**В качестве У берете показатель, по которому строили тренд.**

**Х (не менее 4) выбираете самостоятельно**

**Лабораторная работа № 2**

На уровень инвестиций в строительство влияет большое количество факторов. Попробуем изучить взаимосвязь величины уровень инвестиций в строительство и других экономических явлений, происходящих в Оренбургской области. В корреляционно-регрессионном анализе можно устранить воздействие какого-либо фактора, если зафиксировать воздействие этого фактора на результат и другие, включенные в модель факторы. Данный прием широко применяется в анализе временных рядов, когда тенденция фиксируется через включение фактора времени в модель в качестве независимой переменной.

Для проведения корреляционно-регрессионного анализа используем следующие факторные признаки:

У – уровень инвестиций в строительство;

Х1 – количество строительных организаций;

Х2 – среднесписочная численность работников строительных организаций, тыс.тыс.руб.;

Х3 – среднемесячная номинальная начисленная заработная плата работников занятых на строительстве зданий и сооружений;

Х4 – средний уровень использования производственных мощностей строительных организаций, %;

Х5 – число убыточных строительных организаций, % от общего числа организаций;

Х6 – материальные затраты на производство строительных работ, в % к итогу всех затрат;

Х7 – задолженность заказчиков за выполненные объемы работ, в % от общей задолженности;

Х8 – сальдированный финансовый результат.

Параметры модели с включением фактора времени оцениваются с помощью обычного метода наименьших квадратов (МНК).

С помощью ПК получаем матрицу парных коэффициентов, на основании которых необходимо сделать вывод о факторах, которые могут быть включены в модель множественной регрессии (таблица 6.4). Корреляционная матрица получена с помощью табличного редактора Excel ХР в пакете анализа.

Таблица 6.4 – Корреляционная матрица влияния факторов на уровень инвестиций в строительство Оренбургской области

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | *у* | *Х1* | *Х2* | *Х3* | *Х4* | *Х5* | *Х6* | *Х7* | *Х8* |
| У | 1 |  |  |  |  |  |  |  |  |
| Х1 | **0,784196** | 1 |  |  |  |  |  |  |  |
| Х2 | **0,740948** | 0,472513 | 1 |  |  |  |  |  |  |
| Х3 | 0,170823 | 0,20517 | 0,0347 | 1 |  |  |  |  |  |
| Х4 | **-0,57755** | -0,16364 | -0,42623 | -0,21985 | 1 |  |  |  |  |
| Х5 | -0,15388 | -0,27971 | -0,35876 | 0,479488 | -0,20222 | 1 |  |  |  |
| Х6 | -0,11082 | -0,74114 | -0,51858 | 0,163678 | -0,97561 | 0,446238 | 1 |  |  |
| Х7 | -0,3984 | 0,115731 | -0,12415 | -0,15768 | 0,78703 | -0,42972 | -0,9336 | 1 |  |
| Х8 | **-0,79397** | -0,7641 | -0,84743 | 0,038866 | 0,410725 | 0,532786 | 0,816379 | -0,01735 | 1 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Нижние 95%* | *Верхние 95%* | *Нижние 95,0%* | *Верхние 95,0%* |
| -45178,2 | 10410,8 | -45178,2 | 10410,8 |
| 24,24046 | 94,01987 | 24,24046 | 94,01987 |
| -141,505 | 686,0612 | -141,505 | 686,0612 |
| -9,82623 | -1,28239 | -9,82623 | -1,28239 |
| -698,294 | 355,6697 | -698,294 | 355,6697 |

Из корреляционной матрицы видна достаточно сильная взаимосвязь между результативным (У) и факторными признаками (Х1, Х2, Х4, Х8). Связь очень сильная.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ВЫВОД ИТОГОВ | |  |  |  |
|  |  |  |  |  |
| *Регрессионная статистика* | |  |  |  |
| Множественный R | 0,910875 |  |  |  |
| R-квадрат | 0,829692 |  |  |  |
| Нормированный R-квадрат | 0,772923 |  |  |  |
| Стандартная ошибка | 1451,887 |  |  |  |
| Наблюдения | 17 |  |  |  |
|  |  |  |  |  |
| Дисперсионный анализ | | |  |  |
|  | *df* | *SS* | *MS* | *F* |
| Регрессия | 4 | 1,23E+08 | 30808467 | 14,61519 |
| Остаток | 12 | 25295703 | 2107975 |  |
| Итого | 16 | 1,49E+08 |  |  |
|  |  |  |  |  |
|  | *Коэффициенты* | *Стандартная ошибка* | *t-статистика* | *P-Значение* |
| Y-пересечение | -17383,7 | 12756,71 | -2,36271 | 0,00198001 |
| Х1 | 59,13016 | 16,01317 | 3,692595 | 0,00003077 |
| Х2 | 272,278 | 189,9122 | 3,433704 | 0,00177196 |
| Х4 | -5,55431 | 1,960663 | -2,83287 | 0,00015095 |
| Х8 | 171,312 | 241,8665 | 2,70829 | 0,00492294 |

Проведем регрессионный анализ. По результатам регрессионного анализа получено следующее уравнение регрессии:

 (6.19)

(-2,36) (3,69) (3,43) (-2,83) (0,70)

В скобках указаны значения t-критерия Стьюдента.

В результате построения уравнения регрессии получили следующие результаты (таблица 6.5).

Таблица 6.5 – Результаты построения регрессии

|  |  |
| --- | --- |
| Показатели | Значения |
| Коэффициент корреляции R | 0,910 |
| Коэффициент детерминации R2 | 0,829 |
| Скорректированный коэффициент детерминации R2 | 0,773 |
| Фактическое значении F-критерия Фишера | 14,61 |
| Табличное значении F-критерия Фишера | 2,79 |
| Стандартная ошибка | 5,11 |

Множественный коэффициент регрессии равен 0,910. Это свидетельствует о высокой связи между признаками. Коэффициент детерминации – равен 0,829, следовательно, 82,9% вариации уровня инвестиций в строительство Оренбургской области обусловлено факторами, включенными в модель (6.19).

Анализ полученного уравнения позволяет сделать выводы о том, что с ростом количества строительных организаций – уровень инвестиций в строительство возрастает на 59,13 тыс.руб., с ростом среднесписочной численности работников строительных организаций Оренбургской области на 1 тыс.руб. - уровень инвестиций в строительство увеличивается на 272,3 тыс.руб., с ростом среднего уровня использования производственных мощностей строительных организаций на 1% - уровень инвестиций в строительство увеличивается на 5,55 тыс.тыс.руб., с ростом сальдированного финансового результата на 1 млн.руб. - уровень инвестиций в строительство увеличивается на 171,3 тыс.руб.

Проверка адекватности модели, построенной на основе уравнений регрессии, начинается с проверки значимости каждого коэффициента регрессии. Значимость коэффициента регрессии осуществляется с помощью t-критерия Стьюдента:

 (6.20)

Параметры уравнения все значимы, кроме параметра при факторе времени, так как их расчетные значения меньше табличных ( , уровень значимости = 0,05, )

Проверка адекватности всей модели осуществляется с помощью расчета F-критерия. Если Fp>Fт при α=0,05, то модель в целом адекватна изучаемому явлению.



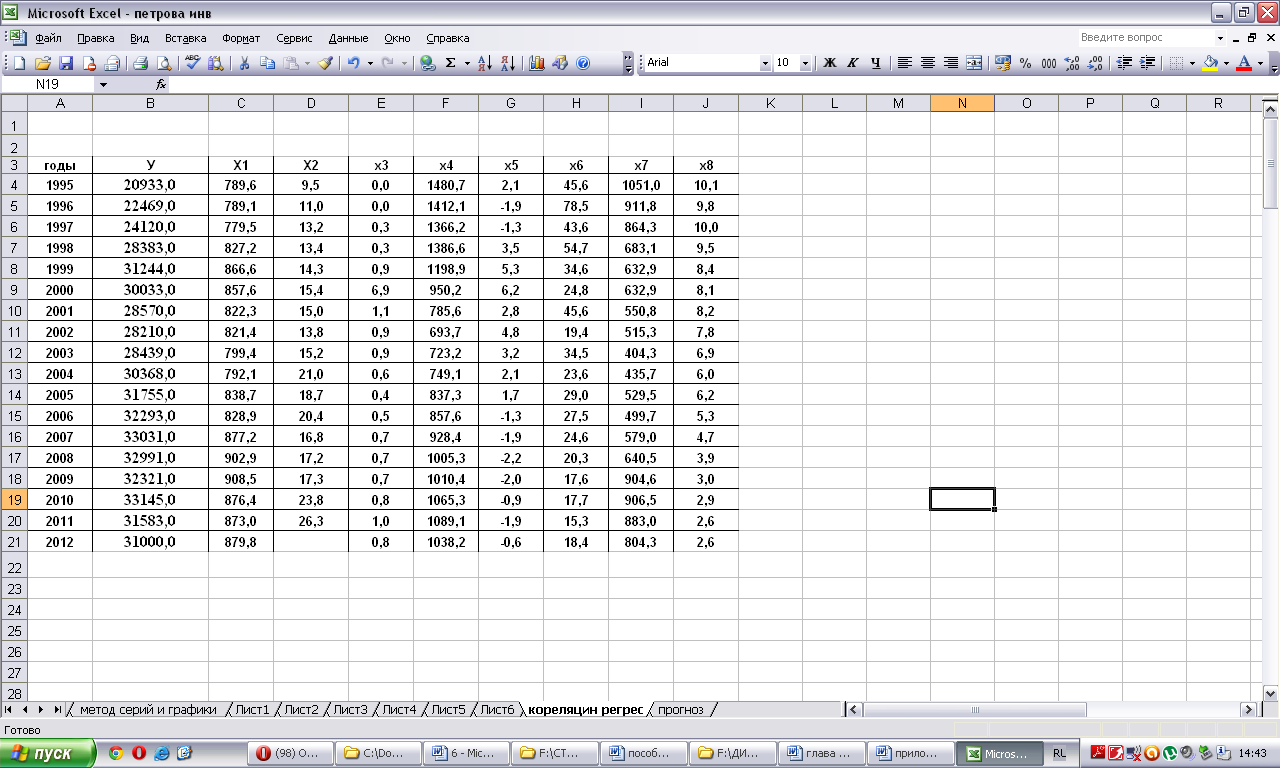
Следовательно, построенная модель на основе её проверки по F-критерию Фишера в целом адекватна, и все коэффициенты регрессии значимы. Такая модель может быть использована для принятия решений и осуществления прогнозов.

**Пример 6.2. Практика построения уравнения множественной регрессии в Excel**

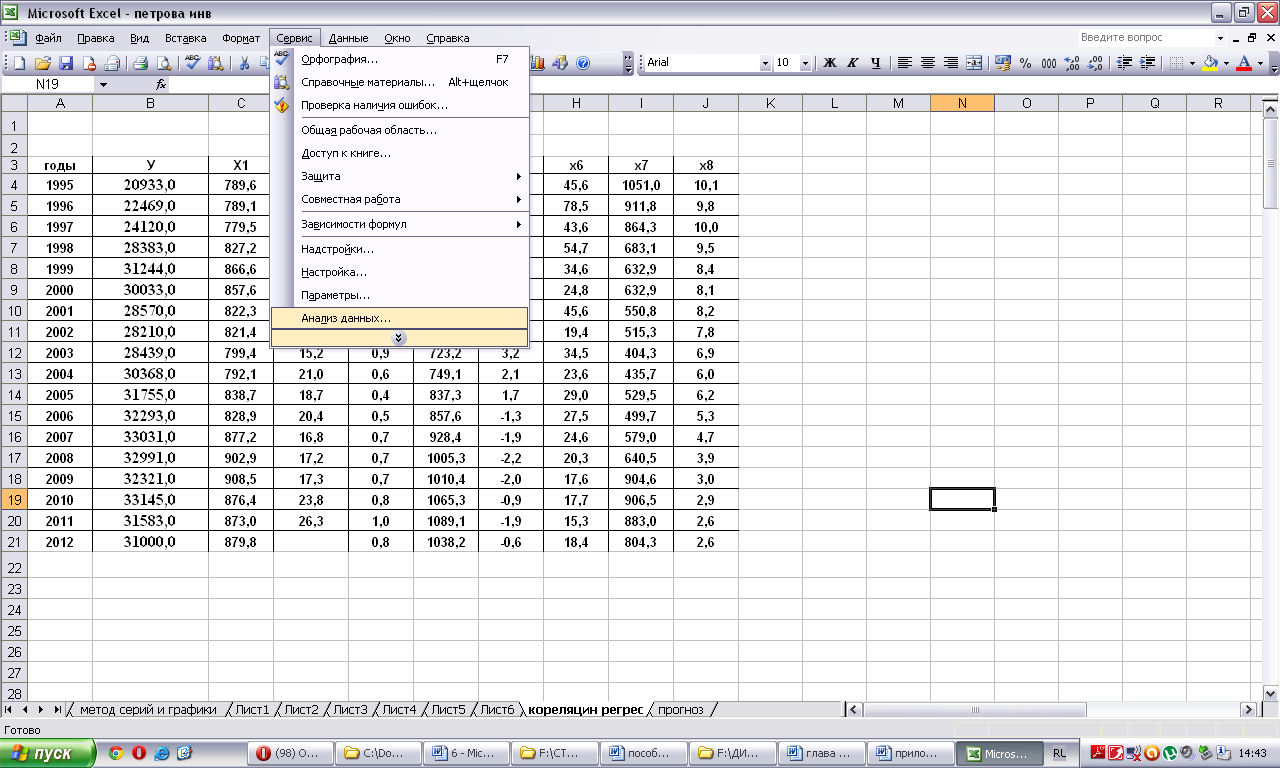
На основании имеющихся данных проведем корреляционно-регрессионный анализ.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| У - | численность умерших на 1000 человек населения; | | |  |  |  |  |  |  |  |  |
| Х1 - | уровень заболеваемости населения; |  |  |  |  |  |  |  |  |  |  |
| Х2 - | число зарегистрированных преступлений по Оренбургской области на 100 000 человек; | | |  |  |  |  |  |  |  |  |
| Х3 - | уровень безработицы, %; | |  |  |  |  |  |  |  |  |  |
| Х4 - | численность больных состоящих на учете в лечебно профилактических  учреждениях с диагнозом алкоголизм и алкогольный психоз на  100 000 человек; | | | | | | | | |  |  |
| Х5 - | прирост миграции на 1000 человек; | | |  |  |  |  |  |  |  |  |
| Х6 - | численность населения с денежными доходами ниже величины  прожиточного минимума, в % от общей численности населения; | | | | | | |  |  |  |  |
| Х7- | выбросы в окружающую среду; |  |  |  |  |  |  |  |  |  |  |
| Х8 - | Численность пострадавших при несчастных случаях на производстве с  утратой трудоспособности на один рабочий день и более и со  смертельным исходом на 1000 работающих | | | | | | | | | | |

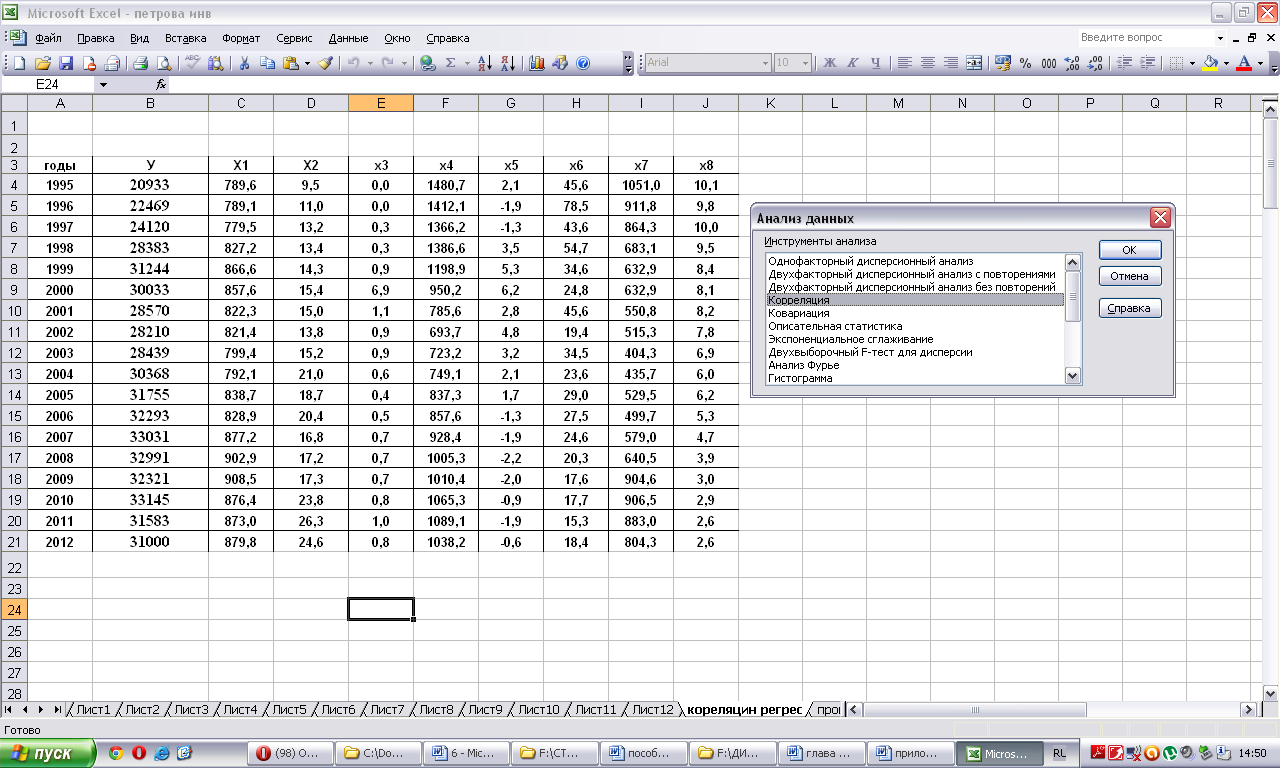
1. Забиваем исходные цифровые данные в файл Excel.



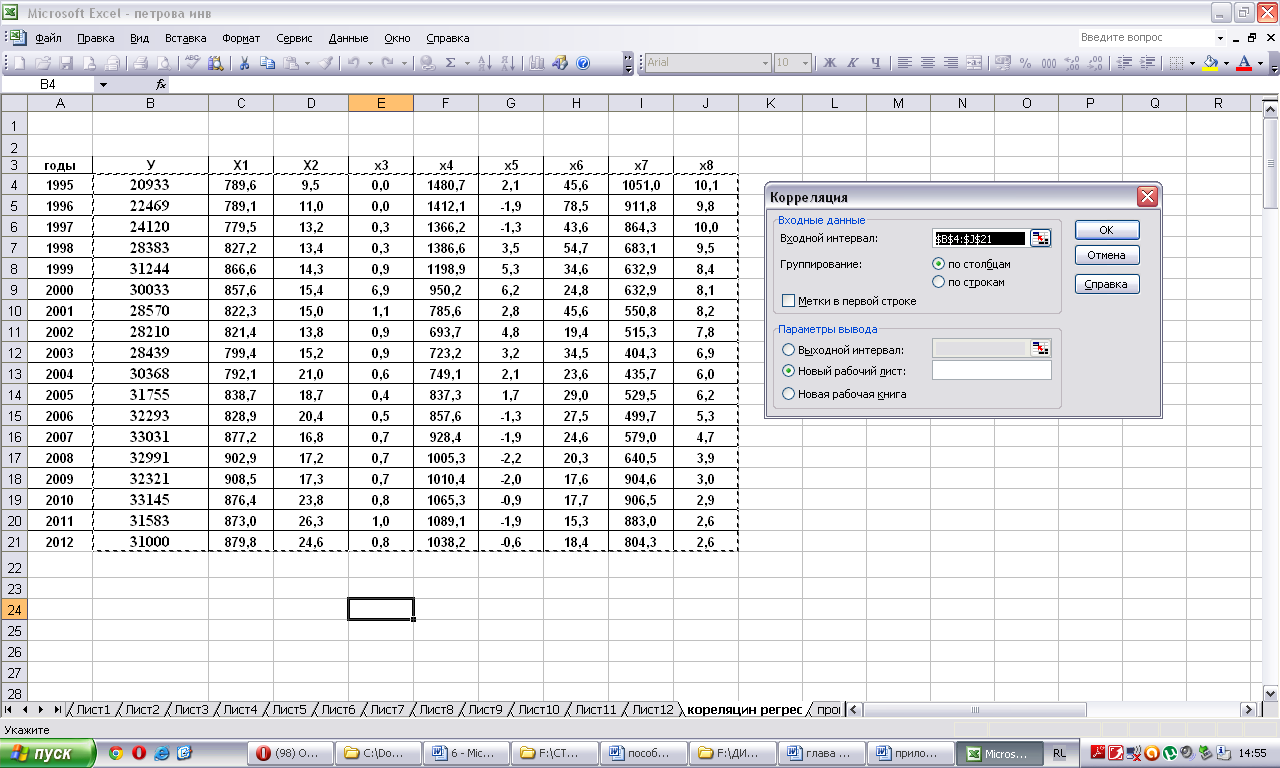
2. Выбираем на панели инструментов закладку Сервис → Анализ данных.



3. В окне Анализа данных выбираем закладку «Корреляция».

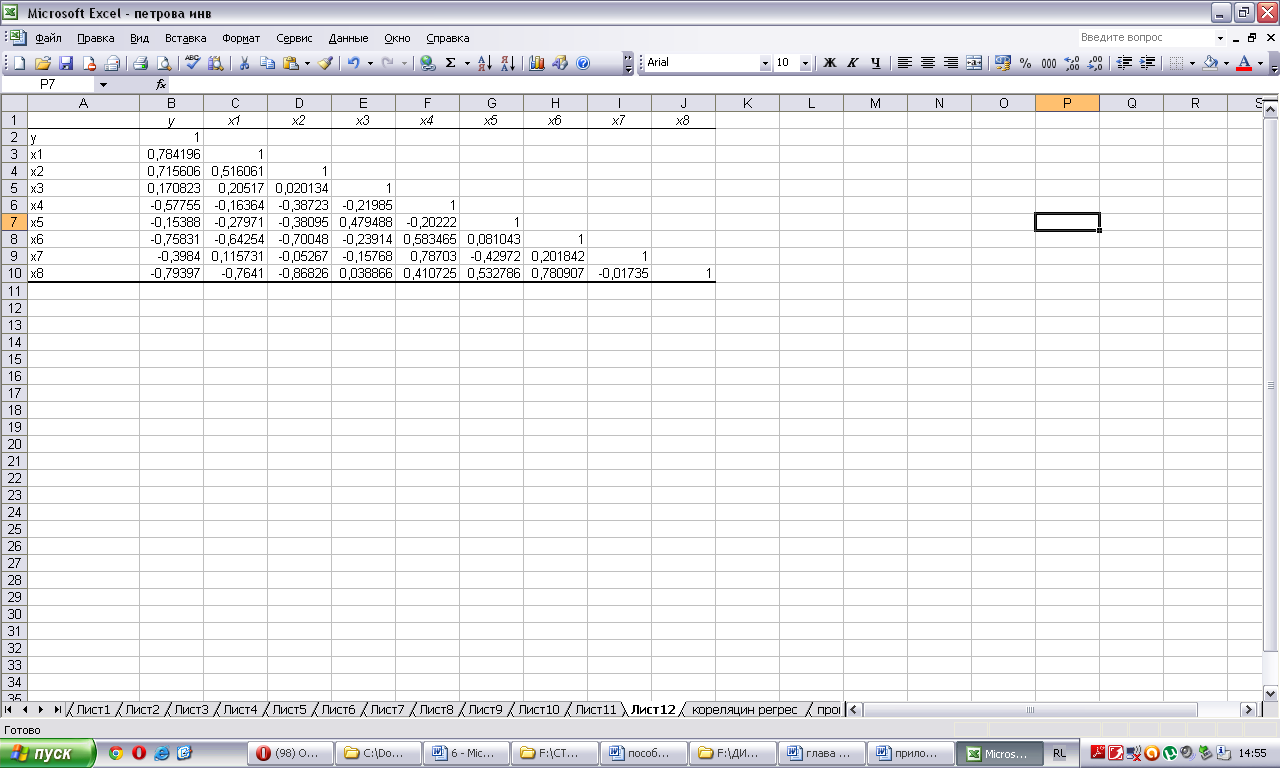


4. В появившемся окне «Корреляция» выбираем входной интервал, ставим галочку в способе группировки (по столбцам или по строкам) и выбираем входной интервал, т.е. то место, где будут отображаться полученные данные (на новом рабочем листе или в конкретно указанном месте).

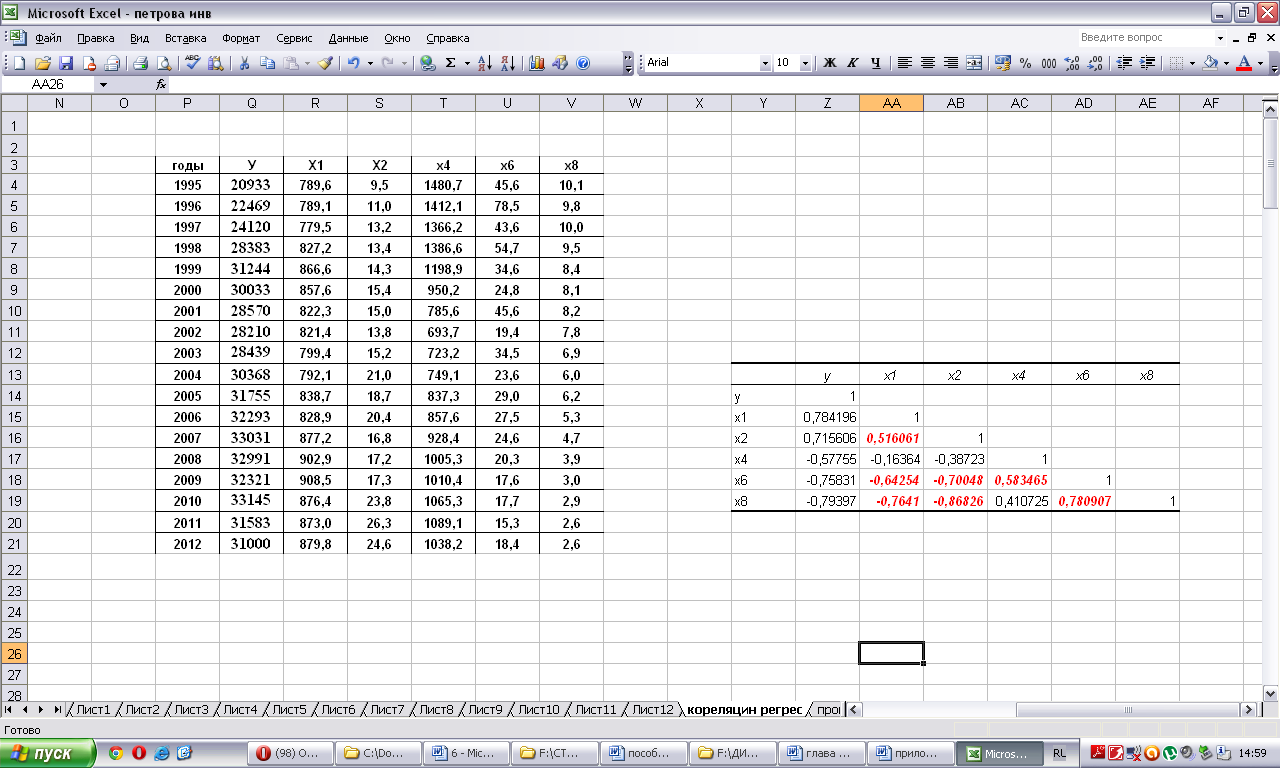


Во входном интервале выбираем цифровые данные у и всех х

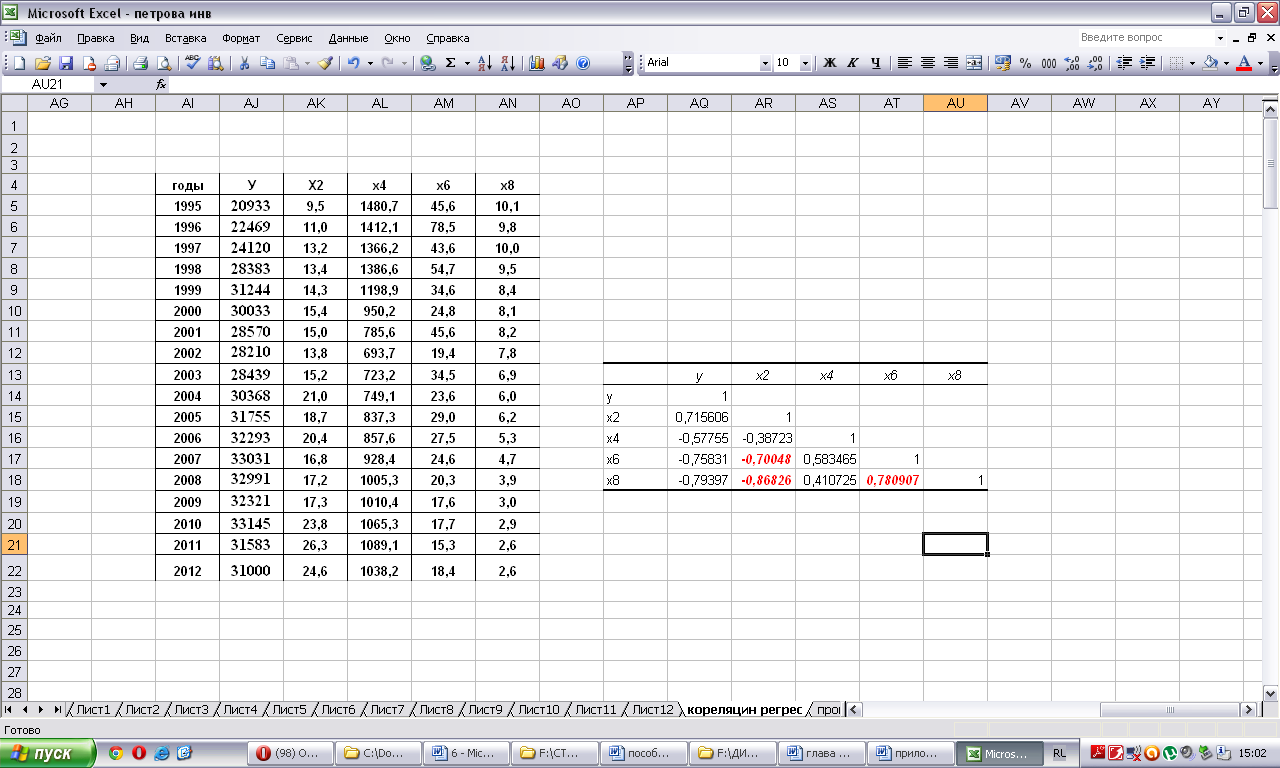
5. Т.к. мы выбрали новый рабочий лист, корреляционная матрица появилась на новом листе. По значениям корреляционной матрицы определяем, какие факторные признаки следует исключить, а какие оставить. Рекомендательно, оставляют те факторные признаки, значения коэффициентов корреляции у которых больше 0,5. В нашем случае в модель могут быть включены Х1, Х2, Х4, Х6, Х8.



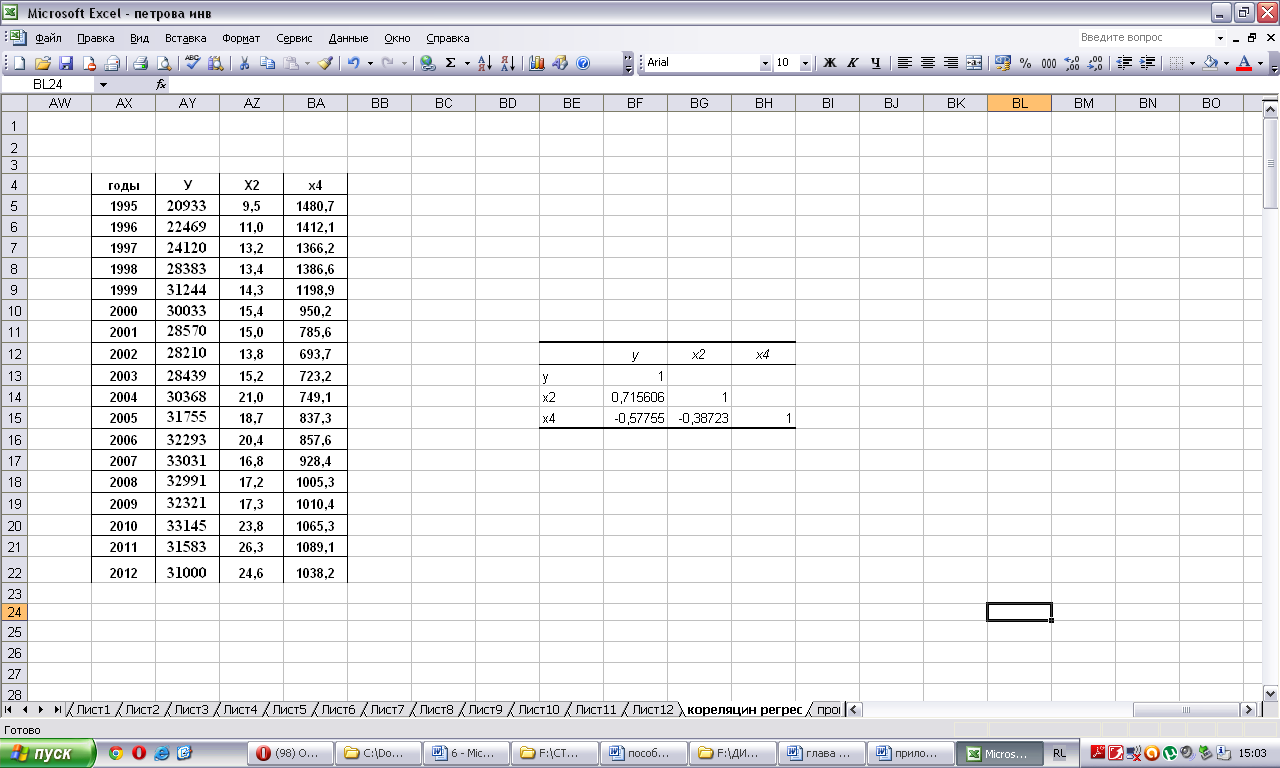
6. Аналогичную процедуру проделываем для оставшихся факторов. По корреляционной матрице проверяем мультиколлинеарность факторов (т.е. есть ли взаимосвязь между самими факторами), в случае, если существует тесная связь между факторами (r ≥ 0,5), то включать их в одну модель нельзя. Из приведенного ниже примера видно, что существует тесная связь между Х1 и Х2, Х6, Х8; Х2 и Х6,Х8; Х4 и Х6; Х6 и Х8. Принимаем решение исключить Х1.



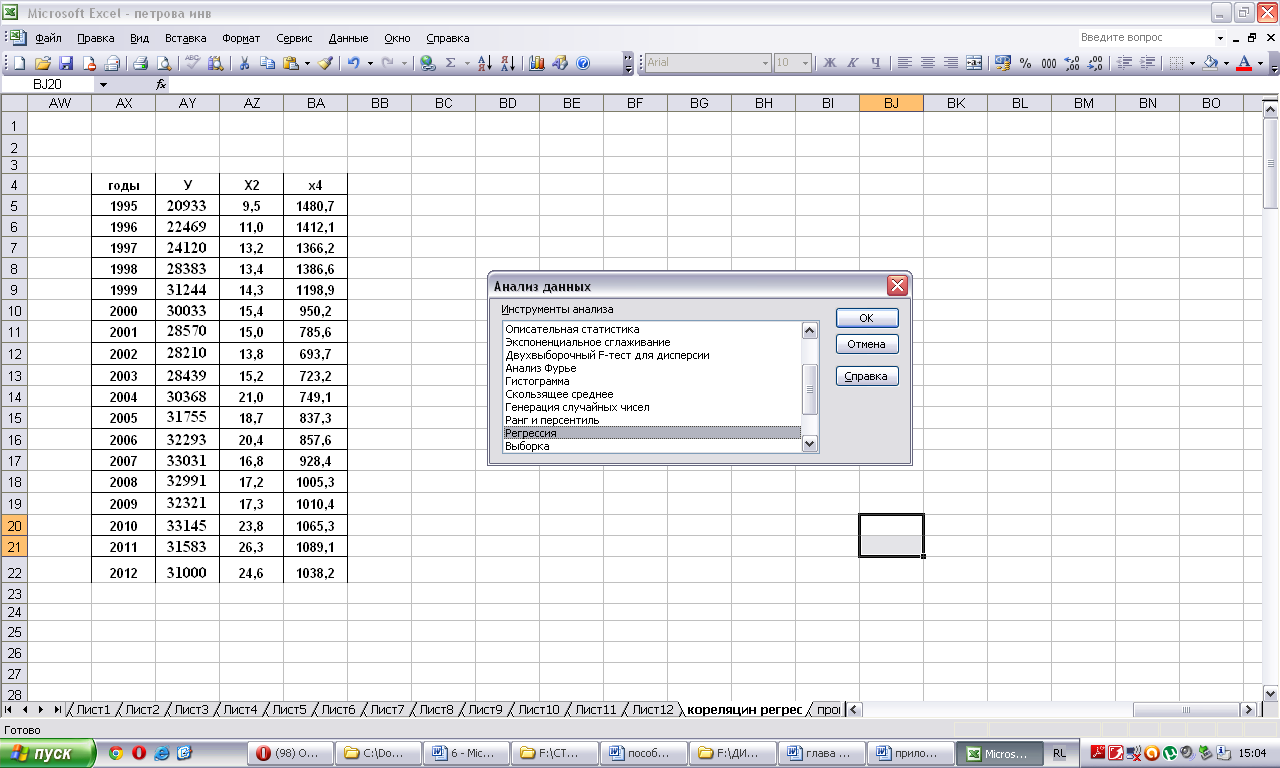
7. Проводим корреляционный анализ. По результатам корреляционной матрицы видно, что существует связь между Х2 и Х6, Х8; Х6 и Х8. Исключать и подбирать факторы необходимо до тех пор, пока не будет устранена мультиколлинеарность.



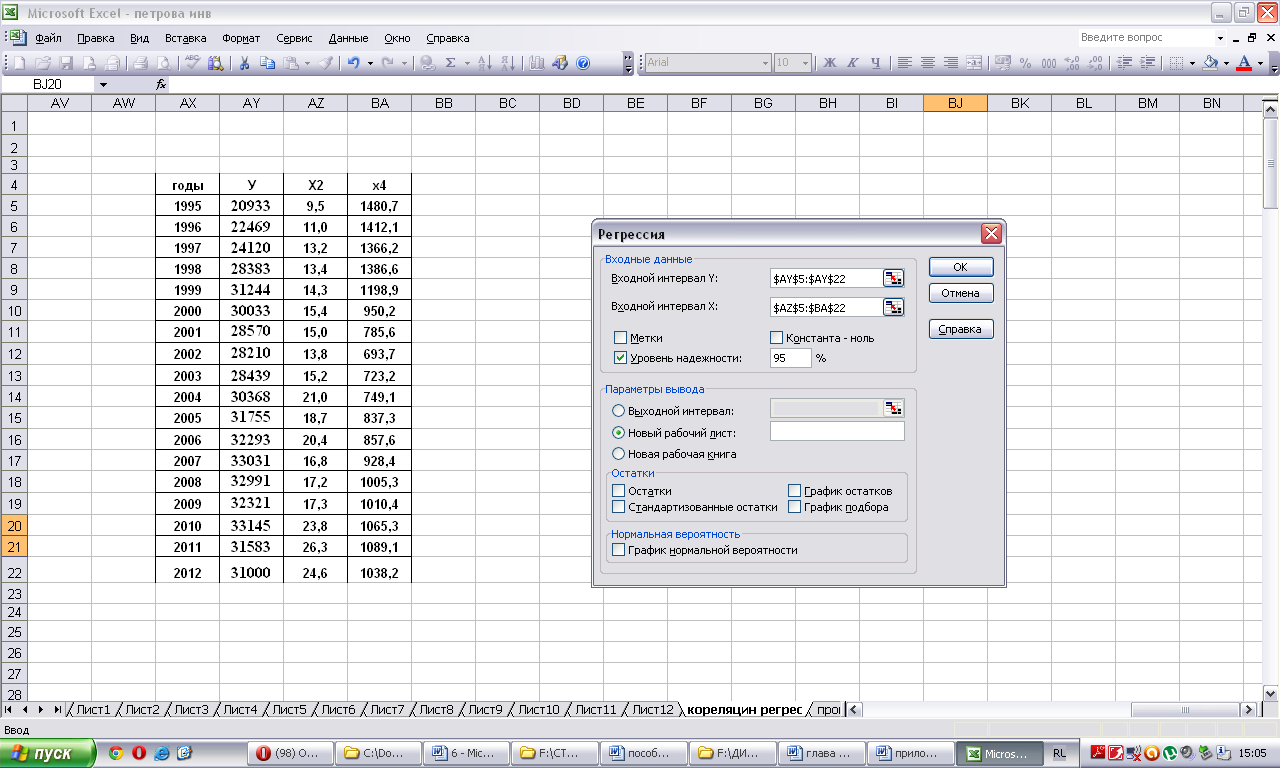
8. После устранения мультиколлинеарности между факторами и получения корреляционной матрицы со значимыми коэффициентами корреляции переходим к построению уравнения регрессии.



9. Выбираем на панели инструментов закладку Сервис → Анализ данных → Регрессия.



10. В окне «Регрессия» выделяем входной интервал для У и входной интервал для Х (выбираем только числовые данные); устанавливаем основные настройки и нажимаем ОК.



Выбираем ячейки c цифровыми значениями х2 и х4

Выбираем ячейки AY5-AY22

1. Получаем результаты регрессионного анализа.

Уравнение регрессии выглядит следующим образом:

*у = 26104,9+453,6Х2-5,2Х4*

