

Вводим фиктивную переменную «ПОЛ»



Таблица исходных данных:

|  |  |  |  |
| --- | --- | --- | --- |
| i | y | x | z |
| 1 | 6 | 10 | 1 |
| 2 | 4 | 6 | 0 |
| 3 | 4 | 8 | 1 |
| 4 | 5 | 8 | 0 |
| 5 | 4 | 6 | 0 |
| 6 | 7 | 7 | 1 |
| 7 | 3 | 6 | 0 |
| 8 | 4 | 7 | 1 |
| 9 | 9 | 9 | 1 |
| 10 | 3 | 6 | 0 |
| 11 | 2 | 5 | 1 |
| 12 | 3 | 7 | 0 |

Для построения регрессии используем Анализ данных – Регрессия:





Результат расчета:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ВЫВОД ИТОГОВ |  |  |  |  |  |  |
| *Регрессионная статистика* |  |  |  |  |  |
| Множественный R | 0,733 |  |  |  |  |  |
| R-квадрат | 0,537 |  |  |  |  |  |
| Нормированный R-квадрат | 0,434 |  |  |  |  |  |
| Стандартная ошибка | 1,488 |  |  |  |  |  |
| Наблюдения | 12,000 |  |  |  |  |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Дисперсионный анализ |  |  |  |  |  |  |
|  | *df* | *SS* | *MS* | *F* | *Значимость F* |  |
| Регрессия | 2,000 | 23,083 | 11,541 | 5,215 | 0,031 |  |
| Остаток | 9,000 | 19,917 | 2,213 |  |  |  |
| Итого | 11,000 | 43,000 |   |   |   |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | *Коэффициенты* | *Стандартная ошибка* | *t-статистика* | *P-Значение* | *Нижние 95%* | *Верхние 95%* |
| Y-пересечение | -2,086 | 2,309 | -0,903 | 0,390 | -7,310 | 3,139 |
| x | 0,885 | 0,343 | 2,582 | 0,030 | 0,110 | 1,660 |
| z | 0,634 | 0,947 | 0,669 | 0,520 | -1,509 | 2,777 |
|  |  |  |  |  |  |  |
| ВЫВОД ОСТАТКА |  |  |  |  |  |  |
| *Наблюдение* | *Предсказанное y* | *Остатки* |  |  |  |  |
| 1,000 | 7,398 | -1,398 |  |  |  |  |
| 2,000 | 3,224 | 0,776 |  |  |  |  |
| 3,000 | 5,628 | -1,628 |  |  |  |  |
| 4,000 | 4,994 | 0,006 |  |  |  |  |
| 5,000 | 3,224 | 0,776 |  |  |  |  |
| 6,000 | 4,743 | 2,257 |  |  |  |  |
| 7,000 | 3,224 | -0,224 |  |  |  |  |
| 8,000 | 4,743 | -0,743 |  |  |  |  |
| 9,000 | 6,513 | 2,487 |  |  |  |  |
| 10,000 | 3,224 | -0,224 |  |  |  |  |
| 11,000 | 2,973 | -0,973 |  |  |  |  |
| 12,000 | 4,109 | -1,109 |  |  |  |  |

Уравнение множественной регрессии примет вид:



Коэффициент детерминации равен 0,537.

Уравнение регрессии значимо по F-критерию на 5% уровне, так как /

Из полученного уравнения регрессии следует, что при одном и том же числе решенных примеров число решенных задач на курсовых экзаменах у студентов юношей на 0,634 балла оценки выше, чем у девушек.

Это утверждение можно проверить на значимость с помощью статистики Стьюдента:

,

то можно сказать, что влияние фактора Z (пол) не значимо влияет на результаты курсовых экзаменов.

**Вывод**: Можно считать, что модель одна и та же для юношей и девушек.