

Оглавление

[ВВЕДЕНИЕ 4](#_Toc105001130)

[Цель практики 4](#_Toc105001131)

[Задачи практики 4](#_Toc105001132)

[ОСНОВНАЯ ЧАСТЬ 5](#_Toc105001133)

[1. Выбор платформы и данных 5](#_Toc105001134)

[1.1 Выбор платформы для разработки и необходимых библиотек 5](#_Toc105001135)

[1.1.1 Обоснование выбора платформы интерпретатора языка Python 5](#_Toc105001136)

[1.1.2 Выбор библиотек для анализа данных 5](#_Toc105001137)

[1.1.2.1 Pandas 6](#_Toc105001138)

[1.1.2.2 Matplotlib 6](#_Toc105001139)

[1.1.2.3 Numpy 6](#_Toc105001140)

[1.2 Получение данных 6](#_Toc105001141)

[1.2.1 Предварительный анализ данных 6](#_Toc105001142)

[1.2.2 Выбор формата хранения данных 7](#_Toc105001143)

[1.2.3 Очистка данных от цифрового мусора 7](#_Toc105001144)

[1.3 Выводы по разделу 7](#_Toc105001145)

[2. КОЛИЧЕСТВЕННЫЙ АНАЛИЗ ДАННЫХ 19](#_Toc105001146)

[2.1 Получение данных Dataset (данные) 19](#_Toc105001147)

[2.2 Рассчёт стандартных числовых характеристик для набора данных Series 19](#_Toc105001148)

[2.2.1 Минимальное значение: 20](#_Toc105001149)

[2.2.2 Подсчет количества повторяющихся элементов: 20](#_Toc105001150)

[2.2.3 Максимальное значение: 20](#_Toc105001151)

[2.2.4 Среднеквадратическое отклонение: 20](#_Toc105001152)

[2.2.5 Сумма значений: 20](#_Toc105001153)

[2.3 Выводы по разделу 21](#_Toc105001154)

[3. ВИЗУАЛИЗАЦИЯ РЕЗУЛЬТАТОВ КОЛИЧЕСТВЕННОГО АНАЛИЗА ДАННЫХ 24](#_Toc105001155)

[3.1 Визуализация наборов данных с помощью стандартных библиотек по заданным критериям 24](#_Toc105001156)

[3.2 Формирование Dataframe из данных Series и добавление столбцов 25](#_Toc105001157)

[3.3 Визуализация данных полученных в результате промежуточного анализа 25](#_Toc105001158)

[3.3.1 Визуализация линейного графика объекта Series 26](#_Toc105001159)

[3.3.2 Визуализация гистограммы объекта Series 26](#_Toc105001160)

[3.3.3 Визуализация линейного графика объекта Dataframe 27](#_Toc105001161)

[3.4 Выводы по разделу 27](#_Toc105001162)

[ЗАКЛЮЧЕНИЕ 29](#_Toc105001163)

[СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ 30](#_Toc105001164)

# ВВЕДЕНИЕ

Учебная практика (ознакомительная практика) относится к разделу Б2.О.01 блока «Практики» базового учебного плана основной профессиональной образовательной программы по направлению подготовки 09.03.03 «Прикладная информатика» и является видом учебной деятельности, направленной на ознакомление, формирование, закрепление, развитие практических навыков и компетенции в процессе выполнения определенных видов работ, связанных с будущей профессиональной деятельностью. Учебная практика проходила в Московском университете им. С.Ю. Витте (далее - Университете) на кафедре информационных систем

## Цель практики

Приобретение первичных профессиональных навыков, практического опыта, закрепление, систематизация и расширение теоретических знаний по дисциплинам учебного плана при прохождении учебной практики в Университете.

## Задачи практики

- Закрепление приобретенных теоретических знаний.

- Приобретение навыков научного поиска и практической работы с информационными источниками данных.

- Овладение практическими методиками формирования цели и задач.

- Приобретение практических навыков использования иностранного языка при анализе зарубежных информационных Интернет-источников.

- Приобретение навыков использования современных информационных технологий и программных средств для решения задач анализа данных.

- Приобретение навыков инсталлирования программного обеспечения.

- Формирование практических навыков по основам алгоритмизации вычислительных процессов и программированию решения задач обработки данных, их отладке и тестирования на выбранной языковой платформе.

- Получение навыков экспериментального исследования больших данных и практического решения информационных задач по обработке данных.

- Получения навыков проведения инженерных расчетов.

- Получение навыков самостоятельной работы.

- Получение навыков подготовки отчетной документации

Источниками информации явились открытые информационные базы, интернет-ресурсы, учебно-методические материалы, размещенные в Электронном Университете, а также информация, представленная на официальном сайте университета в разделе «Сведения об образовательной организации»: https://www.muiv.ru/sveden/.

# ОСНОВНАЯ ЧАСТЬ

# Выбор платформы и данных

## Выбор платформы для разработки и необходимых библиотек

Python (в русском языке распространено название пито́н) — высокоуровневый язык программирования общего назначения, ориентированный на повышение производительности разработчика и читаемости кода. Синтаксис ядра Python минималистичен. В то же время стандартная библиотека включает большой объём полезных функций. Python поддерживает несколько парадигм программирования, в том числе структурное, объектно-ориентированное, функциональное, императивное и аспектно-ориентированное. Основные архитектурные черты — динамическая типизация, автоматическое управление памятью, полная интроспекция, механизм обработки исключений, поддержка многопоточных вычислений и удобные высокоуровневые структуры данных. Код в Python организовывается в функции и классы, которые могут объединяться в модули (они в свою очередь могут быть объединены в пакеты). Python имеет множество полезных библиотек, такие как Matplotlib, pandas, NumPy и другие.

### Обоснование выбора платформы интерпретатора языка Python

В качестве редактора кода был выбран PyCharm. PyCharm делает разработку максимально продуктивной благодаря функциям автодополнения и анализа кода, мгновенной подсветке ошибок и быстрым исправлениям. Автоматические рефакторинги помогают эффективно редактировать код, а удобная навигация позволяет мгновенно перемещаться по проекту.

Умный редактор PyCharm предназначен для максимально продуктивной разработки на Python, JavaScript, CoffeeScript, TypeScript, CSS и популярных языках шаблонов. Функции автодополнения, обнаружения ошибок и быстрые исправления учитывают особенности каждого из поддерживаемых языков.

Умный поиск позволяет быстро перейти к любому классу, файлу или символу, а также к нужному окну или действию IDE. Переход к вышестоящему методу, тесту, объявлению, вхождению или реализации осуществляется в одно нажатие.

PyCharm предоставляет широкие возможности реорганизации кода с помощью рефакторингов Rename и Delete, Extract Method, Introduce Variable, Inline Variable, Inline Method и многих других. Рефакторинги учитывают особенности конкретного языка или фреймворка, помогая вносить изменения по всему проекту.

### Выбор библиотек для анализа данных

На текущий день можно найти множество библиотек для чтения, обработки и визуализации данных на языке программирования Python. Однако среди всего множества библиотек лишь немногие выделяются удобством использования, большим функционалом и поддержкой сообщества. По всем этим пунктам оказались подходящими библиотеки pandas, matplotlib, numpy.

#### 1.1.2.1 Pandas

Pandas это высокоуровневая Python библиотека для анализа данных. В экосистеме Python, pandas является наиболее продвинутой и быстроразвивающейся библиотекой для обработки и анализа данных. Она предоставляет готовые к использованию высокопроизводительные структуры данных и инструменты анализа данных.

#### 1.1.2.2 Matplotlib

Matplotlib – библиотека на языке программирования Python для визуализации данных двумерной (2D) графикой (3D графика также поддерживается). Matplotlib является гибким, легко конфигурируемым пакетом, который поддерживает многие виды графиков и диаграмм:

• Графики (line plot)

• Диаграммы разброса (scatter plot)

• Столбчатые диаграммы (bar chart) и гистограммы (histogram)

• Круговые диаграммы (pie chart)

• Ствол-лист диаграммы (stem plot)

• Контурные графики (contour plot)

• Поля градиентов (quiver)

• Спектральные диаграммы (spectrogram)

#### 1.1.2.3 Numpy

Это библиотека с открытым исходным кодом, некогда отделившаяся от проекта SciPy. NumPy является наследником Numeric и NumArray. Основан NumPy на библиотеке LAPAC, которая написана на Fortran.

В силу того, что NumPy базируется на Fortran это быстрая библиотека. А в силу того, что поддерживает векторные операции с многомерными массивами — крайне удобная.

## Получение данных

### Предварительный анализ данных

В ходе предварительного анализа определяют, соответствуют ли имеющиеся данные требованиям, предъявляемым к ним математическими методами (сопоставимость данных, их полнота, однородность и устойчивость).

Согласно Индивидуальному заданию требуется самостоятельно получить Dataset с помощью генератора случайных чисел, соответствующих характеристикам: целые, в диапазоне от -300 до 300, в количестве 1000.

*for i in range(1000):*

*\_series.append(random.randint(-300, 300))*

### Выбор формата хранения данных

Из полученного самостоятельно датасета необходимо сформировать объект Series. Первоначально набор данных находится во встроенном в Питон типе данных list. Для хранения использован тип данных библиотеки pandas – Series.

*def get\_series() -> pd.Series:*

*"""Создание объекта Series"""*

*\_series = []*

*for i in range(1000):*

*\_series.append(random.randint(-300, 300))*

*return pd.Series(\_series)*

### Очистка данных от цифрового мусора

Поскольку данные получены самостоятельно по чётким критериям, очистка их от цифрового мусора не требуется.

## Выводы по разделу

Я считаю, что выбор платформы является неотъемлемой частью правильного и компактного кода. Ещё одним важным пунктом является выбор минимального количества нужных библиотек. Основываясь на информации в данном разделе можно сделать следующие выводы (табл.1).

Табл. 1.

Выводы по разделу 1

|  |  |  |
| --- | --- | --- |
| Выводы | Код и наименование компетенции выпускника программы бакалавриата | Код и наименование индикатора достижения компетенции |
| - Проведен анализ сред разработки и необходимой и информационно-технологической инфра-структуры  - Выбрана современная платформа Python и библиотеки Pandas, Matplotlib для решения прикладных профессиональных задач в соответствие с индивидуальным заданием  - проведен анализ технической документации к среде разработке и к используемым библиотекам как на русском, так и на иностранном языках | ОПК-2. Способен понимать принципы работы современных информационных технологий и программных средств, в том числе отечественного производства, и использовать их при решении задач профессиональной деятельности; | ОПК-2.1.  Знает принципы работы современные информационных технологий и программных средств, в том числе отечественного производства при решении задач профессиональной деятельности.  ОПК-2.2.  Умеет выбирать современные информационные технологии и программные средства, в том числе отечественного производства при решении задач профессиональной деятельности.  ОПК-2.3.  Владеет навыками применения современных информационных технологий и программных средств, в том числе отечественного производства, при решении задач профессиональной деятельности. |
| ОПК-3. Способен решать стандартные задачи профессиональной деятельности на основе информационной и библиографической культуры с применением информационно-коммуникационных технологий и с учетом основных требований информационной безопасности; | ОПК-3.1.  Знает принципы, методы и средства решения стандартных задач профессиональной деятельности на основе информационной и библиографической культуры с применением информационно коммуникационных технологий и с учетом основных требований информационной безопасности.  ОПК-3.2.  Умеет решать стандартные задачи профессиональной деятельности на основе информационной и библиографической культуры с применением информационно коммуникационных технологий и с учетом основных требований информационной безопасности.  ОПК-3.3.  Владеет навыками подготовки обзоров, аннотаций, составления рефератов, научных докладов, публикаций, и библиографии по научно-исследовательской работе с учетом требований информационной безопасности. |
| УК-4. Способен осуществлять деловую коммуникацию в устной и письменной формах на государственном языке Российской Федерации и иностранном(ых) языке(ах) | УК-4.1.  Знает принципы построения устного и письменного высказывания на государственном и иностранном языках; требования к деловой устной и письменной коммуникации.  УК-4.2.  Умеет применять на практике устную и письменную деловую коммуникацию.  УК-4.3.  Владеет методикой составления суждения в межличностном деловом общении на государственном и иностранном языках, с применением адекватных языковых форм и средств. |
| Проведена установка и подключение необходимых библиотек в среду разработки Python для использования в качестве готовых компонентов при разработке алгоритмов и программ | ОПК-5. Способен инсталлировать программное и аппаратное обеспечение для информационных и автоматизированных систем; | ОПК-5.1.  Знает основы системного администрирования, администрирования СУБД, современные стандарты информационного взаимодействия систем  ОПК-5.2.  Умеет выполнять параметрическую настройку информационных и автоматизированных систем.  ОПК-5.3.  Владеет навыками инсталляции программного и аппаратного обеспечения информационных и автоматизированных систем. |
| Осуществлен поиск внешних ресурсов с данными согласно индивидуальному заданию. Выбраны источники с данными. Проведен предварительный анализ данных из внешних источников. (сформированы наборы случайных данных)  На основе системного подхода и анализа структуры внешних данных выбран формат хранения данных. | УК-1. Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач | УК-1.1.  Знает принципы сбора, отбора и обобщения информации, методики системного подхода для решения профессиональных задач.  УК-1.2.  Умеет анализировать и систематизировать разнородные данные, оценивать эффективность процедур анализа проблем и принятия решений в профессиональной деятельности.  УК-1.3.  Владеет навыками научного поиска и практической работы с информационными источниками; методами принятия решений. |
| - Выполнена постановка задачи, сформулирована цель исследования данных.  - Распределены функциональные обязанности: определен объем работ, в части создания программного кода; сформирован круг задач; заданы временные ограничения на разработку  - Выбраны оптимальные способы решения поставленных задач, исходя из имеющихся возможностей разработчика. | УК-2. Способен определять круг задач в рамках поставленной цели и выбирать оптимальные способы их решения, исходя из действующих правовых норм, имеющихся ресурсов и ограничений | УК-2.1.  Знает необходимые для осуществления профессиональной деятельности правовые нормы и методологические основы принятия управленческого решения.  УК-2.2.  Умеет анализировать альтернативные варианты решений для достижения намеченных результатов; разрабатывать план, определять целевые этапы и основные направления работ.  УК-2.3.  Владеет методиками разработки цели и задач проекта; методами оценки продолжительности и стоимости проекта, а также потребности в ресурсах. |
| УК-3. Способен осуществлять социальное взаимодействие и реализовывать свою роль в команде | УК-3.1.  Знает типологию и факторы формирования команд, способы социального взаимодействия.  УК-3.2.  Умеет действовать в духе сотрудничества; принимать решения с соблюдением этических принципов их реализации; проявлять уважение к мнению и культуре других; определять цели и работать в направлении личностного, образовательного и профессионального роста.  УК-3.3.  Владеет навыками распределения ролей в условиях командного взаимодействия; методами оценки своих действий, планирования и управления временем. |
| УК-6. Способен управлять своим временем, выстраивать и реализовывать траекторию саморазвития на основе принципов образования в течение всей жизни | УК-6.1.  Знает основные принципы самовоспитания и самообразования, исходя из требований рынка труда.  УК-6.2.  Умеет демонстрировать умение самоконтроля и рефлексии, позволяющие самостоятельно корректировать обучение по выбранной траектории.  УК-6.3.  Владеет способами управления своей познавательной деятельностью и удовлетворения образовательных интересов и потребностей. |
| УК-5. Способен воспринимать межкультурное разнообразие общества в социально-историческом, этическом и философском контекстах | УК-5.1.  Знает основные категории философии, законы исторического развития, основы межкультурной коммуникации.  УК-5.2.  Умеет вести коммуникацию в мире культурного многообразия и  демонстрировать взаимопонимание между обучающимися – представителями различных культур с соблюдением этических и межкультурных норм.  УК-5.3.  Владеет практическими навыками анализа философских и исторических фактов, оценки явлений культуры; способами анализа и пересмотра своих взглядов в случае разногласий и конфликтов в межкультурной коммуникации. |
| - обоснованы объемы работы по каждой задаче  - заданы сроки представления результатов создания своего фрагмента кода  - заданы временные траектории совместной работы над общими фрагментами кода с учетом требований безопасных условий труда за компьютером и соблюдения режима труда и отдыха для поддержания должного уровня физической подготовленности. | УК-6. Способен управлять своим временем, выстраивать и реализовывать траекторию саморазвития на основе принципов образования в течение всей жизни | УК-6.1.  Знает основные принципы самовоспитания и самообразования, исходя из требований рынка труда.  УК-6.2.  Умеет демонстрировать умение самоконтроля и рефлексии, позволяющие самостоятельно корректировать обучение по выбранной траектории.  УК-6.3.  Владеет способами управления своей познавательной деятельностью и удовлетворения образовательных интересов и потребностей. |
| УК-7. Способен поддерживать должный уровень физической подготовленности для обеспечения полноценной социальной и профессиональной деятельности | УК-7.1.  Знает виды физических упражнений; научно-практические основы физической культуры и здорового образа и стиля жизни.  УК-7.2.  Умеет применять на практике разнообразные средства физической культуры, спорта и туризма для сохранения и укрепления здоровья, психофизической подготовки и самоподготовки к будущей жизни и профессиональной деятельности; использовать творчески средства и методы физического воспитания для профессионально-личностного развития, физического самосовершенствования, формирования здорового образа и стиля жизни.  УК-7.3.  Владеет средствами и методами укрепления индивидуального здоровья, физического самосовершенствования. |
| УК-8. Способен создавать и поддерживать в повседневной жизни и в профессиональной деятельности безопасные условия жизнедеятельности для сохранения природной среды, обеспечения устойчивого развития, в том числе при угрозе и возникновении чрезвычайных ситуаций и военных конфликтов | УК-8.1.  Знает причины, признаки и последствия опасностей, способы защиты от чрезвычайных ситуаций; основы безопасности жизнедеятельности, телефоны служб спасения.  УК-8.2.  Умеет выявлять признаки, причины и условия возникновения чрезвычайных ситуаций; оценивать вероятность возникновения потенциальной опасности для обучающегося и принимать меры по ее предупреждению в условиях образовательного учреждения; оказывать первую помощь в чрезвычайных ситуациях.  УК-8.3.  Владеет методами прогнозирования возникновения опасных или чрезвычайных ситуаций; навыками поддержания безопасных условий жизнедеятельности. |
| УК-9. Способен принимать обоснованные экономические решения в различных областях деятельности | УК-9.1 Знает основные документы, регламентирующие экономическую деятельность; источники финансирования профессиональной деятельности; принципы планирования экономической деятельности  УК-92. Умеет обосновывать принятие экономических решений, использовать методы экономического планирования для достижения поставленных целей  УК-9.3 Владеет навыками применения экономических инструментов |
| УК-10. Способен формировать нетерпимое отношение к коррупционном поведению | УК-10.1 Знает действующие правовые нормы, обеспечивающие борьбу с коррупцией в различных областях жизнедеятельности; способы профилактики коррупции и формирования нетерпимого отношения к ней  УК-10.2 Умеет планировать, организовывать и проводить мероприятия, обеспечивающие формирование гражданской позиции и предотвращение коррупции в обществе  УК-10.3 Владеет навыками взаимодействия в обществе на основе нетерпимого отношения к коррупции |

# 2. КОЛИЧЕСТВЕННЫЙ АНАЛИЗ ДАННЫХ

## 2.1 Получение данных Dataset (данные)

Согласно Индивидуальному заданию требуется самостоятельно получить Dataset с помощью генератора случайных чисел, соответствующих характеристикам: целые, в диапазоне от -300 до 300, в количестве 1000.

for i in range(1000):

\_series.append(random.randint(-300, 300))

Из полученного самостоятельно датасета необходимо сформировать объект Series. Первоначально набор данных находится во встроенном в Питон типе данных list. Для хранения использован тип данных библиотеки pandas – Series.

def get\_series() -> pd.Series:

"""Создание объекта Series"""

\_series = []

for i in range(1000):

\_series.append(random.randint(-300, 300))

return pd.Series(\_series)

series = get\_series()

## 2.2 Расчёт стандартных числовых характеристик для набора данных Series

Согласно Индивидуальному заданию необходимо определить следующие характеристики полученного датасета:

- минимальное значений

- количество повторяющихся значений

- максимальное значение

- сумму чисел

- среднеквадратическое отклонение

Результирующие данные вывести в консоль с пояснениями.

### 2.2.1 Минимальное значение:

Минимальное значение датасета вычислено с помощью встроенной в Питон функции *min(),* возвращающей наименьшей элемент заданой последовательности.

*series\_min = min(series)*

### 2.2.2 Подсчет количества повторяющихся элементов:

Количество повторяющихся элементов есть разница между длиной списка датасета и полученной при помощи встроенной в Питон функции *unique(),* возвращающей количество уникальных элементов любой последовательности.

*series\_repeats = len(series) - len(series.unique())*

### 2.2.3 Максимальное значение:

Максимальное значение датасета вычислено с помощью встроенной в Питон функции *max(),* возвращающей наибольший элемент заданой последовательности.

*series\_max = max(series)*

### 2.2.4 Среднеквадратическое отклонение:

В теории вероятностей и статистике среднеквадратическое (среднеквадратичное) отклонение — наиболее распространённый показатель рассеивания значений случайной величины относительно её математического ожидания (аналога среднего арифметического с бесконечным числом исходов). Обычно он означает квадратный корень из дисперсии случайной величины, но иногда может означать тот или иной вариант оценки этого значения.

В данном случае отклонение рассчитывается встроенной в библиотеку NumPy функцией .*std()*.

*series\_rms = np.std(series)*

### 2.2.5 Сумма значений:

Сумма значений любой последовательности (в данном случае – датасета Series) можно получить с помощью встроенной в Питон функции *sum()*.

*series\_sum = sum(series)*

## 2.3 Выводы по разделу

Количественный анализ данных — это *неотъемлемая* часть прохождения данной практики. Очень важно визуализировать, делать расчёты и выполнять другие преобразования информации при анализе данных. Основываясь на информации в данном разделе можно сделать следующие выводы (табл.2):

Табл.2

Выводы по разделу 2

|  |  |  |
| --- | --- | --- |
| Выводы | Код и наименование компетенции выпускника программы бакалавриата | Код и наименование индикатора достижения компетенции |
| Согласно полученному заданию были определены требуемые характеристики набора данных, после чего сформирован датасет и преобразован в объект Series | ОПК-1. Способен применять естественнонаучные и общеинженерные знания, методы математического анализа и моделирования, теоретического и экспериментального исследования в профессиональной деятельности; | ОПК-1.1.  Знает основы математики, физики, вычислительной техники и программирования.  ОПК-1.2.  Умеет решать стандартные профессиональные задачи с применением естественнонаучных и общеинженерных знаний, методов математического анализа и моделирования.  ОПК-1.3.  Владеет навыками теоретического и экспериментального исследования объектов профессиональной деятельности. |
| ОПК-2. Способен понимать принципы работы современных информационных технологий и программных средств, в том числе отечественного производства, и использовать их при решении задач профессиональной деятельности; | ОПК-2.1.  Знает принципы работы современные информационных технологий и программных средств, в том числе отечественного производства при решении задач профессиональной деятельности.  ОПК-2.2.  Умеет выбирать современные информационные технологии и программные средства, в том числе отечественного производства при решении задач профессиональной деятельности.  ОПК-2.3.  Владеет навыками применения современных информационных технологий и программных средств, в том числе отечественного производства, при решении задач профессиональной деятельности. |
| ОПК-6. Способен анализировать и разрабатывать организационно-технические и экономические процессы с применением методов системного анализа и математического моделирования; | ОПК-6.1.  Знает основы теории систем и системного анализа, дискретной математики, теории вероятностей и математической статистики, методов оптимизации и исследования операций, нечетких вычислений, математического и имитационного моделирования.  ОПК-6.2.  Умеет применять методы теории систем и системного анализа, математического, статистического и имитационного моделирования для автоматизации задач принятия решений, анализа информационных потоков, расчета экономической эффективности и надежности информационных систем и технологий.  ОПК-6.3.  Владеет навыками проведения инженерных расчетов основных показателей результативности создания и применения информационных систем и технологий. |

# 3. ВИЗУАЛИЗАЦИЯ РЕЗУЛЬТАТОВ КОЛИЧЕСТВЕННОГО АНАЛИЗА ДАННЫХ

## 3.1 Визуализация наборов данных с помощью стандартных библиотек по заданным критериям

Линейный график — также известный как линейный график или линейная диаграмма — это график, который использует линии для соединения отдельных точек данных. Линейный график отображает количественные значения за заданный интервал времени.

Гистограмма — способ представления табличных данных в графическом виде — в виде столбчатой диаграммы. Количественные соотношения некоторого показателя представлены в виде прямоугольников, площади которых пропорциональны. В описательной статистике гистограмма распределения — наглядное представление функции плотности вероятности некоторой случайной величины, построенное по выборке. Иногда её называют частотным распределением, так как гистограмма показывает частоту появления измеренных значений параметров объекта.

Matplotlib – библиотека на языке программирования Python для визуализации данных двумерной (2D) графикой (3D графика также поддерживается). Matplotlib является гибким, легко конфигурируемым пакетом, который поддерживает многие виды графиков и диаграмм:

• Графики (line plot)

• Диаграммы разброса (scatter plot)

• Столбчатые диаграммы (bar chart) и гистограммы (histogram)

• Круговые диаграммы (pie chart)

• Ствол-лист диаграммы (stem plot)

• Контурные графики (contour plot)

• Поля градиентов (quiver)

• Спектральные диаграммы (spectrogram)

*import matplotlib.pyplot as plt*

*plt.show()*

## 3.2 Формирование Dataframe из данных Series и добавление столбцов

Dataframe — это табличная структура данных, напоминающая таблицы из Microsoft Excel. Ее главная задача — позволить использовать многомерные Series. Dataframe состоит из упорядоченной коллекции колонок, каждая из которых содержит значение разных типов (числовое, строковое, булевое и так далее).

Формирование датафрейма, согласно с видением архитектуры разрабатываемого модуля вынесено в отдельную функцию, принимающую на вход объект типа Series и возвращающую объект типа DataFrame.

*def get\_dataframe(series:pd.Series) -> pd.DataFrame:*

*"""Создание объекта dataframe"""*

*\_dataframe = pd.DataFrame({"Series": series})*

*\_dataframe["Ascending"] = series.sort\_values(ascending=True, ignore\_index=True)*

*\_dataframe["Descending"] = series.sort\_values(ascending=False, ignore\_index=True)*

*return \_dataframe*

## 3.3 Визуализация данных полученных в результате промежуточного анализа

Для визуализации данных, полученных во время анализа, потребовалось прибегнуть к использованию библиотеки Matplotlib, а конкретно к функциям *.plot()* и *show()*

*def set\_series\_graph(series:pd.Series) -> None:*

*"""Построение графика объекта series"""*

*\_line = series.plot(kind="line", title="series\_line")*

### 3.3.1 Визуализация линейного графика объекта Series

Линейный график объекта Series, визуализированный средствами библиотеки Matplolib (Рис.1)

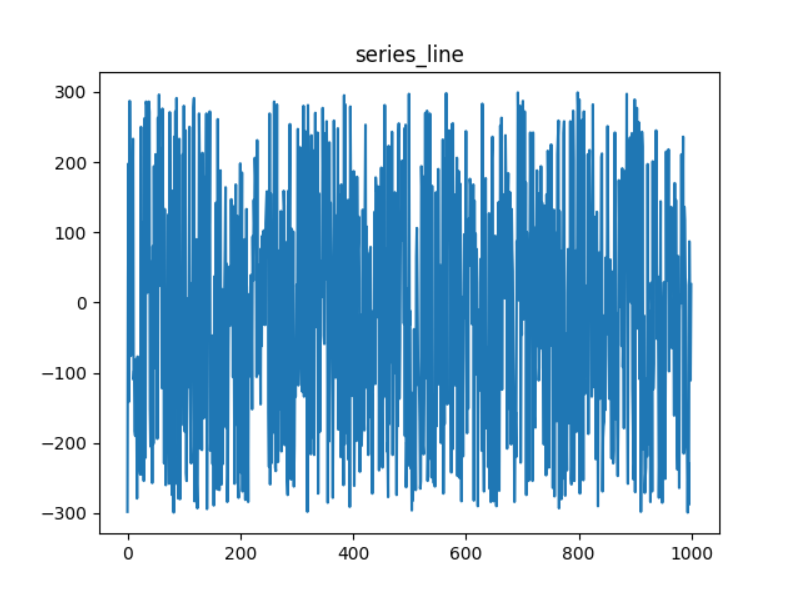


Рис.1

### 3.3.2 Визуализация гистограммы объекта Series

Гистограмма объекта Series, визуализированный средствами библиотеки Matplolib (Рис.2)

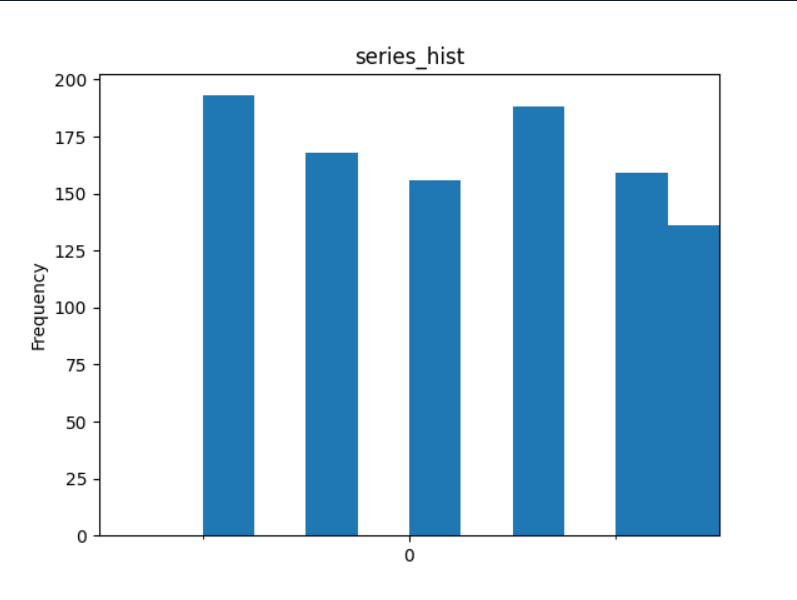


Рис.2

### 3.3.3 Визуализация линейного графика объекта Dataframe

Линейный график объекта Dataframe, визуализированный средствами библиотеки Matplolib (Рис.3)

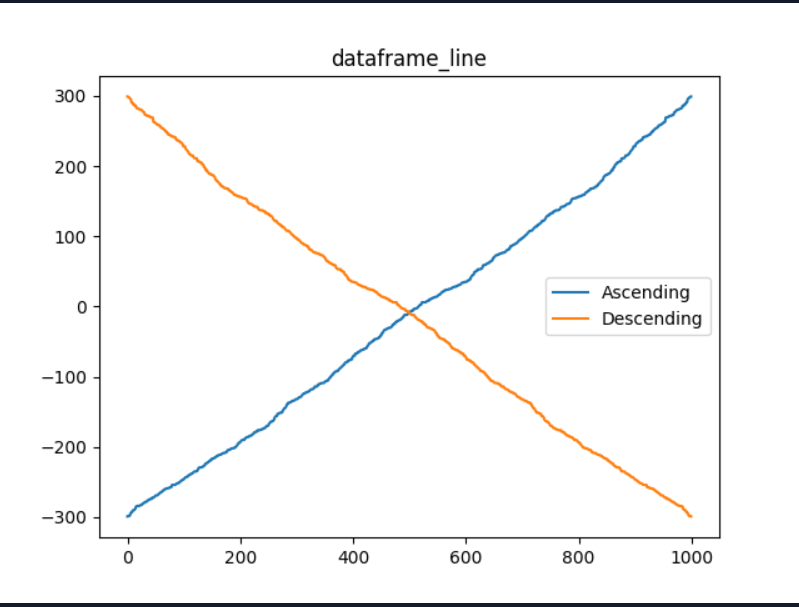


Рис.3

## 3.4 Выводы по разделу

Выводы по разделу представлены в таблице 3:

Табл.3

|  |  |  |
| --- | --- | --- |
| Выводы | Код и наименование компетенции выпускника программы бакалавриата | Код и наименование индикатора достижения компетенции |
| Разработаны алгоритмы и процедуры, использующие стандартные библиотеки для количественного анализа данных.  Эти алгоритмы использовались параллельно с визуализацией.    Для визуализации разработаны процедуры, использующие возможности библиотеки matplotlib    Разработка велась с использованием современных информационных технологий – библиотек для визуализации и обработки данных на языке программирования Python, с использованием документации по использованию библиотек – информационной культуры | ОПК-7. Способен разрабатывать алгоритмы и программы, пригодные для практического применения; | ОПК-7.1.  Знает основные языки программирования и работы с базами данных, операционные системы и оболочки, современные программные среды разработки информационных систем и технологий.  ОПК-7.2.  Умеет применять языки программирования и работы с базами данных, современные программные среды разработки информационных систем и технологий для автоматизации бизнес-процессов, решения прикладных задач различных классов, ведения баз данных и информационных хранилищ.  ОПК-7.3.  Владеет навыками программирования, отладки и тестирования прототипов программно-технических комплексов задач. |

# ЗАКЛЮЧЕНИЕ

В ходе практики я:

– Научился закреплять приобретенные теоретические знания.

– Приобрел навыки научного поиска и практической работы с информационными источниками данных.

– Приобрел практические навыки использования иностранного языка при анализе зарубежных информационных Интернет-источников.

– Приобрел навыки использования современных информационных технологий и программных средств для решения задач анализа данных.

– Приобрел навыки инсталлирования программного обеспечения.

– Сформировал практические навыки по основам алгоритмизации вычислительных процессов и программированию решения задач обработки данных их отладке и тестирования на выбранной языковой платформе.

– Получил навыки экспериментального исследования больших данных и практического решения информационных задач по обработке данных.

– Получил навыки проведения инженерных расчетов.

– Получил навыки самостоятельной работы.

– Получил навыки подготовки отчетной документации.

# СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1 PEP 8 - руководство по написанию кода на Python – Режим доступа: – URL: <https://pythonworld.ru/osnovy/pep-8-rukovodstvo-po-napisaniyu-koda-na-python.html> – Текст : электронный.

2 Полное руководство по PyPlot. – Режим доступа: – URL: <https://pythonru.com/biblioteki/pyplot-uroki> – Текст : электронный.

3 Структуры Pandas Series и DataFrame: создание, методы и функции. – Режим доступа: – URL: <https://pythonru.com/biblioteki/struktury-dannyh-v-pandas> – Текст : электронный.

4 Matplotlib Википедия. – Режим доступа: – URL: <https://ru.wikipedia.org/wiki/Matplotlib> – Текст : электронный.

5 Среднеквадратическое отклонение. – Режим доступа: – URL: <https://ru.wikipedia.org/wiki/Среднеквадратическое_отклонени> – Текст : электронный.