

Описательные статистики

ЛР №1

Задание 1

1. Найдите переменную по варианту
2. Сделайте расчет указанных ниже статистических показателей в Statistica.
3. Результат скопируйте в WS Power Point .
4. Дайте письменную интерпретацию полученных результатов.

Рабочий файл: **Рейтинг агентств поискового маркетинга.xls** Лист **Stat 100**

Переменные по выбору

Форма сдачи задания: Файл MS Power Point с расчетами и интерпретацией результатов

1. МЕРЫ СРЕДНЕГО УРОВНЯ

- Мода
- Медиана
- Среднее

2. МЕРЫ РАССЕЯНИЯ (ДИСПЕРСИИ)

- Дисперсия
- Стандартное отклонение
- Стандартная ошибка
- Размах

3. МЕРЫ РАСПРЕДЕЛЕНИЯ

- Асимметрия
- Эксцесс

Вывод по характеру распределения вариационного ряда



Примеры

Мода ряда распределения объема продаж размеров женских сапог:

Размер	35	36	37	38	39	40	41		
Частота продаж			6	14	22	30	18	7	3

Медиана ряда распределения по уровню ежемесячного дохода 15 и 14 человек:

Респондент	1	2	...	8	...	14	15
Тыс.руб/мес	15	20	...	28	...	35	40

$$\text{Медиана}_{15} = 28$$

Респондент	1	2	...	7	8	...	14
Тыс.руб/мес	15	20	...	26	28	...	35

$$\text{Медиана}_{14} = (26+28)/2 = 27$$



Примеры. Среднее и размах вариации

Пусть дан следующий вариационный ряд:

№	(1)	(2)	(3)	(4)	(5)
(x)	13	15	17	25	30

$$\bar{x} = \frac{13 + 15 + 17 + 25 + 30}{5} = \frac{100}{5} = 20$$

Размах вариации:

$$R = x_{\max} - x_{\min} = 30 - 13 = 17$$



Примеры. Стандартное отклонение

Стандартное отклонение:

$$\begin{aligned}d &= \frac{\sum_{k=1}^n |x_k - \bar{x}|}{n} \\&= \frac{|13 - 20| + |15 - 20| + |17 - 20| + |25 - 20| + |30 - 20|}{5} \\&= \frac{7 + 5 + 3 + 5 + 10}{5} = \frac{30}{5} = 6\end{aligned}$$



Примеры. Дисперсия

Дисперсия вариационного ряда, пример:

$$\begin{aligned}\sigma^2 &= \frac{13^2 + 15^2 + 17^2 + 25^2 + 30^2}{5} - 20^2 \\ &= \frac{2208}{5} - 20^2 = 441,6 - 400 = 41,6\end{aligned}$$

Среднее квадратическое отклонение:

$$\sigma = \sqrt{\frac{\sum_{k=1}^n (x_k - \bar{x})^2}{n}} = \sqrt{41,6} \cong 6,45$$



Примеры. Коэффициент вариации

Коэффициент вариации признака
(относительная величина):

$$CV = \frac{\sigma}{\bar{x}} \times 100\%$$

$$= \frac{6.45}{20} \times 100\% \cong 32.25\%$$



Примеры. Коэфф. асимметрии

Коэффициент асимметрии:

$$A = \frac{\mu_3}{\sigma^3} = \frac{\sum_{l=1}^n (x_l - \bar{x})^3}{n\sigma^3}$$

Если $A = 0$, то распределение имеет симметричную форму

Если $A < 0$, то распределение скошено влево

Если $A > 0$, то распределение скошено вправо



Примеры. Эксцесс

Эксцесс вариационного ряда:

$$E = \frac{\mu_4}{\sigma^4} - 3 = \frac{\sum_{l=1}^n (x_l - \bar{x})^4}{n\sigma^4} - 3$$

Эксцесс нормального распределения равен 0

Если $E < 0$, то распределение имеет более пологую вершину чем нормальное

Если $E > 0$, то распределение более островерхое чем нормальное



Пример вывода

Вывод для формы распределения
вариационного ряда (по примеру выше):

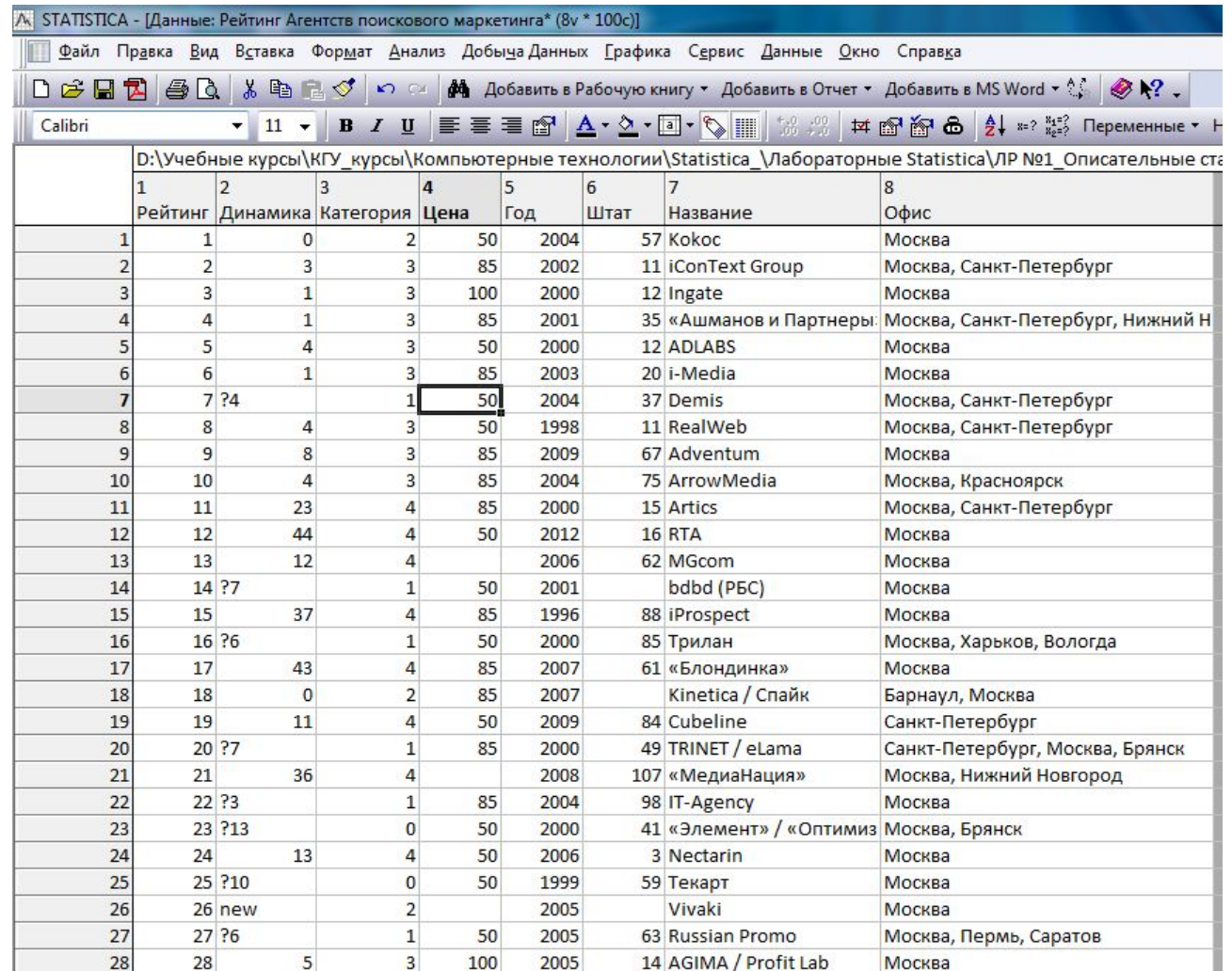
Имеет место:

- ✓ правая скошенность распределения
($A = 0,07 > 0$)
- ✓ вершина полигона более пологая, чем у
нормального распределения
($E = -1,41 < 0$)



Пример анализа в Statistica

Открыт лист
Stat100



STATISTICA - [Данные: Рейтинг Агентств поискового маркетинга* (8v * 100с)]

Файл Правка Вид Вставка Формат Анализ Добыча Данных Графика Сервис Данные Окно Справка

Добавить в Рабочую книгу Добавить в Отчет Добавить в MS Word

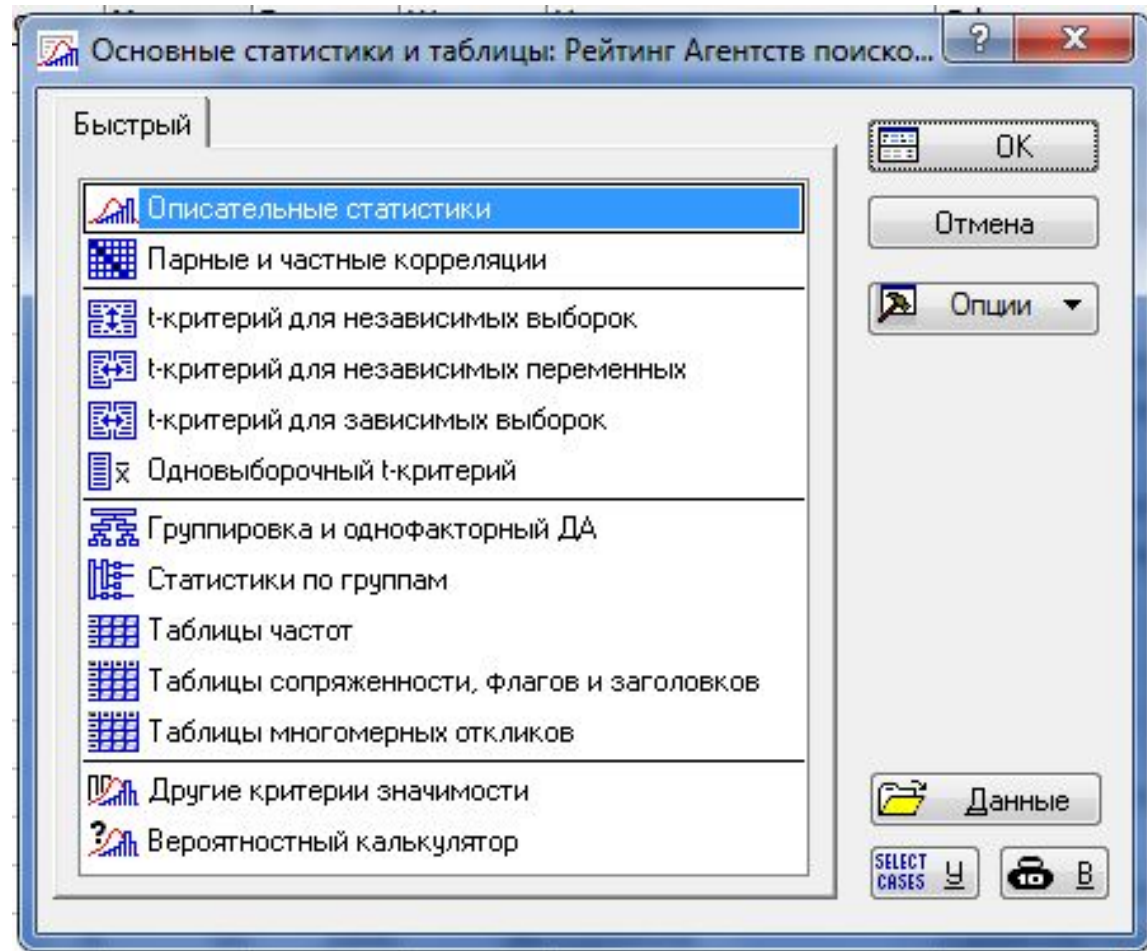
Calibri 11 B I U

D:\Учебные курсы\КГУ_курсы\Компьютерные технологии\Statistica_Лабораторные Statistica\ЛР №1_Описательные ста

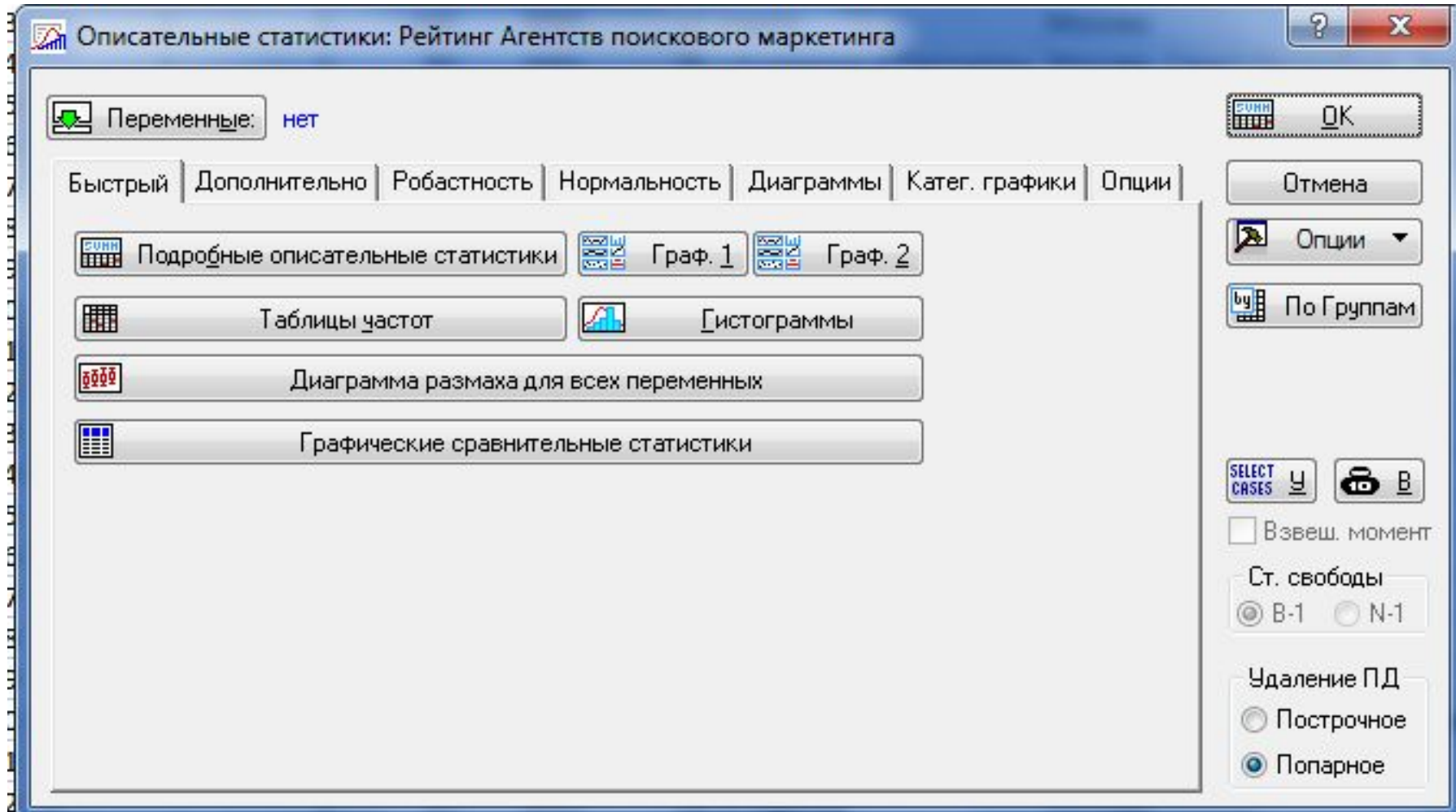
	1	2	3	4	5	6	7	8
	Рейтинг	Динамика	Категория	Цена	Год	Штат	Название	Офис
1	1	0	2	50	2004	57	Kokoc	Москва
2	2	3	3	85	2002	11	iConText Group	Москва, Санкт-Петербург
3	3	1	3	100	2000	12	Ingate	Москва
4	4	1	3	85	2001	35	«Ашманов и Партнеры»	Москва, Санкт-Петербург, Нижний Н
5	5	4	3	50	2000	12	ADLABS	Москва
6	6	1	3	85	2003	20	i-Media	Москва
7	7 ?4		1	50	2004	37	Demis	Москва, Санкт-Петербург
8	8	4	3	50	1998	11	RealWeb	Москва, Санкт-Петербург
9	9	8	3	85	2009	67	Adventum	Москва
10	10	4	3	85	2004	75	ArrowMedia	Москва, Красноярск
11	11	23	4	85	2000	15	Artics	Москва, Санкт-Петербург
12	12	44	4	50	2012	16	RTA	Москва
13	13	12	4		2006	62	MGcom	Москва
14	14 ?7		1	50	2001		bdbd (РБС)	Москва
15	15	37	4	85	1996	88	iProspect	Москва
16	16 ?6		1	50	2000	85	Трилан	Москва, Харьков, Вологда
17	17	43	4	85	2007	61	«Блондинка»	Москва
18	18	0	2	85	2007		Kinetica / Спайк	Барнаул, Москва
19	19	11	4	50	2009	84	Cubeline	Санкт-Петербург
20	20 ?7		1	85	2000	49	TRINET / eLama	Санкт-Петербург, Москва, Брянск
21	21	36	4		2008	107	«МедиаНация»	Москва, Нижний Новгород
22	22 ?3		1	85	2004	98	IT-Agency	Москва
23	23 ?13		0	50	2000	41	«Элемент» / «Оптимиз	Москва, Брянск
24	24	13	4	50	2006	3	Nectarin	Москва
25	25 ?10		0	50	1999	59	Текарт	Москва
26	26 new		2		2005		Vivaki	Москва
27	27 ?6		1	50	2005	63	Russian Promo	Москва, Пермь, Саратов
28	28	5	3	100	2005	14	AGIMA / Profit Lab	Москва

Пример анализа в Statistica

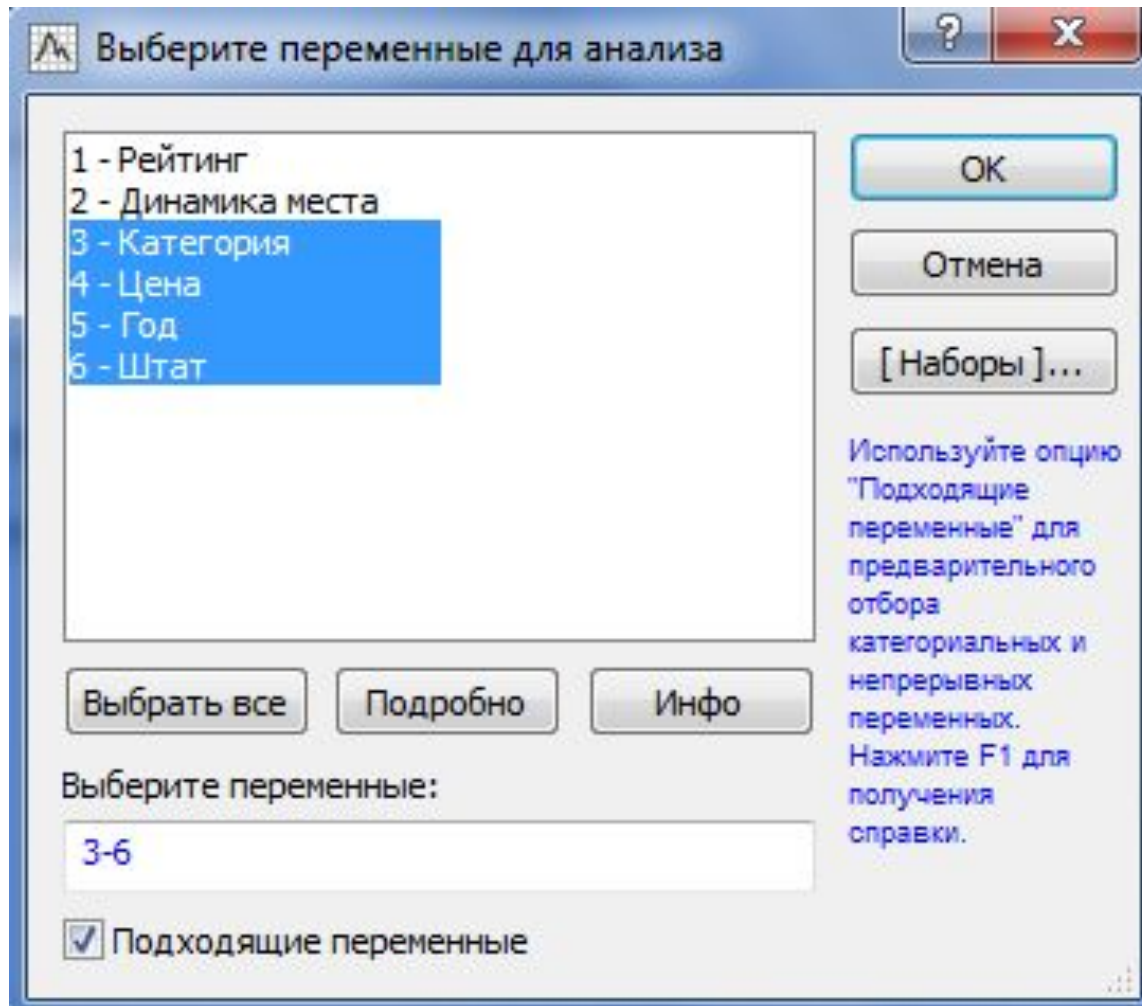
Открыт лист
Stat100



Пример анализа в Statistica



Пример анализа в Statistica



Пример анализа в Statistica

The screenshot shows the 'Descriptive Statistics' dialog box in Statistica. The title bar reads 'Описательные статистики: Рейтинг Агентств поискового маркетинга'. The variable 'Категория-Штат' is selected in the 'Переменные:' field. The 'Быстрые' (Quick) tab is active, and the 'Дополнительно' (Advanced) tab is highlighted with a red circle. The 'Подобные описательные статистики' (Similar Descriptive Statistics) section is expanded, showing three columns of options:

- Парам. положения (Positional Parameters):**
 - N набл.
 - % набл.
 - Среднее
 - Сумма
 - Медиана
 - Мода
 - Геометр. среднее
 - Гармонич. среднее
- Вармация, моменты (Variance, Moments):**
 - Стандартное отклонение
 - Доверит. интервал станд. откл. (Уровень: 95,00 %)
 - Коэффициент вариации
 - Дисперсия
 - Стандартн. ошибка среднего
 - Доверит. интервал среднего (Уровень: 95,00 %)
 - Асимметрия
 - Станд. ош. асимметрии
 - Эксцесс
 - Станд. ош. эксцесса
- Процентили, размахи (Percentiles, Ranges):**
 - Минимум и максимум
 - Нижн., верхн. квартили
 - Границы процентилей (Первая: 10,00 %, Вторая: 90,00 %)
 - Размах
 - Квартильн. размах

Buttons on the right include 'Отмена', 'Опции', 'По Группам', 'SELECT CASES', 'Взвеш. момент', 'Ст. свободы' (B-1, N-1), 'Удаление ПД' (Построчное, Попарное), 'Выбрать все', 'Сброс', and 'Сохранить как умолчания'. The 'OK' button is at the top right.

Пример анализа в Statistica

STATISTICA - [Рабочая книга4* - Описательные статистики (Рейтинг Агентств поискового маркетинга)]

Файл Правка Вид Вставка Формат Анализ Добыча Данных Графика Сервис Данные Рабочая книга Окно Справка

Добавить в Рабочую книгу Добавить в Отчет Добавить в MS Word

Arial 10 **B I U** Переменные Наблюдения

Переменная	Описательные статистики (Рейтинг Агентств поискового маркетинга)													
	N набл.	Сред нее	Медиана	Мода	Частота моды	Минимум	Максим.	Дисперсия	Ст.откл.	Дов.инт. Ст.откл. -95,000%	Дов.инт. Ст.откл. +95,000%	Козф. Вар	Асимметрия	Эксцесс
Категория	100	2,200	2,000	4,000000	25	0,000	4,000	1,980	1,40705	1,23540	1,63454	63,95695	-0,164282	-1,23033
Цена	61	68,197	50,000	50,00000	31	50,000	100,000	360,861	18,99633	16,12237	23,12681	27,85520	0,163994	-1,80323
Год	100	04,530	2005,000	2007,000	11	1996,000	2014,000	16,090	4,01123	3,52189	4,65975	0,20011	0,052922	-0,48584
Штат	62	70,565	65,000	Множест.	2	1,000	149,000	2378,938	48,77436	41,44658	59,27410	69,12024	0,167726	-1,35555

Переменная	Описательные статистики (Рейтинг Агентств поискового маркетинга)						
	N набл.	Сред нее	Медиана	Мода	Частота моды	Минимум	Максим.
Категория	100	2,200	2,000	4,000000	25	0,000	4,000
Цена	61	68,197	50,000	50,00000	31	50,000	100,000
Год	100	04,530	2005,000	2007,000	11	1996,000	2014,000
Штат	62	70,565	65,000	Множест.	2	1,000	149,000

Дисперсия	Ст.откл.	Дов.инт. Ст.откл. -95,000%	Дов.инт. Ст.откл. +95,000%	Козф. Вар	Асимметрия	Эксцесс
1,980	1,40705	1,23540	1,63454	63,95695	-0,164282	-1,23033
360,861	18,99633	16,12237	23,12681	27,85520	0,163994	-1,80323
16,090	4,01123	3,52189	4,65975	0,20011	0,052922	-0,48584
2378,938	48,77436	41,44658	59,27410	69,12024	0,167726	-1,35555

Задание 2

1. Выберите вариант категоризации (регион или отрасль)
2. Постройте частотное распределение по числовой переменной выручка.
3. Результат в виде таблицы скопируйте в Excel.
4. Удалите ненужные данные, подготовьте таблицу и постройте по ней диаграмму и отредактируйте её
5. Скопируйте диаграмму на слайд MS Power Point и дайте ее интерпретацию.

Рабочий файл: Эксперт400.xls

Форма сдачи задания: Файл MS Power Point с расчетами, интерпретацией

На диаграмме должны быть указаны:

- ✓ название диаграммы
 - ✓ названия категорий
 - ✓ названия осей
 - ✓ подписи данных
 - ✓ линии сетки
-

