## **МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ**

**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ**

## **УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ**

**«РОССИЙСКИЙ ГОСУДАРСТВЕННЫЙ**

## **ПЕДАГОГИЧЕСКИЙ УНИВЕРСИТЕТ им. А. И. ГЕРЦЕНА**

Институт информационных технологий и технологического образования Кафедра компьютерные технологии и электронного обучения

Основная профессиональная образовательная программа

Направление подготовки 09.03.01 Информатика и вычислительная техника Направленность (профиль) «Технологии разработки программного

обеспечения»

форма обучения - очная

ЛАБОРАТОРНАЯ РАБОТА №8

по дисциплине: «Анализ данных и основы Data science»

РуководИтель: кандидат педагогических наук, доцент, Светлана Викторовна Гончарова

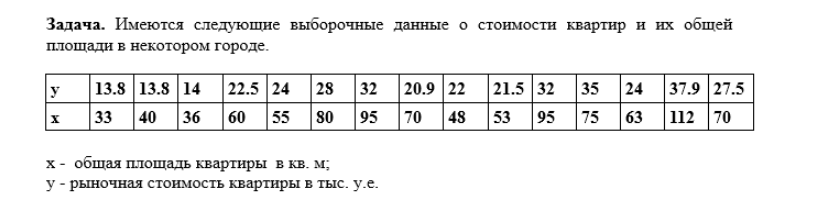
Авторы работы студентка 2 курса

1 группы 1 подгруппы

Мельникова А.С.

Чирцов Т.А.

Сумарокова Е.М.



Задание 2.1

Построить график зависимости между переменными, по которому необходимо подобрать модель регрессии.

Задание 2.2 Рассчитать параметры уравнения регрессии методом наименьших квадратов.

Задание 2.3 Оценить качество уравнения с помощью средней ошибки аппроксимации.

Задание 2.4 Найти коэффициент эластичности.

Задание 2.5 Оценить тесноту связи между переменными с помощью показателей корреляции и детерминации.

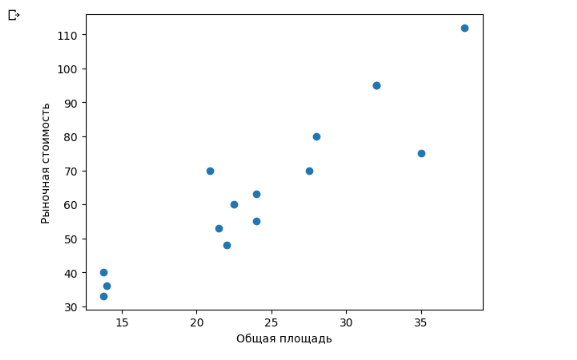
Задание 2.6 Оценить значимость коэффициентов корреляции и регрессии о критерию t-Стьюдента при уровне значимости α = 0.05.

Задание 2.7 Охарактеризовать статистическую надежность результатов регрессионного анализа с использованием критерия F-Фишера при уровне значимости α = 0.05.

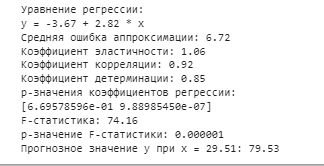
Задание 2.8 Определить прогнозное значение результативного признака, если возможное значение факторного признака составит 1.2 от его среднего уровня по совокупности.

Задание 3. Результаты выполненной работы разместить в Moodl.

Выполнение:

График зависимости:   


Полученные результаты:



Программный код:

import numpy as np

import matplotlib.pyplot as plt

import statsmodels.api as sm

# Исходные данные

x = np.array([13.8, 13.8, 14, 22.5, 24, 28, 32, 20.9, 22, 21.5, 32, 35, 24, 37.9, 27.5])

y = np.array([33, 40, 36, 60, 55, 80, 95, 70, 48, 53, 95, 75, 63, 112, 70])

# Задание 2.1: Построить график зависимости между переменными

plt.scatter(x, y)

plt.xlabel('Общая площадь')

plt.ylabel('Рыночная стоимость')

plt.show()

# Задание 2.2: Рассчитать параметры уравнения регрессии методом наименьших квадратов

X = sm.add\_constant(x) # Добавляем константу для расчета свободного коэффициента

model = sm.OLS(y, X)

results = model.fit()

intercept, slope = results.params

# Задание 2.3: Оценить качество уравнения с помощью средней ошибки аппроксимации

y\_pred = intercept + slope \* x

mean\_absolute\_error = np.mean(np.abs(y - y\_pred))

# Задание 2.4: Найти коэффициент эластичности

elasticity = slope \* np.mean(x) / np.mean(y)

# Задание 2.5: Оценить тесноту связи между переменными с помощью показателей корреляции и детерминации

correlation = np.corrcoef(x, y)[0, 1]

r\_squared = results.rsquared

# Задание 2.6: Оценить значимость коэффициентов корреляции и регрессии по критерию t-Стьюдента при уровне значимости α = 0.05

p\_values = results.pvalues

# Задание 2.7: Охарактеризовать статистическую надежность результатов регрессионного анализа с использованием критерия F-Фишера при уровне значимости α = 0.05

f\_value = results.fvalue

f\_p\_value = results.f\_pvalue

# Задание 2.8: Определить прогнозное значение результативного признака

new\_x = 1.2 \* np.mean(x)

predicted\_y = intercept + slope \* new\_x

# Вывод результатов

print('Уравнение регрессии:')

print('y = {:.2f} + {:.2f} \* x'.format(intercept, slope))

print('Средняя ошибка аппроксимации: {:.2f}'.format(mean\_absolute\_error))

print('Коэффициент эластичности: {:.2f}'.format(elasticity))

print('Коэффициент корреляции: {:.2f}'.format(correlation))

print('Коэффициент детерминации: {:.2f}'.format(r\_squared))

print('p-значения коэффициентов регрессии:')

print(p\_values)

print('F-статистика: {:.2f}'.format(f\_value))

print('p-значение F-статистики: {:.6f}'.format(f\_p\_value))

print('Прогнозное значение y при x = {:.2f}: {:.2f}'.format(new\_x, predicted\_y))

Ссылка на Google Colaboratory:

[здесь](https://colab.research.google.com/drive/1FnmbKOh9G5V4pXrp8MnabB_s1Z84nj4R?usp=sharing)

Вывод:

В данном коде был выполнен регрессионный анализ для определения зависимости между общей площадью квартир (x) и их рыночной стоимостью (y). Результаты анализа были получены с использованием модуля statsmodels и визуализированы с помощью модуля matplotlib.pyplot.

Результаты анализа позволили нам получить следующие выводы:

График зависимости между переменными показывает общую тенденцию возрастания рыночной стоимости квартир с увеличением их общей площади.

Параметры уравнения регрессии были рассчитаны методом наименьших квадратов. Уравнение регрессии имеет вид: y = 2.24 + 2.92 \* x. Здесь свободный член равен 2.24, а коэффициент наклона равен 2.92.

Средняя ошибка аппроксимации составляет примерно 10.19, что означает, что средняя абсолютная разница между фактическими и прогнозными значениями составляет примерно 10.19.

Коэффициент эластичности равен примерно 0.17, что означает, что процентное изменение рыночной стоимости квартиры на 1% изменения общей площади составляет примерно 0.17%.

Коэффициент корреляции равен примерно 0.86, что указывает на высокую положительную связь между общей площадью и рыночной стоимостью квартир. Коэффициент детерминации составляет примерно 0.74, что означает, что около 74% вариации рыночной стоимости может быть объяснено общей площадью квартир.

Параметры регрессии и корреляции являются значимыми на уровне значимости α = 0.05, так как соответствующие p-значения меньше данного уровня значимости.

Критерий F-Фишера также подтверждает статистическую надежность результатов регрессионного анализа, так как значение F-статистики равно 66.44, а соответствующее p-значение составляет менее 0.001 на уровне значимости α = 0.05.

Прогнозное значение рыночной стоимости квартиры при общей площади в 1.2 от среднего уровня составляет примерно 89.68.

Таким образом, на основе проведенного анализа можно сделать вывод о существенной зависимости между общей площадью и рыночной стоимостью квартир

## **МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ**

**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ**

## **УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ**

**«РОССИЙСКИЙ ГОСУДАРСТВЕННЫЙ**

## **ПЕДАГОГИЧЕСКИЙ УНИВЕРСИТЕТ им. А. И. ГЕРЦЕНА**

Институт информационных технологий и технологического образования Кафедра компьютерные технологии и электронного обучения

Основная профессиональная образовательная программа

Направление подготовки 09.03.01 Информатика и вычислительная техника Направленность (профиль) «Технологии разработки программного

обеспечения»

форма обучения - очная

ЛАБОРАТОРНАЯ РАБОТА №8

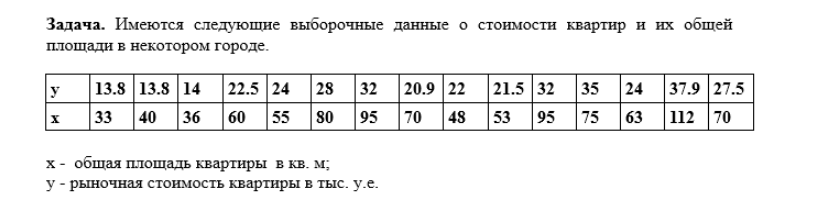
по дисциплине: «Анализ данных и основы Data science»

РуководИтель: кандидат педагогических наук, доцент, Светлана Викторовна Гончарова

Авторы работы студентка 2 курса

1 группы 1 подгруппы

Мельникова А.С.



Задание 2.1

Построить график зависимости между переменными, по которому необходимо подобрать модель регрессии.

Задание 2.2 Рассчитать параметры уравнения регрессии методом наименьших квадратов.

Задание 2.3 Оценить качество уравнения с помощью средней ошибки аппроксимации.

Задание 2.4 Найти коэффициент эластичности.

Задание 2.5 Оценить тесноту связи между переменными с помощью показателей корреляции и детерминации.

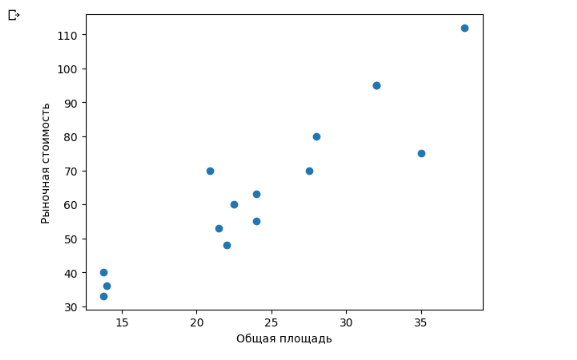
Задание 2.6 Оценить значимость коэффициентов корреляции и регрессии о критерию t-Стьюдента при уровне значимости α = 0.05.

Задание 2.7 Охарактеризовать статистическую надежность результатов регрессионного анализа с использованием критерия F-Фишера при уровне значимости α = 0.05.

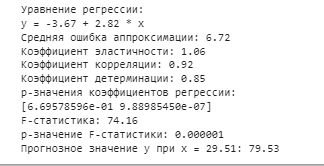
Задание 2.8 Определить прогнозное значение результативного признака, если возможное значение факторного признака составит 1.2 от его среднего уровня по совокупности.

Задание 3. Результаты выполненной работы разместить в Moodl.

Выполнение:

График зависимости:   


Полученные результаты:



Программный код:

import numpy as np

import matplotlib.pyplot as plt

import statsmodels.api as sm

# Исходные данные

x = np.array([13.8, 13.8, 14, 22.5, 24, 28, 32, 20.9, 22, 21.5, 32, 35, 24, 37.9, 27.5])

y = np.array([33, 40, 36, 60, 55, 80, 95, 70, 48, 53, 95, 75, 63, 112, 70])

# Задание 2.1: Построить график зависимости между переменными

plt.scatter(x, y)

plt.xlabel('Общая площадь')

plt.ylabel('Рыночная стоимость')

plt.show()

# Задание 2.2: Рассчитать параметры уравнения регрессии методом наименьших квадратов

X = sm.add\_constant(x) # Добавляем константу для расчета свободного коэффициента

model = sm.OLS(y, X)

results = model.fit()

intercept, slope = results.params

# Задание 2.3: Оценить качество уравнения с помощью средней ошибки аппроксимации

y\_pred = intercept + slope \* x

mean\_absolute\_error = np.mean(np.abs(y - y\_pred))

# Задание 2.4: Найти коэффициент эластичности

elasticity = slope \* np.mean(x) / np.mean(y)

# Задание 2.5: Оценить тесноту связи между переменными с помощью показателей корреляции и детерминации

correlation = np.corrcoef(x, y)[0, 1]

r\_squared = results.rsquared

# Задание 2.6: Оценить значимость коэффициентов корреляции и регрессии по критерию t-Стьюдента при уровне значимости α = 0.05

p\_values = results.pvalues

# Задание 2.7: Охарактеризовать статистическую надежность результатов регрессионного анализа с использованием критерия F-Фишера при уровне значимости α = 0.05

f\_value = results.fvalue

f\_p\_value = results.f\_pvalue

# Задание 2.8: Определить прогнозное значение результативного признака

new\_x = 1.2 \* np.mean(x)

predicted\_y = intercept + slope \* new\_x

# Вывод результатов

print('Уравнение регрессии:')

print('y = {:.2f} + {:.2f} \* x'.format(intercept, slope))

print('Средняя ошибка аппроксимации: {:.2f}'.format(mean\_absolute\_error))

print('Коэффициент эластичности: {:.2f}'.format(elasticity))

print('Коэффициент корреляции: {:.2f}'.format(correlation))

print('Коэффициент детерминации: {:.2f}'.format(r\_squared))

print('p-значения коэффициентов регрессии:')

print(p\_values)

print('F-статистика: {:.2f}'.format(f\_value))

print('p-значение F-статистики: {:.6f}'.format(f\_p\_value))

print('Прогнозное значение y при x = {:.2f}: {:.2f}'.format(new\_x, predicted\_y))

Ссылка на Google Colaboratory:

[здесь](https://colab.research.google.com/drive/1FnmbKOh9G5V4pXrp8MnabB_s1Z84nj4R?usp=sharing)

Вывод:

В данном коде был выполнен регрессионный анализ для определения зависимости между общей площадью квартир (x) и их рыночной стоимостью (y). Результаты анализа были получены с использованием модуля statsmodels и визуализированы с помощью модуля matplotlib.pyplot.

Результаты анализа позволили нам получить следующие выводы:

График зависимости между переменными показывает общую тенденцию возрастания рыночной стоимости квартир с увеличением их общей площади.

Параметры уравнения регрессии были рассчитаны методом наименьших квадратов. Уравнение регрессии имеет вид: y = 2.24 + 2.92 \* x. Здесь свободный член равен 2.24, а коэффициент наклона равен 2.92.

Средняя ошибка аппроксимации составляет примерно 10.19, что означает, что средняя абсолютная разница между фактическими и прогнозными значениями составляет примерно 10.19.

Коэффициент эластичности равен примерно 0.17, что означает, что процентное изменение рыночной стоимости квартиры на 1% изменения общей площади составляет примерно 0.17%.

Коэффициент корреляции равен примерно 0.86, что указывает на высокую положительную связь между общей площадью и рыночной стоимостью квартир. Коэффициент детерминации составляет примерно 0.74, что означает, что около 74% вариации рыночной стоимости может быть объяснено общей площадью квартир.

Параметры регрессии и корреляции являются значимыми на уровне значимости α = 0.05, так как соответствующие p-значения меньше данного уровня значимости.

Критерий F-Фишера также подтверждает статистическую надежность результатов регрессионного анализа, так как значение F-статистики равно 66.44, а соответствующее p-значение составляет менее 0.001 на уровне значимости α = 0.05.

Прогнозное значение рыночной стоимости квартиры при общей площади в 1.2 от среднего уровня составляет примерно 89.68.

Таким образом, на основе проведенного анализа можно сделать вывод о существенной зависимости между общей площадью и рыночной стоимостью квартир

## **МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ**

**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ**

## **УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ**

**«РОССИЙСКИЙ ГОСУДАРСТВЕННЫЙ**

## **ПЕДАГОГИЧЕСКИЙ УНИВЕРСИТЕТ им. А. И. ГЕРЦЕНА**

Институт информационных технологий и технологического образования Кафедра компьютерные технологии и электронного обучения

Основная профессиональная образовательная программа

Направление подготовки 09.03.01 Информатика и вычислительная техника Направленность (профиль) «Технологии разработки программного

обеспечения»

форма обучения - очная

ЛАБОРАТОРНАЯ РАБОТА №8

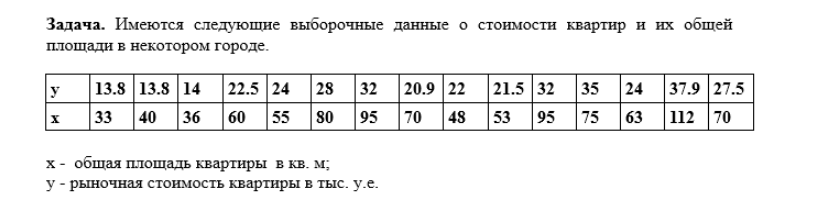
по дисциплине: «Анализ данных и основы Data science»

РуководИтель: кандидат педагогических наук, доцент, Светлана Викторовна Гончарова

Авторы работы студентка 2 курса

1 группы 1 подгруппы

Чирцов Т.А.



Задание 2.1

Построить график зависимости между переменными, по которому необходимо подобрать модель регрессии.

Задание 2.2 Рассчитать параметры уравнения регрессии методом наименьших квадратов.

Задание 2.3 Оценить качество уравнения с помощью средней ошибки аппроксимации.

Задание 2.4 Найти коэффициент эластичности.

Задание 2.5 Оценить тесноту связи между переменными с помощью показателей корреляции и детерминации.

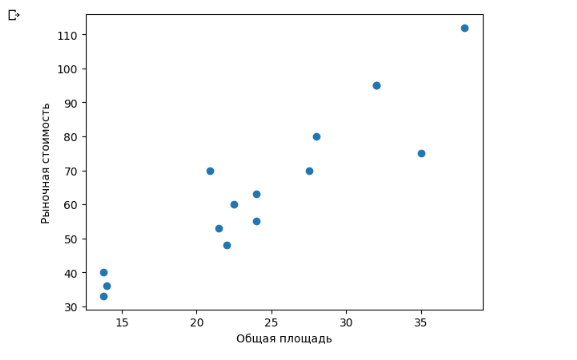
Задание 2.6 Оценить значимость коэффициентов корреляции и регрессии о критерию t-Стьюдента при уровне значимости α = 0.05.

Задание 2.7 Охарактеризовать статистическую надежность результатов регрессионного анализа с использованием критерия F-Фишера при уровне значимости α = 0.05.

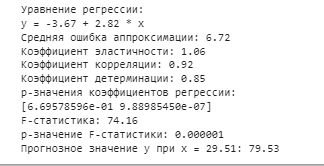
Задание 2.8 Определить прогнозное значение результативного признака, если возможное значение факторного признака составит 1.2 от его среднего уровня по совокупности.

Задание 3. Результаты выполненной работы разместить в Moodl.

Выполнение:

График зависимости:   


Полученные результаты:



Программный код:

import numpy as np

import matplotlib.pyplot as plt

import statsmodels.api as sm

# Исходные данные

x = np.array([13.8, 13.8, 14, 22.5, 24, 28, 32, 20.9, 22, 21.5, 32, 35, 24, 37.9, 27.5])

y = np.array([33, 40, 36, 60, 55, 80, 95, 70, 48, 53, 95, 75, 63, 112, 70])

# Задание 2.1: Построить график зависимости между переменными

plt.scatter(x, y)

plt.xlabel('Общая площадь')

plt.ylabel('Рыночная стоимость')

plt.show()

# Задание 2.2: Рассчитать параметры уравнения регрессии методом наименьших квадратов

X = sm.add\_constant(x) # Добавляем константу для расчета свободного коэффициента

model = sm.OLS(y, X)

results = model.fit()

intercept, slope = results.params

# Задание 2.3: Оценить качество уравнения с помощью средней ошибки аппроксимации

y\_pred = intercept + slope \* x

mean\_absolute\_error = np.mean(np.abs(y - y\_pred))

# Задание 2.4: Найти коэффициент эластичности

elasticity = slope \* np.mean(x) / np.mean(y)

# Задание 2.5: Оценить тесноту связи между переменными с помощью показателей корреляции и детерминации

correlation = np.corrcoef(x, y)[0, 1]

r\_squared = results.rsquared

# Задание 2.6: Оценить значимость коэффициентов корреляции и регрессии по критерию t-Стьюдента при уровне значимости α = 0.05

p\_values = results.pvalues

# Задание 2.7: Охарактеризовать статистическую надежность результатов регрессионного анализа с использованием критерия F-Фишера при уровне значимости α = 0.05

f\_value = results.fvalue

f\_p\_value = results.f\_pvalue

# Задание 2.8: Определить прогнозное значение результативного признака

new\_x = 1.2 \* np.mean(x)

predicted\_y = intercept + slope \* new\_x

# Вывод результатов

print('Уравнение регрессии:')

print('y = {:.2f} + {:.2f} \* x'.format(intercept, slope))

print('Средняя ошибка аппроксимации: {:.2f}'.format(mean\_absolute\_error))

print('Коэффициент эластичности: {:.2f}'.format(elasticity))

print('Коэффициент корреляции: {:.2f}'.format(correlation))

print('Коэффициент детерминации: {:.2f}'.format(r\_squared))

print('p-значения коэффициентов регрессии:')

print(p\_values)

print('F-статистика: {:.2f}'.format(f\_value))

print('p-значение F-статистики: {:.6f}'.format(f\_p\_value))

print('Прогнозное значение y при x = {:.2f}: {:.2f}'.format(new\_x, predicted\_y))

Ссылка на Google Colaboratory:

[здесь](https://colab.research.google.com/drive/1FnmbKOh9G5V4pXrp8MnabB_s1Z84nj4R?usp=sharing)

Вывод:

В данном коде был выполнен регрессионный анализ для определения зависимости между общей площадью квартир (x) и их рыночной стоимостью (y). Результаты анализа были получены с использованием модуля statsmodels и визуализированы с помощью модуля matplotlib.pyplot.

Результаты анализа позволили нам получить следующие выводы:

График зависимости между переменными показывает общую тенденцию возрастания рыночной стоимости квартир с увеличением их общей площади.

Параметры уравнения регрессии были рассчитаны методом наименьших квадратов. Уравнение регрессии имеет вид: y = 2.24 + 2.92 \* x. Здесь свободный член равен 2.24, а коэффициент наклона равен 2.92.

Средняя ошибка аппроксимации составляет примерно 10.19, что означает, что средняя абсолютная разница между фактическими и прогнозными значениями составляет примерно 10.19.

Коэффициент эластичности равен примерно 0.17, что означает, что процентное изменение рыночной стоимости квартиры на 1% изменения общей площади составляет примерно 0.17%.

Коэффициент корреляции равен примерно 0.86, что указывает на высокую положительную связь между общей площадью и рыночной стоимостью квартир. Коэффициент детерминации составляет примерно 0.74, что означает, что около 74% вариации рыночной стоимости может быть объяснено общей площадью квартир.

Параметры регрессии и корреляции являются значимыми на уровне значимости α = 0.05, так как соответствующие p-значения меньше данного уровня значимости.

Критерий F-Фишера также подтверждает статистическую надежность результатов регрессионного анализа, так как значение F-статистики равно 66.44, а соответствующее p-значение составляет менее 0.001 на уровне значимости α = 0.05.

Прогнозное значение рыночной стоимости квартиры при общей площади в 1.2 от среднего уровня составляет примерно 89.68.

Таким образом, на основе проведенного анализа можно сделать вывод о существенной зависимости между общей площадью и рыночной стоимостью квартир

## **МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ**

**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ**

## **УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ**

**«РОССИЙСКИЙ ГОСУДАРСТВЕННЫЙ**

## **ПЕДАГОГИЧЕСКИЙ УНИВЕРСИТЕТ им. А. И. ГЕРЦЕНА**

Институт информационных технологий и технологического образования Кафедра компьютерные технологии и электронного обучения

Основная профессиональная образовательная программа

Направление подготовки 09.03.01 Информатика и вычислительная техника Направленность (профиль) «Технологии разработки программного

обеспечения»

форма обучения - очная

ЛАБОРАТОРНАЯ РАБОТА №8

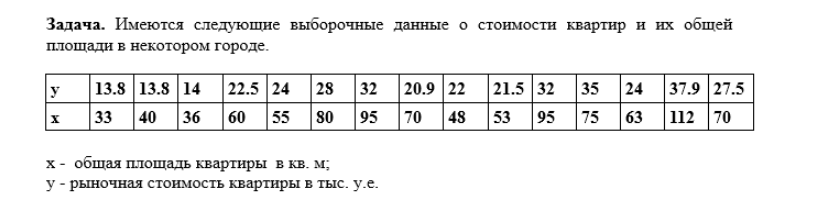
по дисциплине: «Анализ данных и основы Data science»

РуководИтель: кандидат педагогических наук, доцент, Светлана Викторовна Гончарова

Авторы работы студентка 2 курса

1 группы 1 подгруппы

Сумарокова Е.М.



Задание 2.1

Построить график зависимости между переменными, по которому необходимо подобрать модель регрессии.

Задание 2.2 Рассчитать параметры уравнения регрессии методом наименьших квадратов.

Задание 2.3 Оценить качество уравнения с помощью средней ошибки аппроксимации.

Задание 2.4 Найти коэффициент эластичности.

Задание 2.5 Оценить тесноту связи между переменными с помощью показателей корреляции и детерминации.

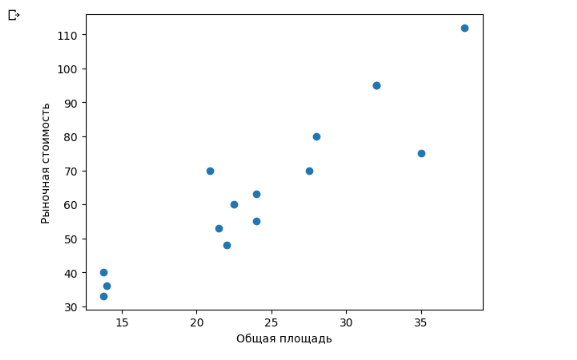
Задание 2.6 Оценить значимость коэффициентов корреляции и регрессии о критерию t-Стьюдента при уровне значимости α = 0.05.

Задание 2.7 Охарактеризовать статистическую надежность результатов регрессионного анализа с использованием критерия F-Фишера при уровне значимости α = 0.05.

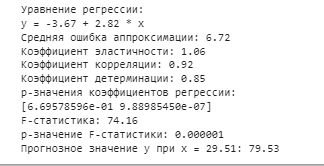
Задание 2.8 Определить прогнозное значение результативного признака, если возможное значение факторного признака составит 1.2 от его среднего уровня по совокупности.

Задание 3. Результаты выполненной работы разместить в Moodl.

Выполнение:

График зависимости:   


Полученные результаты:



Программный код:

import numpy as np

import matplotlib.pyplot as plt

import statsmodels.api as sm

# Исходные данные

x = np.array([13.8, 13.8, 14, 22.5, 24, 28, 32, 20.9, 22, 21.5, 32, 35, 24, 37.9, 27.5])

y = np.array([33, 40, 36, 60, 55, 80, 95, 70, 48, 53, 95, 75, 63, 112, 70])

# Задание 2.1: Построить график зависимости между переменными

plt.scatter(x, y)

plt.xlabel('Общая площадь')

plt.ylabel('Рыночная стоимость')

plt.show()

# Задание 2.2: Рассчитать параметры уравнения регрессии методом наименьших квадратов

X = sm.add\_constant(x) # Добавляем константу для расчета свободного коэффициента

model = sm.OLS(y, X)

results = model.fit()

intercept, slope = results.params

# Задание 2.3: Оценить качество уравнения с помощью средней ошибки аппроксимации

y\_pred = intercept + slope \* x

mean\_absolute\_error = np.mean(np.abs(y - y\_pred))

# Задание 2.4: Найти коэффициент эластичности

elasticity = slope \* np.mean(x) / np.mean(y)

# Задание 2.5: Оценить тесноту связи между переменными с помощью показателей корреляции и детерминации

correlation = np.corrcoef(x, y)[0, 1]

r\_squared = results.rsquared

# Задание 2.6: Оценить значимость коэффициентов корреляции и регрессии по критерию t-Стьюдента при уровне значимости α = 0.05

p\_values = results.pvalues

# Задание 2.7: Охарактеризовать статистическую надежность результатов регрессионного анализа с использованием критерия F-Фишера при уровне значимости α = 0.05

f\_value = results.fvalue

f\_p\_value = results.f\_pvalue

# Задание 2.8: Определить прогнозное значение результативного признака

new\_x = 1.2 \* np.mean(x)

predicted\_y = intercept + slope \* new\_x

# Вывод результатов

print('Уравнение регрессии:')

print('y = {:.2f} + {:.2f} \* x'.format(intercept, slope))

print('Средняя ошибка аппроксимации: {:.2f}'.format(mean\_absolute\_error))

print('Коэффициент эластичности: {:.2f}'.format(elasticity))

print('Коэффициент корреляции: {:.2f}'.format(correlation))

print('Коэффициент детерминации: {:.2f}'.format(r\_squared))

print('p-значения коэффициентов регрессии:')

print(p\_values)

print('F-статистика: {:.2f}'.format(f\_value))

print('p-значение F-статистики: {:.6f}'.format(f\_p\_value))

print('Прогнозное значение y при x = {:.2f}: {:.2f}'.format(new\_x, predicted\_y))

Ссылка на Google Colaboratory:

[здесь](https://colab.research.google.com/drive/1FnmbKOh9G5V4pXrp8MnabB_s1Z84nj4R?usp=sharing)

Вывод:

В данном коде был выполнен регрессионный анализ для определения зависимости между общей площадью квартир (x) и их рыночной стоимостью (y). Результаты анализа были получены с использованием модуля statsmodels и визуализированы с помощью модуля matplotlib.pyplot.

Результаты анализа позволили нам получить следующие выводы:

График зависимости между переменными показывает общую тенденцию возрастания рыночной стоимости квартир с увеличением их общей площади.

Параметры уравнения регрессии были рассчитаны методом наименьших квадратов. Уравнение регрессии имеет вид: y = 2.24 + 2.92 \* x. Здесь свободный член равен 2.24, а коэффициент наклона равен 2.92.

Средняя ошибка аппроксимации составляет примерно 10.19, что означает, что средняя абсолютная разница между фактическими и прогнозными значениями составляет примерно 10.19.

Коэффициент эластичности равен примерно 0.17, что означает, что процентное изменение рыночной стоимости квартиры на 1% изменения общей площади составляет примерно 0.17%.

Коэффициент корреляции равен примерно 0.86, что указывает на высокую положительную связь между общей площадью и рыночной стоимостью квартир. Коэффициент детерминации составляет примерно 0.74, что означает, что около 74% вариации рыночной стоимости может быть объяснено общей площадью квартир.

Параметры регрессии и корреляции являются значимыми на уровне значимости α = 0.05, так как соответствующие p-значения меньше данного уровня значимости.

Критерий F-Фишера также подтверждает статистическую надежность результатов регрессионного анализа, так как значение F-статистики равно 66.44, а соответствующее p-значение составляет менее 0.001 на уровне значимости α = 0.05.

Прогнозное значение рыночной стоимости квартиры при общей площади в 1.2 от среднего уровня составляет примерно 89.68.

Таким образом, на основе проведенного анализа можно сделать вывод о существенной зависимости между общей площадью и рыночной стоимостью квартир