТТРИБУТЫ (ФАКТОРЫ)	УРОВНИ АТТРИБУТОВ (ЗНАЧЕНИЯ ФАКТОРОВ)
объем упаковки	100 мл
	150 мл
форма упаковки	с острыми краями
	с закругленными краями
основной цвет тюбика	белый
	голубой
материал, из которого изготовлена упаковка	пластик
	стекло

- Каждый из этих наборов значений факторов называется **ПРОФИЛЕМ** например, пластиковая овальная бутылка объемом 150 мл белого цвета.
- **Задача** – определить оптимальный профиль, наиболее предпочитаемый потребителями.

Ограничения

- Участники эксперимента должны быть действительными или потенциальными пользователями исследуемого товара или услуги.
- Конджоинт-анализ невозможно применять для оценки товаров, атрибуты которых взаимосвязаны либо если товар или услуга не могут быть подвергнуты декомпозиции до элементарных атрибутов.
- «Сложные» товары, то есть обладающие большим количеством значимых для принятия решений атрибутов, могут генерировать слишком большое количество альтернатив, так что респондент оказывается не в состоянии обработать настолько большое число вариантов