

Лабораторная работа

Анализ зависимостей и прогнозирование экономических показателей в MS Excel

Цель работы: изучение зависимостей экономических показателей и выполнение прогнозов с использованием инструментов MS Excel; приобретение практических навыков построения линий тренда, расчета и анализа регрессионных моделей с помощью встроенных функций категории «Статистические».

Теоретические сведения

Большинство явлений и процессов в экономике находятся в постоянной взаимной и всеохватывающей связи. Исследование зависимостей и взаимосвязей между существующими явлениями и процессами играет большую роль в экономике. Оно дает возможность глубже понять сложный механизм причинно-следственных отношений между явлениями. Для исследования интенсивности, вида и формы зависимостей широко применяется регрессионный анализ, который является методическим инструментарием при решении задач прогнозирования, планирования и анализа хозяйственной деятельности предприятий.

В регрессионном анализе рассматривается односторонняя зависимость случайной переменной Y от одной (или нескольких) неслучайной независимой переменной X . Односторонняя статистическая зависимость выражается с помощью функции, которая называется *регрессией*.

Относительно формы зависимости различают *линейную* и *нелинейную* регрессию (в последнем случае возможно дальнейшее уточнение: квадратичная, экспоненциальная, степенная и т. д.). В зависимости от числа переменных различают *парную* регрессию (регрессию между двумя переменными) и *множественную* регрессию (регрессию между зависимой переменной и несколькими независимыми переменными).

В MS Excel операции регрессионного анализа позволяют реализовать встроенные функции категории *Статистические*. К их числу относятся такие функции, как ЛИНЕЙН, НАКЛОН, ОТРЕЗОК, ТЕНДЕНЦИЯ, ЛГРФПРИБЛ, ПРЕДСКАЗ и другие. Они обеспечивают

расчет и оценку регрессий *линейной* и *экспоненциальной* структуры, которые могут быть представлены уравнениями вида:

- $y = a_0 + a_1 x$ – линейная парная регрессия;
- $y = a_0 + a_1 x_1 + a_2 x_2 + \dots + a_m x_m$ – линейная множественная регрессия;
- $y = a_0 a_1^x$ – экспоненциальная (показательная) парная регрессия;
- $y = a_0 a_1^{x_1} \dots a_m^{x_m}$ – экспоненциальная множественная регрессия.

Линейные регрессии являются наиболее распространенным и простым видом зависимости между экономическими переменными.

Функции для расчета линейных регрессий:

1. **ЛИНЕЙН** (известные_значения_y, известные_значения_x, конст; статистика) – рассчитывает массив данных, описывающих уравнение линейной множественной (или парной) регрессии на основе метода наименьших квадратов.

Аргументы функции:

известные_значения_y – множество наблюдаемых значений Y ;

известные_значения_x – множество наблюдаемых значений $X_1, X_2, \dots, X_m$;

конст – логическое значение, которое указывает, должен ли коэффициент a_0 быть равным 0. Если аргумент *конст*=1 (или опущен), то a_0 вычисляется обычным образом, если аргумент *конст*=0, то a_0 полагается равным 0.

статистика – логическое значение, которое указывает, требуется ли рассчитывать дополнительные статистические характеристики по регрессии. Если аргумент *статистика* = 0 (или опущен), то функция вычисляет только коэффициенты уравнения регрессии. Если аргумент *статистика* = 1, то функция рассчитывает дополнительную регрессионную статистику. При этом возвращаемый массив будет иметь такой вид, как показано в табл. 1.

Таблица 1 – Результаты, возвращаемые функцией **ЛИНЕЙН**

a_m	a_{m-1}	...	a_2	a_1	a_0
se_m	se_{m-1}	...	se_2	se_1	se_0
R^2	se_y				
F	d_f				
$SS_{req.}$	$SS_{resid.}$				

Здесь $a_0, a_1, a_2, \dots, a_m$ – коэффициенты уравнения регрессии. Все остальные показатели – это дополнительная статистика по регрессии:

$se_0, se_1, se_2, \dots, se_m$ – стандартные ошибки для коэффициентов $a_0, a_1, a_2, \dots, a_m$;

R^2 – коэффициент детерминации, который показывает, как близко уравнение описывает исходные данные;

S_y – стандартная ошибка для Y ;

F – F -статистика Фишера, которая используется для определения того, является ли наблюдаемая взаимосвязь между зависимой и независимой переменными случайной или нет;

d_f – степень свободы;

$SS_{reg.}$ – регрессионная сумма квадратов;

SS_{resid} – остаточная сумма квадратов.

2. НАКЛОН (известные_значения_u; известные_значения_x) – определяет наклон прямой линии для парной линейной регрессии. В уравнении парной регрессии вида $y = a_0 + a_1 x$ наклон линии определяется коэффициентом a_1 и показывает, насколько изменится в среднем значение признака Y при увеличении X на единицу.

3. ОТРЕЗОК (известные_значения_u; известные_значения_x) – рассчитывает значение, соответствующее точке пересечения линии парной линейной регрессии с осью OY .

4. ТЕНДЕНЦИЯ (известные_значения_u; известные_значения_x; новые_значения_x; конст) – эта функция на основании линейной регрессии вычисляет будущее значение зависимой переменной Y , соответствующее заданному массиву значений X .

5. ПРЕДСКАЗ (X ; известные_значения_u; известные_значения_x) – определяет для парной линейной регрессии прогнозируемое значение Y для фиксированного значения X . Функция ПРЕДСКАЗ является частным случаем функции ТЕНДЕНЦИЯ, когда последняя применяется к парной регрессии.

Функции для расчета экспоненциальных регрессий:

1. ЛГРФПРИБЛ (известные_значения_u; известные_значения_x; конст; статистика) – рассчитывает массив данных, описывающих уравнение экспоненциальной регрессии. Технология работы с функцией ЛГРФПРИБЛ аналогична функции ЛИНЕЙН.

2. РОСТ (известные_значения_u; известные_значения_x; новые_значения_x; конст) – эта функция аналогична функции

ТЕНДЕНЦИЯ и используется для расчета прогнозируемого экспоненциального роста на основании имеющихся данных. Она возвращает значения $\hat{Y}$ для последовательности новых значений X , задаваемых с помощью существующих значений X и $\hat{Y}$.

Другие виды парных регрессий (полиномиальная, логарифмическая, степенная) могут быть получены путем построения графика зависимости изучаемых показателей с последующим подбором подходящей аппроксимирующей функции – *линии тренда*.

MS Excel предоставляет следующие виды линий трендов:

- *линейная* – $y = a_0 + a_1 x$, отражающая рост и убывание данных с постоянной скоростью.

- *полиномиальная* – $y = a_0 + a_1 x + a_2 x^2 + \dots + a_k x^k$ ($k \leq 6$) – описывает попеременно возрастающие и убывающие данные.

- *логарифмическая* – $y = a_0 + a_1 \ln x$ – описывает быстро возрастающие (убывающие) данные, которые затем стабилизируются.

- *степенная* – $y = a_0 x^{a_1}$ – отражает данные с постоянно увеличивающейся (убывающей) скоростью роста.

- *экспоненциальная* – $y = a_0 e^{a_1 x}$, (e – основание натурального логарифма).

Степень точности аппроксимации экспериментальных данных выбранной функцией оценивается коэффициентом детерминации R^2 . Чем ближе этот коэффициент к 1, тем точнее приближение.

Задание 1

1. Имеются данные о прибыли предприятий, величине оборотных средств и стоимости основных фондов (рис. 1).

	A	B	C	D
	Номер предприятия	Прибыль Y , млн. руб.	Величина оборотных средств X_1 , млн. руб.	Стоимость основных фондов X_2 , млн. руб.
1				
2	1	18,8	12,9	51,0
3	2	7,8	6,4	19,0
4	3	9,3	6,9	24,0
5	4	15,2	8,7	47,0
6	5	5,5	4,7	11,0
7	6	16,1	10,2	42,0

Рисунок 1 – Исходные данные

2. Проведите анализ зависимости прибыли предприятий от величины оборотных средств и стоимости основных фондов.

3. Спрогнозируйте прибыль для предприятия, величина оборотных средств которого составляет 9,5 млн. руб. и стоимость основных фондов – 38,0 млн. руб.

Технология выполнения задания

1. Разместите таблицу с исходными данными в ячейках A1:D7 рабочего листа Excel.

2. Используя функцию ЛИНЕЙН, постройте уравнение, выражающее зависимость прибыли предприятий (Y) от величины оборотных средств (X_1) и стоимости основных фондов (X_2), в виде $y = a_0 + a_1 x_1 + a_2 x_2$. Для этого:

а) Выделите блок ячеек F2:H6 под массив результатов.

б) Для вызова функции ЛИНЕЙН выполните команду: вкладка *Формулы* → группа *Библиотека функций* → кнопка *Вставить функцию*. В открывшемся окне выберите категорию *Статистические* и в списке статистических функций найдите функцию ЛИНЕЙН.

в) Заполните диалоговое окно *Аргументы функции* (рис. 2.84).

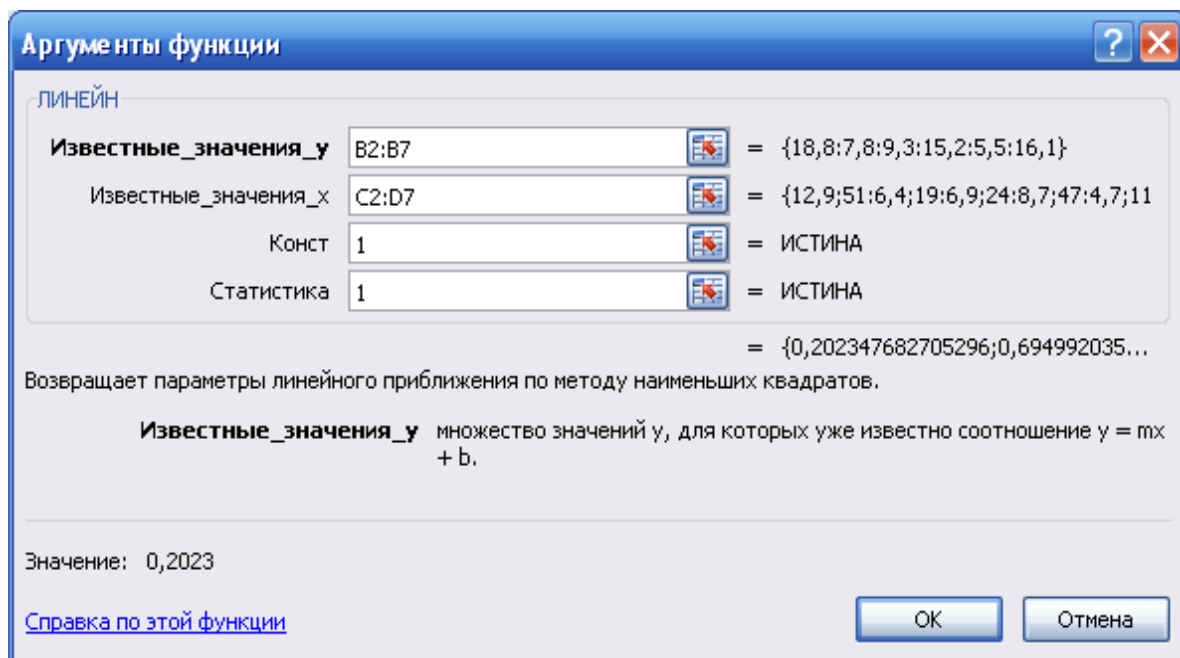


Рисунок 2 –Диалоговое окно функции ЛИНЕЙН

г) После заполнения полей аргументов функции завершите процедуру нажатием комбинации клавиш Ctrl+Shift+Enter.

а) В результате в диапазоне F2:H6 будет выведен массив значений (рис. 3):

■ ячейка F2 – коэффициент $a_2 = 0,202$;

■ ячейка G2 – коэффициент $a_1 = 0,695$;

■ ячейка H2 – коэффициент $a_0 = -0,194$;

■ диапазон ячеек F3:H3 – стандартные ошибки коэффициентов;

■ ячейка F4 – коэффициент детерминации $R^2 = 0,995$.

■ ячейка F5 – критерий Фишера $F = 273,741$.

Таким образом, уравнение зависимости прибыли предприятий от величины оборотных средств и стоимости основных фондов примет вид: $y = -0,194 + 0,695x_1 + 0,202x_2$. Поскольку коэффициент детерминации ($R^2 = 0,99$) близок к 1, уравнение регрессии следует признать адекватным и использовать для прогнозирования.

3. С помощью функции ТЕНДЕНЦИЯ спрогнозируйте прибыль для предприятия, если $x_1 = 9,5$, $x_2 = 38,0$. Для этого:

а) Внесите в ячейку C9 значение 9,5, в ячейку D9 – значение 38.

б) В ячейку B9 с помощью Мастера функций (рис. 4) введите формулу =ТЕНДЕНЦИЯ(B2:B7;C2:D7;C9:D9;1).

	F	G	H
2	0,202	0,695	-0,194
3	0,035	0,197	0,763
4	0,995	0,505	#Н/Д
5	273,741	3,000	#Н/Д
6	139,623	0,765	#Н/Д

Рисунок 3 – Результат вычислений

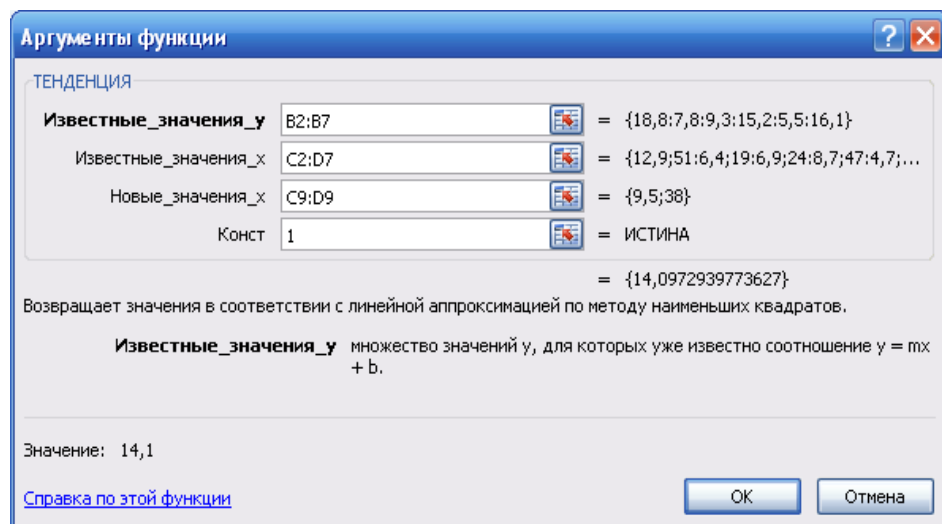


Рисунок 4 – Диалоговое окно функции ТЕНДЕНЦИЯ

В результате прогнозируемое значение прибыли составит 14,1 млн. руб.

Задание 2

Используя графический способ, по данным табл. 2, определите зависимость прибыли от вложенных в производство средств.

Таблица 2 – Прибыль и средства, вложенные в производство

Вложенные средства, тыс. руб.	3000	4000	5000	6000	7000	8000
Прибыль, тыс. руб.	850	900	1000	1300	2000	4000

Технология выполнения задания

1. Введите исходные данные в ячейки диапазона A1:G2 рабочего листа Excel.

2. Выполните графическое изображение зависимости прибыли от вложенных средств, построив *точечную диаграмму* на основе данных диапазона B1:G2 (рис. 5). По расположению точек на диаграмме можно предположить, что имеет место нелинейная зависимость.



Рисунок 5 – Фрагмент рабочего листа с построенной точечной диаграммой

3. Добавьте к диаграмме линию тренда полиномиального вида, выполнив следующую последовательность действий:

а) Щелкните на диаграмме, чтобы выделить ее. Командой: вкладка *Работа с диаграммами / Макет* → группа *Анализ* → кнопка *Линии тренда* → пункт *Дополнительные параметры линии тренда*, откройте диалоговое окно *Формат линия тренда* (рис. 6).

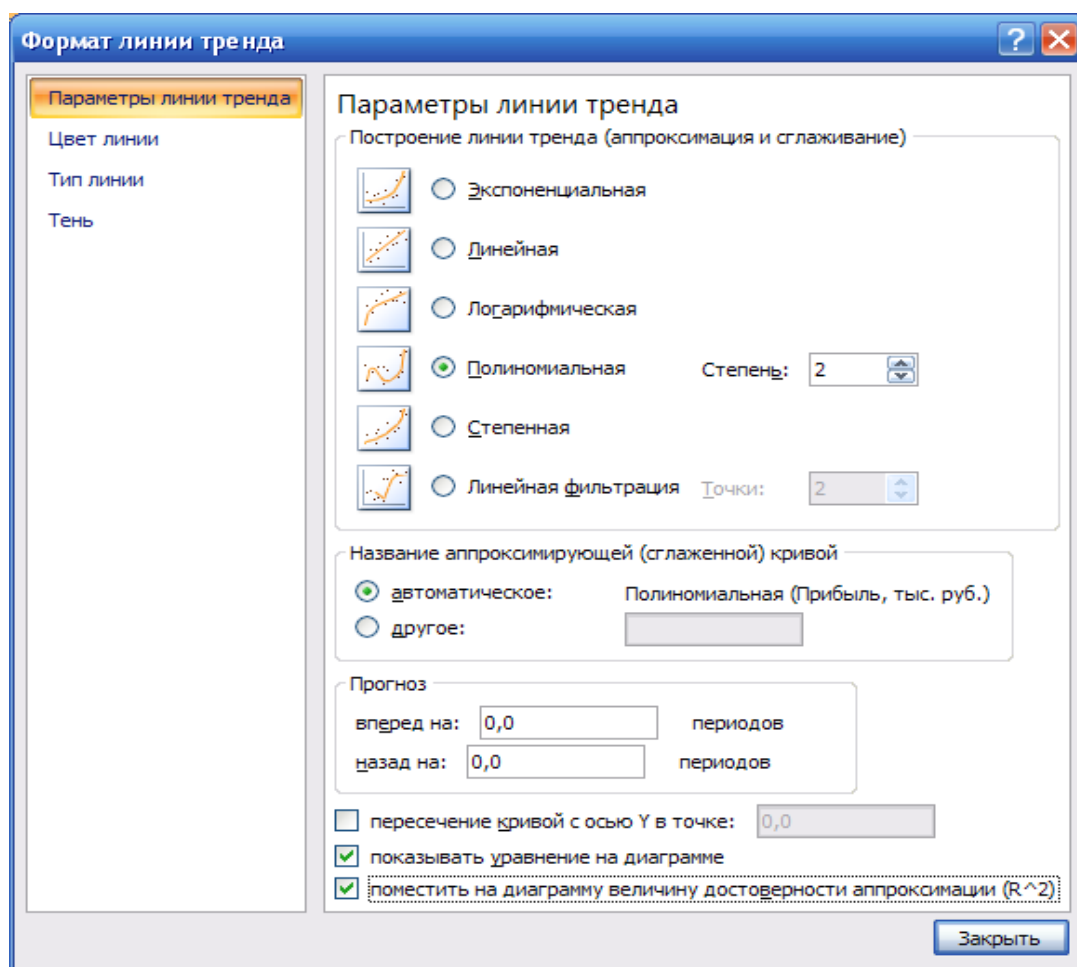


Рисунок 6 – Диалоговое окно *Формат линии тренда*

б) В окне *Формат линии тренда* выберите тип линии тренда – *полиномиальная (степень: 2)* и установите флажки в поля:

- ✓ *Показывать уравнение на диаграмме;*
- ✓ *Поместить на диаграмму величину достоверности аппроксимации (R^2).*

Щелкните на кнопке *Заккрыть*.

в) В результате на диаграмме отобразится линия тренда, уравнение линии тренда и коэффициент детерминации R^2 (рис. 7).

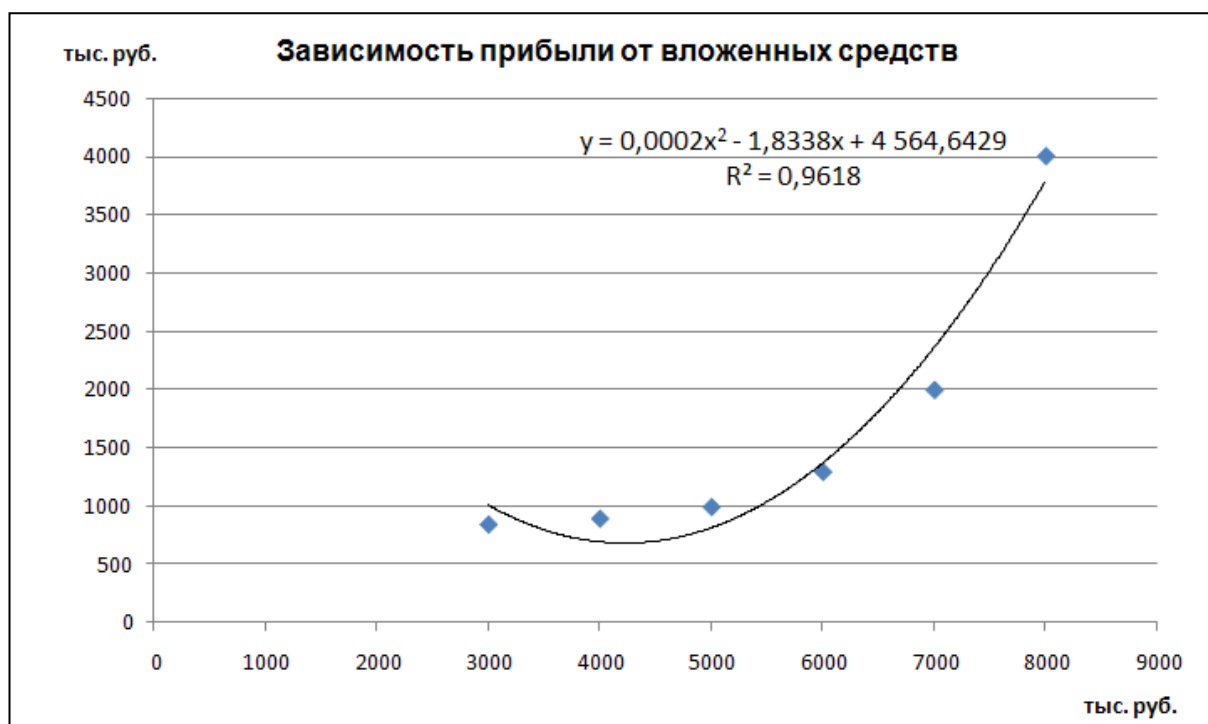


Рисунок 7 – Диаграмма с полиномиальной линией тренда

Задание 3 (для самостоятельного выполнения)

В табл. 3 приведены данные об урожайности зерновых культур и количестве органических удобрений, внесенных на 1 га посева.

Таблица 3 – Урожайность и количество удобрений, внесенных на 1 га посева

Номер участка	1	2	3	4	5	6	7	8	9	10
Внесено удобрений на 1 га, т	5,0	5,5	5,6	6,5	7,0	8,0	7,0	7,3	6,5	7,5
Урожайность, ц/га	13,8	14,0	14,3	14,8	15,1	15,6	15,8	16,1	16,4	16,6

Номер участка	11	12	13	14	15	16	17	18	19	20
Внесено удобрений на 1 га, т	8,3	8,5	6,0	7,5	8,5	6,0	8,0	9,5	10,0	10,9
Урожайность, ц/га	16,8	17,0	17,2	17,6	17,9	18,1	18,3	19,5	20,3	20,8

1. Выполните графическое изображение зависимости урожайности зерновых культур (Y) от количества органических удобрений, внесенных на 1 га их посева (X).

2. Используя построение линии тренда, получите аналитическое выражение функции, наилучшим образом отражающей зависимость между X и Y .

3. Используя полученное уравнение тренда, спрогнозируйте урожайность при внесении 9 т удобрений на 1 га посева.

Краткая справка. Для нахождения наиболее адекватного уравнения тренда необходимо перебрать все возможные линии и выбрать ту, для которой коэффициент детерминации R^2 имеет наибольшее значение. Результаты подбора уравнения можно оформить в виде табл. 2.28.

Таблица 4 – Уравнения трендов

Вид уравнения	Уравнение	Коэффициент R^2
Линейное		
Логарифмическое		
Полином 2-го порядка		
Степенное		
Экспоненциальное		

Задание 4 (для самостоятельного выполнения)

В табл. 5 приведены данные о потребительских расходах в среднем на душу населения и среднедушевых денежных доходах населения по Приволжскому федеральному округу РФ в 2016 г.

Используя функции категории *Статистические*:

1. Получите линейное уравнение зависимости потребительских расходов от среднедушевых денежных доходов.

2. Постройте экспоненциальное уравнение регрессии, оценивающее влияние среднедушевых денежных доходов на величину потребительских расходов.

3. Выберите из полученных уравнений наиболее адекватное уравнение, по которому рассчитайте прогнозное значение потребительских расходов при уровне среднедушевых денежных доходов, равном 25 тыс. руб.

Таблица 5 – Исходные данные

Приволжский федеральный округ	Среднедушевые де- нежные доходы (в месяц), тыс. руб.	Потребительские рас- ходы в среднем на душу населения (в месяц), тыс. руб.
Республика Башкортостан	28,1	22,2
Республика Марий Эл	18,7	12,8
Республика Мордовия	17,7	11,5
Республика Татарстан	32,6	24,1
Удмуртская Республика	23,9	15,8
Чувашская Республика	17,9	13,1
Пермский край	28,4	20,8
Кировская область	21,3	15,6
Нижегородская область	30,6	21,8
Оренбургская область	22,0	15,7
Пензенская область	21,8	16,3
Самарская область	26,8	20,8
Саратовская область	19,4	14,3
Ульяновская область	22,5	15,1

Контрольные вопросы

1. Какие средства для изучения зависимостей экономических показателей и прогнозирования содержит MS Excel?
2. Что такое регрессия? Какие виды регрессий различают?
3. Какие графические средства MS Excel применяются для получения прогнозов?
4. Опишите процедуру построения линии тренда.
5. Какие статистические функции применяется для нахождения коэффициентов линейной регрессии?
6. Какие статистические функции можно использовать для прогнозирования на основе линейной зависимости?
7. В чем различие между функциями ПРЕДСКАЗ и ТЕНДЕНЦИЯ?
8. С помощью какой функции можно найти прогнозные значения, если функция прогнозирования экспоненциально зависит от нескольких факторов?