

Эконометрика
Лабораторная работа №1
Парная линейная регрессия

Порядок выполнения работы

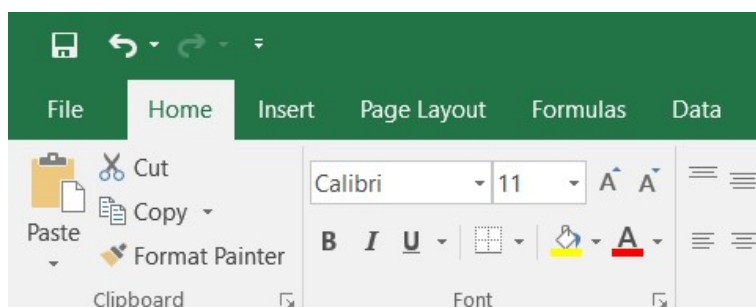
1. Создайте новый документ Excel и сохраните в рабочем каталоге для студентов, выбрав уникальное информативное имя файла.

2. Сгенерируйте исходные данные (по столбцам).

Таблица 1

Обозначение	Содержание	Заполнение
N	Порядковый номер	Целые числа от 1 до 15
XX	Вспомогательный столбец	Равномерно распределённые случайные числа в диапазоне от 150 до 200
X	Рост (см).	Округлить XX до целого
E	Случайная ошибка	Нормально распределённые случайные числа с параметрами: среднее = 0, стандартное отклонение = 1.
YY	Вспомогательный столбец	Вычисляется по формуле: $YY = 0,95 X - 110 + 5 E$
Y	Вес (кг)	Округлить YY до целого

Чтобы включить надстройку «Анализ данных», вызовите File -> Options -> Add-ins -> Manage -> Excel Add-ins -> Go -> Analysis Toolpak -> OK.





Info

New

Open

Save

Save As

Print

Share

Export

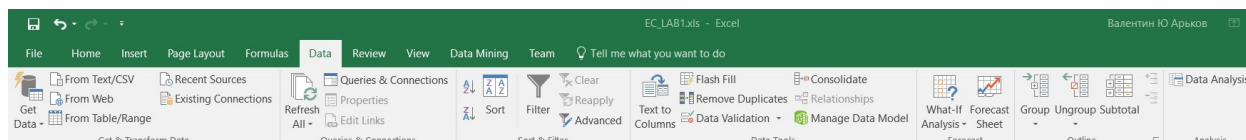
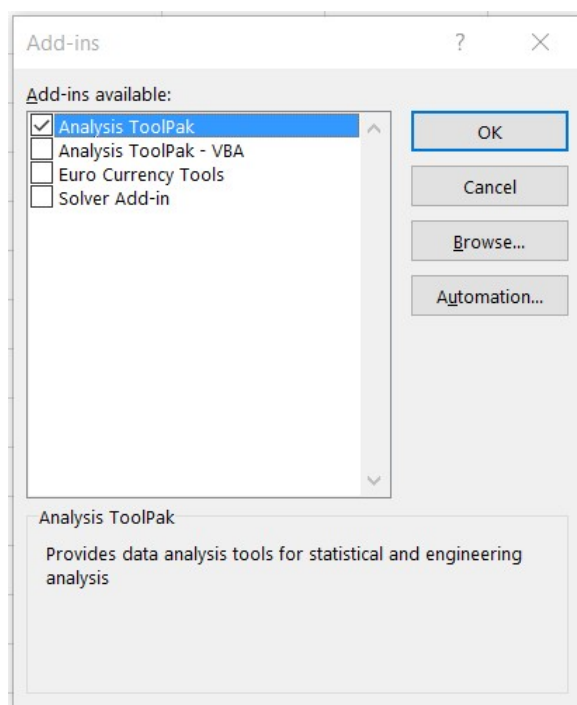
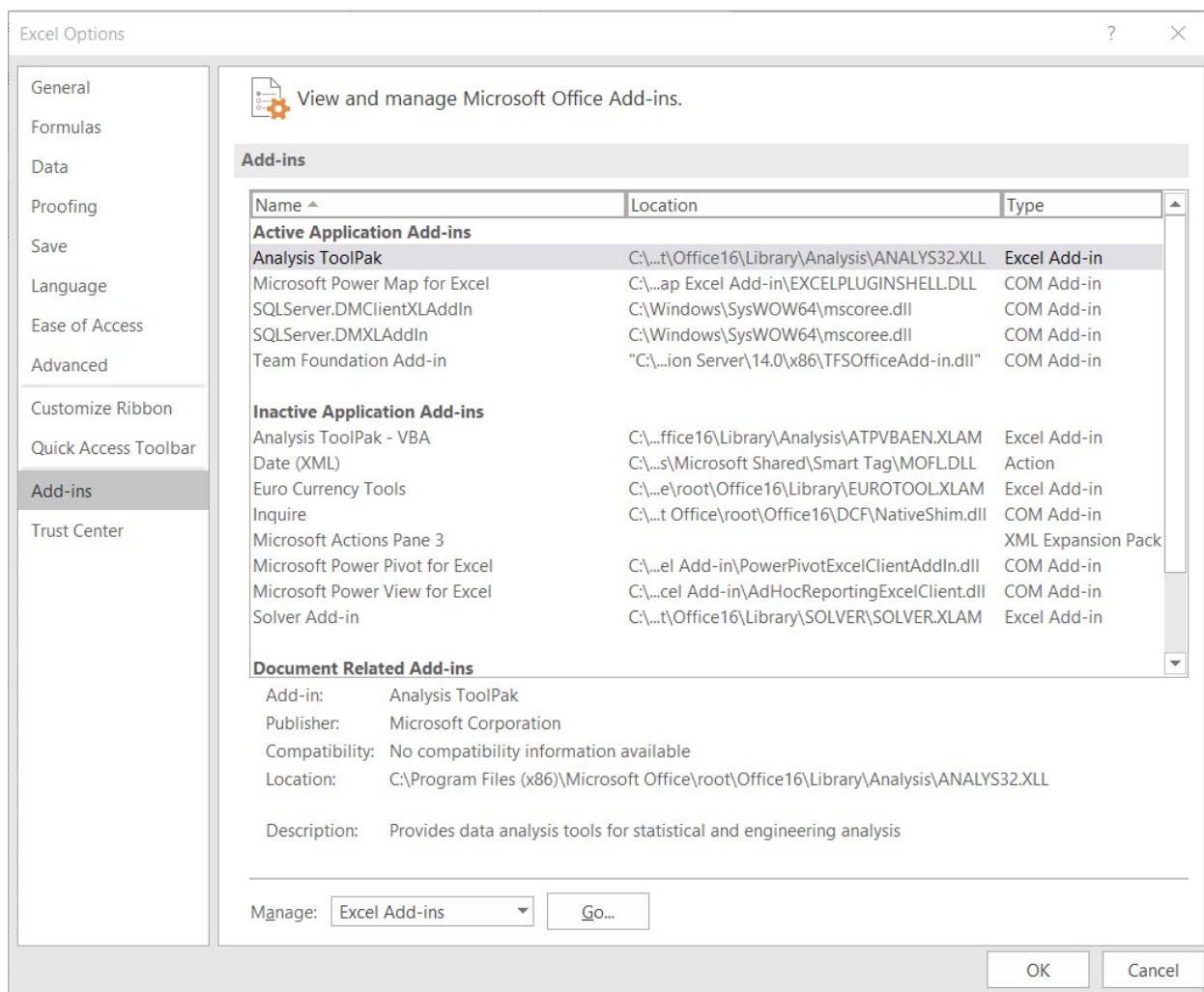
Publish

Close

Account

Feedback

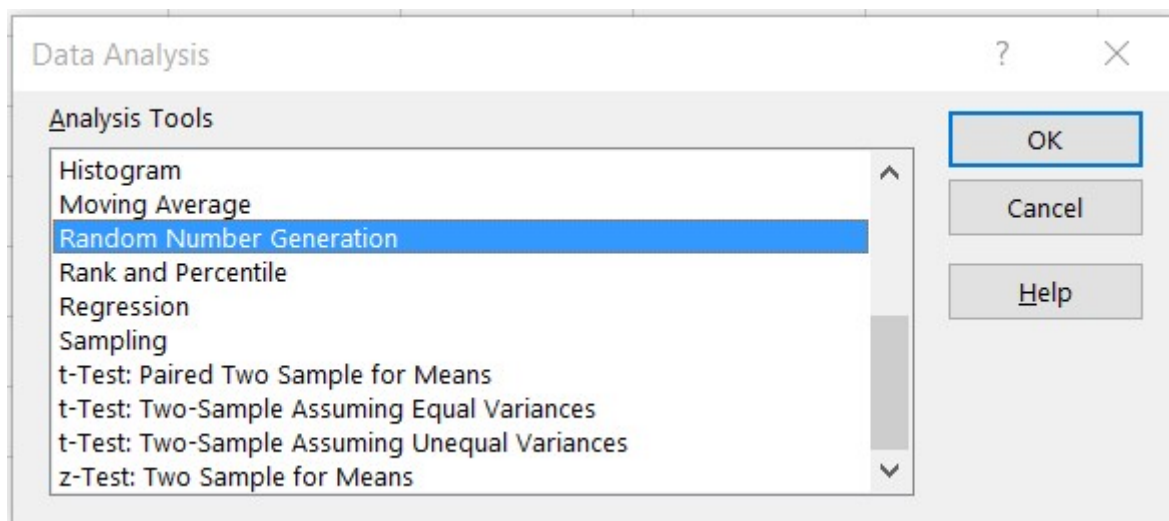
Options



Заполните колонку XX случайными числами с помощью генератора надстройки «Анализ данных».

Вызовите генератор случайных чисел:

Data -> Analysis -> Data Analysis -> Random Number Generation -> OK



Установите следующие параметры генератора:

Number of Variables – Число переменных – число столбцов случайных чисел = 1

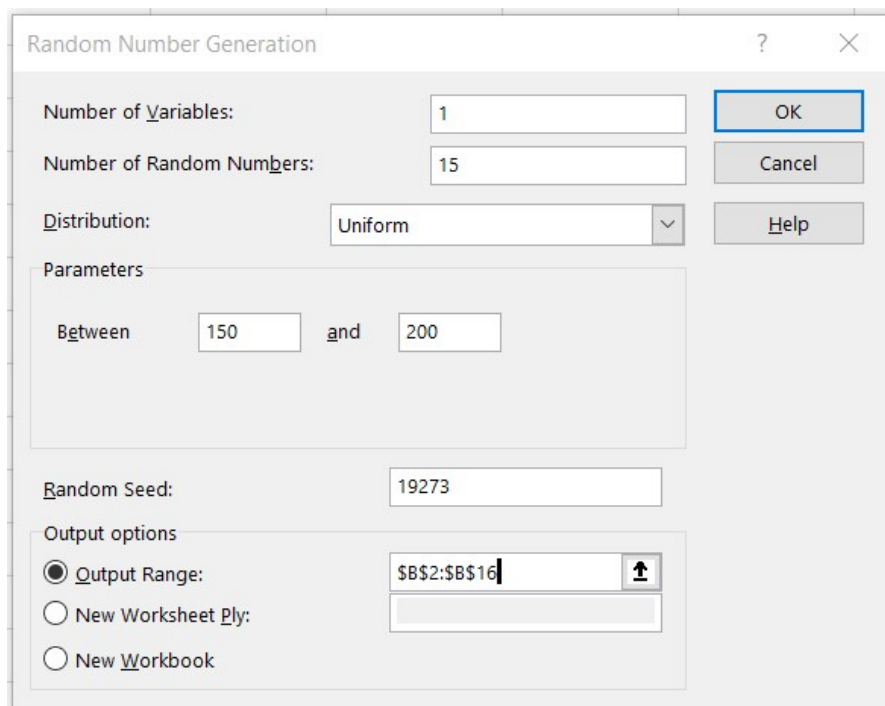
Number of Random Numbers – Количество случайных чисел = 15

Distribution (Распределение) -> Uniform (Равномерное)

Random Seed – начальная настройка генератора: -> введите последние 5 цифр номера зачетки

Output Options (Варианты вывода) -> Output Range – диапазон адресов ячеек для вывода.

Выберите диапазон ячеек в соответствующей колонке электронной таблицы.



Заполните колонку X, округлив значения XX до целого с помощью функции ROUND (ОКРУГЛ).

Заполните колонку E нормально распределёнными случайными числами. В качестве начальной настройки генератора используйте пять первых цифр номера зачётки.

Разные настройки генератора позволят создать независимые случайные числа в разных колонках.

Random Number Generation

Number of Variables: 1

Number of Random Numbers: 15

Distribution: Normal

Parameters

Mean = 0

Standard deviation = 1

Random Seed: 15026

Output options

☒ Output Range: \$D\$2:\$D\$16

☐ New Worksheet Ply:

☐ New Workbook

OK Cancel Help

Заполните колонку Y, округлив значения YY до целого с помощью функции ROUND (ОКРУГЛ).

В полученной таблице скройте вспомогательные столбцы. Укажите на заголовок столбца и нажмите правую кнопку мыши. Выберите Hide (Скрыть).

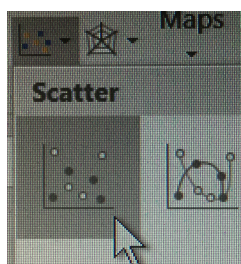
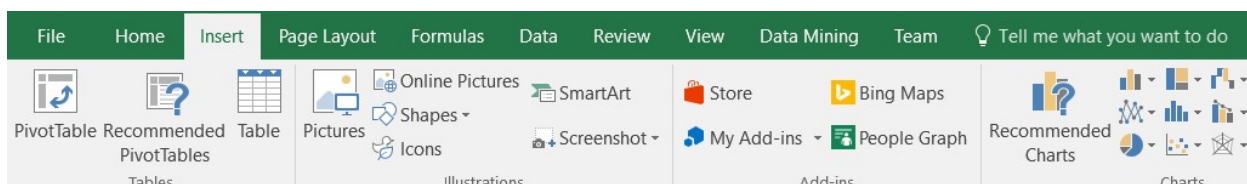
	A	B	C	D	E	F
1	N	tmp_XX	X	tmp_E	tmp_YY	Y
2	1	196.0952	196	-0.0034	76.1830	76
3	2	153.7523	154	2.5097	48.8483	49
4	3	186.6802	187	-0.2759	66.2704	66
5	4	164.0049	164	-0.0541	45.5296	46
6	5	197.2427	197	-1.3294	70.5031	71
7	6	196.3118	196	2.3271	87.8355	88
8	7	194.5463	195	-0.0203	75.1484	75
9	8	163.1901	163	0.4194	46.9468	47
10	9	193.1562	193	0.9400	78.0498	78
11	10	192.7122	193	0.5070	75.8851	76
12	11	159.2669	159	-0.3713	39.1935	39
13	12	188.2153	188	-1.5941	60.6293	61
14	13	167.6473	168	1.9456	59.3282	59
15	14	199.1577	199	-0.1400	78.3502	78
16	15	164.3132	164	-3.0792	30.4041	30

	A	C	F
1	N	X	Y
2	1	196	76
3	2	154	49
4	3	187	66
5	4	164	46
6	5	197	71
7	6	196	88
8	7	195	75
9	8	163	47
10	9	193	78
11	10	193	76
12	11	159	39
13	12	188	61
14	13	168	59
15	14	199	78
16	15	164	30

3. Постройте диаграмму разброса Y(X)

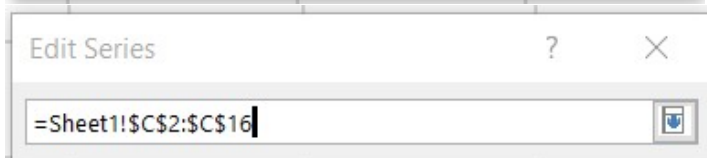
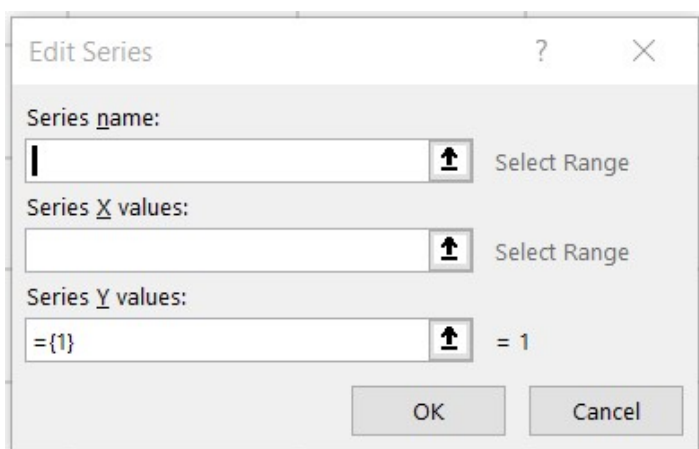
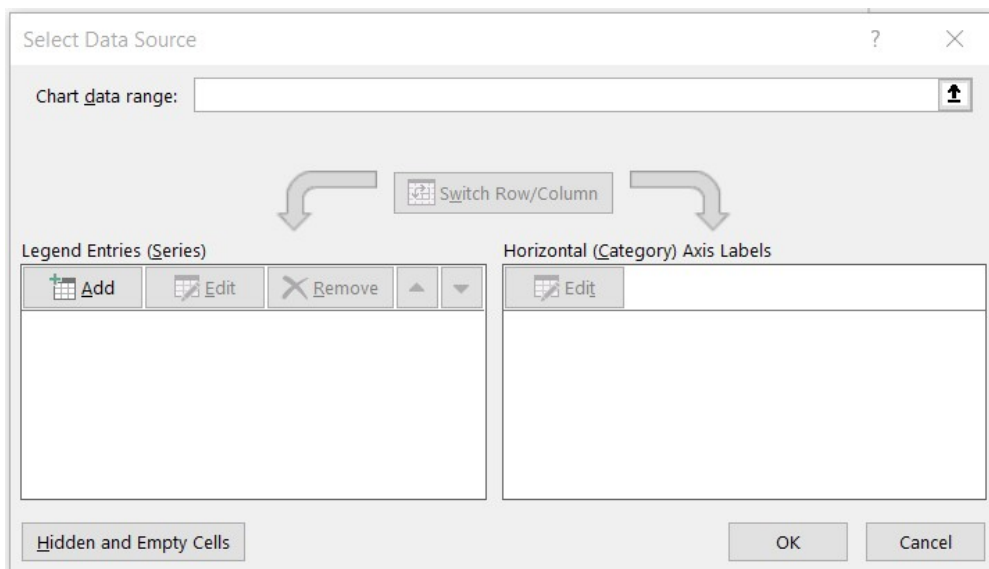
Выберите ячейку для вывода графика.

Вызовите Insert -> Charts -> Scatter -> Scatter



Щёлкните по графику правой кнопкой мыши и выберите Select Data (Выбрать данные).

В диалоговом окне выберите Legend Entries / Series (Легенда графика либо ряды данных) -> Add (Добавить).

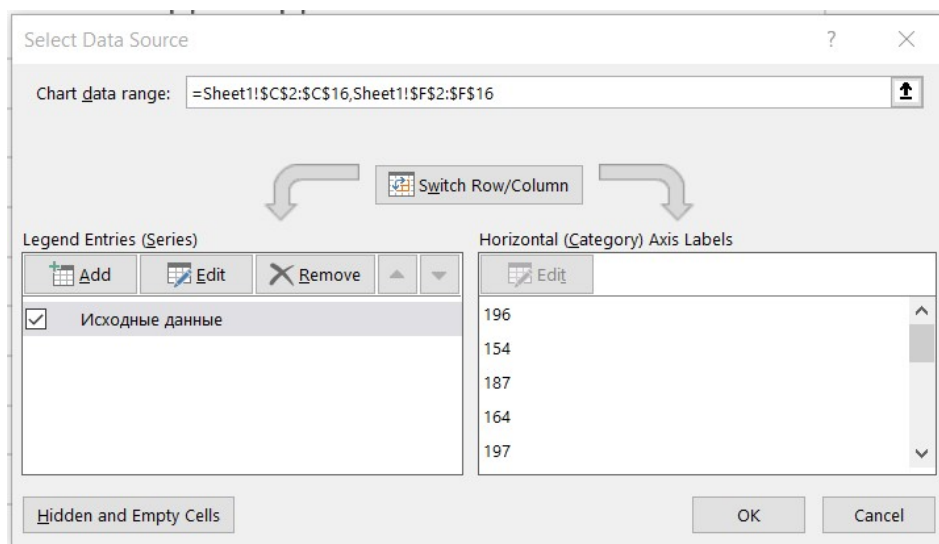


В меню Edit Series (Редактирование ряда данных) установите следующие настройки:

Series Name (Название ряда) – Исходные данные

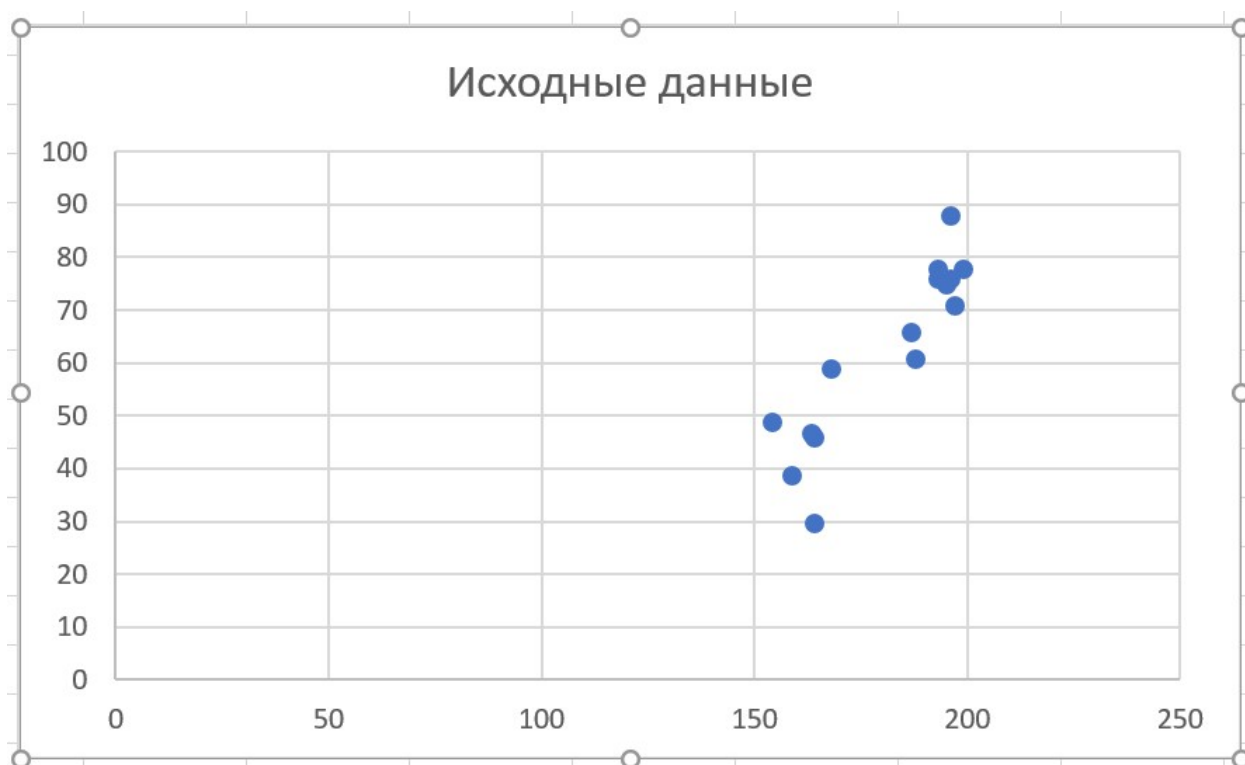
Series X values (Ряд X) – нажмите кнопку , выделите колонку значений X и нажмите кнопку  в окне Edit Series (Редактирование ряда)

Series Y values (Ряд Y) – выделите колонку значений Y



Обратите внимание на изменение в диалоговом окне Select Data Source (Выберите источник данных). В списке рядов данных Legend Entries (Series) появился ряд «Исходные данные». В разделе Horizontal (Category) Axis Labels (Метки горизонтальной оси) выводятся значения переменной X.

После нажатия OK появится окно графика.



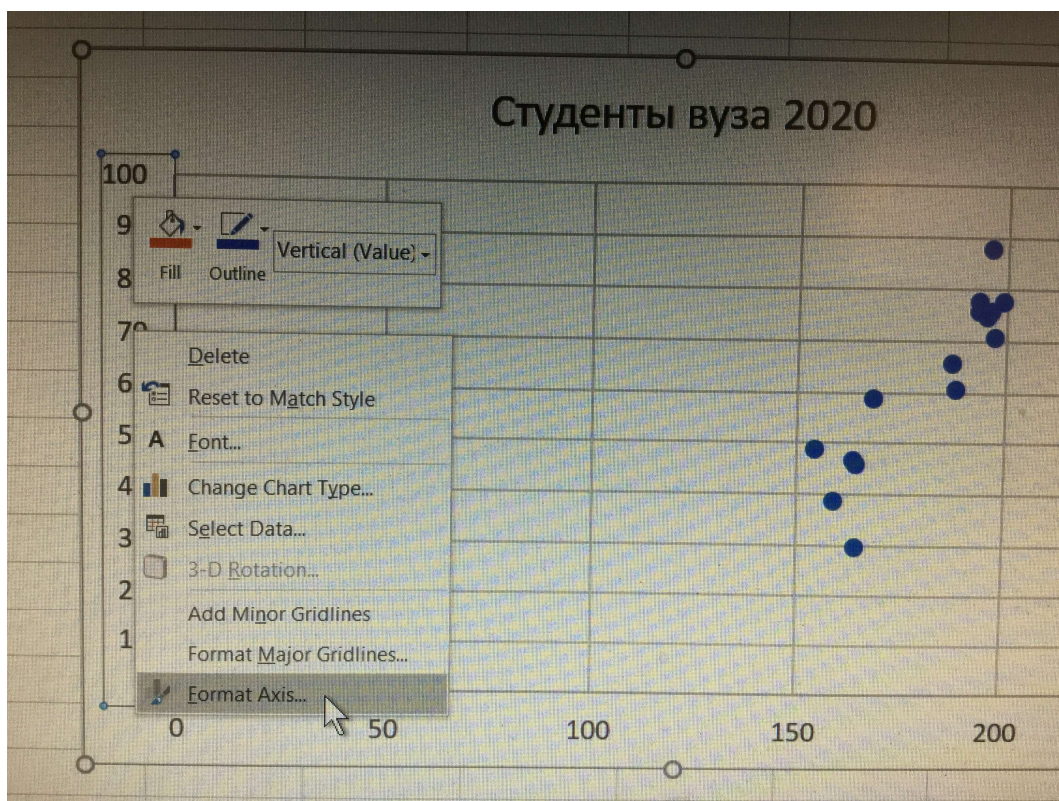
Можно видеть настройки графика по умолчанию: заголовок графика – имя ряда данных, масштаб по осям начинается от нуля. Названия осей отсутствуют.

Настройте график.

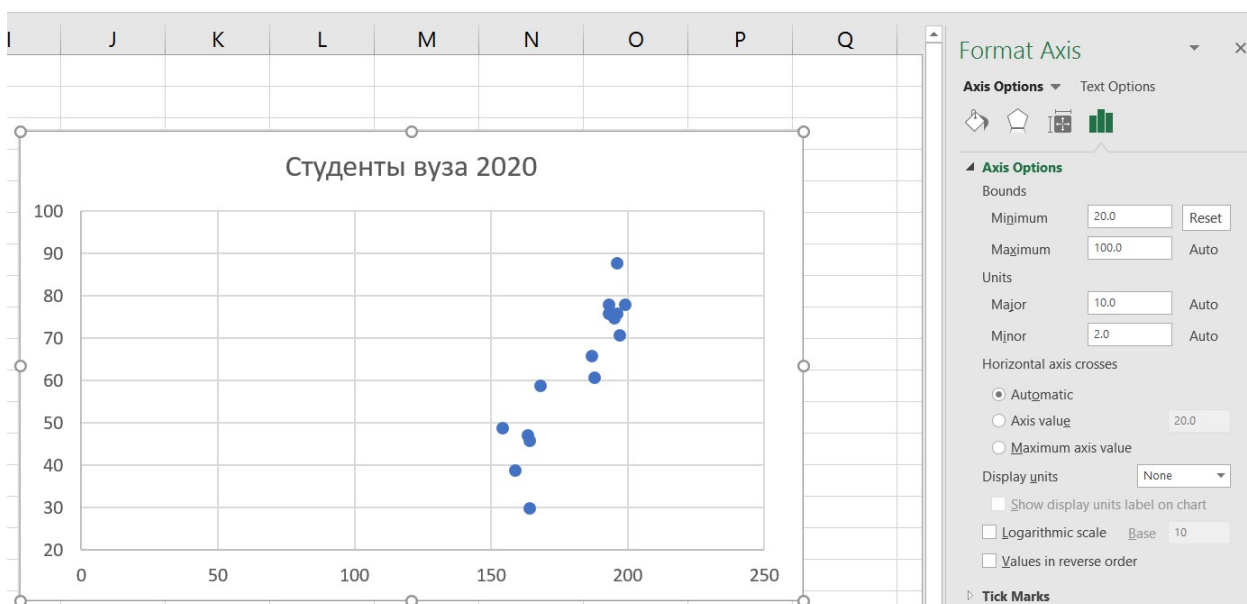
Измените заголовок, сделав двойной щелчок по заголовку.

Настройте масштаб по осям.

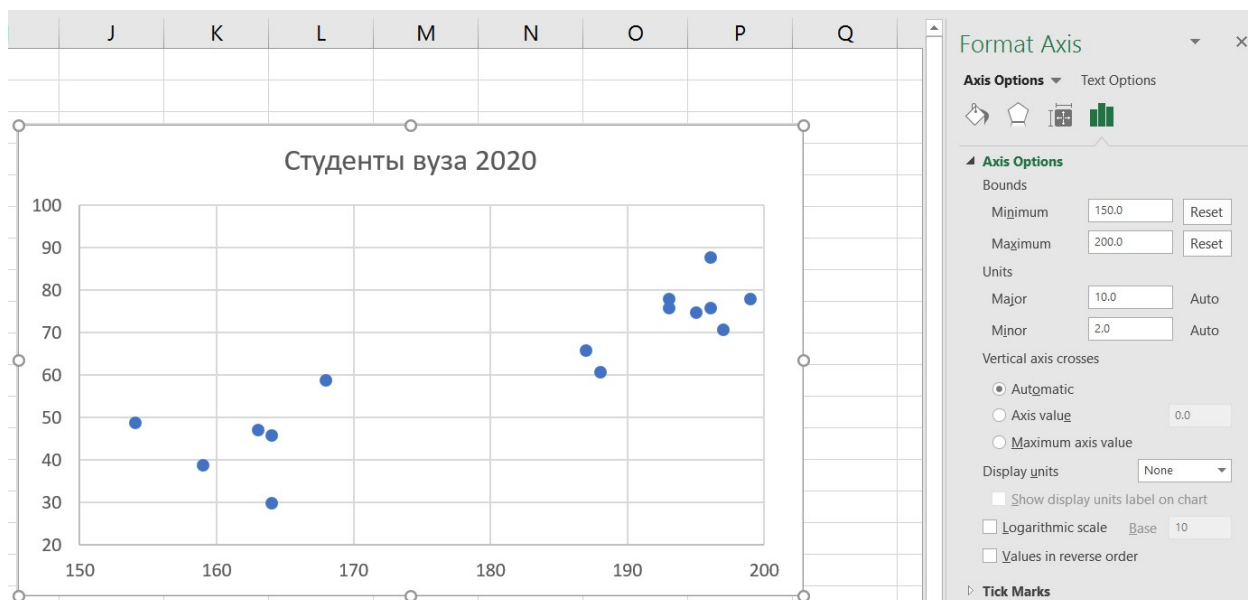
Укажите на ось, щёлкните правой кнопкой по горизонтальной оси и выберите в контекстном меню Format Axis (Форматирование оси).



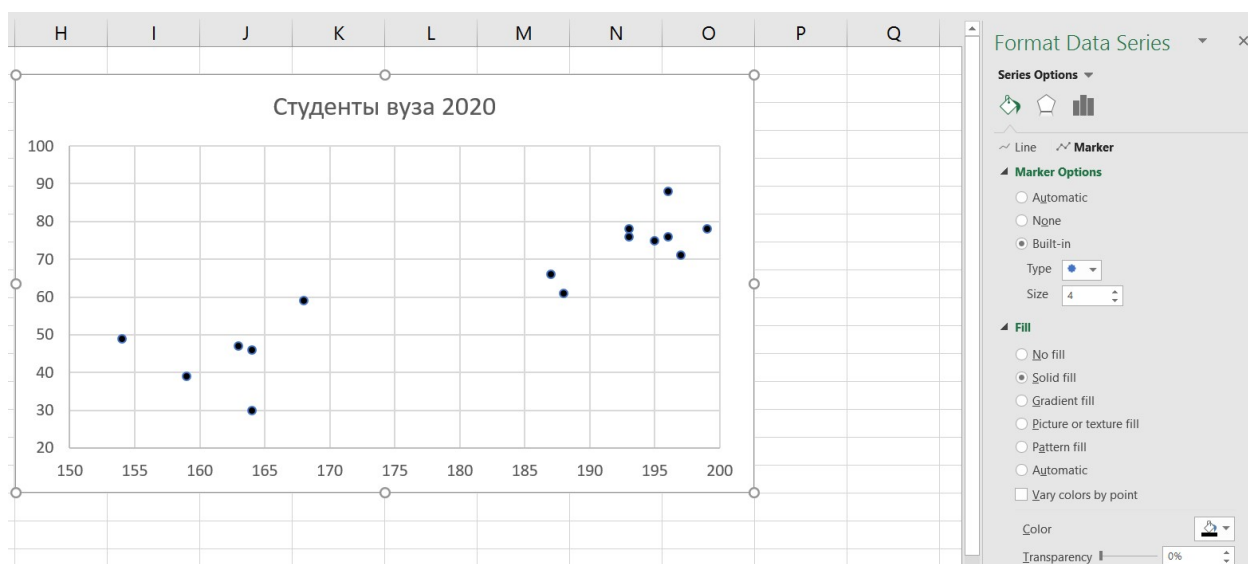
В правой части окна Excel появится окно Format Axis. Выберите масштаб по оси так, чтобы занять всё поле графика исходными данными, но оставить место по краям графика: Axis Options (Настройка осей) -> Bounds (Граничные значения) -> Minimum (Минимум) / Maximum (Максимум).



Аналогично настройте масштаб по вертикальной оси.

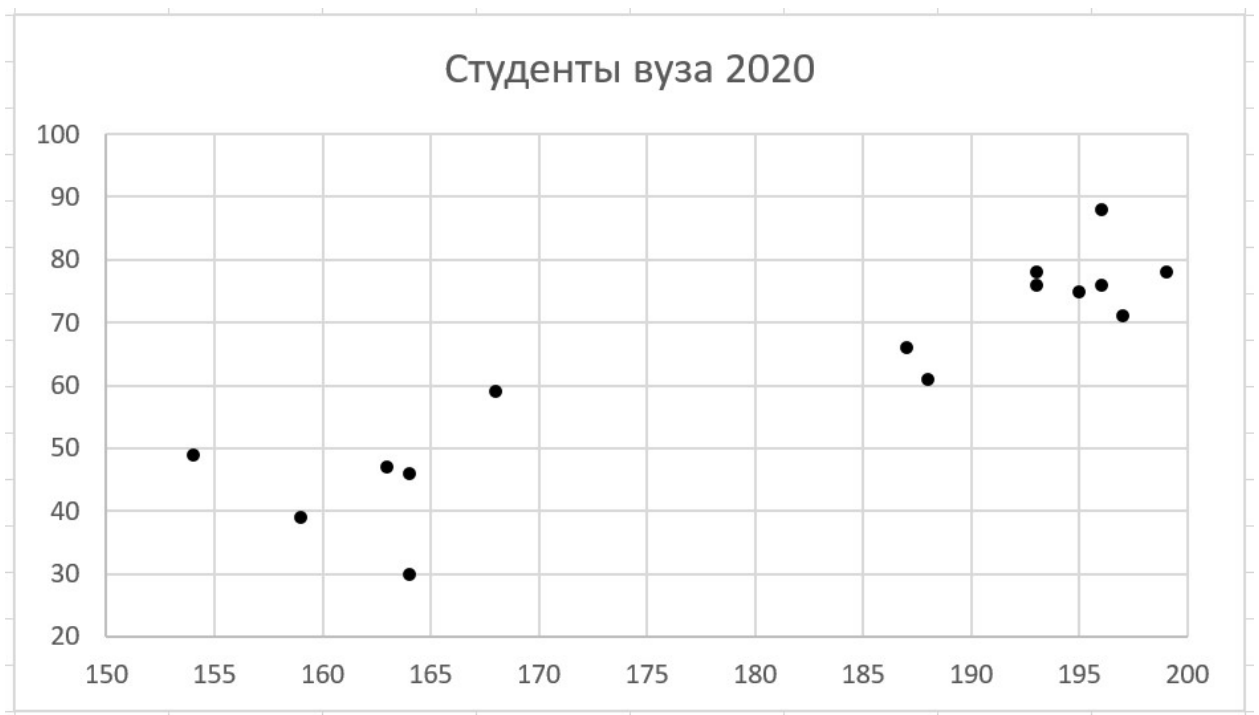


Настройте маркеры для точек на графике. Щёлкните по любому из маркеров. Появится меню Format Data Series (Форматировать ряд данных). Выберите Marker (Маркер) -> Marker Options (Настройка маркера) -> Built-in (Встроенный) -> Type (Тип) и Size (Размер).



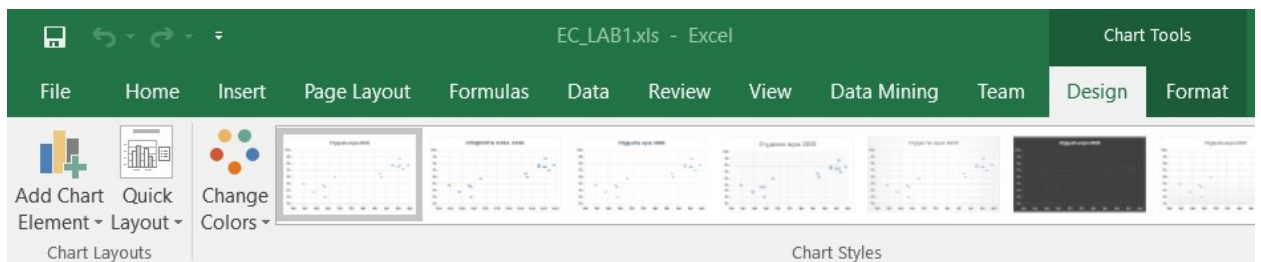
Настройте заливку маркера:

Format Data Series (Форматировать ряд данных). Выберите Marker (Маркер) -> Fill (Заливка) -> Solid Fill (Сплошная заливка) и Color (Цвет), а также Border (Оформление) -> No line (Без оформления).



Установите названия осей.

Щёлкните по графику. В верхнем меню появится раздел Chart Tools (Инструменты диаграмм) -> Design (Конструктор).

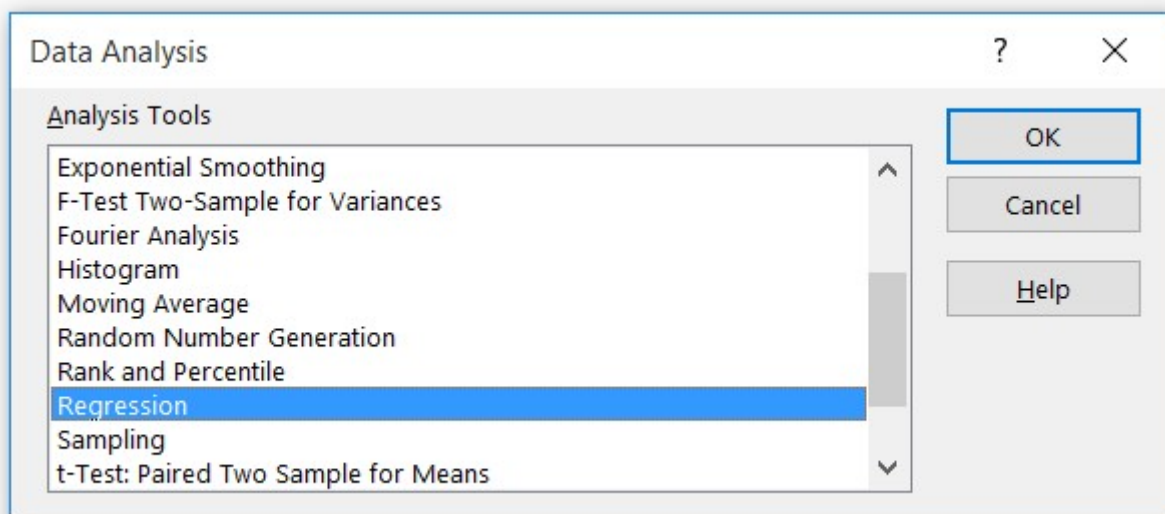


Выберите Add Chart Element (Добавить элемент диаграммы) -> Axis Titles (Заголовки осей) -> Primary horizontal (Основной заголовок горизонтальной оси).

Постройте регрессию Y на X.

Запустите надстройку Анализ данных:

Data (Данные) -> Data Analysis (Анализ данных) -> Regression (Регрессия)



Выберите факторную и результирующую переменные.

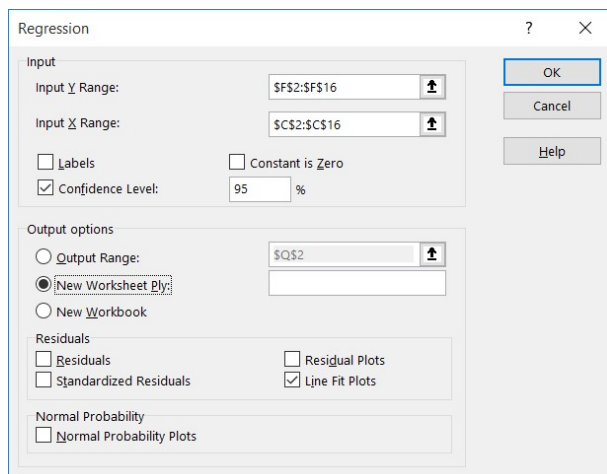
Input (Исходные данные) -> Input Y Range (Диапазон исходных данных по Y)

Input (Исходные данные) -> Input X Range (Диапазон исходных данных по X)

Confidence Level (Доверительная вероятность) = 95%

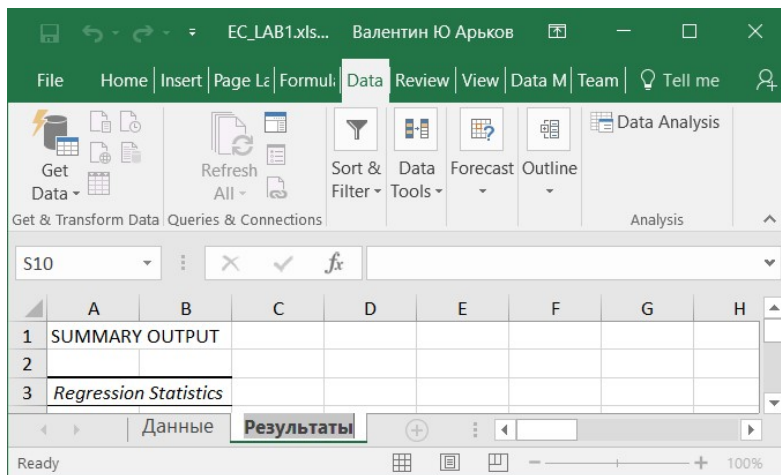
Output Options (Параметры вывода результатов) -> New Worksheet Ply (Вывод результатов на новую страницу рабочей книги Excel)

Residuals (Остатки) -> Line Fit Plots (График линейной регрессии)



После нажатия ОК будет сформирована новая страница с результатами регрессионного анализа.

Переименуйте листы рабочей книги, дважды щёлкнув по соответствующей закладке.



Перетащите лист исходных данных на первое место.

Настройте оформление графика.

Установите приемлемые размеры графика.

Настройте заголовок графика и названия осей, дважды щёлкнув по соответствующим надписям.

Установите масштаб по осям. Всё поле графика должно быть занято данными. При этом точки не должны оказаться на границе или за пределами графика.

Настройте маркеры для исходных данных.

Настройте вывод на график линии регрессии.

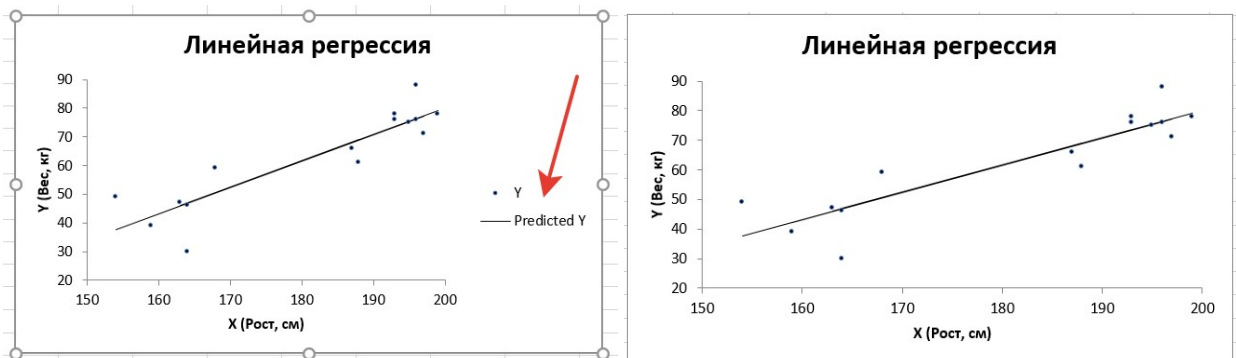
Format Data Series (Формат вывода данных на график) -> Series Options (Настройки вывода для ряда данных) -> Fill & Line (Заливка и линии)  -> Marker (Маркер) -> None (Без маркера)

Line (Линия) -> Solid Line (Сплошная линия)

Color (Цвет линии)

Width (Толщина линии)

Удалите легенду (условные обозначения) с графика



Найдите таблицу коэффициентов уравнения регрессии.

	<i>Coefficients</i>	<i>Standard Error</i>
Intercept	-103.9860586	22.4059666
X Variable 1	0.920026097	0.123258642

В таблице приводятся следующие оценки.

Coefficients (Коэффициенты)

Standard Error (Стандартные ошибки коэффициентов)

Intercept (Свободный член уравнения)

X Variable 1 (Коэффициент регрессии при 1-й переменной X)

Запишите результаты регрессионного анализа в следующем виде:

$$y = a + bx$$

$\begin{matrix} (b_a) & (b_b) \end{matrix}$

Количество значащих разрядов должно быть не меньше, чем в исходных данных. Значащие разряды относятся к представлению числа с плавающей точкой, а не к знакам после десятичной запятой.

В нашем примере значения роста содержат три значащих разряда, значения веса – два. Поэтому следует оставить хотя бы три значащих разряда.

$$y = -104 + 0,920x$$

$\begin{matrix} (22,4) & (0,123) \end{matrix}$

Запишите уравнение регрессии с указанием названия переменных.

$$\text{Вес} = -104 + 0,920 \cdot \text{Рост}$$

Проведите интерпретацию уравнения регрессии:

При увеличении роста на 10 сантиметров вес увеличивается в среднем на 9,2 килограмма.

Загрузите реальные данные с любого сайта компьютерного магазина, например <http://www.nix.ru>

Выберите компьютерные компоненты, например, жёсткие диски.

Проведите регрессионный анализ.

Постройте новую таблицу исходных данных, оставив только средние показатели по каждой группе. Например, можно определить средние цены всех моделей жёстких дисков ёмкостью 1, 2, 4, 6, 10 и 12 Тб.

Проведите регрессионный анализ.

Оформите отчет по анализу цен на выбранные компьютерные компоненты.

Рассмотрите график и выберите компоненты, сильнее всего отличающиеся от остальных. Запишите возможные объяснения.

Отправьте отчет преподавателю.