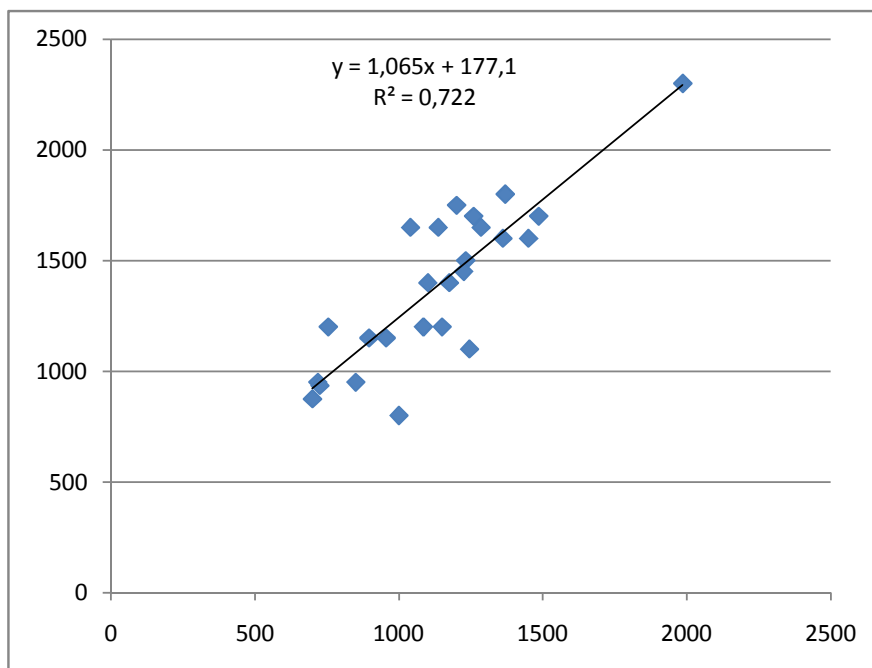


X Y

850	950
1450	1600
1085	1200
1232	1500
718	950
1485	1700
1136	1650
726	935
700	875
956	1150
1100	1400
1285	1650
1985	2300
1369	1800
1175	1400
1225	1450
1245	1100
1259	1700
1150	1200
896	1150
1361	1600
1040	1650
755	1200
1000	800
1200	1750



1. Используя регрессионный модуль ответить на вопросы:
 - а) Определить уравнение линии регрессии $Y=b_0+b_1X$
 - б) Интерпретировать коэффициенты регрессии b_0 и b_1
 - в) Рассчитать коэффициент детерминации и интерпретировать
 - г) Проверить гипотезу, что коэффициент направления регрессии генерального множества равен 0 . $\alpha=0.05$
 - д) Определить 95% зоны регрессии (рассчитать левые и правые границы для всех X значений выборки) Результат изобразить графически с помощью EXCEL
 - у) Используя графики Residual Plot, Line Fit Plot, Normal Probability Plot проверить исполняются ли приемы анализов регрессии.

1, Найдем коэффициенты линии регрессии
С помощью встроенных функций

b_1 = (наклон линии регрессии)
 b_0 =(отрезок отсекаемый по оси Y)

1,0651
177,12

Уравнение регрессии: $y=177,1208+1,095144X$

2, Интерпретировать коэффициенты регрессии

Коэффициент b_0 - является коэффициентом постоянного изменения, как бы переменная X не менялась. При $X = 0$.

$Y = 177,1208$

Коэффициент b_1 - является коэффициентом переменного изменения. При изменении переменной X на 1 единицу,

3, Коэффициент детерминации

Найдем сперва коэффициент корреляции:

$R = 0,8501$

Коэффициент детерминации - это коэффициент корреляции в квадрате

$R^2 = 0,7226$
 $0,7226$

Коэффициент детерминации является мерой вариации переменной Y , которая объясняется

Другими словами на 72,26% переменная X влияет на переменную Y

Что свидетельствует о сильной связи между переменными X и Y

г) Проверить гипотезу, что коэффициент направления регрессии генерального множества равен 0. $\alpha = 0.05$

Для проверки гипотезы проведем сперва регрессионный анализ с помощью встроенного модуля EXCEL

ВЫВОД ИТОГОВ

Регрессионная статистика

Множественный коэффициент корреляции: 0,8501

R-квадрат: 0,7226

Нормированный коэффициент корреляции: 0,7105

Стандартная ошибка: 194,6

Наблюдено: 25

Дисперсионный анализ

	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Значимость F</i>
Регрессия	1	2268776,545	2268776,545	59,914	8E-08
Остаток	23	870949,4547	37867,3676		
Итого	24	3139726			

	Коэффициенты	Стандартная ошибка	t-статистика	P-Значение				
				Значение	Нижние 95%	Верхние 95%	Нижние 95,0%	Верхние 95,0%
Y-пересечение	177,1208	161,0042766	1,100100096	0,28267	-155,94	510,1835404	-155,942	510,18354
Переменная X1	1,065144	0,137608412	7,740398215	7,52E-08	0,78048	1,349808593	0,780479	1,34980859

Хотим проверить гипотезу о том что регрессия существует

Другими словами выдвинем 2 гипотезы:

H_0 : $b_1 = 0$ регрессия отсутствует

H_1 : $b_1 \neq 0$ регрессия имеет место быть

Нам потребуется следующие данные

b1= 1,065143906
Sb1= 0,137608412

сам коэффициент
стандартная ошибка

Статистика $t=b1/Sb1=$ 7,740398215

Критическое значение пр 0,05

$Tkr(0,05; n=25-2)$ 2,068657599

Так как наблюдаемое значение статистики $7,740398215 > 2,068657599$ больше критического значения, то гипотезу о равенстве коэффициента B1 нулю

РЕГРЕССИЯ ЕСТЬ!!! Гипотеза H1 верна!

д) Определить 95% зоны регрессии (рассчитать левые и правые границы для всех X значений выборки) Результат изобразить графически с помощью EXCEL

X	Y	Yprog	V(Yvid)	V(Yvid)	V(Yvid)	V(Yind)	V(Yind)	V(Yind)right
850	950	1082,493141	15,70557867	1074,3	1091	37883,073	679,86	1485,128
1450	1600	1721,579484	17,41979138	1712,9	1730	37884,787	1318,9	2124,223
1085	1200	1332,801959	8,030214632	1326,9	1339	37875,398	930,21	1735,396
1232	1500	1489,378113	8,693374399	1483,3	1495	37876,061	1086,8	1891,976
718	950	941,894145	24,73093192	931,61	952,2	37892,099	539,21	1344,577
1485	1700	1758,859521	19,68250155	1749,7	1768	37887,050	1356,2	2161,515
1136	1650	1387,124298	7,78386078	1381,4	1393	37875,151	984,53	1789,717
726	935	950,4152962	24,08740931	940,26	960,6	37891,455	547,74	1353,095
700	875	922,7215547	26,22439887	912,13	933,3	37893,592	520,03	1325,412
956	1150	1195,398395	10,91288283	1188,6	1202	37878,280	792,79	1598,008
1100	1400	1348,779117	7,905210224	1343	1355	37875,273	946,19	1751,372
1285	1650	1545,83074	9,963958666	1539,3	1552	37877,332	1143,2	1948,435
1985	2300	2291,431474	78,03735371	2273,2	2310	37945,405	1888,5	2694,397
1369	1800	1635,302828	13,09755818	1627,8	1643	37880,465	1232,7	2037,924
1175	1400	1428,66491	7,93703059	1422,8	1434	37875,305	1026,1	1831,258
1225	1450	1481,922106	8,566431576	1475,9	1488	37875,934	1079,3	1884,519
1245	1100	1503,224984	8,954425951	1497	1509	37876,322	1100,6	1905,824
1259	1700	1518,136998	9,272341566	1511,8	1524	37876,640	1115,5	1920,738
1150	1200	1402,036313	7,804786319	1396,3	1408	37875,172	999,44	1804,629
896	1150	1131,48976	13,35715397	1123,9	1139	37880,725	728,87	1534,112
1361	1600	1626,781677	12,73995566	1619,4	1634	37880,108	1224,2	2029,401
1040	1650	1284,870483	8,667964815	1278,8	1291	37876,036	882,27	1687,468
755	1200	981,3044695	21,85905346	971,63	991	37889,227	578,64	1383,972
1000	800	1242,264727	9,565707819	1235,9	1249	37876,933	839,66	1644,867
1200	1750	1455,293508	8,190912342	1449,4	1461	37875,559	1052,7	1857,888

Xsr= 1135,3
n= 25

Левая граница для индивидуального значения Y рассчитывается по формуле

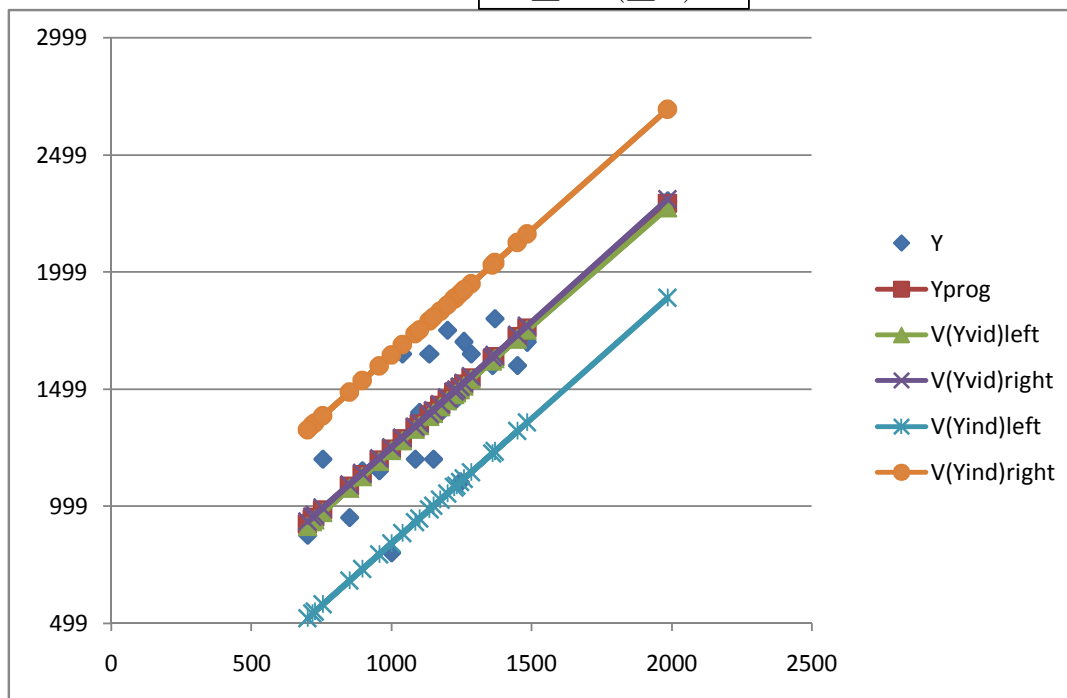
$$y = b_0 + b_1 x_0 - t_{\alpha; n-2} * S * \sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum x_i^2 - (\sum x_i)^2 / n}}$$

Правая граница для индивидуального значения Y рассчитывается по формуле

$$y = b_0 + b_1 x_0 + t_{\alpha; n-2} * S * \sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum x_i^2 - (\sum x_i)^2 / n}}$$

В колонки рассчитываем величину

$$\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum x_i^2 - (\sum x_i)^2 / n}$$



Как видно из графика все наблюдения находятся в заданных границах

у) Используя графики Residual Plot, Line Fit Plot, Normal Probability Plot проверить исполняются ли приемы анализов
ВЫВОД ИТОГОВ

ессионная статистика

Множеств	0,8501
R-квадрат	0,7226
Нормиров	0,7105
Стандарт	194,6
Наблюден	25

Дисперсионный анализ

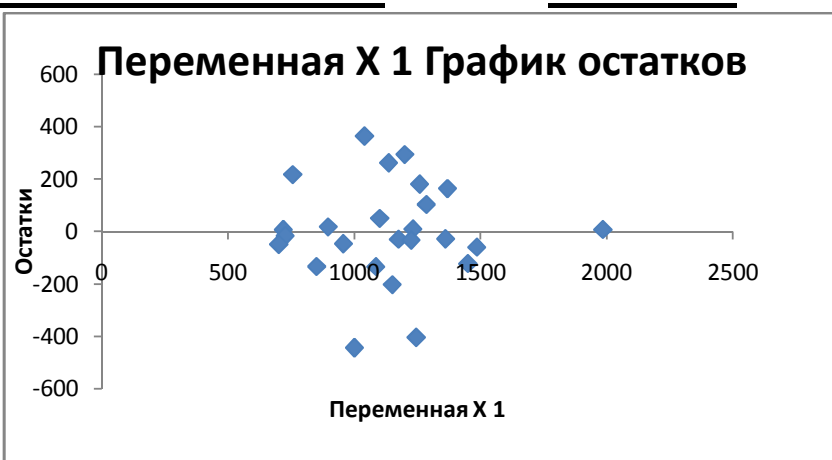
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>значимость F</i>
Регрессия	1	2268776,545	2268776,545	59,914	8E-08
Остаток	23	870949,4547	37867,3676		
Итого	24	3139726			

	<i>Коэффициент</i>	<i>Стандартная ошибка</i>	<i>Статистика</i>	<i>Значение</i>	<i>Верхние 95% границы</i>	<i>Нижние 95% границы</i>	<i>Верхние 95,0%</i>	<i>Нижние 95,0%</i>
Y-пересеч	177,12	161,0042766	1,100100096	0,2827	-155,9	510,18354	-155,9	510,1835
Переменная	1,0651	0,137608412	7,740398215	8E-08	0,78	1,34980859	0,7805	1,349809

ВЫВОД ОСТАТКА

ВЫВОД ВЕРОЯТНОСТИ

<i>Наблюдения</i>	<i>Значения</i>	<i>Остатки</i>	<i>Персенти</i>	<i>Y</i>
1	1082,5	-132,4931406	2	800
2	1721,6	-121,5794845	6	875
3	1332,8	-132,8019586	10	935
4	1489,4	10,62188713	14	950
5	941,89	8,105855007	18	950
6	1758,9	-58,85952118	22	1100
7	1387,1	262,8757021	26	1150
8	950,42	-15,41529624	30	1150
9	922,72	-47,72155468	34	1200
10	1195,4	-45,39839471	38	1200
11	1348,8	51,22088277	42	1200
12	1545,8	104,1692601	46	1400
13	2291,4	8,568525633	50	1400
14	1635,3	164,697172	54	1450
15	1428,7	-28,6649102	58	1500
16	1481,9	-31,92210552	62	1600
17	1503,2	-403,2249837	66	1600
18	1518,1	181,8630017	70	1650
19	1402	-202,0363125	74	1650
20	1131,5	18,51023967	78	1650
21	1626,8	-26,78167679	82	1700
22	1284,9	365,1295172	86	1700
23	981,3	218,6955305	90	1750
24	1242,3	-442,2647266	94	1800
25	1455,3	294,7064921	98	2300



Переменная X 1 График подбора



График нормального распределения

