МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное

образовательное учреждение высшего образования

«ЧЕРЕПОВЕЦКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»

Институт (факультет) \_\_\_\_\_\_\_\_\_Институт информационных технологий\_\_\_\_\_\_\_\_\_

Кафедра \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_МПО ЭВМ\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

КУРСОВАЯ РАБОТА

по дисциплине Прикладная статистика

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

на тему Статистический анализ социально-экономических показателей

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Выполнил студент группы

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1ПИб-02-1оп-22

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

направления подготовки (специальности)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Программная инженерия

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

*шифр, наименование*

\_\_\_\_Микуцких Григорий Андреевич\_\_\_\_

*фамилия, имя, отчество*

Руководитель

Гонтарева Ирина Борисовна

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

*фамилия, имя, отчество*

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

доцент

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

*должность*

Дата представления работы

«\_\_\_\_\_\_» \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_20 \_\_\_ г.

Заключение о допуске к защите

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Оценка \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_, \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

количество баллов

Подпись преподавателя\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Череповец, \_2023\_

*год*

Оглавление

[Введение 3](#_Toc154149358)

[1. Описание предметной области 4](#_Toc154149359)

[2. Расчётная часть 5](#_Toc154149360)

[Заключение 14](#_Toc154149361)

[Список литературы 15](#_Toc154149362)

[Приложение 1 16](#_Toc154149363)

# Введение

В последнее время программирование и кодирование являются одними из наиболее перспективных и быстроразвивающихся направлений. С помощью компьютерных языков разного рода и назначений можно написать web-приложения и сайты, мобильные приложения, разрабатывать игры, работать с базами данных и искусственным интеллектом, управлять системами, сетями и прочее. Для того, чтобы начать изучение какого-либо языка, необходимо знать какие из них наиболее популярны и востребованы на сегодняшний день.

Целью работы является статистический анализ самых популярных компьютерных языков с помощью методов прикладной статистики. Объектом исследования являются интервальные относительные показатели популярности языка, страны, года выпуска и типизации его переменных за третьи кварталы 2023, 2020, 2017 и 2014 годов двадцати пяти самых используемых языков за третий квартал 2023 на сервисах GitHub и Stack Overflow.

Прикладная статистика – это наука о том, как обрабатывать данные. Методы прикладной статистики активно применяются в технических исследованиях, экономике, теории и практике управления, социологии, медицине, геологии, истории и т.д. С результатами наблюдений, измерений, испытаний, опытов, с их анализом имеют дело специалисты во всех отраслях практической деятельности, почти во всех областях теоретических исследований [1].

Предмет статистики –общие методы сбора и обработки данных.

В рамках курсовой работы будут исследованы зависимость популярности языка от года его выпуска, в каких странах в большем количестве выпускают популярные языки, зависимость типизации переменных языка от временного периода его выпуска, построены ряды динамики для некоторых самых популярных языков.

Источником данных является сайт «Languish — Тенденции в области языков программирования» [3].

# Описание предметной области

Единица наблюдения – это часть объекта наблюдения (генеральной совокупности), имеющая свои признаки, подлежащие учёту, сравнению.

Признаки – свойства единиц наблюдения.

Описательные (качественный) признаки – признаки, выражающиеся словесно (например, форма собственности предприятия, вид покупаемой продукции, направление подготовки и т. д.). Описательные признаки подразделяют на номинальные, которые нельзя упорядочить, ранжировать (например, национальность, отраслевая принадлежность предприятия и др.), и порядковые, которые можно ранжировать (например, балл успеваемости студента, тарифный разряд, рейтинги компаний и др.) [2].

Количественные признаки – признаки, отдельные значения которых имеют числовое выражение (например, площадь территории области, стоимость фондов предприятия, цена товара и т. д.) [2].

Данные брались с сайта «https://tjpalmer.github.io/languish/». Сайт «в настоящее время собирает данные из сервисов GitHub и Stack Overflow, измеряет текущую активность / скорость, а не общую (включая прошлую) популярность, по умолчанию используется простой "Средний балл"» [3] компьютерных языков.

С ресурса были взяты первые двадцать пять языков, наиболее популярные в 2023Q3, и их относительные показатели популярности за третий квартал 2014, 2017, 2020 и 2023 годов. Из глобальной сети Интернет были взяты год, страна выпуска и типизация переменных данных языков. Таблица с исходными данными представлена на рис. 1.



Рис. 1. Самые популярные компьютерные языки в 2023Q3

Единицами наблюдения в исследовании являются двадцать пять самых популярных компьютерных языков. Их качественными признаками – страна выпуска и типизация переменных, количественными – относительные показатели популярности и год выпуска языков.

# Расчётная часть

Чтобы узнать в каких странах выпускаются самые популярные языки, была составлена сводная таблица, показывающая суммарную популярность выпущенных языков в странах их выпуска в заданные интервалы (рис. 2 - 5).



Рис. 2. Страны-лидеры выпуска популярных языков в 2014Q3



Рис. 3. Страны-лидеры выпуска популярных языков в 2017Q3



Рис. 4. Страны-лидеры выпуска популярных языков в 2020Q3



Рис. 5. Страны-лидеры выпуска популярных языков в 2023Q3

Метод группировки – разбиение общей совокупности единиц объекта наблюдения по одному или нескольким существенным признакам на однородные группы, различающиеся между собой в количественном и качественном отношении и позволяющие выделить социально-экономические типы, изучить структуру совокупности и проанализировать связи между отдельными признаками [4].

В исследовании метод группировки был применён для определения популярности языков в третьем квартале 2014, 2017, 2020 и 2023 годов, выпущенных с 1971 по 2020, гг. с шагом в 10 лет, (рис. Х-Х).



Рис. 6. Популярность языков, выпущенных в разные периоды, в 2014Q3



Рис. 7. Популярность языков, выпущенных в разные периоды, в 2017Q3



Рис. 8. Популярность языков, выпущенных в разные периоды, в 2020Q3



Рис. 9. Популярность языков, выпущенных в разные периоды, в 2023Q3

Можно сделать вывод, что языки, выпущенные в период 1991-2000 гг. являются наиболее популярными с 2014Q3 по 2023Q3, практически не теряя своих позиций. Более «молодые» языки также являются достаточно популярными и активно «отбивают» рынок, поэтому можно говорить о продолжении роста их популярности в будущем.

Статистическое изучение взаимосвязей относится к анализу и изучению связей между переменными и факторами на основе статистических методов и моделей. Оно позволяет определить, есть ли статистически значимая зависимость между двумя или более переменными, и в какой степени они взаимосвязаны.

В работе была исследована взаимосвязь между периодом выпуска языка и типизацией его переменных (рис. 10 -13).



Рис. 10. Страны-лидеры выпуска языков разных типизаций



Рис. 11. Популярность типизаций языков в странах-лидерах



Рис. 12. Показатели дисперсии



Рис. 13. Связь между типизацией и периодом выпуска

Ряды динамики – это последовательность упорядоченных во времени количественных статистических величин, характеризующих развитие изучаемого явления или процесса. Анализ рядов динамики позволяет выделить тенденции и закономерности развития (т. е. тренд) [2].

В исследовании ряды динамики были составлены для пяти языков: Python, JavaScript, TypeScript, C и C# (рис. 14-23).



Рис. 14. Динамика роста «Python»



Рис. 15. График роста «Python»



Рис. 16. Динамика роста «JavaScript»



Рис. 17. График роста «JavaScript



Рис. 18. Динамика роста «TypeScript»



Рис. 19. График роста «TypeScript»



Рис. 20. Динамика роста «C»



Рис. 21. График роста «C»



Рис. 22. Динамика роста «C#»



Рис. 23. График роста «C#»

Можно сделать вывод, что популярность языков может как очень сильно меняться, так и оставаться примерно на одном уровне на протяжении многих лет. Подобное поведение является следствием поддержки языка разработчиками, спектром задач, для решения которых и создавался язык, а также появлением на рынке более удобных и быстрых аналогов.

Например, синтаксис языка «C» является негласным стандартом для создания многих других языков (обращение к элементу массива через квадратные скобки, форма вызова функции, знак разделения инструкций и т.д.), поэтому их синтаксис называют C-подобными (например «R», «C#», «JavaScript», «PHP» и т.д.). Сам же «C», несмотря на свою «старость», до сих пор используется для написания драйверов устройств, баз данных, open source-проектов и является фундаментом для огромного числа программного обеспечения.

# Заключение

В ходе выполнения курсовой работы был проведён статистический анализ самых популярных компьютерных языков. Были сделаны выводы:

* популярность языка не зависит от года его выпуска;
* самой популярной страной для выпуска ведущих компьютерных языков является США;
* есть взаимосвязь между временным периодом выпуска языка и типизацией его переменных;
* построены ряды динамики популярности для языков «Python», «JavaScript», «TypeScript», «C» и «C#».

Также можно отметить, что популярность языка зависит от множества факторов, включая удобство синтаксиса языка, тех задач, для решения которых он и создавался, мощность, доступность обучения и простота использования. Также популярность может быть обусловлена спросом на определенные технологии и приложения, которые могут быть разработаны с использованием конкретных языков программирования.

На данный момент самым популярным компьютерным языком является Python. Он отличается простым и понятным синтаксисом, широким спектром применения в различных областях, большим количеством библиотек и фреймворков, а также большим сообществом разработчиков, что сегодня делает его одним из наиболее востребованных языков программирования.

# Список литературы

[1] Орлов А.И. Прикладная статистика. Учебник. / А.И.Орлов.- М.: Издательство «Экзамен», 2004. - 656 с.

[2] Полякова, В. В. Прикладная статистика: методы анализа эмпирической

информации : учебно-методическое пособие / В. В. Полякова,

Н. В. Шаброва ; Министерство науки и высшего образования

Российской Федерации, Уральский федеральный университет

им. первого Президента России Б. Н. Ельцина. – Екатеринбург :

Изд-во Урал. ун-та, 2020. – 188 с. : ил. – 30 экз. – ISBN 978-5-7996-

3021-8. – Текст : непосредственный.

[3] https://tjpalmer.github.io/languish/

[4] https://studfile.net/preview/5739423/page

# Приложение 1

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное

образовательное учреждение высшего образования

«ЧЕРЕПОВЕЦКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»

­­\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_Институт информационных технологий\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

наименование института (факультета)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_МПО ЭВМ\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

наименование кафедры

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_Прикладная статистика\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

наименование дисциплины в соответствии с учебным планом

УТВЕРЖДАЮ

Зав. кафедрой МПО ЭВМ,

д.т.н., профессор \_\_\_\_\_\_\_\_\_\_ Ершов Е.В.

«\_\_\_» \_\_\_\_\_\_\_\_\_\_\_\_\_ 20\_\_ г.

Статистический анализ социально-экономических показателей

Техническое задание на курсовую работу

Листов \_5\_

Руководитель Гонтарева И. Б.\_\_\_\_\_\_\_\_

Ф.И.О. преподавателя

Исполнитель

студент студент гр. 1ПИб-02-1оп-22

группа

\_\_\_\_\_\_Микуцких Г.А.\_\_\_\_

Фамилия, имя, отчество

\_2023\_ год

Введение

Проблема изучения компьютерных языков в последние десятилетия очень актуальна. По разным данным насчитывают от семисот до десяти тысяч различных языков за всё время существования вычислительной техники. Поэтому важно понимать, выбор какого языка для изучения наиболее выгодно и перспективно на сегодняшний день.

1. Основания для исследования

Основанием для исследования является задание на курсовую работу по дисциплине "Прикладная статистика", выданное на кафедре МПО ЭВМ ИИТ ЧГУ.

Дата утверждения: 13 декабря 2023 года.

Наименование темы исследования: «Статистический анализ компьютерных языков».

1. Назначение исследования

В ходе исследования будут определены зависимость популярности компьютерного языка от года его выпуска, в каких странах выпускают самые популярные языки, зависимость типизации переменных языка от временного периода его выпуска, динамика популярности актуальных языков.

3. Требование к документации

3.1. Содержание расчётно-пояснительной записки

Документация исследования должна содержать расчётно-пояснительную записку, которая содержит:

Титульный лист

Оглавление

Введение

1. Описание предметной области

2. Расчётная часть

Заключение

Список литературы

Приложения.

3.2. Требования к оформлению

Элементы курсовой работы оформлены в соответствии с параметрами ниже (табл. П1.1).

Таблица П1.1

Требования к оформлению

|  |  |
| --- | --- |
| Элемент | Требования |
| Документ | Печать на отдельных листах формата А4 (210х297 мм); оборотная сторона не заполняется; листы нумеруются. Печать возможна ч/б. |
| Страницы | 1. Ориентация – книжная; отдельные страницы, при необходимости, альбомная.  2. Поля: верхнее, нижнее – по 2 см, левое – 3 см, правое – 1 см. |
| Абзацы | Межстрочный интервал – 1.5, перед и после абзаца – 0. |
| Шрифты | Кегль – 14. В таблицах шрифт 12. |
| Рисунки | Подписывается под ним по центру: Рис. Х. Название  В приложениях: Рис. П1.3. Название. |
| Таблицы | 1. Подписывается: над таблицей, выравнивание по правому: «Таблица Х».  2. В следующей строке по центру Название  3. Надписи в «шапке» (имена столбцов, полей) – по центру.  4. В теле таблицы (записи) текстовые значения – выровнены по левому краю, числа, даты – по правому. |

4. Стадии и этапы исследования

Исследование будет проводится в течение следующих этапов (табл. П1.2):

Таблица П1.2

Стадии и этапы исследования

|  |  |  |  |
| --- | --- | --- | --- |
| Наименование  этапа разработки | Сроки разработки | Результат выполнения | Отметка о выполнении |
| Оформление технического задания | 13 декабря 2023 г. | Составлено техническое задание |  |
| Проведение исследования | 14 – 20 декабря 2023 г. | Получены результаты исследования |  |
| Оформление расчётно-поясни-тельной записки | 21 – 26 декабря 2023 г. | Составлена расчётно-поясни-тельная записка |  |

5. Порядок контроля и приемки

Курсовая работа будет оформляться в течение этапов в табл. П1.3.

Таблица П1.3

Порядок контроля и приемки

|  |  |  |  |
| --- | --- | --- | --- |
| Наименование  контрольного этапа  выполнения  курсовой работы | Сроки  контроля | Результат выполнения | Отметка о приемке  результата  контрольного этапа |
| 1 | 2 | 3 | 4 |
| 1 | 2 | 3 | 4 |
| Сдача техничес-кого задания | 13 декабря 2023 г. | Техническое задание принято |  |
| Сдача расчётно-пояснительной записки | 26 декабря 2023 г. | Расчётно-пояснительная записка принята |  |
| Защита курсовой работы | 27 декабря 2023 г. | Курсовая работа защищена |  |