

Министерство науки и высшего образования Российской Федерации
 федеральное государственное автономное
 образовательное учреждение высшего образования
 «Национальный исследовательский Томский политехнический университет» (ТПУ)

Инженерная школа информационных технологий и робототехники
 Направление подготовки 09.04.04 Программная инженерия
 Отделение школы (НОЦ) Информационных технологий

МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ

Тема работы
Разработка информационного портала для оценки уровня диверсификации моногородов и визуализации данных

УДК 004.738.5:004.774:352.07

Студент

Группа	ФИО	Подпись	Дата
8ПМ9И	Копасов Дмитрий Владимирович		21.06.2021 г.

Руководитель ВКР

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОИТ ИШИТР	Соколова В.В.	к.т.н.		21.06.2021 г.

КОНСУЛЬТАНТЫ ПО РАЗДЕЛАМ:

По разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОСГН ШБИП	Гончарова Н. А.	к.э.н.		

По разделу «Социальная ответственность»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ООД ШБИП	Антоневич О. А.	к.б.н.		

ДОПУСТИТЬ К ЗАЩИТЕ:

Руководитель ООП	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОИТ ИШИТР	Савельев А.О.	к.т.н.		

ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОСВОЕНИЯ ООП
по направлению 09.04.04 «Программная инженерия»

Код компетенции	Наименование компетенции
Универсальные компетенции	
УК(У)-1	Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий
УК(У)-2	Способен управлять проектом на всех этапах его жизненного цикла
УК(У)-3	Способен организовывать и руководить работой команды, вырабатывая командную стратегию для достижения поставленной цели
УК(У)-4	Способен применять современные коммуникативные технологии, в том числе на иностранном (-ых) языке (-ах), для академического и профессионального взаимодействия
УК(У)-5	Способен анализировать и учитывать разнообразие культур в процессе межкультурного взаимодействия
УК(У)-6	Способен определять и реализовывать приоритеты собственной деятельности и способы ее совершенствования на основе самооценки
Общепрофессиональные компетенции	
ОПК(У)-1	Способен самостоятельно приобретать, развивать и применять математические, естественно-научные, социально-экономические и профессиональные знания для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте
ОПК(У)-2	Способен разрабатывать оригинальные алгоритмы и программные средства, в том числе с использованием современных интеллектуальных технологий, для решения профессиональных задач
ОПК(У)-3	Способен анализировать профессиональную информацию, выделять в ней главное, структурировать, оформлять и представлять в виде аналитических обзоров с обоснованными выводами и рекомендациями
ОПК(У)-4	Способен применять на практике новые научные принципы и методы исследований
ОПК(У)-5	Способен разрабатывать и модернизировать программное и аппаратное обеспечение информационных и автоматизированных систем
ОПК(У)-6	Способен самостоятельно приобретать с помощью информационных технологий и использовать в практической деятельности новые знания

	и умения, в том числе в новых областях знаний, непосредственно не связанных со сферой деятельности
ОПК(У)-7	Способен применять при решении профессиональных задач методы и средства получения, хранения, переработки и трансляции информации посредством современных компьютерных технологий, в том числе, в глобальных компьютерных сетях
ОПК(У)-8	Способен осуществлять эффективное управление разработкой программных средств и проектов
Профессиональные компетенции	
ПК(У)-1	Способен к созданию вариантов архитектуры программного средства
ПК(У)-2	Способен разрабатывать и администрировать системы управления базам данных
ПК(У)-3	Способен управлять процессами и проектами по созданию (модификации) информационных ресурсов
ПК(У)-4	Способен проектировать и организовывать учебный процесс по образовательным программам с использованием современных образовательных технологий
ПК(У)-5	Способен осуществлять руководство разработкой комплексных проектов на всех стадиях и этапах выполнения работ

Министерство науки и высшего образования Российской Федерации
 федеральное государственное автономное
 образовательное учреждение высшего образования
 «Национальный исследовательский Томский политехнический университет» (ТПУ)

Инженерная школа информационных технологий и робототехники
 Направление подготовки (специальность) 09.04.04 Программная инженерия
 Отделение школы (НОЦ) Информационных технологий

УТВЕРЖДАЮ:
 Руководитель ООП
 _____ Савельев А.О.
 (подпись) (дата) (Ф.И.О.)

ЗАДАНИЕ
на выполнение выпускной квалификационной работы

В форме:

Магистерской диссертации (бакалаврской работы, дипломного проекта/работы, магистерской диссертации)

Студенту:

Группа	ФИО
8ПМ9И	Копасову Дмитрию Владимировичу

Тема работы:

Разработка информационного портала для оценки уровня диверсификации моногородов и визуализации данных	
Утверждена приказом директора (дата, номер)	№ 40-5/с от 09.02.2021

Срок сдачи студентом выполненной работы:	15.06.2021
--	------------

ТЕХНИЧЕСКОЕ ЗАДАНИЕ:

Исходные данные к работе <i>(наименование объекта исследования или проектирования; производительность или нагрузка; режим работы (непрерывный, периодический, циклический и т. д.); вид сырья или материал изделия; требования к продукту, изделию или процессу; особые требования к особенностям функционирования (эксплуатации) объекта или изделия в плане безопасности эксплуатации, влияния на окружающую среду, энергозатратам; экономический анализ и т. д.).</i>	Объектом исследования являются монопрофильные муниципальные образования (моногорода) Российской Федерации. В качестве исходных данных использовались данные отчётности экономической деятельности предприятий исследуемых моногородов.
--	---

<p>Перечень подлежащих исследованию, проектированию и разработке вопросов</p> <p><i>(аналитический обзор по литературным источникам с целью выяснения достижений мировой науки техники в рассматриваемой области; постановка задачи исследования, проектирования, конструирования; содержание процедуры исследования, проектирования, конструирования; обсуждение результатов выполненной работы; наименование дополнительных разделов, подлежащих разработке; заключение по работе).</i></p>	<ol style="list-style-type: none"> 1. Обработка исходных данных. 2. Расчёт экономических показателей. 3. Разработка веб-приложения для визуализации рассчитанных данных моногородов. 4. Исследование эффективности регрессионных моделей для оценки доли моноотрасли в моногородах. 5. Работа над разделом по финансовому менеджменту, ресурсоэффективности и ресурсосбережения. 6. Работа над разделом по социальной ответственности. 7. Работа над разделом на английском языке.
--	---

<p>Перечень графического материала</p> <p><i>(с точным указанием обязательных чертежей)</i></p>	<ol style="list-style-type: none"> 1. Фрагмент исходного набора данных. 2. Примеры рассчитанных экономических показателей. 3. Гистограммы распределения рассчитанных экономических показателей. 4. Скриншоты интерфейса разработанного веб-приложения. 5. Календарный план выполнения проекта.
--	---

Консультанты по разделам выпускной квалификационной работы
(с указанием разделов)

Раздел	Консультант
Основная часть	Доцент ОИТ ИШИТР, к.т.н., доцент Соколова В. В.
Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	Доцент ОСГН ШБИП, к.э.н., доцент Гончарова Н. А.
Социальная ответственность	Доцент ООД ШБИП, к.б.н., доцент Антоневиц О. А.
Английский язык	Доцент ОИЯ ШБИП, к.ф.н., доцент Сидоренко Т. В.

Названия разделов, которые должны быть написаны на русском и иностранном языках:

Глава 1. Подготовка данных и расчёт экономических показателей

Дата выдачи задания на выполнение выпускной квалификационной работы по линейному графику	01.03.2021
---	------------

Задание выдал руководитель:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОИТ ИШИТР	Соколова В. В.	к.т.н., доцент		01.03.2021

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8ПМ9И	Копасов Дмитрий Владимирович		01.03.2021

Министерство науки и высшего образования Российской Федерации
 федеральное государственное автономное
 образовательное учреждение высшего образования
 «Национальный исследовательский Томский политехнический университет» (ТПУ)

Инженерная школа информационных технологий и робототехники
 Направление подготовки (специальность) 09.04.04 Программная инженерия
 Уровень образования магистратура
 Отделение школы (НОЦ) Информационных технологий
 Период выполнения весенний семестр 2020 /2021 учебного года

Форма представления работы:

Магистерская диссертация

(бакалаврская работа, дипломный проект/работа, магистерская диссертация)

КАЛЕНДАРНЫЙ РЕЙТИНГ-ПЛАН выполнения выпускной квалификационной работы

Срок сдачи студентом выполненной работы:	15.06.2021
--	------------

Дата контроля	Название раздела (модуля) / вид работы (исследования)	Максимальный балл раздела (модуля)
01.06.2021	Основная часть	70
01.06.2021	Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	10
01.06.2021	Социальная ответственность	10
01.06.2021	Английский язык	10

СОСТАВИЛ:

Руководитель ВКР

Должность	ФИО	Ученая степень, звание	Подпись	Дата
доцент ОИТ ИШИТР	Соколова В. В.	К.Т.Н.		

СОГЛАСОВАНО:

Руководитель ООП

Должность	ФИО	Ученая степень, звание	Подпись	Дата
доцент ОИТ ИШИТР	Савельев А. О.	К.Т.Н.		

**ЗАДАНИЕ ДЛЯ РАЗДЕЛА
«ФИНАНСОВЫЙ МЕНЕДЖМЕНТ, РЕСУРСОЭФФЕКТИВНОСТЬ И
РЕСУРСОСБЕРЕЖЕНИЕ»**

Студенту:

Группа	ФИО
8ПМ9И	Копасову Дмитрию Владимировичу

Школа	ИШИТР	Отделение (НОЦ)	ОИТ
Уровень образования	Магистратура	Направление/специальность	09.04.04 Программная инженерия

Исходные данные к разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»:

1. <i>Стоимость ресурсов научного исследования (НИ): материально-технических, энергетических, финансовых, информационных и человеческих</i>	1. <i>Затраты на зарплату –145994.51 рублей</i> 2. <i>Бюджет проекта – 230457.29 рублей</i>
2. <i>Нормы и нормативы расходования ресурсов</i>	1. <i>Стоимость элект-энергии –1105.92 рублей</i>
3. <i>Используемая система налогообложения, ставки налогов, отчислений, дисконтирования и кредитования</i>	1. <i>Отчисления во внебюджетные фонды – 39564.51 рублей</i> 2. <i>Накладные расходы – 43792.35 рублей</i>

Перечень вопросов, подлежащих исследованию, проектированию и разработке:

1. <i>Анализ конкурентных технических решений (НИ)</i>	<i>Расчет конкурентоспособности SWOT-анализ</i>
2. <i>Формирование плана и графика разработки и внедрения (НИ)</i>	<i>Структура работ</i> <i>Определение трудоемкости</i> <i>Разработка графика проведения исследования</i>
3. <i>Составление бюджета (НИ)</i>	<i>Расчёт бюджетной стоимости проекта</i>
4. <i>Оценка ресурсной, финансовой, бюджетной эффективности (НИ)</i>	<i>Интегральные финансовые показатели разработки и ресурсоэффективности</i>

Перечень графического материала (с точным указанием обязательных чертежей):

1. <i>Оценка конкурентоспособности проекта</i>
2. <i>Матрица SWOT</i>
3. <i>Календарный график работ проекта</i>
4. <i>Бюджет проекта</i>
5. <i>Основные показатели эффективности проекта</i>

Дата выдачи задания для раздела по линейному графику	01.03.2021
---	-------------------

Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОСГН ШБИП	Гончарова Наталья Александровна	к.э.н.		

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8ПМ9И	Копасов Дмитрий Владимирович		

ЗАДАНИЕ ДЛЯ РАЗДЕЛА «СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ»

Студенту:

Группа	ФИО
8ПМ9И	Копасову Дмитрию Владимировичу

Школа	ИШИТР	Отделение (НОЦ)	ОИТ
Уровень образования	Магистратура	Направление/специальность	09.04.04 Программная инженерия

Тема ВКР:

Разработка информационного портала для оценки уровня диверсификации моногородов и визуализации данных	
Исходные данные к разделу «Социальная ответственность»:	
1. Характеристика объекта исследования (вещество, материал, прибор, алгоритм, методика, рабочая зона) и области его применения	Объект исследования: разрабатываемый информационный портал Область применения: оценка состояния моногородов Рабочая зона: сидячее место за компьютером в офисном помещении
Перечень вопросов, подлежащих исследованию, проектированию и разработке:	
1. Правовые и организационные вопросы обеспечения безопасности: <ul style="list-style-type: none"> – специальные (характерные при эксплуатации объекта исследования, проектируемой рабочей зоны) правовые нормы трудового законодательства; – организационные мероприятия при компоновке рабочей зоны. 	<ul style="list-style-type: none"> – ГОСТ 12.2.032-78 – требования к рабочему месту при выполнении работ сидя – ТК РФ ФЗ-197 – организация труда в течение рабочего времени – СП 2.4.3648-20 – "Санитарно-эпидемиологические требования к организациям воспитания и обучения, отдыха и оздоровления детей и молодежи"
2. Производственная безопасность: 2.1. Анализ выявленных вредных и опасных факторов 2.2. Обоснование мероприятий по снижению воздействия	Вредные факторы: <ul style="list-style-type: none"> – отклонение показателей микроклимата; – повышенный уровень шума; – недостаточная освещенность рабочей зоны; – умственное переутомление, в том числе вызванное информационной нагрузкой. Опасные факторы: <ul style="list-style-type: none"> – повышенное значение напряжения в электрической цепи, замыкание которой может произойти через тело человека.
3. Экологическая безопасность:	Атмосфера, гидросфера, литосфера: возможны негативные последствия вследствие неправильной утилизации используемой электронной техники (используемого в работе компьютера и его комплектующих)

	Требования к утилизации определены в ГОСТ Р 55102-2012
4. Безопасность в чрезвычайных ситуациях:	Потенциальные чрезвычайные ситуации: пожар

Дата выдачи задания для раздела по линейному графику	01.03.2021
---	-------------------

Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ООД ШБИП	Антоневич Ольга Алексеевна	к.б.н.		

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
8ПМ9И	Копасов Дмитрий Владимирович		

Реферат

Выпускная квалификационная работа содержит 102 страницы, 22 рисунка, 21 таблицу, 18 источников и 4 приложения.

Ключевые слова: моногород, уровень диверсификации, веб-приложение, визуализация, регрессионный анализ, линейная регрессия, «случайный лес».

Цель работы: разработка информационного портала для оценки уровня диверсификации моногородов и визуализации данных.

Результатом работы является веб-портал, функционал которого позволяет визуализировать моногорода и их рассчитанные показатели на карте, производить поиск и фильтрацию моногородов по параметрам. Также разработка содержит модуль построения регрессионной модели для предсказания доли моноотрасли на основании ряда рассчитанных входных параметров.

Практической значимостью выпускной квалификационной работы являются упрощение процесса проведения расчётов показателей диверсификации экономики моногородов, визуализация данных с помощью интерактивной карты и использование регрессионной модели для оценки долей моноотрасли для моногородов, в которых недостаточно данных для расчёта данных показателей.

Оглавление

Реферат	10
Оглавление	11
Список терминов и сокращений	13
Введение	16
Глава 1. Подготовка данных и расчёт экономических показателей.....	17
1.1. Описание исходных данных	17
1.2. Использование веб-скрейпинга для извлечения дополнительных данных	20
1.3. Расчёт экономических показателей моногородов	22
1.4. Выводы по разделу	29
Глава 2. Разработка информационного портала.....	30
2.1. Разработка модуля визуализации моногородов.....	31
2.2. Разработка модуля визуализации распределения моногородов по параметрам	34
2.3. Выводы по разделу	36
Глава 3. Разработка модуля построения регрессионной модели.....	37
3.1. Модуль построения регрессионной модели в веб-приложении.....	37
3.2. Сравнение качества построенных регрессионных моделей.....	39
3.3. Выводы по разделу	43
Глава 4. Финансовый менеджмент, ресурсоэффективность и ресурсосбережение.....	45
4.1. Анализ конкурентных технических решений.....	45
4.2. SWOT-анализ	46
4.3. Формирование проекта.....	47
4.4. Бюджет проекта	50
4.4.1. Расчёт материальных затрат проекта	51
4.4.2. Расчёт заработной платы участников проекта	52
4.4.3. Дополнительная заработная плата	53
4.4.4. асчёт расходов на отчисления во внебюджетные фонды.....	54
4.4.5. Накладные расходы	54
4.4.6. Формирование бюджета проекта.....	55
4.5. Определение ресурсной (ресурсосберегающей), финансовой, бюджетной, социальной и экономической эффективности исследования.....	55
4.6. Заключение по разделу.....	58
Глава 5. Социальная ответственность	59
5.1. Правовые и организационные вопросы обеспечения безопасности.....	59

5.1.1. Специальные правовые нормы трудового законодательства	59
5.1.2. Организационные мероприятия при компоновке рабочей зоны.....	60
5.2. Производственная безопасность	61
5.2.1. Анализ вредных и опасных факторов	61
5.2.1.1. Отклонение показателей микроклимата	62
5.2.1.2. Повышенный уровень шума.....	63
5.2.1.3. Недостаток необходимого естественного или искусственного освещения.....	66
5.2.1.4. Описание показателей освещённости рабочего помещения.....	67
5.2.1.5. Умственное переутомление.....	68
5.2.1.6. Повышенное значение напряжения в электрической цепи.....	69
5.3. Экологическая безопасность.....	70
5.4. Безопасность в чрезвычайных ситуациях.....	71
Заключение.....	74
Список публикаций и научных достижений	75
Список используемых источников	76
Приложение А.....	78
Приложение Б	89
Приложение В.....	92
Приложение Г	102

Список терминов и сокращений

1. Django – фреймворк для разработки веб-приложений на языке Python.
2. Excel-файл – файловый формат хранения электронных таблиц, используемый в среде офисных приложений Microsoft.
3. Firebase – комплексная платформа для разработки и размещения веб-приложений от компании Google.
4. JS – JavaScript – мультипарадигменный язык программирования, преимущественно используемый при разработке веб-приложений.
5. Python – высокоуровневый язык программирования общего назначения.
6. PythonAnywhere – облачная платформа, позволяющая разрабатывать и размещать веб-приложения, написанные на языке программирования Python и ряде его веб-фреймворков.
7. React – JavaScript-библиотека с открытым исходным кодом, предназначенная для разработки пользовательских интерфейсов веб-приложений.
8. Веб-приложение – клиент-серверное приложение, в котором клиент взаимодействует с веб-сервером при помощи браузера.
9. Веб-скрейпинг – Web scraping – технология получения веб-данных путем извлечения их со страниц веб-ресурсов, обычно относится к автоматизированным процессам, реализованным с помощью кода, который выполняет GET-запросы на целевой сайт.
10. Диверсификации – мера разнообразия видов деятельности.
11. Индекс Херфиндаля-Хиршмана – экономический показатель, использующийся для оценки степени монополизации отрасли.
12. Корреляция – статистическая взаимосвязь двух или более величин, выражаемая в закономерном взаимном изменении этих величин.

13. Коэффициент детерминации (R-squared) – доля дисперсии зависимой переменной, объясняемая моделью зависимости от набора независимых величин.

14. Линейная регрессия – регрессионная модель зависимости одной переменной от другой или нескольких других переменных с использованием линейной функцией зависимости.

15. Машинное обучение – обширный подраздел искусственного интеллекта, изучающий методы построения алгоритмов, способных обучаться.

16. Моногород – поселение городского типа с численностью населения более 3 000 человек, из которых не менее 20% работают на предприятиях одной компании.

17. Моноотрасль – основная экономическая отрасль градообразующих предприятий моногорода.

18. ОКАТО – общероссийский классификатор объектов административно-территориального деления, который используется для обеспечения сопоставимости, достоверности и удобства автоматизированной обработки информации в статистике и экономике в разрезах административно-территориального деления.

19. ОКВЭД – общероссийский классификатор видов экономической деятельности, который определяет основные и дополнительные виды экономической деятельности, осуществляемых предприятиями и хозяйствующими субъектами.

20. ОКПО – общероссийский классификатор предприятий и организаций, который представляет единый список юридических лиц и их филиалов и представительств, индивидуальных предпринимателей, а также организации, осуществляющие деятельность без образования юридического лица.

21. Парсинг – процесс сбора и извлечения данных из теста с определённой структурой с последующей обработкой и анализом.

22. Регрессионная модель – статистическая модель (структура), основанная на уравнении регрессии, или системе регрессионных уравнений, связывающих входные (независимые) величины с выходными (зависимыми) переменными.

23. Регрессия – математическая зависимость (формула), описывающая взаимосвязь набора независимых переменных и зависящей от них конечной величины.

24. «Случайный лес» (Random Forest) – алгоритм машинного обучения, суть которого заключается в использовании большого набора деревьев решений, каждый из которых строится независимо друг от друга, для решения задач классификации, регрессии и кластеризации.

25. Средняя абсолютная ошибка (MAE) – ошибка результатов построения регрессионной модели, рассчитываемая как среднее значение суммы модулей разности предсказанного значения и исходного.

26. Средняя квадратичная ошибка (MSE) – ошибка результатов построения регрессионной модели, рассчитываемая как среднее значение суммы квадратов разности предсказанного значения и исходного.

27. Фреймворк – программное обеспечение, облегчающее разработку и объединение разных компонентов большого программного проекта.

28. Хостинг – процесс размещения разработанного веб-приложения в сети Интернет.

Введение

В Российской Федерации существует множество так называемых моногородов [1], которые являются относительно небольшими населёнными пунктами, развитие которых тесно связано с деятельностью близлежащих градообразующих предприятий. По большей части население таких городов обеспечивает трудовыми ресурсами градообразующее предприятие или группу предприятий, в случаях если они образуют производственные цепочки. В результате этого различные параметры, характеризующие деятельность градообразующего предприятия, оказывают сильное влияния на развитие моногорода в целом.

Важным фактором развития моногородов является расширение сфер экономической деятельности предприятий, действующих в моногороде, т.е. увеличение уровня диверсификации экономики [2]. Разработка инструмента для расчёта и визуализации параметров, позволяющих оценить уровень развития экономики моногородов является целью данной работы.

Для достижения описанной цели поставлены следующие задачи:

- произвести первоначальную обработку исходных данных;
- дополнить исходные данные рядом показателей, необходимых для расчётов и визуализации;
- рассчитать основные экономические показатели;
- разработать функционал визуализации моногородов и их показателей экономической эффективности на карте;
- разработать модуль построения регрессионной модели для оценки доли моноотрасли моногородов и сравнить эффективность двух моделей на разных наборах входных данных.

Глава 1. Подготовка данных и расчёт экономических показателей

В рамках данной главы описываются исходные данные моногородов, которые будут использоваться для анализа, процесс обработки этих данных для расчёта необходимых экономических показателей и процесс сбора дополнительных данных, отсутствующих в исходном наборе данных, но потенциально полезных для анализа.

1.1. Описание исходных данных

Первоначальные данные состоят из набора Excel-файлов, которые содержат информацию о предприятиях в моногородах (рисунок 1). Файлы включают в себя более сотни свойств, описывающих множество факторов деятельности предприятий, таких как наименование предприятия, адрес расположения, дата регистрации, характеристики экономической деятельности за 2013-2017 года (уплаченные налоги, страховые взносы, основные средства, выручка, затраты на оплату труда, дивиденды) и прочие.

№	Наименование полное	Адрес (место нахождения)	Дата регистрации	Возраст компании	Дата ликвидации	Статус	Код налогоплательщика	Код статистики	Регион регистрации	Код ОКАТО	Наименование ОКАТО
1	ОТКРЫТОЕ АКЦИОНЕРНОЕ ОБЩЕСТВО "АЛТАЙ-КОКС"	659107, Алтайский кр. г. Заринск, ул. Притаенная, д. 2	28.10.1992	26,0		Действующая	2205001753	00188110	Алтайский край	01406000000	Алтайский край, Заринск
2	АКЦИОНЕРНОЕ ОБЩЕСТВО АЛТАЙСКОГО ВАГОНСТРОЕНИЯ	658087, Алтайский кр. г. Новоалтайск, ул. 22 Партсъезда, д. 16	30.11.1992	26,0		Действующая	2208000010	00729244	Алтайский край	01413000000	Алтайский край, Новоалтайск
3	ЗАКРЫТОЕ АКЦИОНЕРНОЕ ОБЩЕСТВО "АЛЕРЬСВЕРНОПРОДУКТ" ИМЕНИ С.Н.СТАРОВОЙТОВА	658130, Алтайский кр. г. Алейск, ул. Первомайская, д. 81	30.04.1996	22,5		Действующая	2201000766	00932382	Алтайский край	01403000000	Алтайский край, Алейск
4	ОТКРЫТОЕ АКЦИОНЕРНОЕ ОБЩЕСТВО "КУЧУКСУЛЬФАТ"	Благовещенский район, рабочий пос. Степное Озеро	18.12.1992	26,0		Действующая	2235001430	05762306	Алтайский край	01205557000	Благовещенский район, Степное Озеро
5	ОБЩЕСТВО С ОГРАНИЧЕННОЙ ОТВЕТСТВЕННОСТЬЮ "ХОЛОД"	659100, Алтайский кр. г. Заринск, ул. Зеленая, д. 51	15.11.1999	19,0		Действующая	2205006790	50665727	Алтайский край	01406000000	Алтайский край, Заринск
6	ГОСУДАРСТВЕННОЕ УНИТАРНОЕ ПРЕДПРИЯТИЕ ДОРОЖНОГО ХОЗЯЙСТВА АЛТАЙСКОГО КРАЯ "СЕВЕРО-ВОСТОЧНОЕ ДОРОЖНО-СТРОИТЕЛЬНОЕ УПРАВЛЕНИЕ"	658083, Алтайский кр. г. Новоалтайск, ул. Дорожная, д. 7	23.07.2003	15,0		Действующая	2208010989	14723085	Алтайский край	01413000000	Алтайский край, Новоалтайск
7	ГОСУДАРСТВЕННОЕ УНИТАРНОЕ ПРЕДПРИЯТИЕ ДОРОЖНОГО ХОЗЯЙСТВА АЛТАЙСКОГО КРАЯ "ЮЖНОЕ ДОРОЖНО-СТРОИТЕЛЬНОЕ УПРАВЛЕНИЕ"	658131, Алтайский кр. г. Алейск, ул. Дорожная, д. 6 корп. А	23.07.2003	15,0		Действующая	2201006430	14725913	Алтайский край	01403000000	Алтайский край, Алейск
8	ОБЩЕСТВО С ОГРАНИЧЕННОЙ ОТВЕТСТВЕННОСТЬЮ "ЭКОДОМСТРОЙ"	658084, Алтайский кр. г. Новоалтайск, ул. Вагоностроительная, д. 9	14.01.2015	4,0		Действующая	2208032830	31491546	Алтайский край	01413000000	Алтайский край, Новоалтайск
9	ОБЩЕСТВО С ОГРАНИЧЕННОЙ ОТВЕТСТВЕННОСТЬЮ "ЗАРИНСКИЙ МЯСОПЕРЕРАБАТЫВАЮЩИЙ ЗАВОД"	659100, Алтайский кр. г. Заринск, ул. Победы, д. 43	09.01.2008	11,0		Действующая	2205011014	83924925	Алтайский край	01406000000	Алтайский край, Заринск
10	ОТКРЫТОЕ АКЦИОНЕРНОЕ ОБЩЕСТВО "АЛТАЙСКИЙ ХИМПРОМ" ИМ. ВЕРЕЩАГИНА	658837, Алтайский кр. г. Яровое, пл. Предзаводская, д. 2	27.09.2012	6,0		Действующая	2211005485	10022712	Алтайский край	01430000000	Алтайский край, Яровое
11	ОБЩЕСТВО С ОГРАНИЧЕННОЙ ОТВЕТСТВЕННОСТЬЮ "ЗАРИНСКАЯ СЕТЕВАЯ КОМПАНИЯ"	659100, Алтайский кр. г. Заринск, ул. Молодежная, д. 17	29.12.2011	7,0		Действующая	2205012650	30893142	Алтайский край	01406000000	Алтайский край, Заринск
12	ОБЩЕСТВО С ОГРАНИЧЕННОЙ ОТВЕТСТВЕННОСТЬЮ "АЛТАЙСКИЙ БЕКОН"	658087, Алтайский кр. г. Новоалтайск, ул. Заготзерно, д. 5А	28.12.2007	11,0		Действующая	2277010771	83927415	Алтайский край	01413000000	Алтайский край, Новоалтайск
13	АКЦИОНЕРНОЕ ОБЩЕСТВО "АЛТАЙСКИЙ ЗАВОД СЕЛЬСКОХОЗЯЙСТВЕННОГО МАШИНОСТРОЕНИЯ"	658083, Алтайский кр. г. Новоалтайск, ул. Дорожная, д. 74 офис 68	06.09.2010	8,0		Действующая	2222789799	67104890	Алтайский край	01413000000	Алтайский край, Новоалтайск
14	МУНИЦИПАЛЬНОЕ УНИТАРНОЕ ПРЕДПРИЯТИЕ ГОРОДА НОВОАЛТАЙСКА "НОВОАЛТАЙСКИЕ ТЕПЛОВЫЕ СЕТИ"	658080, Алтайский кр. г. Новоалтайск, ул. 7 Микрорайон, д. 25	12.03.1992	26,5		Действующая	2208002579	21454691	Алтайский край	01413000000	Алтайский край, Новоалтайск

Рисунок 1. Фрагмент исходных данных о предприятиях

Среди этих параметров был выделен набор необходимых для проведения дальнейшего анализа и оценки экономического состояния моногородов параметров, а именно:

- наименование предприятия – полное наименование предприятия, которое не понадобится для конечного анализа, но может быть использовано на промежуточных этапах;
- дата регистрации и дата ликвидации – фактически сокращаются до года и служат для определения того, действует ли предприятие на определённый момент времени или оно уже ликвидировано;
- код статистики – уникальный код предприятия, также называемый кодом ОКПО, используется для идентификации отдельных предприятий в городе;
- код ОКАТО – уникальный код города или района города, который используется для группировки нескольких предприятий по городам;
- код основного вида деятельности – код из единого классификатора видов экономической деятельности также называемый кодом ОКВЭД, который позволит определять, относится ли предприятие к группе предприятий основной отрасли моногорода;
- основные средства – экономический показатель, определяющий совокупную стоимость всех материальных объектов, которые длительно используются предприятием в процессе производства продукции либо для управленческих нужд;
- выручка – экономический показатель, определяющий совокупную сумму требований, предъявленных предприятием в результате реализации произведённой продукции за определённый период;
- оплата труда – экономический показатель, определяющий совокупные затраты предприятий на содержание сотрудников.

Таким образом, после извлечения описанных показателей по всем предприятиям в моногородах, которые предполагается анализировать в

данной работе (список этих моногородов задаётся отдельным файлом), формируется единый обработанный набор данных. На этом же этапе подготовки данных используется заранее подготовленный список моногородов (рисунок 2), в котором для каждого моногорода определён код ОКАТО, список кодов ОКПО одного или нескольких градообразующих предприятий и список кодов ОКВЭД отраслей, соответствующих основному виду деятельности указанных градообразующих предприятий.

Моногород	ОКАТО	ОКПО	ОКВЭД
Анжеро-Судженск	32404	70622045, 165480, 45064664, 46492054, 00165480	05,46,28,14
Белово	32407	65159881, 70605762, 72292636, 57629640	10,05
Гурьевск	32413000	186223	24
Калтан	32415	92757343, 95371667	05,10,35
Киселевск	32416	22928033, 14792230, 164925, 10901721, 48622573, 93121541	05
Ленинск-Кузнецкий	32419	80298858	05
Мариинск	32422	337923	11
Междуреченск	32425	26635687	05
Мундыбаш (посёлок)	32227557	57622135	13,07
Мыски	32428	93137358	51
Новокузнецк	32431	82727542	26,27
Осинники	32434	160815	05
Полысаево	32435	160815	05

Рисунок 2. Фрагмент списка моногородов и их сопутствующие характеристики, предполагаемых для использования в данной работе

Используя упомянутые файлы, в исходных данных производится поиск только тех моногородов, чьи коды ОКАТО совпадают с указанными. По коду ОКПО и коду ОКВЭД внутри каждого из городов производится выборка предприятий, которые формируют основную экономическую отрасль в городе – моноотрасль. Однако, в дополнение к указанным данным о моногородах также необходимо собрать ряд дополнительных сведений с открытого веб-ресурса, что можно выполнить с помощью технологии веб-скрейпинг.

1.2. Использование веб-скрейпинга для извлечения дополнительных данных

Для анализа моногородам помимо уже имеющихся и описанных данных необходимо дополнить сведения о моногородах рядом параметров доступных на открытом веб-ресурсе Википедия [1]. Данная веб-страница содержит таблицу (рисунок 3) со всеми моногородами России и включает название моногорода, ссылку на страницу с информацией о моногороде, численность населения моногорода, список градообразующих предприятий и область расположения данного моногорода. В связи с тем, что данная страница содержит набор хорошо структурированных данных, было принято решение воспользоваться подходом веб-скрейпинг, который позволяет написать программу для извлечения данных с веб-ресурса при помощи парсинга страницы.

Название	Численность населения (тыс. человек)	Градообразующее предприятие
Алтайский край		
Яровое	18,05	ОАО «Алтайхимпром»
Заринск	46,25	ОАО «Алтай-кокс»
Алейск	29,2	АО «Алейскзернопродукт» имени С. Н. Старовойтова

Рисунок 3. Фрагмент таблицы, содержащей дополнительную информацию о моногородах

Например, следующий код (листинг 1) позволяет извлечь всю представленную информацию обо всех моногородах в таблице в удобную структуру.

Листинг 1. Алгоритм извлечения данных с веб-ресурса.

```
browser = webdriver.Chrome('chromedriver.exe')
browser.get('https://ru.wikipedia.org/wiki/Список_моногородов_России')
html = bs.BeautifulSoup(browser.page_source, 'lxml')
towns_stats, towns, region = dict(), [], None
getText = lambda item: item.text.replace('\n', '').replace('\xa0', ' ')
```

```

').strip()
for table_block in html.select('tr')[1:]:
    row = table_block.select('td')
    if len(row) == 1:
        if region is not None:
            towns_stats[region] = towns
            towns = []
            region = getText(row[0])
        else:
            population = getText(row[1])
            if population != '':
                population = int(round(float(population.replace(',', '.'))*1000))
            link = row[0].select_one('a')
            if link is not None:
                link = 'https://ru.wikipedia.org' + link.get('href')
            towns.append({
                'name': getText(row[0]),
                'region': region,
                'population': population,
                'company': getText(row[2]).split('[')[0],
                'link': link
            })
    if region is not None:
        towns_stats[region] = towns

```

Пример структуры извлечённых данных представлен в листинге 2.

Листинг 2. Структура данных, извлечённых с веб-ресурса.

```

{
  "Алтайский край": [
    {
      "name": "Яровое",
      "region": "Алтайский край",
      "population": 18050,
      "company": "ОАО «Алтайхимпром»",
      "link": "https://ru.wikipedia.org/wiki/Яровое"
    },
    {
      "name": "Заринск",
      "region": "Алтайский край",
      "population": 46250,
      "company": "ОАО «Алтай-кокс»",
      "link": "https://ru.wikipedia.org/wiki/Заринск"
    },
    {
      "name": "Алейск",
      "region": "Алтайский край",
      "population": 29200,
      "company": "АО «Алейскзернопродукт» имени С. Н. Старовойтова",
      "link": "https://ru.wikipedia.org/wiki/Алейск"
    }
  ]
}

```

Для каждого моногорода была извлечена ссылка на страницу с его информацией. Парсинг каждой из этих страниц дополнительно позволяет произвести поиск данных с координатами географического расположения этих городов для конечной визуализации на карте.

Используя представленный подход, можно автоматизировать сбор большого объёма данных с веб-ресурсов, что особенно актуально в случаях периодического обновления таких данных.

По итогам выполнения веб-скрейпинга, были внесены данные по принадлежности моногорода к определённому субъекту Российской Федерации (область, край), список градообразующих предприятий и координаты (широта и долгота) моногорода, численности населения моногорода. Численность населения позволит произвести расчёт показателей на душу населения и эффективно сравнить показатели крупных и малых моногородов.

1.3. Расчёт экономических показателей моногородов

Для оценки экономической ситуации моногородов и их уровня диверсификации в данной работе производится расчёт ряда экономических показателей. Полный код расчёта всех показателей приведён в приложении Б.

1. Суммарная выручка, сумма основных средств и суммарная оплата труда всех предприятий моногорода за 2013-2017, а также соответствующие удельные показатели (в расчёте на одного жителя моногорода).

Соответствующие шесть рассчитанных показателей за 2017 год для трёх моногородов отображены на рисунке 4.

Моногород	Выручка 2017	Основные средства 2017	Оплата труда 2017	Удельный показатель выручка 2017	Удельный показатель основные средства 2017	Удельный показатель оплата труда 2017
Алейск	9.949636e+09	1.327005e+09	8.974520e+08	3.407410e+05	45445.376712	30734.657534
Заринск	7.539299e+10	4.399851e+09	2.912246e+09	1.630119e+06	95131.913514	62967.481081
Яровое	1.888970e+09	5.438510e+08	2.015820e+08	1.046521e+05	30130.249307	11167.977839

Рисунок 4. Пример рассчитанных показателей суммарной выручки, основных средств, оплаты труда и их удельных значений за 2017 год

2. Количество действующих, открытых и ликвидированных предприятий в моногороде в течении заданного года.

Соответствующие три показателя за 2017 год для трёх моногородов отображены на рисунке 5.

Моногород	Открытые 2017	Закрытые 2017	Действующие 2017
Алейск	27	89	241
Заринск	15	40	426
Яровое	4	7	120

Рисунок 5. Пример рассчитанных показателей количества открытых, закрытых и действующих предприятий за 2017 год

3. Доля моноотрасли по одному из параметров предприятия (основные средства, выручка, оплата труда) рассчитывается по формуле (1).

$$CR_P = \sum_i^n \frac{P_i^{осн}}{P}, \quad (1)$$

где $P_i^{осн}$ – значение одного из параметров i -ого предприятия в моногороде, относящегося к группе градообразующих предприятий (код ОКПО находится в списке, указанном в справочной таблице для данного моногорода) или предприятий со схожим видом экономической деятельности (первый уровень ОКВЭДа предприятия соответствует одному из значений в справочной таблице для данного моногорода), P – суммарное значение одного из параметров всех предприятий моногорода.

Соответствующие рассчитанные доли моноотрасли по трём параметрам за 2017 год для трёх моногородов отображены на рисунке 6.

Моногород	Доля моноотрасли ОС 2017	Доля моноотрасли Выручка 2017	Доля моноотрасли Оплата труда 2017
Алейск	0.623040	0.686640	0.601925
Заринск	0.643623	0.866848	0.743472
Яровое	0.662209	0.268194	0.515577

Рисунок 6. Пример рассчитанных показателей доли моноотрасли по основным средствам, выручке, оплате труда за 2017 год

Доля моноотрасли является одним из основных показателей уровня диверсификации экономики в моногороде. На основе того, как распределены рассчитанные значения (рисунок 7), можно утверждать, что исходные данные недостаточно достоверны, поскольку наблюдается существенное количество моногородов с долей моноотрасли ниже 10% по всем трём параметрам, хотя моногорода должны характеризоваться высоким вкладом градообразующих предприятий в экономику города.

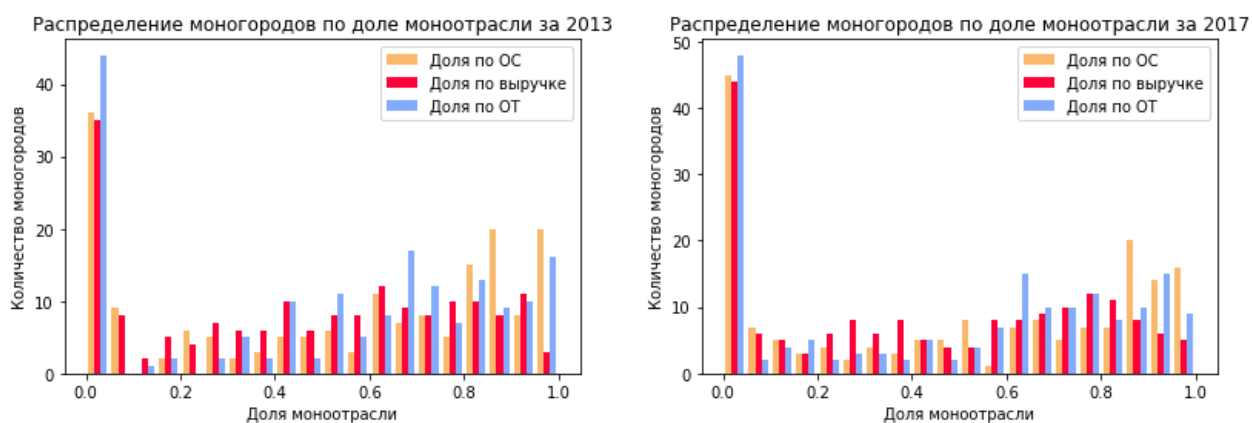


Рисунок 7. Гистограммы распределения рассчитанной доли моноотрасли в моногородах по трём параметрам в 2013 и 2017 годах

4. Индекс Херфиндаля-Хиршмана (*HHI*) по одному из параметров предприятия рассчитывается по формуле (2).

$$HNI_P = \sum_i^n \left(\frac{P_i}{P}\right)^2, \quad (2)$$

где P_i – значение одного из параметров i -ого предприятия в моногороде, P – суммарное значение одного из параметров всех предприятий моногорода.

Соответствующие рассчитанные значения HNI по трём параметрам за 2017 год для трёх моногородов отображены на рисунке 8.

Моногород	HNI ОС 2017	HNI Выручка 2017	HNI Оплата труда 2017
Алейск	0.379528	0.424643	0.384302
Заринск	0.426396	0.755491	0.570445
Яровое	0.468525	0.148235	0.443292

Рисунок 8. Пример рассчитанных показателей HNI по основным средствам, выручке, оплате труда за 2017 год

Индекс Херфиндаля-Хиршмана принимает значения от 0 (в идеальном случае совершенной конкуренции, когда на рынке бесконечно много предприятий с незначительной долей вклада в экономику города) до 1 (в ситуации, когда на рынке присутствует только одно доминирующее предприятие). Основным преимуществом данного показателя является способность чутко реагировать на перераспределение долей между предприятиями, действующими на рынке.

Оценив распределение моногородов по рассчитанному значению HNI (рисунок 9), можно отметить, что данный показатель распределён более равномерно, чем предыдущий, но при этом заметно смещение основной массы моногородов в сторону 0, что также связано с описанной проблемой.

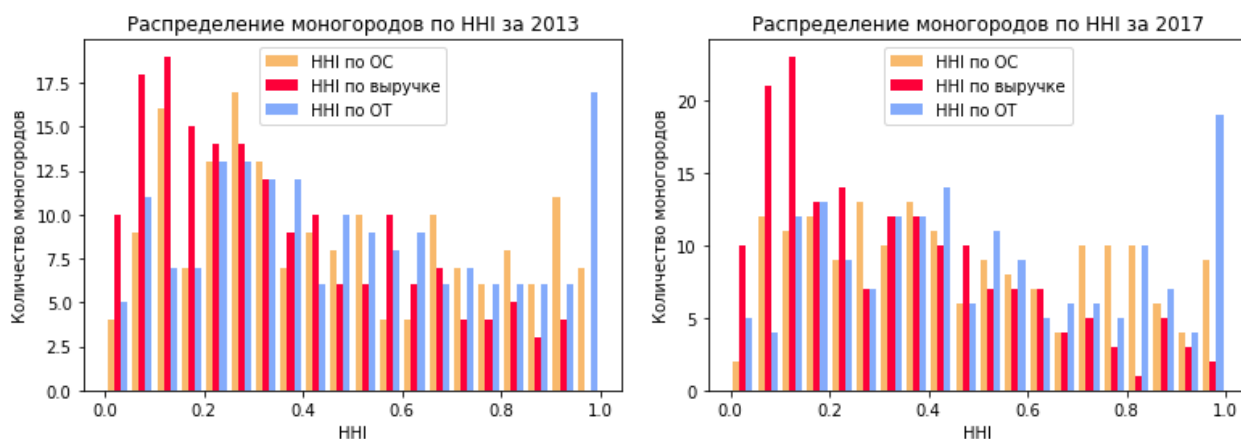


Рисунок 9. Гистограммы распределения рассчитанного индекса Херфиндаля-Хиршмана в моногородах по трём параметрам в 2013 и 2017 годах

5. Модифицированный индекс Херфиндаля-Хиршмана (HHI^*) по одному из параметров предприятия рассчитывается по формуле (3).

$$HHI^*_P = \frac{\sum_i^n (P_i^{неосн})^2}{(\sum_i^n P_i^{неосн})^2}, \quad (3)$$

где $P_i^{неосн}$ – значение одного из параметров i -ого предприятия в моногороде, не относящегося к группе градообразующих предприятий или предприятий со схожим видом экономической деятельности.

Соответствующие рассчитанные значения HHI^* по трём параметрам за 2017 год для трёх моногородов отображены на рисунке 10.

Моногород	HHI^* ОС 2017	HHI^* Выручка 2017	HHI^* Оплата труда 2017
Алейск	0.050914	0.030125	0.091545
Заринск	0.012145	0.004065	0.017695
Яровое	0.030004	0.076307	0.177472

Рисунок 10. Пример рассчитанных показателей HHI^* по основным средствам, выручке, оплате труда за 2017 год

Модифицированный индекс Херфиндаля-Хиршмана аналогичен обычному индексу Херфиндаля-Хиршмана, однако использование данного индекса может помочь избежать искажений, связанных с недостатком данных об основной отрасли моногородов. HHI^* имеет обратный эффект – чем ниже данный показатель, тем выше монополизации отдельных предприятий в городе. Согласно распределению моногородов по рассчитанному значению HHI^* (рисунок 11), можно отметить, что данный показатель показывает себя лучше в сравнении с обычным HHI – наблюдается высокая доля моногородов с низким значением HHI^* , что свидетельствует о низком уровне диверсификации.

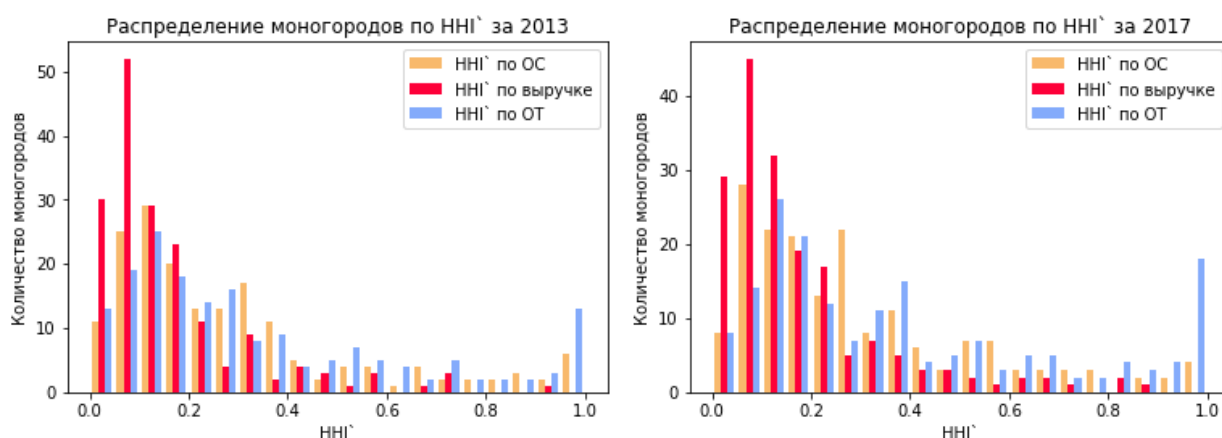


Рисунок 11. Гистограммы распределения рассчитанного модифицированного индекса Херфиндаля-Хиршмана в моногородах по трём параметрам в 2013 и 2017 годах

6. Энтропия (E) по одному из параметров предприятия рассчитывается по формуле (4).

$$E_P = \sum_i^n \left(\ln \left(\frac{P}{P_i} \right) * \frac{P_i}{P} \right), \quad (4)$$

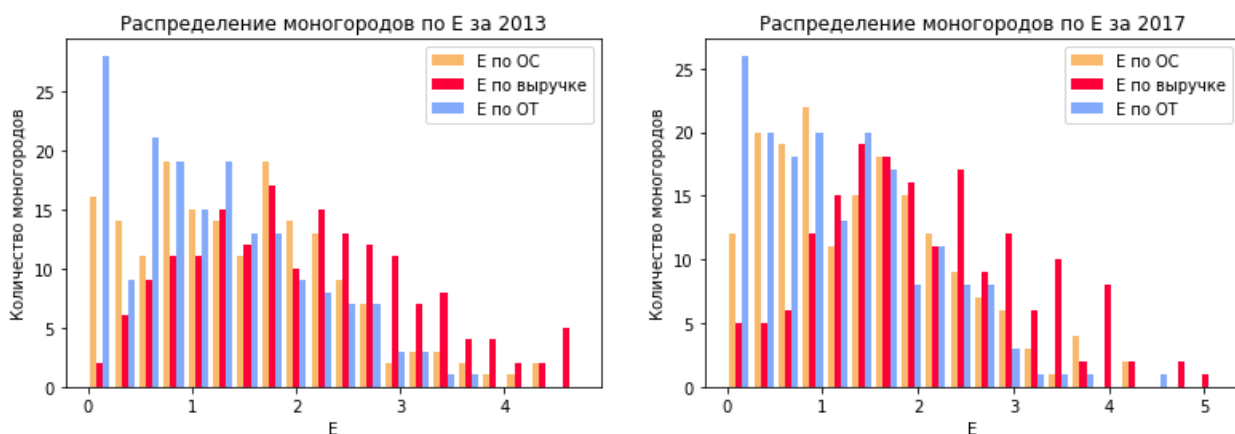
где P_i – значение одного из параметров i -ого предприятия в моногороде, P – суммарное значение одного из параметров всех предприятий моногорода.

Соответствующие рассчитанные значения E по трём параметрам за 2017 год для трёх моногородов отображены на рисунке 12.

Моногород	Е ОС 2017	Е Выручка 2017	Е Оплата труда 2017
Алейск	1.567217	1.644075	1.325883
Заринск	1.709132	0.751810	1.004074
Яровое	1.238305	2.444869	0.890618

Рисунок 12. Пример рассчитанных показателей Е по основным средствам, выручке, оплате труда за 2017 год

Энтропия представляет собой показатель, обратной концентрации: чем выше её значение, тем ниже концентрация предприятий на рынке. Энтропия измеряет неупорядоченность распределения долей между предприятиями рынка: чем выше показатель энтропии, тем ниже возможности продавцов влиять на рыночную цену. При оценке распределения энтропии (рисунок 13) заметно сильное различие между энтропией по выручке и энтропией по основным средствам и оплатой труда. Моногорода с первым показателем в основном сконцентрированы в центре, что свидетельствует о том, что большинство моногородов имеют средний уровень диверсификации по данному параметру. При этом энтропия по основным средствам и оплате труда показывает большую долю моногородов по с низким уровнем диверсификации, что в большей степени свойственно моногородам.



1.4. Выводы по разделу

В рамках данного раздела были рассмотрены исходные данные предприятий, которые предполагается использовать для проводимого анализа. Также рассмотрены рассчитываемые параметры моногородов, которые позволяют оценить уровень диверсификации:

- доля моноотрасли;
- индекс Херфиндаля-Хиршмана (HHI);
- модифицированный индекс Херфиндаля-Хиршмана (HHI');
- энтропия (E).

По рассчитанным показателям видно, что часть моногородов имеют недостаточное количество данных для корректной оценки их уровня диверсификации, что выражается в несвойственном значении этих показателей.

Дополнительно использовался способ веб-скрейпинга для сбора вспомогательных данных моногородов для их визуализации.

Глава 2. Разработка информационного портала

В данной главе рассматривается разработанный функционал информационного портала для оценки уровня диверсификации моногородов и визуализации данных по моногородам.

Всего разработано три модуля веб-приложения, два из которых рассмотрены в этой главе:

- карта с простой визуализацией маркеров-моногородов с отображением ряда рассчитанных параметров по этим моногородам;
- карта с визуализацией распределения моногородов по выбранному параметру;
- модуль построения регрессионной модели для предсказания доли моноотрасли в моногородах на основании ряда заданных параметров (данный модуль рассмотрен в отдельной главе).

Для разработки описанного функционала создано веб-приложение, состоящее из клиентского приложения, с которым пользователь взаимодействует напрямую через браузер, и серверного приложения, которое хранит все рассчитанные данные и производит обработку клиентских запросов по их получению.

Для разработки клиентского приложения использовался JS фреймворк React, а для размещения завершённого проекта в сети были использованы услуги сервиса Firebase (размещённое веб-приложение доступно по адресу <https://monotowns.web.app>) [3].

Разработка серверного приложения производилась посредством языка программирования Python и фреймворка Django, а для хостинга был использован ресурс PythonAnywhere [4].

2.1. Разработка модуля визуализации моногородов

Для предоставления возможности ознакомиться с проанализированными моногородами и оценить их уровень диверсификации, посредством изучения и сравнения рассчитанных экономических показателей была разработана визуализация с помощью карты. В данном модуле визуализации производится проставление маркеров моногородов на карте с использованием соответствующих им координат (рисунок 14).

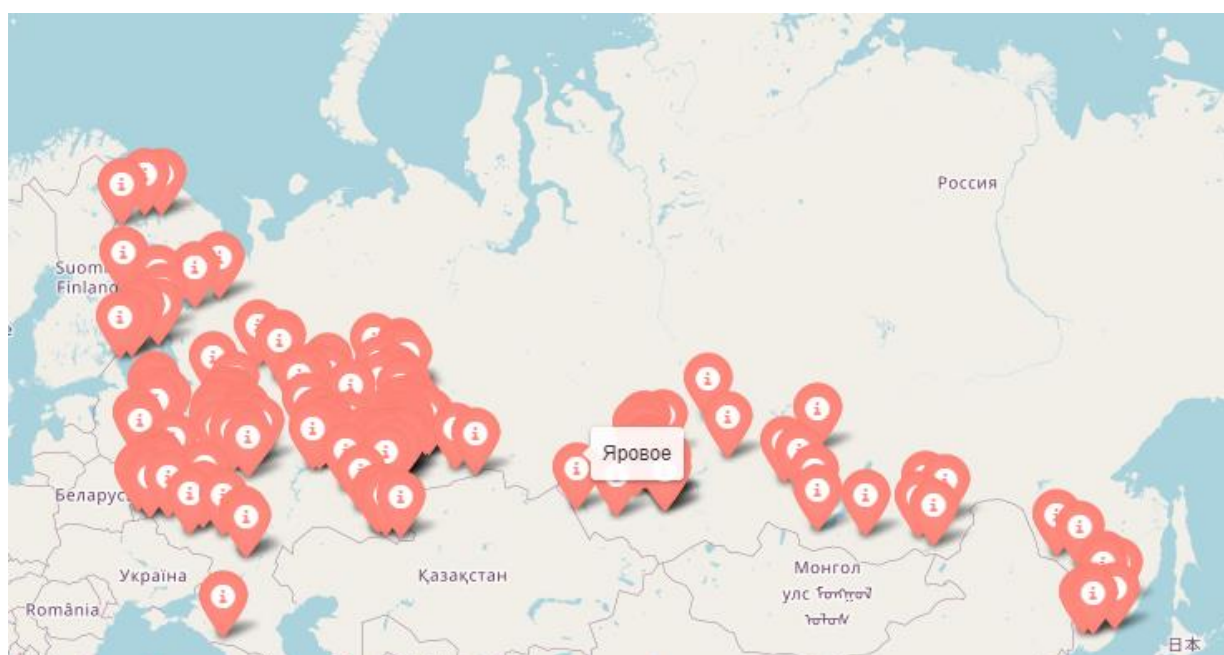


Рисунок 14. Отображение маркеров моногородов на карте

При наведении на каждый маркер отображается название моногорода, а при нажатии можно ознакомиться с небольшим перечнем показателей за определённым год (рисунок 15). Более подробные сведения о моногороде можно получить, нажав кнопку «Подробнее» на всплывающей панели.

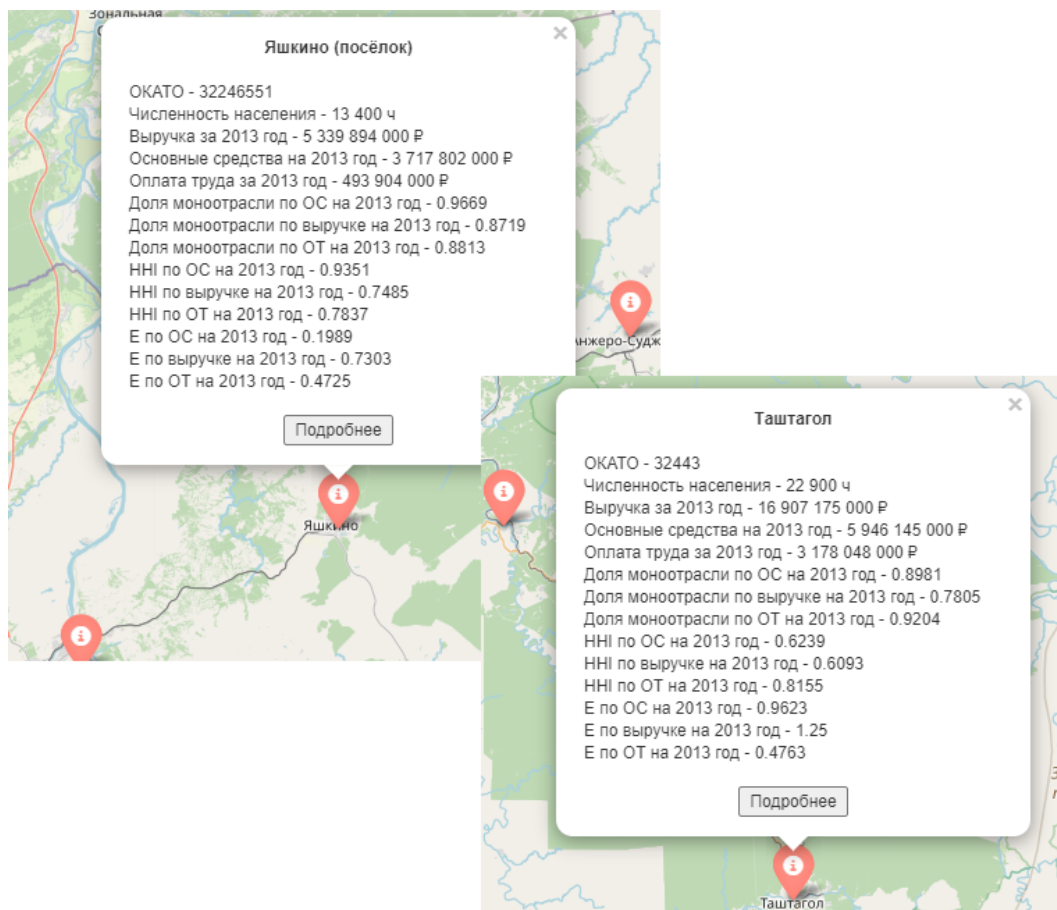


Рисунок 15. Информация о моногороде, выводимая во всплывающей панели

В подробных сведениях о моногородах приводится информация о численности населения, количество предприятий, использованных при анализе моногорода, наименование градообразующих предприятий и все рассчитанные экономические показатели за выбранный год. На этой странице предусмотрена возможность выбора целевого моногорода и года, по которым и выводится информация. Например, основываясь на показателях, рассчитанных для моногорода Яшкино (рисунок 16), видно, что он имеет достаточно высокую долю моноотрасли (больше 85%), высокий индекс Херфиндаля-Хиршмана (больше 0.75) и низкую энтропию (меньше 1), поэтому можно сказать, что ООО «КДВ Яшкино» оказывает существенное влияние на весь моногород, и в целом моногород имеет низкий уровень диверсификации экономики.

Моногород: <input type="text" value="Яшкино (посёлок)"/>	Год: <input type="text" value="2013"/>	
Население города: 13 400		
Количество предприятий, использованных для анализа: 304		
Градообразующие предприятия: ООО «КДВ Яшкино»		
Суммарные показатели моногорода в 2013 году		ННИ в 2013 году
Основные средства: 3 717 802 000 Р	Удельные основные средства: 277 448 Р/ч	По основным средствам 0.9351
Выручка: 5 339 894 000 Р	Удельная выручка: 398 500 Р/ч	По выручке 0.7485
Оплата труда: 493 904 000 Р	Удельная оплата труда: 36 859 Р/ч	По оплате труда 0.7837
Доля моноотрасли в 2013 году		Е в 2013 году
По основным средствам 0.9669		По основным средствам 0.1989
По выручке 0.8719		По выручке 0.7303
По оплате труда 0.8813		По оплате труда 0.4725

Рисунок 16. Пример выводимой подробной информации и экономических показателей о моногороде Яшкино за 2013 год

Описанная карта также имеет функционал поиска по ряду категорий:

- по названию моногорода;
- по номеру ОКАТО;
- фильтрация по округам и областям;
- фильтрация по одному из рассчитанных числовых параметров с возможностью указать условие сравнения (допускаются варианты сравнения больше, меньше, равно или в диапазоне).

Используя подобную фильтрацию, можно произвести поиск моногородов в Кемеровской области с численностью населения от 50 до 100 тысяч человек. Задав соответствующие параметры, видно, что в Кемеровской области присутствует шесть моногородов, удовлетворяющих данным требованиям (рисунок 17).

Поиск по округу: Поиск по области:

Искать по параметру:

Выбор группы параметров для поиска: Выбор параметра для поиска:

Выбор условия поиска: Значение условия: Второе значение условия:

Конечное условие поиска: **50000 < Численность населения < 100000**

Найдено городов: 6

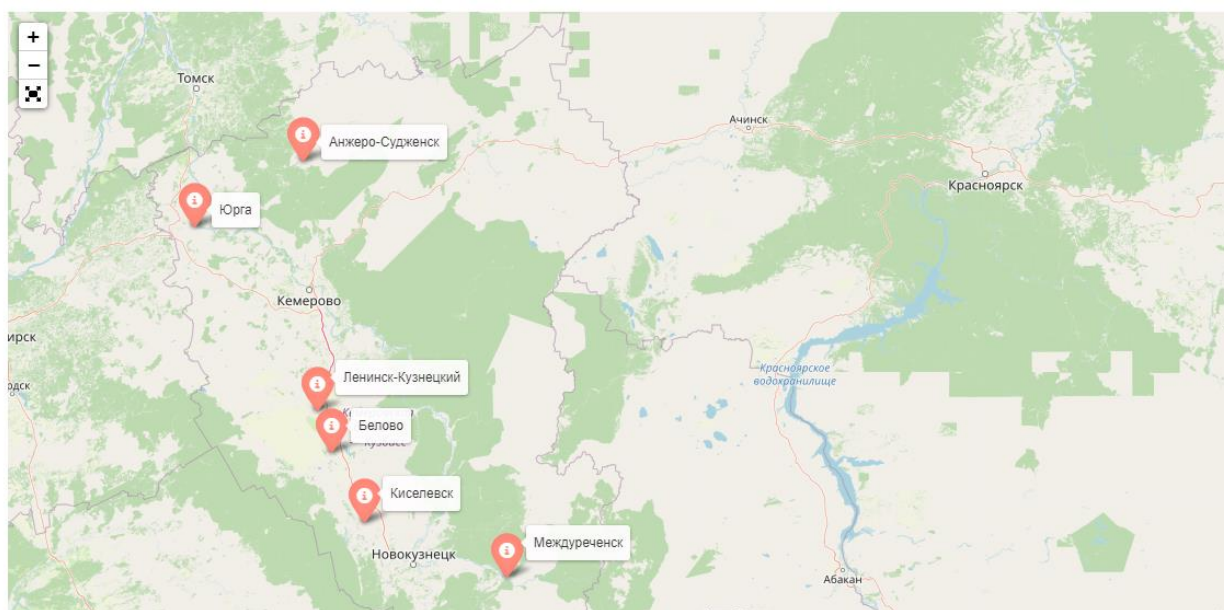


Рисунок 17. Пример использования поиска моногородов в Кемеровской области с численностью населения от 50 до 100 тысяч человек

2.2. Разработка модуля визуализации распределения моногородов по параметрам

При анализе моногородов по рассчитанным показателям также важно иметь возможность быстро сравнить значения разных моногородов. Для этой цели был разработан дополнительный модуль визуализации распределения моногородов по заданному параметру. Подразумевается, что в этом варианте визуализации каждый моногород представляется в виде круга определённого радиуса, который зависит от величины показателя: чем выше значение показателя, тем больше радиус круга. Также предусмотрен вариант, что некоторые из рассчитанных показателей могут иметь значение 0, что можно интерпретировать, как ошибку ввиду отсутствия достаточных для расчёта

показателя данных, поэтому моногорода со значением любого показателя равного 0 отмечаются чёрным кругом с малым радиусом.

Данный модуль визуализации позволяет выбрать группу параметров, в которую входят: группа основных показателей (численность населения, суммарная выручка моногорода и прочие подобные показатели), группа доли моноотрасли по трём параметрам (основные средства, оплата труда и выручка), группа показателей индексов Херфиндаля-Хиршмана и энтропии также по трём параметрам. Например, построив карту распределения моногородов по оплате труда за 2017 год (рисунок 18), можно отметить, что наибольшая суммарная оплата труда в Нижнекамске, Набережных Челнах и Новокузнецке.

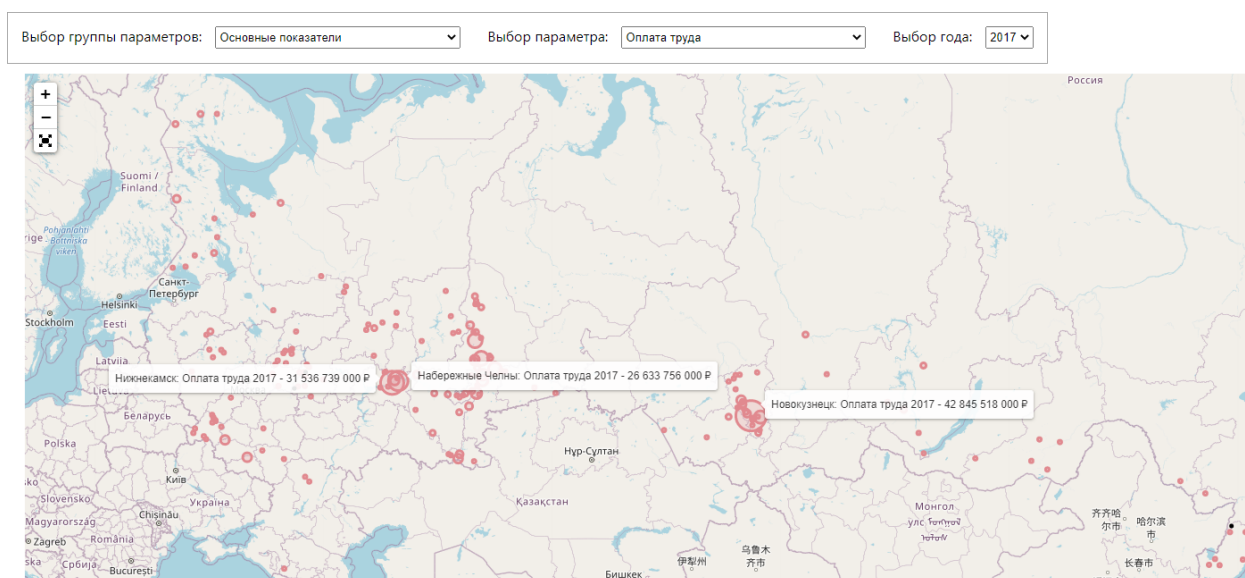


Рисунок 18. Распределение моногородов по суммарной оплате труда за 2017 год

Также построение подобной карты по доле моноотрасли позволяет оценить количество моногородов, в которых данный показатель оказался равен нулю (рисунок 19). Можно сделать вывод, что присутствует большое количество моногородов с неизвестной долей моноотрасли, которую можно предсказать на основе других характеристик моногорода с помощью регрессионной модели.

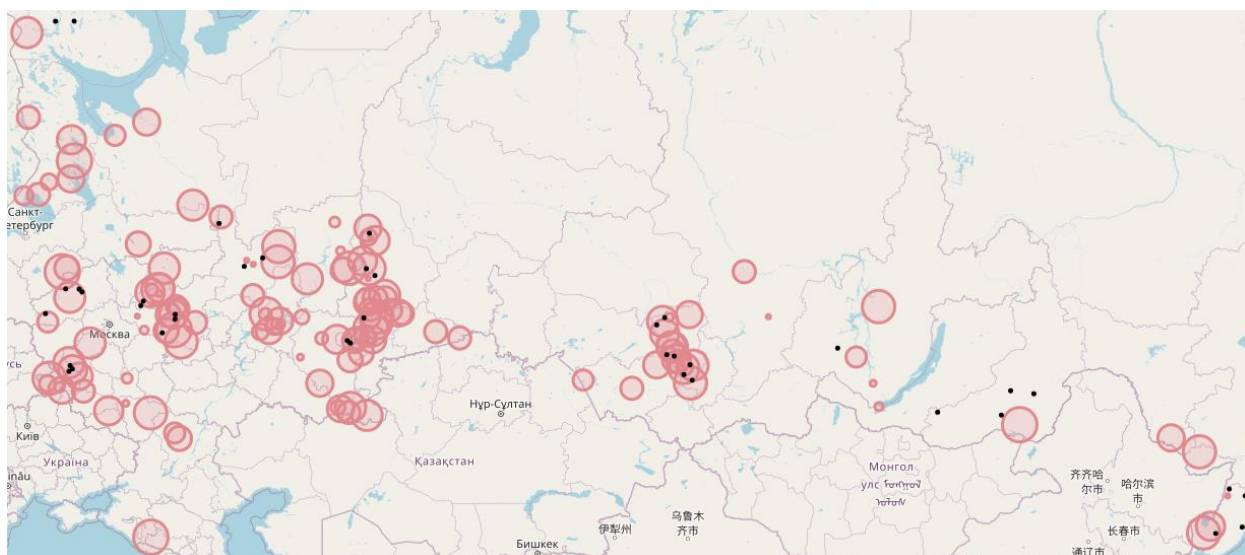


Рисунок 19. Распределение долей моноотрасли по оплате труда за 2017 год

2.3. Выводы по разделу

В рамках данного раздела был рассмотрен разработанный функционал по визуализации моногородов и их данных на карте. Два представленных модуля позволяют по-разному оценить экономическое положение моногородов и могут применяться в разных ситуациях.

Полный код, написанный для построения визуализации на серверном приложении приведён в приложении В.

Глава 3. Разработка модуля построения регрессионной модели

В данной главе рассматривается модуль веб-приложения по построению регрессионной модели для оценки доли моноотрасли на основании нескольких входных параметров.

Предполагается использовать две достаточно распространённые и простые в реализации модели машинного обучения [5]: линейную регрессию и «случайный лес» (Random Forest). Рассмотрев эффективность двух моделей на разном наборе данных, можно определить, как более подходящую модель для поставленной задачи, так и выбрать лучший набор входных параметров.

3.1. Модуль построения регрессионной модели в веб-приложении

В уже описанном веб-приложении также был разработан модуль, позволяющий построить регрессионную модель на основе нескольких входных рассчитанных параметров, которые можно выбрать произвольным образом. Также предусмотрен выбор вида модели (линейная регрессия или «случайный лес»). Внутренняя реализация построения моделей строится на общих принципах регрессионного анализа [6], а код описывающий этот процесс приведён в приложении В.

По выбранным параметрам происходит построение модели, результаты предсказания которой сопоставляются с рассчитанными значениями долей моноотрасли на диаграмме рассеивания (рисунок 20), где каждому моногороду на горизонтальной оси соответствует две точки по вертикальной оси: рассчитанное значение и предсказанное построенной моделью. Рассчитанные значения обозначаются точками синего цвета, если они имеют значение отличное от нуля, и красного цвета, если значение равно 0, то есть должно восприниматься в качестве ошибки. Предсказанным же значениям соответствуют точки жёлтого цвета. В приведённом примере видно, что рассчитанная доля моноотрасли по основным средствам в моногороде Выкса

равна 0, а предсказанное значение – 1.02, что с одной стороны является демонстрацией эффективности применения регрессионной модели для предсказания значений. С другой стороны, данный пример показывает недостаток применения линейной регрессии, поскольку очевидно, что данное значение не может превышать 1 или быть ниже 0 (в приведённом примере также видны случаи с отрицательным предсказанным значением). Регрессионная модель «случайный лес» является более гибким инструментом, и, например, данная модель, построенная на тех же входных параметрах, предсказывает значение для того же моногорода равным 0.992.

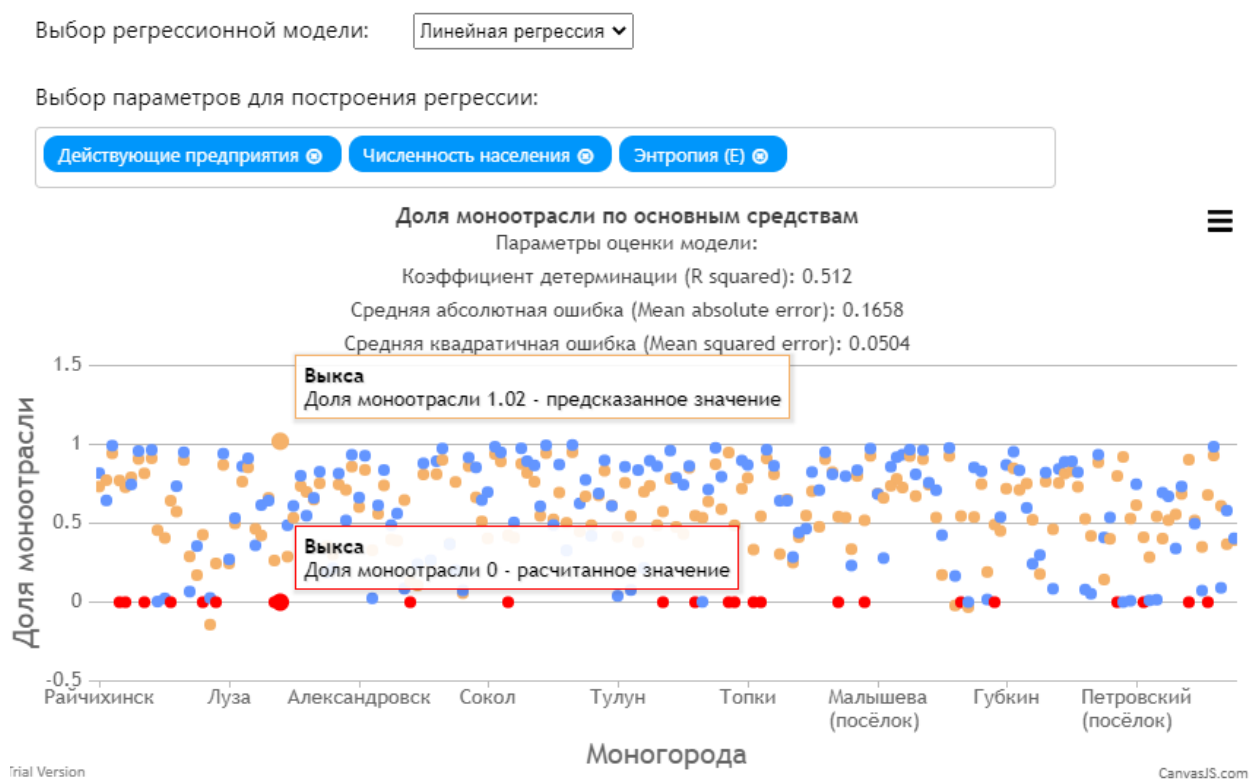


Рисунок 20. Пример построения регрессионной модели на основании параметров «Количество действующих предприятий», «Численность населения», «Энтропия» и визуализации предсказанных значений

После построения регрессионной модели можно также оценить её качество на основании трёх показателей: коэффициент детерминации (R-squared), средняя абсолютная ошибка (MAE) и средняя квадратичная ошибка (MSE). Пример отображения этих показателей также можно найти в представленном примере выше.

При выборе независимых переменных, по которым будет строиться модель необходимо оценивать степень зависимости между параметрами – корреляцию. Для информирования о том, какие параметры имеют высокую степень корреляции и соответственно наилучшим образом подходят для использования в модели, в веб-приложение была добавлена матрица корреляции, фрагмент которой изображён на рисунке 21, а полная версия размещена в приложении Г.

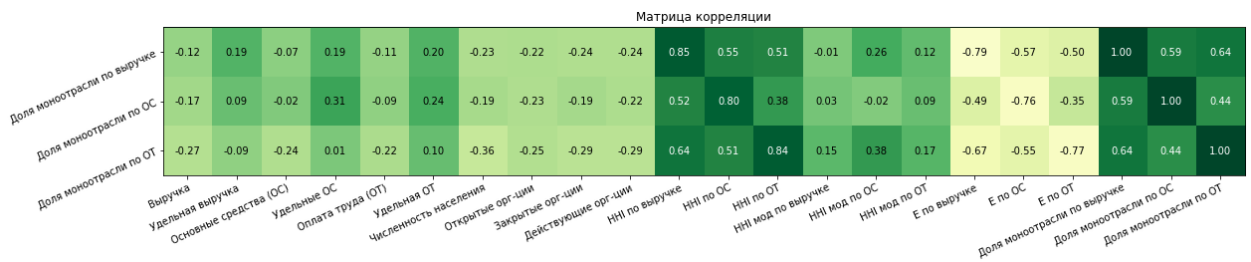


Рисунок 21. Фрагмент матрицы корреляции для доли моноотрасли

Наибольшую корреляцию с долей моноотрасли показывают параметры энтропия и *NNI*, поэтому данные параметры будут предпочтительными для построения регрессионной модели.

3.2. Сравнение качества построенных регрессионных моделей

Для определения того, какая из регрессионных моделей лучше подходит для поставленной задачи, необходимо произвести оценку качества двух моделей и сравнить результаты. Подбирая разные доступные входные параметры можно построить регрессионную модель и по предсказанным значениям получить оценку по трём параметрам.

Таким образом, было проведено несколько замеров точности двух моделей, построенных с использованием как одного входного параметра, так и двух. Результаты замеров представлены в таблицах 1, 2 и 3, в каждой из которых в качестве выходного параметра использовались доли моноотрасли по выручке, основным средствам и оплате труда соответственно. В каждой таблице введено сокращение регрессионных моделей: «ЛР» соответствует линейной регрессии, а «СЛ» – «случайному лесу».

Таблица 1. Сравнение регрессионных моделей при предсказании доли моноотрасли по выручке

Входные параметры	R-squared		MAE		MSE	
	ЛР	СЛ	ЛР	СЛ	ЛР	СЛ
Выручка	0.01	-0.518	0.183	0.213	0.046	0.07
Удельная выручка	0.024	-0.321	0.183	0.197	0.045	0.061
Основные средства	0.003	-0.284	0.185	0.198	0.046	0.059
Удельные основные средства	0.018	-0.276	0.184	0.195	0.045	0.059
Оплата труда	0.001	-0.3	0.185	0.203	0.046	0.006
Удельная оплата труда	0.058	-0.447	0.176	0.211	0.043	0.067
Численность населения	0.04	0.561	0.179	0.089	0.044	0.02
Количество открытых предприятий	0.033	-0.016	0.181	0.182	0.045	0.047
Количество закрытых предприятий	0.011	-0.28	0.183	0.201	0.046	0.059
Количество действующих предприятий	0.023	-0.516	0.181	0.217	0.045	0.07
ННІ по выручке	0.739	0.657	0.071	0.079	0.012	0.016
ННІ по основным средствам	0.353	0.055	0.139	0.163	0.03	0.044
ННІ по оплате труда	0.283	-0.018	0.15	0.172	0.033	0.047
ННІ модифицированный по выручке	-0.003	-0.362	0.186	0.206	0.046	0.063
ННІ модифицированный по основным средствам	0.037	-0.232	0.183	0.194	0.044	0.057
ННІ модифицированный по оплате труда	-0.003	-0.497	0.186	0.218	0.046	0.069
Е по выручке	0.605	0.468	0.1	0.113	0.018	0.025
Е по основным средствам	0.373	0.235	0.135	0.14	0.029	0.035
Е по оплате труда	0.26	-0.012	0.153	0.169	0.034	0.047
Численность населения, Количество действующих предприятий	0.023	0.425	0.181	0.115	0.045	0.026
Численность населения, ННІ по выручке	0.742	0.82	0.071	0.055	0.012	0.008
Численность населения, Е по выручке	0.676	0.777	0.088	0.063	0.015	0.01

Таблица 2. Сравнение регрессионных моделей при предсказании доли моноотрасли по основным средствам

Входные параметры	R-squared		MAE		MSE	
	ЛР	СЛ	ЛР	СЛ	ЛР	СЛ
Выручка	0.038	-0.278	0.191	0.204	0.05	0.066
Удельная выручка	-0.037	-0.145	0.198	0.189	0.053	0.059
Основные средства	-0.001	-0.128	0.194	0.183	0.052	0.058
Удельные основные средства	0.05	-0.158	0.191	0.191	0.049	0.06
Оплата труда	0.003	-0.217	0.194	0.19	0.051	0.063
Удельная оплата труда	0.058	-0.188	0.186	0.183	0.049	0.061
Численность населения	0.059	0.65	0.188	0.08	0.049	0.018
Количество открытых предприятий	0.039	-0.146	0.19	0.202	0.05	0.059
Количество закрытых предприятий	0.043	-0.059	0.19	0.192	0.049	0.055
Количество действующих предприятий	0.062	-0.409	0.188	0.212	0.048	0.073
ННІ по выручке	0.318	0.118	0.154	0.166	0.035	0.045
ННІ по основным средствам	0.712	0.688	0.083	0.079	0.015	0.016
ННІ по оплате труда	0.236	0.042	0.16	0.173	0.039	0.049
ННІ модифицированный по выручке	-0.006	-0.462	0.194	0.213	0.052	0.075
ННІ модифицированный по основным средствам	-0.002	-0.36	0.194	0.212	0.052	0.07
ННІ модифицированный по оплате труда	-0.007	-0.374	0.195	0.206	0.052	0.071
Е по выручке	0.309	0.069	0.152	0.169	0.036	0.048
Е по основным средствам	0.629	0.476	0.097	0.11	0.019	0.027
Е по оплате труда	0.232	-0.036	0.161	0.178	0.04	0.053
Численность населения, Количество действующих предприятий	0.06	0.577	0.188	0.099	0.048	0.022
Численность населения, ННІ по основным средствам	0.704	0.852	0.082	0.05	0.015	0.008
Численность населения, Е по основным средствам	0.627	0.834	0.09	0.058	0.019	0.009

Таблица 3. Сравнение регрессионных моделей при предсказании доли моноотрасли по оплате труда

Входные параметры	R-squared		MAE		MSE	
	ЛР	СЛ	ЛР	СЛ	ЛР	СЛ
Выручка	0.03	-0.37	0.165	0.186	0.038	0.053
Удельная выручка	-0.031	-0.25	0.17	0.177	0.04	0.049
Основные средства	0.027	-0.36	0.165	0.179	0.038	0.053
Удельные основные средства	-0.024	-0.407	0.17	0.186	0.04	0.055
Оплата труда	0.009	-0.382	0.168	0.187	0.038	0.054
Удельная оплата труда	0.022	-0.376	0.165	0.183	0.038	0.053
Численность населения	0.075	0.587	0.161	0.079	0.036	0.016
Количество открытых предприятий	0.019	0.032	0.166	0.16	0.038	0.038
Количество закрытых предприятий	0.028	-0.081	0.165	0.169	0.038	0.042
Количество действующих предприятий	0.045	-0.268	0.163	0.176	0.037	0.049
ННИ по выручке	0.506	0.227	0.105	0.126	0.019	0.03
ННИ по основным средствам	0.38	0.063	0.123	0.15	0.024	0.036
ННИ по оплате труда	0.699	0.579	0.068	0.08	0.012	0.016
ННИ модифицированный по выручке	-0.002	-0.477	0.167	0.197	0.039	0.057
ННИ модифицированный по основным средствам	0.074	-0.384	0.153	0.185	0.036	0.054
ННИ модифицированный по оплате труда	-0.001	-0.463	0.166	0.193	0.039	0.057
Е по выручке	0.516	0.391	0.102	0.116	0.019	0.024
Е по основным средствам	0.407	0.114	0.119	0.147	0.023	0.034
Е по оплате труда	0.547	0.407	0.098	0.105	0.018	0.023
Численность населения, Количество действующих предприятий	0.075	0.51	0.161	0.1	0.036	0.019
Численность населения, ННИ по оплате труда	0.711	0.827	0.065	0.050	0.011	0.007
Численность населения, Е по оплате труда	0.608	0.731	0.087	0.069	0.015	0.01

По полученным расчётам можно сделать вывод, что при использовании одиночных входных параметрах обе модели имеют очень низкое качество (R -squared близок к нулю, а иногда даже ниже нуля), и линейная регрессия показывает лучшие результаты по сравнению со «случайным лесом». Но при использовании двух параметров, ситуация меняется, и в результате «случайный лес» выдаёт лучший результат: коэффициент детерминации в лучших случаях достигает 0.85 против 0.74 у линейной регрессии. При этом стоит отметить, что комбинирование четырёх параметров, выдающих наилучший результат, «Численность населения», «Количество действующих предприятий», «НИИ» и «Энтропия» способствует ещё большему улучшению моделей. Однако поскольку при построении любой регрессионной модели необходимо стараться сократить количество используемых переменных, то увеличение количества входных параметров является недопустимым в данной ситуации. К тому же одновременное использование параметра «НИИ» и «Энтропии» в рамках одной модели является недопустимым, поскольку они имеют достаточно высокую корреляцию между собой (около 0.9).

3.3. Выводы по разделу

В рамках данного раздела был рассмотрен разработанный модуль построения регрессионной модели для оценки доли моноотрасли, который включает в себя построение матрицы корреляции, позволяющей узнать степень зависимости входных параметров, и визуализацию предсказаний регрессионной модели в виде диаграммы рассеивания, показывающей рассчитанные значения долей моноотрасли и предсказанные моделью.

По результатам сравнения двух регрессионных моделей было выявлено, что наилучший результат показывает модель «случайный лес», входными параметрами при этом являются «НИИ» и «Численность населения»,

коэффициент детерминации в лучшем случае равен 0.85, средняя абсолютная ошибка – 0.05 и средняя квадратичная ошибка – 0.007.

Глава 4. Финансовый менеджмент, ресурсоэффективность и ресурсосбережение

Целью проекта является разработка информационного портала для визуализации ряда характеристик моногородов, а также разработка регрессионной модели для оценки уровня диверсификации экономики моногорода. Перед выводом разработанного продукта на рынок информационных технологий, необходимо оценить данную разработку с точки зрения востребованности, перспективности, планирования финансовой и коммерческой ценности. Данный раздел нацелен на рассмотрение следующих задач:

- оценка коммерческого потенциала разработки;
- планирование научно-исследовательской работы;
- расчёт бюджета научно-исследовательской работы;
- определение ресурсной, финансовой, бюджетной эффективности исследования.

4.1. Анализ конкурентных технических решений

В качестве одного из представителей конкурентных технических решений можно представить аналогичный разработанный информационный портал, который вместо использования регрессионной модели для заполнения пустых данных будет использоваться среднее значение по моногородам со схожим видом деятельности. В качестве второго конкурентного решения можно рассмотреть продукт, использующий в качестве регрессионной модели линейную функцию первого порядка, которая является более простой моделью по сравнению с используемой в данной работе регрессионной моделью случайный лес. Результаты проведённого сравнения конкурентных технических решений представлены в таблице 4.

Таблица 4. Оценочная карта для сравнения конкурентных технических решений, где К1 – конкурент номер 1, К2 – конкурент номер 2 и Ф – текущая разработка

Критерии оценки	Вес критерия	Баллы			Конкурентоспособность		
		Бк1	Бк2	БФ	Кк1	Кк2	КФ
Критерии оценки эффективности							
Повышение качества данных	0.2	4	6	8	0.8	1.2	1.6
Точность вычислений	0.3	6	7	8	1.8	2.1	2.4
Простота вычислений	0.15	8	6	4	1.2	0.9	0.6
Экономические критерии оценки эффективности							
Финансирование научной разработки	0.15	4	4	4	0.6	0.6	0.6
Стоимость разработки	0.1	8	7	7	0.8	0.7	0.7
Быстрота разработки и внедрения	0.1	6	5	3	0.6	0.5	0.3
Итого	1	38	39	34	5.8	6	6.2

По результатам оценки конкурентоспособности проектов можно сделать вывод, что перспективность текущей разработки выше, чем у конкурентных решений. За счёт большей сложности используемых алгоритмов и вычислений текущая разработка уступает по некоторым показателям, однако стоит отметить, что данные критерии имеют небольшую значимость.

4.2. SWOT-анализ

SWOT-анализ применяется для исследования внешней и внутренней среды проекта. Целью использования SWOT-анализа, результаты которого указаны в таблице 5, для данной разработки является определение возможной эффективности и прогнозирование направлений для дальнейшего развития

реализуемого продукта, что позволит более наглядно оценить проект и рассмотреть его особенности с учётом существующей обстановки.

Таблица 5. Сводная матрица SWOT-анализа

	Сильные стороны разработки (Strengths):	Слабые стороны (Weaknesses):
	<p>S1. Предсказание доли моноотрасли для моногородов при отсутствии данных о градообразующем предприятии.</p> <p>S2. Единое представление об оценке и анализе диверсификации моногородов.</p> <p>S3. Выявление закономерности развития моногородов.</p>	<p>W1. При большом объёме данных возрастают требования к вычислительной мощности компьютера.</p> <p>W2. Высокая доля исходных данных с пустыми значениями.</p>
Возможности (Opportunities):	<p>O1S3. Составление более эффективной программы развития моногородов.</p> <p>O2S2. Анализ и корректировка программы развития моногородов.</p>	<p>O1W2. Предоставление филиалами отчетности муниципалитетов моногородов</p>
Угрозы (Threads):	<p>T1S3. Обеспечение поддержки областей, в состав которых входят моногорода.</p>	<p>T2W1. Привлечение новых источников финансирования.</p>
	<p>T1. Отсутствие заинтересованности участников и организаторов в использовании разработки.</p> <p>T2. Ограничения доступа к исходным сторонним данным, на которых базируется данный проект.</p>	

4.3. Формирование проекта

При проектировании и разработке любого продукта необходимо сформировать общее представление о целях проекта, а также о лицах, которые

будут наиболее заинтересованы в его успешной реализации. Заинтересованные в данной разработке стороны представлены в таблице 6.

Таблица 6. Заинтересованные стороны проекта

Заинтересованные стороны	Ожидания заинтересованных сторон
Муниципалитет моногорода	Удобный инструмент для оценки качества развития «собственных» моногородов
Инвесторы	Возможность оценить перспективность развития той или иной отрасли или предприятия в моногородах

Таблица 7 описывает цели данной разработки и ожидаемые результаты при успешной реализации.

Таблица 7. Цели и результаты разработки

Цели проекта:	Разработка информационного ресурса для оценки уровня диверсификации моногородов и визуализации данных
Ожидаемые результаты проекта:	1) Подготовленные и рассчитанные данные, характеризующие основные параметры моногородов 2) Многофункциональная визуализация данных
Критерии приёмки результатов проекта:	1) Информативность рассчитанных показателей 2) Простота и интуитивная понятность для использования
Требования к результату проекта:	1) Визуализация должна содержать пространственное сравнение моногородов с помощью карты 2) Функционал ресурса должен быть достаточно широк, чтобы предоставлять максимально много полезных функций для анализа

Для выполнения данного проекта необходимо сформировать рабочую группу, в состав которой входят научный руководитель (НР), научный консультант (НК) и исполнитель (И). Для каждого участника определены функции в проекте и рабочее время, затраченное на проект, что отражено в таблице 8.

Таблица 8. Участники проекта

№	Участник	Роль в проекте	Функции	Трудозатраты, часы
1	Соколова Вероника Валерьевна	Научный руководитель	Утверждение основных разделов, координирование деятельности исполнителя	20
2	Антонова Ирина Сергеевна	Научный консультант	Выдача заданий к исполнению, консультация по специфическим вопросам предметной	30

			области, оценка качества разработанного проекта	
3	Копасов Дмитрий Владимирович, магистрант гр.8ПМ9И	Исполнитель	Исполнение поставленных задач	280

Данный проект не имеет строго регламентированный бюджет, источники финансирования и сроки реализации, поэтому этап с описанием ограничений проекта пропущен.

Чтобы отразить ключевые события по ведению разработки, необходимо составить календарный план, который представлен в таблице 9.

Таблица 9. Календарный план проекта

Название вида работ	Продолжительность работы, рабочие дни	Дата начала	Дата окончания	Участники
Формирование требований	2	12.02	15.02	НР, НК
Анализ требований	1	16.02	16.02	И
Обсуждение возможных путей реализации	1	17.02	17.02	НР, НК, И
Подготовка исходных данных для проекта	3	18.02	22.02	НК
Предварительная обработка данных и расчёт необходимых показателей	5	24.02	02.03	И
Оценка правильности и пригодности рассчитанных данных	3	03.03	05.03	НК
Разработка визуализации	30	09.03	19.04	И
Разработка регрессионной модели	20	20.04	19.05	И
Подготовка отчёта с предварительными результатами по производственной практике	5	20.05	26.05	И
Проверка работы и формирование списка замечаний	2	27.05	28.05	НК, НР
Исправление всех замечаний и финальная полировка проекта	5	31.05	04.06	И

Подготовка финального отчёта	4	07.06	10.06	И
Сдача работы	1	11.06	11.06	НР, НК, И

На основании приведённых данных по трудозатратам каждого из участников проекта можно построить диаграмму Гантта, представленную на рисунке 22, учитывающую различных исполнителей и их параллельное исполнение работ.

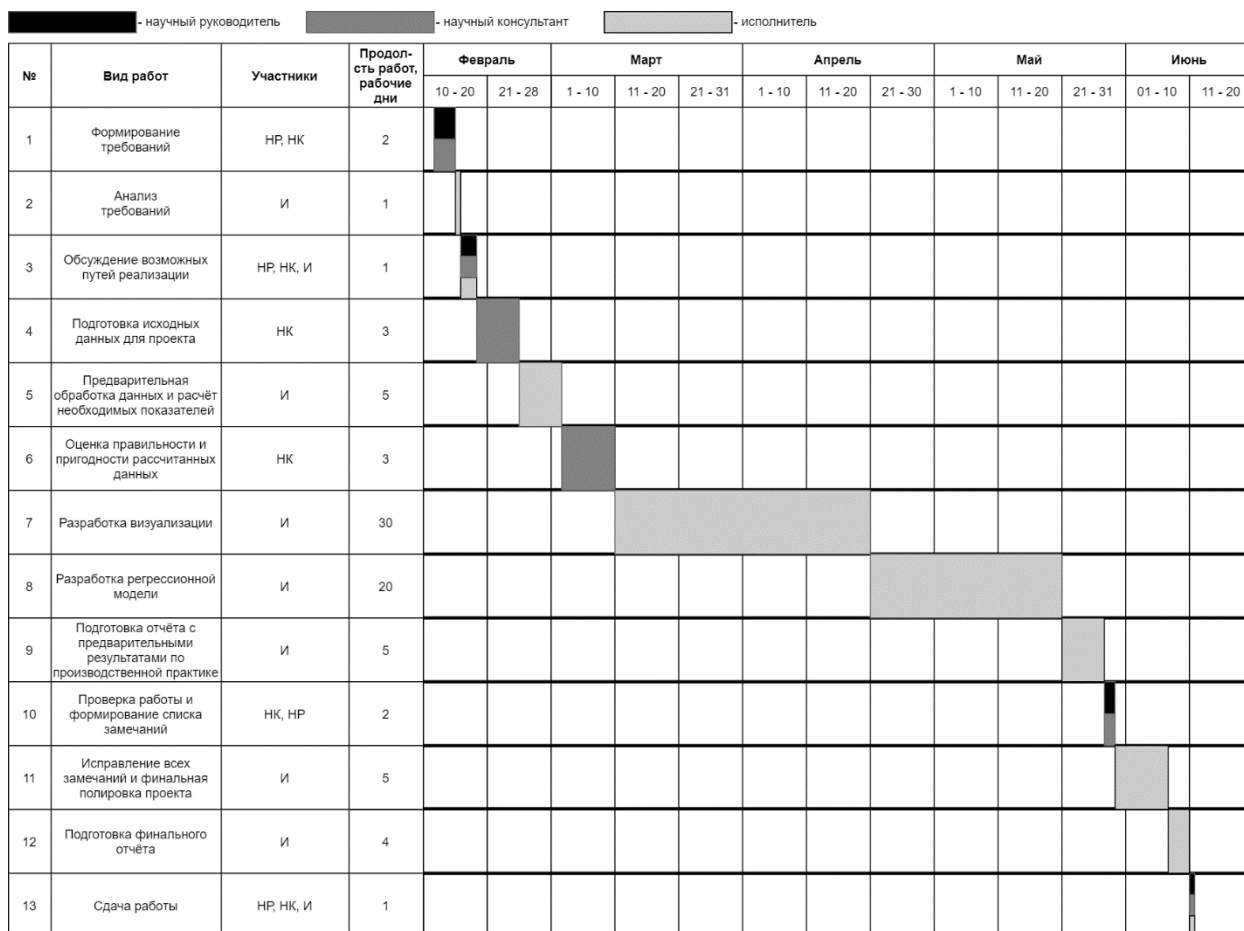


Рисунок 22. Календарный план выполнения проекта

4.4. Бюджет проекта

Для оценки бюджета проекта нужно оценить все возможные расходы, возникающие при реализации. Возможные расходы разделяются на следующие категории:

- материальные затраты проекта;
- затраты на оборудование для проведения работ;

- заработная плата участников проекта;
- отчисления во внебюджетные фонды (страховые отчисления);
- накладные расходы.

4.4.1. Расчёт материальных затрат проекта

Для данной статьи расходов необходимо учесть всевозможные расходы на различные материалы и ресурсы, которые используются и расходуются в процессе разработки. В целом это включает расходы на канцелярию – распечатка отчётов, закупка тетрадей и письменных принадлежностей, но данная работа проводится полностью с использованием компьютера, то есть нет никакой необходимости в использовании данных материалов.

При этом единственным расходуемым ресурсом в данной пункте можно считать на электроэнергию, поскольку это единственный необходимый для компьютера ресурс. Расчёт затрат на электроэнергию можно рассчитать по формуле:

$$C_{Эл} = P_{Эл} * P * T,$$

где $C_{Эл}$ – затраты на электроэнергию, $P_{Эл}$ – тариф за использование 1 кВт*ч электроэнергии (текущий тариф в Томске 2.56 рубля/(кВт*ч)), P – суммарная мощность электроприборов, используемых при разработке проекта, T – время работы электроприборов.

Предположив, что компьютеры трёх участников проекта потребляют 600 Вт/ч (3 компьютера по 200 Вт каждый), а также учитывая, что каждый участник использует свой компьютер независимо друг от друга, даже когда работа в соответствии с планом ведётся параллельно, то суммарное время работы компьютеров составляет 90 рабочих дней или 720 часов (при 8-часовом рабочем дне). Исходя из этого, общие затраты на электроэнергию составляет $C_{Эл} = P_{Эл} * P * T = 2.56 * 600 * 720/1000 = 1105.92$ рублей.

4.4.2. Расчёт заработной платы участников проекта

В данном проекте участвует три участника: научный руководитель, научный консультант и исполнитель (студент). Поскольку научный руководитель и научный консультант имеют одинаковую должность, поэтому их базовую зарплату можно считать идентичной по отношению друг к другу. Согласно статистике, в России за 2021 средняя зарплата доцента составляет 44346 рублей. Ежемесячной зарплатой исполнителя можно считать зарплату, предполагая, что он трудоустроен на должность программиста. Текущий базовый оклад на представленной должности – 12972 рублей.

Базовая зарплата участника рассчитывается по формуле:

$$S_B = S_D * T,$$

где S_B – базовая зарплата участника проекта, S_D – средняя оплата труда за один рабочий день, T – трудозатраты участника, выраженные в рабочих днях.

Средняя зарплата за день рассчитывается по формуле:

$$S_D = \frac{S_M * M}{F},$$

где S_D – средняя оплата труда за один рабочий день, S_M – ежемесячная зарплата, M – количество месяцев работы без отпуска в течение года (при отпуске в 48 дней и 6 дневной рабочей неделе $M=10.4$, при отпуске в 24 дня и 5 дневной рабочей неделе $M=11.2$), F – фактическое время работы в течении года (рабочие дни).

Таблица 10. Фактическое время работы в течении года

Индикаторы рабочего времени	Количество дней (5-дневная рабочая неделя)	Количество дней (6-дневная рабочая неделя)
Календарное число дней	365	
Количество нерабочих дней: – выходные – праздничные дни	122	70

Потери рабочего времени: – отпуск – невыходы по болезни	34	58
Фактическое время работы	209	237

Ежемесячная зарплата рассчитывается по формуле:

$$S_M = S_B * (k_p + k_b) * k_r,$$

где, S_M – ежемесячная заработная плата участника проекта, S_B – базовая ставка оплаты труда (оклад), k_p – ставка надбавки (можно считать, что данный проект выполняется без надбавок и $k_p = 1$), k_b – бонусная ставка (можно принять, что бонус к основной зарплате будет составлять 20%, то есть $k_b = 0.2$), k_r – региональный коэффициент (для Томска $k_r = 1.3$).

Расчёт итоговой зарплате всех участников произведён в таблице 11.

Таблица 11. Расчёт основной зарплата

Участник	S, руб.	k_p	k_b	k_r	S_M , руб.	S_d , руб./день	T, дни	S_B , руб
Научный руководитель	44346	1	0.2	1.3	69179.76	3035.74	6	18214.44
Научный консультант	44346				69179.76	3035.74	12	36428.88
Исполнитель	12972				20236.32	1084.43	72	78078.96

4.4.3. Дополнительная заработная плата

Дополнительная заработная плата определяется по формуле:

$$S_{Доп} = S_B * k_{Доп},$$

где $S_{Доп}$ – дополнительная зарплата участника проекта, S_B – итоговая зарплата участника проекта, рассчитанная в предыдущем пункте, $k_{Доп}$ – коэффициент дополнительной зарплате (можно принять равным 10%).

Дополнительная заработная плата научного руководителя:

$$S_a = 18214.44 * 0.1 = 1821.44 \text{ рублей.}$$

Дополнительная заработная плата научного консультанта:

$$S_a = 36428.88 * 0.1 = 3642.89 \text{ рублей.}$$

Дополнительная заработная плата исполнителя:

$$S_a = 78078.96 * 0.1 = 7807.9 \text{ рублей.}$$

4.4.4. Расчёт расходов на отчисления во внебюджетные фонды

Помимо выплат заработной платы также производятся выплаты во внебюджетные фонды: пенсионный фонд, страхование и т.д. Размер данного типа расходов определяется по формуле:

$$P = k_b * (S_B + S_{\text{Доп}}),$$

где P – расходы на дополнительные отчисления, k_b – коэффициент выплат (равен 27.1% от заработной платы для образовательной и научной деятельности), S_B – зарплата участника проекта, $S_{\text{Доп}}$ – дополнительная зарплата участника проекта.

В соответствии с этой формулой расходы на отчисления во внебюджетные фонды для каждого участника рассчитаны в таблице 12.

Таблица 12. Отчисления во внебюджетные фонды

	Научный руководитель	Научный консультант	Исполнитель
Коэффициент отчислений	0.271		
Полная зарплата, руб.	20035.88	40071.77	85886.86
Размер отчислений, руб.	5429.72	10859.45	23275.34

4.4.5. Накладные расходы

Накладные расходы учитывают прочие затраты при разработке, не попавшие в предыдущие статьи расходов: ремонт и обслуживание оборудования, содержание рабочего помещения, оплата услуг интернета и связи, электроэнергии и прочее. Для упрощения можно рассчитать данный вид расходов по следующей формуле:

$$C = k_{\text{Накл}} * (S_B + S_{\text{Доп}}),$$

где C – накладные расходы, $k_{\text{Накл}}$ – коэффициент накладных расходов (можно принять равным 30%), $S_{\text{Б}}$ – зарплата участника проекта, $S_{\text{Доп}}$ – дополнительная зарплата участника проекта.

Накладные расходы для каждого участника рассчитаны в таблице 13.

Таблица 13. Накладные расходы

	Научный руководитель	Научный консультант	Исполнитель
Коэффициент накладных расходов	0.3		
Полная зарплата, руб.	20035.88	40071.77	85886.86
Накладные расходы, руб.	6010.76	12021.53	25760.06

4.4.6. Формирование бюджета проекта

Теперь необходимо подвести итог по общим затратам на проект, складываемым из всех вышеперечисленных категорий. Соответствующие результаты занесены в таблицу 14.

Таблица 14. Расчёт бюджета проекта

Наименование категории	Сумма, руб.
Материальные затраты	1105.92
Основная зарплата	132722.28
Дополнительная зарплата	13272.23
Отчисления во внебюджетные фонды	39564.51
Накладные расходы	43792.35
Итоговые затраты:	230457.29

4.5. Определение ресурсной (ресурсосберегающей), финансовой, бюджетной, социальной и экономической эффективности исследования

Определение эффективности происходит на основе расчёта интегрального показателя эффективности научного исследования. Его

нахождение связано с определением двух средневзвешенных величин: финансовой эффективности и ресурсоэффективности.

Интегральный показатель финансовой эффективности научного исследования получается в результате оценки бюджетных затрат нескольких вариантов реализации проекта. Для этого наибольший интегральный показатель реализации технической задачи принимается за базу расчёта (знаменатель), с которым соотносятся финансовые значения по всем вариантам исполнения.

Интегральный финансовый показатель разработки определяется как:

$$I_{\text{финр } i} = \frac{\Phi_{pi}}{\Phi_{\text{max}}},$$

где $I_{\text{финр}}$ – интегральный финансовый показатель разработки, Φ_{pi} – стоимость i -го варианта исполнения, Φ_{max} – максимальная стоимость исполнения проекта (в т.ч. аналоги).

Полученная величина интегрального финансового показателя разработки отражает соответствующее численное увеличение стоимости разработки, либо соответствующее численное уменьшение стоимости разработки.

Интегральный показатель ресурсоэффективности вариантов исполнения проекта можно определить следующим образом:

$$I_{pi} = \sum a_i * b_i,$$

где I_{pi} – интегральный показатель ресурсоэффективности для i -го варианта исполнения разработки, a_i – весовой коэффициент i -го варианта исполнения разработки, b_i – бальная оценка i -го варианта исполнения разработки, устанавливается экспертным путем по выбранной шкале оценивания.

Расчёт интегрального показателя ресурсоэффективности приведён в таблице 15.

Таблица 15. Расчёт интегрального показателя ресурсоэффективности

Критерий	Вес критерия (сумма = 1)	Исп1	Исп2	Исп3
1. Количество рассчитанных показателей экономической эффективности	0.4	4	3	4
2. Требуется наличие большого количества данных	0.1	3	4	4
3. Простота применения	0.2	2	2	3
4. Конкурентоспособность	0.1	5	3	2
5. Возможность применения любыми муниципалитетами	0.2	4	5	3

$$I_{p-исп1} = 1.6 + 0.3 + 0.4 + 0.5 + 0.8 = 3.6$$

$$I_{p-исп2} = 1.2 + 0.4 + 0.4 + 0.3 + 1 = 3.3$$

$$I_{p-исп3} = 1.6 + 0.4 + 0.6 + 0.2 + 0.6 = 3.4$$

Интегральный показатель эффективности вариантов исполнения разработки ($I_{испi}$) определяется на основании интегрального показателя ресурсоэффективности и интегрального финансового показателя по формуле:

$$I_{исп i} = \frac{I_{p-исп i}}{I_{финр i}}$$

Сравнение интегрального показателя эффективности вариантов исполнения разработки позволит определить сравнительную эффективность проекта и выбрать наиболее целесообразный вариант из предложенных. Сравнительная эффективность проекта (\mathcal{E}_{cp}):

$$\mathcal{E}_{cp} = \frac{I_{исп1}}{I_{исп2}}$$

В таблице 16 приведены результаты сравнения эффективности разработки.

Таблица 16. Сравнительная эффективность разработки

Показатели	Исп1	Исп2	Исп3
Интегральный финансовый показатель разработки	1	1	1
Интегральный показатель ресурсоэффективности разработки	3.6	3.3	3.4

Интегральный показатель эффективности	3.6	3.3	3.4
Сравнительная эффективность вариантов исполнения	1.09	0.97	1.03

Сравнение значений интегральных показателей эффективности позволяет понять и выбрать более эффективный вариант решения поставленной в диссертации технической задачи с позиции финансовой и ресурсной эффективности.

4.6. Заключение по разделу

В результате выполнения данного раздела были получены следующие результаты:

1. Проведён анализ конкурентных технических решений, в результате которого выявлено, что текущая разработка по совокупности показателей является конкурентоспособной по сравнению с аналогичными решениями.
2. Проведён SWOT-анализ, в результате которого были выявлены сильные и слабые стороны, возможности и угрозы проекта.
3. При проведении планирования был разработан план-график проведения работ для научного руководителя, научного консультанта и исполнителя (студента). Было определено общее количество рабочих дней необходимых для выполнения данного проекта – 82 дней.
4. Сформирован бюджет проекта, который составил 230457.29 рублей.
5. Определены интегральный финансовый показатель, интегральный показатель эффективности, интегральный показатель эффективных вариантов исполнения. Рассчитанные показатели показали, что проект эффективен с позиции финансовой и ресурсной эффективности.

Глава 5. Социальная ответственность

В данном разделе проведён анализ опасных факторов труда, определён комплекс мер организационного, технического и правового характера, который должен препятствовать возникновению или снижению возможности негативных последствий в процессе осуществления работы разработчика программного обеспечения.

Выпускная квалификационная работа по разработке информационного портала для оценки уровня диверсификации моногородов и визуализации данных выполняется главным образом на рабочем месте в офисном помещении и представляет собой в просторную комнату со следующими параметрами:

- ширина рабочего помещения – 7 м, длина – 8.5 м, высота – 3 м;
- площадь помещения – 59.5 м²;
- в помещении установлено искусственное освещение (набор люминесцентных ламп по всему периметру помещения), а также присутствует естественное освещение;
- число одновременно работающих человек в помещении от 4 до 8.

5.1. Правовые и организационные вопросы обеспечения безопасности

5.1.1. Специальные правовые нормы трудового законодательства

Трудовой кодекс РФ 197-ФЗ [7] предусматривает рациональную организацию труда во время рабочей смены, согласно которой:

- продолжительность рабочей недели не должна превышать 40 часов (8 часов при пятидневной рабочей неделе);
- должны быть предусмотрены два перерыва – не менее 20 минут после 1-2 часов работы или не менее 30 минут после 2 часов работы;
- обеденный перерыв должен быть не менее 30 минут.

Также, Трудовым кодексом закреплён обязательный предварительный медицинский осмотр при приеме на работу и периодические медицинские осмотры.

Каждый сотрудник должен пройти инструктаж по технике безопасности перед приёмом на работу и в дальнейшем, должен быть пройден инструктаж по электробезопасности и охране труда. Каждому работнику обязательно должна быть предоставлена рабочая инструкция, с описанием входящих в его должность функций и рабочих моментов, а также конкретным описанием границ ответственности.

5.1.2. Организационные мероприятия при компоновке рабочей зоны

Рабочее место должно быть организовано с учётом требований ГОСТ 12.2.032-78 «ССБТ. Рабочее место при выполнении работ сидя. Общие эргономические требования» [8], а также СП 2.4.3648-20 «Санитарно-эпидемиологические требования к организациям воспитания и обучения, отдыха и оздоровления детей и молодежи» [9].

Согласно ГОСТ 12.2.032-78 при организации работы рабочая зона должна соответствовать ряду требований:

- рабочее место при выполнении работ сидя организуется при легком виде работе, не требующем свободного и постоянного передвижения работника по помещению;
- конструкция рабочего места и взаимное расположение всех его элементов должны соответствовать антропометрическим, физиологическим и психологическим требованиям, а также характеру работы;
- в соответствии с категорией работ (работа за компьютером) при отсутствии механизмов регулирования высоты рабочей

поверхности, высота рабочей поверхности, составляет (для мужчин) 680 мм. При этом высота сиденья должна составлять 430 мм.

- В соответствии с СП 2.4.3648-20 были выявлены следующие правила организации рабочего помещения:
- окна в помещениях должны быть оборудованы регулируемыми солнцезащитными устройствами (жалюзи, шторы), а окна, открываемые в весенний, летний или осенний период должны быть оборудованы москитными сетками;
- в системе освещения помещения допускается использование разрядных, люминесцентных и светодиодных ламп, при этом в одном помещении не допускается комбинировать разные типы ламп;
- в учебных помещениях уровень искусственного освещения должен быть выше 300 люкс.

5.2. Производственная безопасность

5.2.1. Анализ вредных и опасных факторов

Согласно ГОСТ 12.0.003-2015 [10] в таблице 17 указаны все потенциальные вредные факторы и опасные факторы, а также нормативные документы, регулирующие их допустимые границы на рабочем месте.

Таблица 17. Возможные опасные и вредные производственные факторы

Факторы (ГОСТ 12.0.003-2015)	Этапы работ			Нормативные документы
	Проектирование	Разработка	Тестирование	
1. Отклонение показателей микроклимата	+	+	+	СанПиН 2.2.4.548-96 – Гигиенические требования к микроклимату производственных помещений

2. Повышенный уровень шума	+	+	+	СанПиН 1.2.3685-21 – Гигиенические нормативы и требования к обеспечению безопасности и (или) безвредности для человека факторов среды обитания
3. Отсутствие или недостаток необходимого естественного освещения	+	+	+	СП 52.13330.2016 – Естественное и искусственное освещение
4. Умственное переутомление	+	+	+	ТОИ Р-45-084-01 – Типовая инструкция по охране труда при работе на персональном компьютере
5. Повышенное значение напряжения в электрической цепи, замыкание которой может произойти через тело человека	+	+	+	ГОСТ 12.1.019-2017 – Система стандартов безопасности труда (ССБТ). Электробезопасность. Общие требования

5.2.1.1. Отклонение показателей микроклимата

Под микроклиматом рабочего помещения понимают климат внутренней среды помещения, в котором находятся сотрудники в течение рабочего времени. Микроклимат определяется совокупностью показателей, действующих на организм работника – температуры воздуха и поверхностей, относительной влажности воздуха, скорости движения воздуха и интенсивности теплового облучения.

Нормативные показатели микроклимата регламентируются СанПиН 2.2.4.548-96 «Гигиенические требования к микроклимату производственных помещений» [11]. Санитарные нормы устанавливают оптимальные и допустимые значения показателей в рабочей зоне, что позволяет создавать благоприятные условия работы, соответствующие физиологическим потребностям организма человека.

Работа, выполняемая разработчиком программного обеспечения, относится к категории Ia, так как она является мало подвижной и мало интенсивной, проводится сидя с минимальными физическими трудозатратами. В таблице 18 представлены оптимальные и допустимые значения показателей микроклимата на рабочих местах для данной категории.

Таблица 18. Оптимальные величины показателей микроклимата на рабочих местах производственных помещений

Период года	Температура воздуха, °С	Температура поверхностей, °С	Относительная влажность воздуха, %	Скорость движения воздуха, м/с
Холодный	22 – 24	21 – 25	40 – 60	0.1
Тёплый	23 – 25	22 – 26	40 – 60	0.1

Значения показателей, полученные при измерении на рабочем месте:

- температура воздуха 25 °С – оптимальное значение;
- температура поверхностей 22 °С – оптимальное значение;
- относительная влажность воздуха 40 % – оптимальное значение.

5.2.1.2. Повышенный уровень шума

Превышения уровня шума является вредным фактором на рабочем месте. Постоянный шум, превышающий допустимые значения, не только воздействует на органы слуха, но и влияет на общее самочувствие работника, способствует ослаблению организма, а также снижает работоспособность.

Возможными источниками шума на рабочем месте являются компьютеры, находящиеся в помещении, а именно кулеры системы охлаждения процессора, кулеры обдува корпуса компьютера, клавиатура и мышь. Для снижения уровня шума от перечисленного оборудования можно при необходимости заменить наиболее шумные из них на устройства с низким уровнем шума. Также может помочь внедрение шумоизоляции или перекомпоновка рабочего места, чтобы изолировать системный блок. В таблице 19 описаны допустимые значения уровня шума на рабочем месте в соответствии с СанПиНом 1.2.3685-21 [12].

Предполагается [13], что издаваемый компьютером шум может иметь различную частоту вплоть до 4 кГц в зависимости от уровня нагрузки и интенсивности работы. При данной частоте допустимый уровень шума

составляет 30 дБА, но при меньших частотах допустим шум вплоть до 79 дБА (СанПиН 1.2.3685-21).

Таблица 19. Нормируемые параметры шума в октавных полосах частот, эквивалентных и максимальных уровней звука проникающего шума в помещениях

Наименование помещения	Для источников постоянного шума										Для источников непостоянного шума	
	Уровни звукового давления в октавных полосах со среднегеометрическими частотами, Гц									Уровни звука L(A), дБА	Эквивалентные уровни звук L(Aэкв.), дБА	Максимальные уровни звука L(Aмакс.), дБА
	31.5	63	125	250	500	1000	2000	4000	8000			
Классные помещения, учебные кабинеты, учительские комнаты, аудитории образовательных организаций, конференц-залы, читальные залы библиотек	79	63	52	45	39	35	32	30	28	40	40	55

5.2.1.3. Недостаток необходимого естественного или искусственного освещения

Отсутствие естественного света в тёмное время суток является естественным явлением, актуальным для Томска в течение рабочего дня в зимний период в силу короткого светового дня. Однако в случае, если основным источником освещения является естественный свет через окна, освещение должно соответствовать нормам согласно СП 52.13330.2016 [14]. В таблице 20 представлены требования к освещению помещений промышленных предприятий при выполнении работ, требующих малой точности, поскольку наименьшим различимым объектом на рабочем месте является символы на клавиатуре, размер которых составляет 2-4 мм.

Таблица 20. Требования к освещению помещений промышленных предприятий при выполнении работ, требующих малой точности

Характер зрительной работы	Наименьший или эквивалентный размер объекта различения, мм	Разряд зрительной работы	Искусственное освещение			Естественное освещение	Совмещенное освещение		
			При системе общего освещения	Сочетание нормируемых величин объединенного показателя дискомфорта UGR и коэффициента пульсации		КЕО e_n , %			
				UGR, не более	K_L , %, не более	При верхнем или комбинированном освещении	При боковом освещении	При верхнем или комбинированном освещении	При боковом освещении
Малая точность	От 1 до 5	V	200	25	20	3	1	1.8	0.6

5.2.1.4. Описание показателей освещённости рабочего помещения

Исходными данными являются размеры помещения длина $A = 7$ м, ширина $B = 8.5$ м и высота $H = 3$ м, высота рабочей поверхности $h_{рп} = 0.7$ м. Норма освещённости согласно СП 52.13330.2016 для данного типа работ составляет 200 лк.

Индекс помещения определяется по формуле:

$$I = \frac{A * B}{(H - h_{рп}) * (A + B)},$$

где I – индекс помещения; h_1 – расстояние от пола до источника света; h_2 – высота рабочей поверхности; a – длина помещения; b – ширина помещения.

Для представленного рабочего помещения рассчитаем:

$$I = \frac{7 * 8.5}{(3 - 0.7) * (7 + 8.5)} = 1.67$$

По полученному индексу помещения получается, что коэффициент использования светового потока $\eta = 0.55$ (для предполагаемых коэффициентов отражения стен $R_c = 0.3$ и отражения потолка $R_{п} = 0.5$).

Определено, что помещение оборудовано набором из 9 потолочных светильников по 4 люминесцентных лампы в каждом светильнике. Предполагается, что для освещения используются распространённые люминесцентные лампы ЛБ 40, световой поток которых составляет $\Phi = 2800$ лм. Коэффициент запаса $K_3 = 1.5$, коэффициент неравномерности $Z = 1.1$.

Оптимальный световой поток, создаваемый лампой, определяется формулой:

$$\Phi = \frac{E_{н} * S * K_3 * Z}{N_{л} * \eta},$$

где $E_{н}$ – нормативное значение освещённости в помещении, S – площадь помещения, K_3 – коэффициент запаса, учитывающий загрязнение светильника и воздуха в помещении, Z – коэффициент неравномерности освещения, $N_{л}$ –

общее число ламп в помещении, η – коэффициент использования светового потока.

Предполагая, что световой поток, создаваемый лампой известен (определён для конкретной лампы), то, преобразовав формулу выше, можно найти величину освещённости, создаваемое в данном помещении, по формуле:

$$E_{\Pi} = \frac{\Phi * N_{\text{Л}} * \eta}{S * K_3 * Z},$$

где E_{Π} – искомое значение освещённости в данном помещении.

Отсюда следует, что:

$$E_{\Pi} = \frac{\Phi * N_{\text{Л}} * \eta}{S * K_3 * Z} = \frac{2800 * 36 * 0.55}{7 * 8.5 * 1.5 * 1.1} = 565 \text{ лк}$$

Полученное значение освещённости рабочего места равно 565 лк, что больше требуемых 200 лк для данного типа работ, поэтому общую освещённость можно уменьшить либо за счёт замены ламп на лампы с более низким световым потоком, либо сократить число используемых для освещения ламп.

5.2.1.5. Умственное переутомление

Работая за компьютером, работник находится под влиянием вредного производственного фактора, относящегося к нервно-психической группе – умственного переутомление.

Согласно ТООИ Р-45-084-01 [15], виды трудовой деятельности разделяются на 3 группы: группа А – работа по считыванию информации с экрана компьютера с предварительным запросом; группа Б – работа по вводу информации; группа В – творческая работа в режиме диалога с компьютером. При выполнении в течение рабочей смены работ, относящихся к различным видам трудовой деятельности, за основную работу с компьютером следует принимать такую, которая занимает не менее 50% времени в течение рабочей смены или рабочего дня. Уровень нагрузки представлен в таблице 21.

Таблица 21. Уровень нагрузки за рабочую смену при видах работ с компьютером

Категория работ	Уровень нагрузки за рабочую смену при видах работ с компьютером		
	группа А, количество знаков	группа Б, количество знаков	группа В, часов
III	До 60000	До 40000	До 6.0

Для предотвращения и снижения переутомления работников рекомендуется устраивать периодические перерывы во время рабочего дня, а вне рабочего времени важно соблюдать режим сна и поддерживать умеренную физическую активность.

5.2.1.6. Повышенное значение напряжения в электрической цепи

К опасным факторам на рабочем месте относится поражение электрическим током, поскольку работник взаимодействует с электрооборудованием – различные разъёмы подключения к электросети (розетки, удлинители или сетевые фильтры), компьютер, его комплектующие и периферийные устройства.

Согласно ГОСТ 12.1.019-2017 ССБТ [16], во избежание поражения работника электрическим током следует соблюдать следующие базовые принципы:

1. Проводящие части, находящиеся под опасным рабочим, наведенным, остаточным напряжением, не должны быть доступными, а доступные проводящие части не должны находиться под опасным напряжением при нормальных условиях (при отсутствии повреждения), а также в случае единичного повреждения.

2. Защиту при нормальных условиях (защиту от прямого прикосновения) обеспечивают посредством основной защиты, а защиту при условиях единичного повреждения (защиту при косвенном прикосновении) обеспечивают посредством защиты при повреждении.

Степень опасного воздействия на человека электрического тока электрической дуги зависит от:

1. величины напряжения прикосновения, электрического сопротивления тела человека, силы тока, протекающей через него, а также величины падающей энергии электрической дуги;
2. рода (постоянный, переменный, выпрямленный) тока и частоты переменного электрического тока;
3. пути протекания тока через тело человека и площади контакта электрической дуги с поверхностью тела человека;
4. продолжительности воздействия электрического тока и электрической дуги на организм человека;
5. индивидуальных особенностей организма человека;
6. условий внешней среды.

5.3. Экологическая безопасность

Утилизация компьютерной и организационной техники регулируется законодательно согласно ГОСТ Р 55102-2012 [17]. Переработка компьютерного оборудования происходит путём обязательного извлечения компонентов, их сортировки и последующей отправки для повторного использования. Такая утилизация обязательно производится на оборудованных полигонах с привлечением квалифицированного персонала, поскольку без должного обращения тяжёлые металлы и прочие опасные материалы, содержащиеся в технике, могут нанести существенный вред окружающей среде и живым организмам при попадании в водоёмы и грунт.

Используемые на рабочем месте люминесцентные лампы содержат ртуть, негативно влияющую на здоровье человека и окружающую среду, поэтому люминесцентные лампы и все прочие ртутьсодержащие отходы и

вышедшие из строя приборы, содержащих ртуть, подлежат сбору и возврату для последующей регенерации ртути в специализированных организациях согласно ГОСТ 12.3.031-83 [18]. Главным условием при замене и сборе отработанных ртутьсодержащих ламп является сохранение их герметичности. Факт сдачи ртутьсодержащих отходов подтверждается возвращением паспорта на вывоз отходов с отметкой о приёме представителя специализированного предприятия.

Утилизация мусорных отходов, таких как бумажная макулатура, отходы от канцелярских принадлежностей, отходы от продуктов питания, производится через сбор, обязательную сортировку и утилизацию. Отходы, которые можно использовать повторно, после сортировки отправляют на переработку через организации, занимающиеся сбором таких отходов.

Используя такую систему утилизации отходов работы можно значительно уменьшить воздействие на экологию, так как качественная утилизация отходов исключает отравление опасными веществами и попадание тяжелых металлов в окружающую среду.

5.4. Безопасность в чрезвычайных ситуациях

Наиболее вероятной чрезвычайной ситуацией для представленного рабочего помещения является пожар. Данная чрезвычайная ситуация может произойти в случае несоблюдения мер пожаробезопасности, нарушения техники использования электрических приборов и ПК, нарушениях разводки электрических сетей и ряда других комплексных причин.

Данное рабочее помещение относится к категории В1 – пожароопасное помещение, характеризующееся наличием горючих и трудногорючих жидкостей, веществ и материалов.

В качестве возможных причин возникновения пожара можно указать следующее:

- короткое замыкание;

- опасная перегрузка сетей, которая ведёт за собой сильный нагрев токоведущих частей и загорание изоляции;
- пуск оборудования после некорректного и неквалифицированного ремонта.

Для предотвращения чрезвычайных ситуаций необходимо соблюдать правила пожарной безопасности, чтобы обеспечить состояние защищенности работников и оборудования от пожара.

Для предупреждения возникновения пожаров необходимо исключить образование горючей среды, следить за применением при строительстве и отделке зданий невоспламеняющихся или трудновоспламеняющихся материалов.

Необходимо проводить следующие пожарно-профилактические мероприятия:

- организационные мероприятия, касающиеся технического процесса с учётом пожарной безопасности объекта (инструктаж персонала, обучение правилам техники безопасности, издание инструкций, планов эвакуации);
- эксплуатационные мероприятия, рассматривающие эксплуатацию используемого оборудования (соблюдение эксплуатационных норм оборудования, обеспечение свободного подхода к оборудованию, поддержание исправности изоляции проводников);
- технические и конструктивные мероприятия, связанные с правильным размещением и монтажом электрооборудования и отопительных приборов (соблюдение противопожарных мероприятий при устройстве электропроводок, оборудования, систем отопления, вентиляции и освещения).

Для повышения устойчивости рабочего помещения к чрезвычайным ситуациям необходимо произвести установку систем противопожарной

сигнализации, реагирующих на дым и другие продукты горения, установку огнетушителей.

В случае угрозы возникновения пожара необходимо вызвать противопожарную службу, отключить питание электрооборудования и, следуя плану эвакуации, эвакуировать находящихся в помещении людей и покинуть помещение. В случае, если очаг возгорания является небольшим, и нет угрозы поражения электрическим током, можно использоваться углекислотными огнетушители, расположение которых можно найти на плане эвакуации.

Заключение

В результате выполнения выпускной квалификационной работы были рассчитаны основные показатели диверсификации экономики моногородов, разработано веб-приложение (доступно по адресу <https://monotowns.web.app>), которое включает в себя модули визуализации моногородов и их экономических показателей на карте, а также модуль построения регрессионной модели для оценки доли моноотрасли в моногородах, где для расчёта данного показателя недостаточно исходных данных о деятельности градообразующих предприятий в этих моногородах.

Проведено сравнение эффективности использования регрессионных моделей линейной регрессии и случайного леса, в результате чего было выявлено, что лучшей моделью в рамках данной работы является случайный лес, построенный на наборе входных данных численности населения моногородов и индекса Херфиндаля-Хиршмана.

Научной новизной данной работы можно назвать использование оригинальной методики расчёта показателей уровня диверсификации экономики моногородов, оригинального подхода применения регрессионной модели случайный лес для оценки доли моноотрасли и оригинального алгоритма визуализации моногородов на карте веб-приложения.

Список публикаций и научных достижений

1. Стипендия Президента РФ по ПНР на 2020-2021 учебный год.
2. Академический обмен по программе FIRST+ (Finnish–Russian Student and Teacher Exchange Programme) между ТПУ и VAMK в городе Вааса, Финляндия.
3. По результатам дипломной работы составлена статья для участия в «VI Международном симпозиуме по региональной экономике» (пока не опубликована).

Список используемых источников

1. Список моногородов России – [Электронный ресурс]. URL: https://ru.wikipedia.org/wiki/Список_моногородов_России (дата обращения 10.06.2021).
2. Колесникова, О. А. Рынок труда монопрофильных территорий: проблемы и решения: монография / О. А. Колесникова; под редакцией И. Я. Львовича. – Воронеж: ВИВТ, 2016. – 172 с. – ISBN 978-5-4446-0771-8. – Текст: электронный // Лань: электронно-библиотечная система. – URL: <https://e.lanbook.com/book/157477> (дата обращения: 15.06.2021). – Режим доступа: для авториз. пользователей.
3. Государев, И. Б. Введение в веб-разработку на языке JavaScript: учебное пособие / И. Б. Государев. – Санкт-Петербург: Лань, 2019. – 144 с. – ISBN 978-5-8114-3539-5. – Текст: электронный // Лань: электронно-библиотечная система. – URL: <https://e.lanbook.com/book/118648> (дата обращения: 15.06.2021). – Режим доступа: для авториз. пользователей.
4. Тарланов, А. Т. Основы языка программирования Python: учебно-методическое пособие / А. Т. Тарланов, Ш. Г. Магомедов. – Москва: РТУ МИРЭА, 2019. – 107 с. – Текст: электронный // Лань: электронно-библиотечная система. – URL: <https://e.lanbook.com/book/171465> (дата обращения: 16.06.2021). – Режим доступа: для авториз. пользователей.
5. Рашка, С. Python и машинное обучение: крайне необходимое пособие по новейшей предсказательной аналитике, обязательное для более глубокого понимания методологии машинного обучения: руководство / С. Рашка; перевод с английского А. В. Логунова. – Москва: ДМК Пресс, 2017. – 418 с. – ISBN 978-5-97060-409-0. – Текст: электронный // Лань: электронно-библиотечная система. – URL: <https://e.lanbook.com/book/100905> (дата обращения: 17.06.2021). – Режим доступа: для авториз. пользователей.
6. Круценюк, К. Ю. Корреляционно-регрессионный анализ в эконометрических моделях: учебное пособие / К. Ю. Круценюк. – Норильск:

НГИИ, 2018. – 108 с. – ISBN 978-5-89009-698-2. – Текст: электронный // Лань: электронно-библиотечная система. – URL: <https://e.lanbook.com/book/155915> (дата обращения: 15.06.2021). – Режим доступа: для авториз. пользователей.

7. Трудовой кодекс Российской Федерации от 30.12.2001 N 197-ФЗ.
8. ГОСТ 12.2.032-78 ССБТ. Рабочее место при выполнении работ сидя. Общие эргономические требования.
9. СП 2.4.3648-20. Санитарно-эпидемиологические требования к организациям воспитания и обучения, отдыха и оздоровления детей и молодежи.
10. ГОСТ 12.0.003-2015. Система стандартов безопасности труда (ССБТ). Опасные и вредные производственные факторы.
11. СанПиН 2.2.4.548-96. Гигиенические требования к микроклимату производственных помещений.
12. СанПиН 1.2.3685-21. Гигиенические нормативы и требования к обеспечению безопасности и (или) безвредности для человека факторов среды обитания.
13. Шум современных компьютерных систем охлаждения – [Электронный ресурс] URL: <https://www.ixbt.com/cpu/shum1metod.shtml> (дата обращения 29.05.2021).
14. СП 52.13330.2016. Естественное и искусственное освещение.
15. ТОИ Р-45-084-01. Типовая инструкция по охране труда при работе на персональном компьютере.
16. ГОСТ 12.1.019-2017. Система стандартов безопасности труда (ССБТ). Электробезопасность. Общие требования и номенклатура видов защиты.
17. ГОСТ Р 55102-2012. Обращение с отходами. Руководство по безопасному сбору, хранению, транспортированию и разборке отработавшего электротехнического и электронного оборудования.
18. ГОСТ 12.3.031-83. Работы со ртутью. Требования безопасности.

Приложение А

Data preparation and economic indicators calculation

Студент:

Группа	ФИО	Подпись	Дата
8ПМ9И	Копасов Дмитрий Владимирович		

Руководитель ИШИТР ОИТ:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОИТ	Соколова Вероника Валерьевна	к.т.н.		

Консультант – лингвист ШБИП ОИЯ:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОИЯ	Сидоренко Татьяна Валерьевна	к.п.н.		

Chapter 1. Data preparation and economic indicators calculation

1.1. Initial data definition

The initial data consists of a set of Excel files that contain information about enterprises in monotowns (figure 1). Those files include more than one hundred features that describe different business aspects, such as company name, address, registration and liquidation dates, and economic activity indicators during 2013-2017 (taxes, insurance payments, fixed assets, revenue, labor costs, dividends, etc.).

№	Наименование полное	Адрес (место нахождения)	Дата регистрации	Возраст компани	Дата ликвидации	Статус	Код налогоплательщика	Код статистики	Регион регистрации	Код ОКATO	Наименование ОКATO
1	ОТКРЫТОЕ АКЦИОНЕРНОЕ ОБЩЕСТВО "АЛТАЙ-НОКС"	659107, Алтайский кр, г. Заринск, ул. Приважная, д. 2	28.10.1992	26,0		Действующая	2205001753	00188110	Алтайский край	01406000000	Алтайский край, Заринск
2	АКЦИОНЕРНОЕ ОБЩЕСТВО АЛТАЙСКОГО ВАГОНОСТРОЕНИЯ	658087, Алтайский кр, г. Новоалтайск, ул. 22 Партсъезда, д. 16	30.11.1992	26,0		Действующая	2208000010	00729244	Алтайский край	01413000000	Алтайский край, Новоалтайск
3	ЗАКРЫТОЕ АКЦИОНЕРНОЕ ОБЩЕСТВО "АЛЕРКСВЕРНОПРОДУКТ" ИМЕНИ С.Н.СТАРОВОЙТОВА	658130, Алтайский кр, г. Алейск, ул. Первомайская, д. 81	30.04.1996	22,5		Действующая	2201000766	00932382	Алтайский край	01403000000	Алтайский край, Алейск
4	ОТКРЫТОЕ АКЦИОНЕРНОЕ ОБЩЕСТВО "КУЧУКСУЛЬФАТ"	Благовещенский район, рабочий пос. Степное Озеро	18.12.1992	26,0		Действующая	2235001430	05762306	Алтайский край	01205557000	Благовещенский район, Степное Озеро
5	ОБЩЕСТВО С ОГРАНИЧЕННОЙ ОТВЕТСТВЕННОСТЬЮ "ХОЛОД"	659100, Алтайский кр, г. Заринск, ул. Зеленая, д. 51	15.11.1999	19,0		Действующая	2205006790	50665727	Алтайский край	01406000000	Алтайский край, Заринск
6	ГОСУДАРСТВЕННОЕ УНИТАРНОЕ ПРЕДПРИЯТИЕ ДОРОЖНОГО ХОЗЯЙСТВА АЛТАЙСКОГО КРАЯ "СЕВЕРО-ВОСТОЧНОЕ ДОРОЖНО-СТРОИТЕЛЬНОЕ УПРАВЛЕНИЕ"	658083, Алтайский кр, г. Новоалтайск, ул. Дорожная, д. 7	23.07.2003	15,0		Действующая	2208010989	14723085	Алтайский край	01413000000	Алтайский край, Новоалтайск
7	ГОСУДАРСТВЕННОЕ УНИТАРНОЕ ПРЕДПРИЯТИЕ ДОРОЖНОГО ХОЗЯЙСТВА АЛТАЙСКОГО КРАЯ "ЮЖНОЕ ДОРОЖНО-СТРОИТЕЛЬНОЕ УПРАВЛЕНИЕ"	658131, Алтайский кр, г. Алейск, ул. Дорожная, д. 6 юрп. А	23.07.2003	15,0		Действующая	2201006430	14725913	Алтайский край	01403000000	Алтайский край, Алейск
8	ОБЩЕСТВО С ОГРАНИЧЕННОЙ ОТВЕТСТВЕННОСТЬЮ "ЭКОДОМСТРОЙ"	658084, Алтайский кр, г. Новоалтайск, ул. Вагоностроительная, д. 9	14.01.2015	4,0		Действующая	2208032830	31491546	Алтайский край	01413000000	Алтайский край, Новоалтайск
9	ОБЩЕСТВО С ОГРАНИЧЕННОЙ ОТВЕТСТВЕННОСТЬЮ "ЗАРИНСКИЙ МАСОПЕРЕРАБАТЫВАЮЩИЙ ЗАВОД"	659100, Алтайский кр, г. Заринск, ул. Победы, д. 43	09.01.2008	11,0		Действующая	2205011014	83924925	Алтайский край	01406000000	Алтайский край, Заринск
10	ОТКРЫТОЕ АКЦИОНЕРНОЕ ОБЩЕСТВО "АЛТАЙСКИЙ ХИМПРОМ" ИМ. ВЕРЕЩАГИНА	658837, Алтайский кр, г. Яровое, пл. Предазводская, д. 2	27.09.2012	6,0		Действующая	2211005435	10022712	Алтайский край	01430000000	Алтайский край, Яровое
11	ОБЩЕСТВО С ОГРАНИЧЕННОЙ ОТВЕТСТВЕННОСТЬЮ "ЗАРИНСКАЯ СЕТЕВАЯ КОМПАНИЯ"	659100, Алтайский кр, г. Заринск, ул. Молодежная, д. 17	29.12.2011	7,0		Действующая	2205012650	30893142	Алтайский край	01406000000	Алтайский край, Заринск
12	ОБЩЕСТВО С ОГРАНИЧЕННОЙ ОТВЕТСТВЕННОСТЬЮ "АЛТАЙСКИЙ БЕКОН"	658087, Алтайский кр, г. Новоалтайск, ул. Заготзерно, д. 5А	28.12.2007	11,0		Действующая	2277010771	83927415	Алтайский край	01413000000	Алтайский край, Новоалтайск
13	АКЦИОНЕРНОЕ ОБЩЕСТВО "АЛТАЙСКИЙ ЗАВОД СЕЛЬСКОХОЗЯЙСТВЕННОГО МАШИНОСТРОЕНИЯ"	658083, Алтайский кр, г. Новоалтайск, ул. Дорожная, д. 74 офис 68	06.09.2010	8,0		Действующая	2222789799	67104890	Алтайский край	01413000000	Алтайский край, Новоалтайск
14	МУНИЦИПАЛЬНОЕ УНИТАРНОЕ ПРЕДПРИЯТИЕ ГОРОДА НОВОАЛТАЙСКА "НОВОАЛТАЙСКИЕ ТЕПЛОВЫЕ СЕТИ"	658080, Алтайский кр, г. Новоалтайск, ул. 7 Микрорайон, д. 25	12.03.1992	26,5		Действующая	2208002579	21454691	Алтайский край	01413000000	Алтайский край, Новоалтайск

Figure 1. Fragment of initial companies data

Among those features, a set of parameters for further analysis and monotowns economic condition calculation was selected:

- Company's name – full name of the enterprise, which is not needed for final analysis but can be used in intermediate steps;
- Registration and liquidation dates, which reduced to the year part and serve to identify if the enterprise has opened or closed at a certain year;
- Code of statistics – unique enterprise code also called OKPO code, used to identify certain companies in the town;

- Code of OKATO – unique town or region code used to group several companies within one town;
- Main activity code – a code from a classifier of economic activities also called OKVED code, which will make it possible to determine whether the company belongs to the group of main town enterprises;
- Fixed assets – long-term assets that a company has purchased and is using for the production of its’ goods and services;
- Revenue – economic indicator, that aggregates the income generated from whole business operations and services;
- The labor cost – the sum of all wages paid to employees, as well as the cost of employee benefits and payroll taxes paid by an employer.

Thus, after extracting all described companies’ indicators in monotowns (the list of necessary monotowns is set by a separate file), the whole collected data combined into one dataset. At the same time, a specially prepared file with the list of monotowns (figure 2) is used. This file defines for each monotown its’ OKATO code, a list of OKPO codes of one or more town-forming enterprises, and a list of OKVED codes of the main town industries activities types.

Моногород	ОКАТО	ОКПО	ОКВЭД
Анжоро-Судженск	32404	70622045, 165480, 45064664, 46492054, 00165480	05,46,28,14
Белово	32407	65159881, 70605762, 72292636, 57629640	10,05
Гурьевск	32413000	186223	24
Калтан	32415	92757343, 95371667	05,10,35
Киселевск	32416	22928033, 14792230, 164925, 10901721, 48622573, 93121541	05
Ленинск-Кузнецкий	32419	80298858	05
Мариинск	32422	337923	11
Междуреченск	32425	26635687	05
Мундыбаш (посёлок)	32227557	57622135	13,07
Мыски	32428	93137358	51
Новокузнецк	32431	82727542	26,27
Осинники	32434	160815	05
Полысаево	32435	160815	05
Прокопьевск	32437	70614749	10,05
Таштагол	32443	80299390	07

Figure 2. Fragment of the monotowns list and corresponding features used for analysis

Using those files, there were found monotown companies in the initial dataset with OKATO codes equals to those, which are pointed in the reference file. Group of main enterprises within each town is selected according to the OKPO code, and the OKVED code, which form the main economic industry in the town –

monoindustry. However, in addition to the prepared data, some additional information is needed to be collected from an open web-resource. This process could be done using the web-scraping approach.

1.2. Usage of web-scraping approach to collecting additional data

For the monotowns analysis, in addition to the already prepared data, it is necessary to complete the data with a number of parameters available on the open web-resource Wikipedia. This web page contains a table (figure 3) with all monotowns in Russia and includes their full name, links to a page with information about monotown, the population, a list of town-forming enterprises, and the location coordinates of this monotown. This page contains a set of well-structured data and because of this reason; it was decided to use the web-scraping approach, which allows writing a script to extract data from a web-resource using page parsing process.

Название	Численность населения (тыс. человек)	Градообразующее предприятие
Алтайский край		
Яровое	18,05	ОАО «Алтайхимпром»
Заринск	46,25	ОАО «Алтай-кокс»
Алейск	29,2	АО «Алейскзернопродукт» имени С. Н. Старовойтова

Figure 3. Fragment of the table with extra information about monotowns

By using the following code, it possible to extract all presented information about all monotowns in the table and convert it into an easy to use structure:

```

browser = webdriver.Chrome('chromedriver.exe')
browser.get('https://ru.wikipedia.org/wiki/Список_моногородов_России')
html = bs.BeautifulSoup(browser.page_source, 'lxml')
towns_stats, towns, region = dict(), [], None
getText = lambda item: item.text.replace('\n', '').replace('\xa0', ' ').strip()
for table_block in html.select('tr')[1:]:
    row = table_block.select('td')
    if len(row) == 1:
        if region is not None:
            towns_stats[region] = towns
            towns = []
        region = getText(row[0])
    else:
        population = getText(row[1])
        if population != '':
            population = int(round(float(population.replace(',', '.'))*1000))
        link = row[0].select_one('a')
        if link is not None:
            link = 'https://ru.wikipedia.org' + link.get('href')
        towns.append({
            'name': getText(row[0]),
            'region': region,
            'population': population,
            'company': getText(row[2]).split('[')[0],
            'link': link
        })
if region is not None:
    towns_stats[region] = towns

```

The structure extracted in this way has the following content:

```

{
  "Алтайский край": [
    {
      "name": "Яровое",
      "region": "Алтайский край",
      "population": 18050,
      "company": "ОАО «Алтайхимпром»",
      "link": "https://ru.wikipedia.org/wiki/Яровое"
    },
    {
      "name": "Заринск",
      "region": "Алтайский край",
      "population": 46250,
      "company": "ОАО «Алтай-кокс»",
      "link": "https://ru.wikipedia.org/wiki/Заринск"
    },
    {
      "name": "Алейск",
      "region": "Алтайский край",
      "population": 29200,
      "company": "АО «Алейскзернопродукт» имени С. Н. Старовойтова",
      "link": "https://ru.wikipedia.org/wiki/Алейск"
    }
  ]
}

```

Additionally, a link of each collected monotown was used to find the geographical coordinates of those towns for final visualization on the map. Using

this approach, it is possible to automatize the process of collecting a large amount of data from web-resources, which is especially relevant in situations when content and data can be updated periodically.

The result of web-scraping includes data about town affiliation to a specific entity of the Russian Federation and region, a list of town-forming enterprises, geographic coordinates (latitude and longitude) of monotowns and population. The town population will allow calculating indicators per capita, which will provide an effective comparison of large and small monotowns.

1.3. Monotowns economic indicators calculation

In order to rate monotowns economic status and diversification level, a number of the following economic indicators have to be calculated:

1. Total revenue, fixed assets and labor costs for all monotown enterprises during 2013-2017, as well as corresponding unit values (per capita).

The corresponding six calculated indicators for 2017 for the three monotowns are shown in figure 4.

Моногород	Выручка 2017	Основные средства 2017	Оплата труда 2017	Удельный показатель выручка 2017	Удельный показатель основные средства 2017	Удельный показатель оплата труда 2017
Алейск	9.949636e+09	1.327005e+09	8.974520e+08	3.407410e+05	45445.376712	30734.657534
Заринск	7.539299e+10	4.399851e+09	2.912246e+09	1.630119e+06	95131.913514	62967.481081
Яровое	1.888970e+09	5.438510e+08	2.015820e+08	1.046521e+05	30130.249307	11167.977839

Figure 4. Sample of calculated total revenue, fixed assets, labor costs, and same values per capita for 2017

2. Amount of operating, opened, and liquidated companies in a monotown during the specific year.

The corresponding three calculated indicators for 2017 for the three monotowns are shown in figure 5.

Моногород	Открытые 2017	Закрытые 2017	Действующие 2017
Алейск	27	89	241
Заринск	15	40	426
Яровое	4	7	120

Figure 5. Sample of calculated amount of opened, closed, and operating companies for 2017

3. Monoindustry share for each of the parameters (revenue, fixed assets, and labor costs) is calculated according to the formula:

$$CR_P = \sum_i^n \frac{P_i^{main}}{P},$$

where P_i^{main} –parameter value of the i-th company in monotown belonging to the group of main enterprises (the OKPO code is in the list of a reference table for a given monotown) or enterprises with a similar economic activity type (the first level of OKVED corresponds to one of the values in the reference table for a given monotown), P – total value of the parameter for all monotown enterprises.

The corresponding calculated monoindustry shares of three parameters for three monotowns for 2017 are shown in figure 6.

Моногород	Доля моноотрасли ОС 2017	Доля моноотрасли Выручка 2017	Доля моноотрасли Оплата труда 2017
Алейск	0.623040	0.686640	0.601925
Заринск	0.643623	0.866848	0.743472
Яровое	0.662209	0.268194	0.515577

Figure 6. Sample of monoindustry shares calculated for fixed assets, revenue, and labour costs for 2017

The monoindustry share is one of the main indicators of diversification level in monotowns but according to calculated values distribution (Figure 7), it is obvious that the initial data are not reliable. By its nature, monotowns should have a high level of monoindustry share, but a significant number of calculated values has a very low result for all three parameters (less than 10%).



Figure 7. Distribution histograms of calculated monoindustry shares in monotowns by three parameters for 2013 and 2017

4. The Herfindahl-Hirschman index (HHI) for one of parameters is calculated by the following formula:

$$HHI_P = \sum_i^n \left(\frac{P_i}{P} \right)^2,$$

where P_i –parameter value of the i -th company in monotown, P – total value of the parameter for all monotown enterprises.

The corresponding values of HHI for 2017 for three monotowns are shown in figure 8.

Моногород	HHI ОС 2017	HHI Выручка 2017	HHI Оплата труда 2017
Алейск	0.379528	0.424643	0.384302
Заринск	0.426396	0.755491	0.570445
Яровое	0.468525	0.148235	0.443292

Figure 8. Sample of calculated HHI values for fixed assets, revenue, and labour costs for 2017

The Herfindahl-Hirschman index values lay in range starting from 0 (in case of perfect competition, with infinite enterprises in the market where each one has a tiny share of the town's business) to 1 (in case with only one dominating company in the market). The main advantage of this indicator is the ability to react sensitively to redistribution of shares among enterprises in the market.

According to the monotowns distribution by calculated HHI values (Figure 9), it can be noticed that this indicator is more smoothly distributed than the previous

one, but there is an obvious shift of the part of monotowns to 0, which is also related to the problem described before.

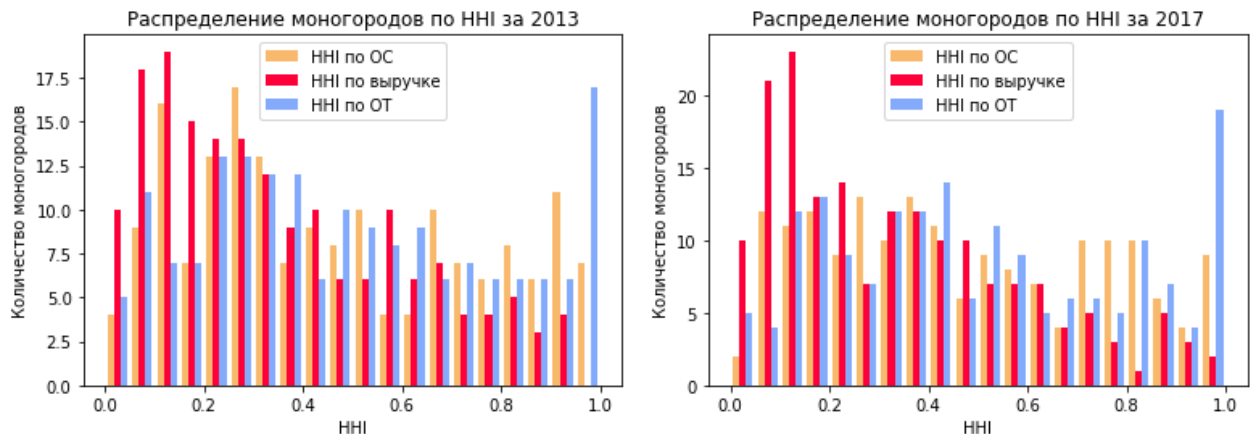


Figure 9. Distribution histograms of calculated Herfindahl-Hirschman index in monotowns by three parameters for 2013 and 2017

5. The modified Herfindahl-Hirschman index (HHI[∞]) for one of parameters is calculated by the following formula:

$$HHI^{\infty}_P = \frac{\sum_i^n (P_i^{secondary})^2}{(\sum_i^n P_i^{secondary})^2},$$

where $P_i^{secondary}$ – parameter value of the i-th company in monotown not belonging to the group of main town enterprises.

The corresponding values of HHI[∞] for 2017 for three monotowns are shown in figure 10.

Моногород	HHI [∞] ОС 2017	HHI [∞] Выручка 2017	HHI [∞] Оплата труда 2017
Алейск	0.050914	0.030125	0.091545
Заринск	0.012145	0.004065	0.017695
Яровое	0.030004	0.076307	0.177472

Figure 10. Sample of calculated HHI[∞] values for fixed assets, revenue, and labour costs for 2017

The modified Herfindahl-Hirschman index is similar to the regular index, but this index may help to avoid errors related to missing town-forming industry data. HHI[∞] has the inverse effect – the lower this indicator, the higher the monopolization of individual enterprises in the town. According to the monotowns distribution by calculated HHI[∞] values (Figure 11), it has better results than regular HHI – there is

a high portion of monotowns with low HHI' value, which indicates low diversification level.

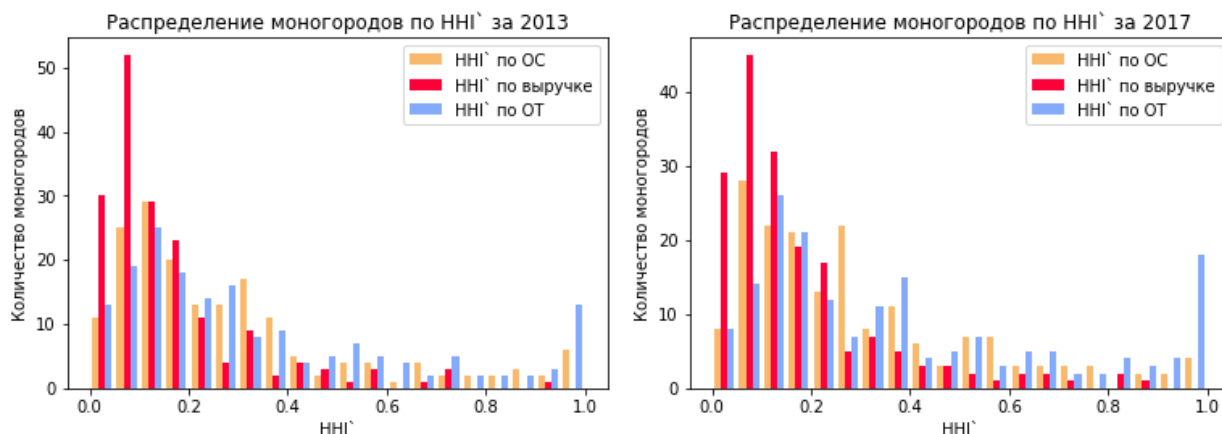


Figure 11. Distribution histograms of calculated modified Herfindahl-Hirschman index in monotowns by three parameters for 2013 and 2017

6. Entropy (E) for one of parameters is calculated by the following formula:

$$E_P = \sum_i^n \left(\ln \left(\frac{P}{P_i} \right) * \frac{P_i}{P} \right),$$

where P_i – parameter value of the i -th company in monotown, P – total value of the parameter for all monotown enterprises.

The corresponding values of E for 2017 for three monotowns are shown in figure 12.

Моногород	E OC 2017	E Выручка 2017	E Оплата труда 2017
Алейск	1.567217	1.644075	1.325883
Заринск	1.709132	0.751810	1.004074
Яровое	1.238305	2.444869	0.890618

Figure 12. Sample of calculated entropy values for fixed assets, revenue, and labour costs for 2017

Entropy is a measure of inverse concentration: the higher its value, the lower the company concentration in the market. Entropy measures the disorderly distribution shares among companies in the market: the higher the entropy, the lower the sellers' ability to influence the market price. According to the entropy

distribution (figure 13), there is an explicit difference between entropy for revenue and entropy for fixed assets and labor costs. Monotowns with the first indicator are mostly concentrated at the center, which indicates that most monotowns have a medium diversification level by this parameter. At the same time, entropy for fixed assets and labor costs shows a high monotown portion with a low diversification level, which is more typical for monotowns.

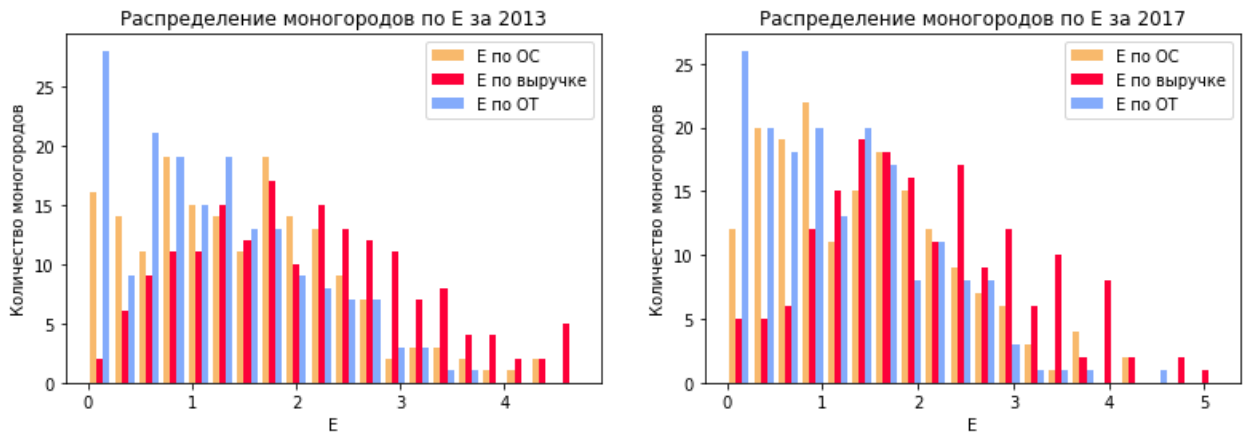


Figure 13. Distribution histograms of calculated entropy in monotowns by three parameters for 2013 and 2017

1.4. Conclusion

This chapter describes initial data about monotown companies, which are used for analysis. There is presented the list of calculated monotown indicators, which allows to rate diversification level. This list includes:

- monoindustry share;
- Herfindahl-Hirschman index (HHI);
- modified Herfindahl-Hirschman index (HHI^{*});
- entropy (E).

Calculated indicators show that part of the monotowns has insufficient initial data to properly rate their diversification level, which is expressed in the untypical indicators values.

Additionally, web-scraping approach has reviewed. This approach provides access to extract extra monotowns' data for visualization.

Приложение Б

Листинг кода для расчёта экономических показателей.

```
import pandas as pd
import numpy as np

data = pd.read_csv('data.csv', low_memory=False, converters={'ОКАТО': str,
'ОКПО': str, 'ОКВЭД':str})
data['ОКПО'] = data['ОКПО'].fillna('').apply(lambda x: str(x).replace(' ',
''))
data['Год регистрации'] = data['Год регистрации'].astype(int)
data['Год ликвидации'] = data['Год ликвидации'].fillna(0)
data['Год ликвидации'] = data['Год ликвидации'].astype(int)

years = [2013, 2014, 2015, 2016, 2017]
temp = pd.DataFrame()
temp['Город'] = data.groupby('ОКАТО').first()['Наименование ОКАТО']
temp['Предприятия'] = data.groupby('ОКАТО').count()['Наименование']

ОКАТО = pd.read_csv('ОКАТО.csv', converters={'ОКАТО*': str, 'ОКАТО': str,
'ОКПО': str, 'ОКВЭД':str})
ОКАТО['ОКПО'] = ОКАТО['ОКПО'].apply(lambda x: [str(i) for i in
str(x).replace(' ', '').split(',') if i != ''])
ОКАТО['ОКВЭД'] = ОКАТО['ОКВЭД'].apply(lambda x: [str(i) for i in
str(x).replace(' ', '').split(',') if i != ''])
ОКАТО = ОКАТО.set_index(ОКАТО['ОКАТО*'])[['Моногород', 'Широта', 'Долгота',
'ОКАТО', 'ОКПО', 'ОКВЭД', 'Численность населения', 'Градообразующие
предприятия', 'Ссылка', 'Регион', 'Федеральный округ', 'hidden']]

temp = temp.join(ОКАТО).drop('Город', axis=1)
temp['ОКАТО'] = temp['ОКАТО'].astype(str)

# Выручка по городу
columns = ['Выручка ' + str(year) for year in years]
Выручка = data.groupby('ОКАТО')[columns].sum()
Выручка.columns = columns
temp = temp.join(Выручка)

# Удельный показатель выручка по городу
for year in years:
    temp['Удельный показатель выручка ' + str(year)] = temp['Выручка ' +
str(year)]/temp['Численность населения']

# ОС по городу
columns = ['Основные средства ' + str(year) for year in years]
ОсновнСредства = data.groupby('ОКАТО')[columns].sum()
ОсновнСредства.columns = columns
temp = temp.join(ОсновнСредства)

# Удельный показатель основные средства по городу
for year in years:
    temp['Удельный показатель основные средства ' + str(year)] =
temp['Основные средства ' + str(year)]/temp['Численность населения']

# Оплата труда по городу
columns = ['Оплата труда ' + str(year) for year in years]
ОплатаТруда = data.groupby('ОКАТО')[columns].sum()
ОплатаТруда.columns = columns
temp = temp.join(ОплатаТруда)
```

```

# Удельный показатель оплата труда по городу
for year in years:
    temp['Удельный показатель оплата труда ' + str(year)] = temp['Оплата
труда ' + str(year)]/temp['Численность населения']

# Созданные, Ликвидированные, Действующие
for year in years:
    temp['Открытые ' + str(year)] = data.groupby('ОКАТО').apply(lambda df:
len(df[df['Год регистрации'] == year]))
    temp['Закрытые ' + str(year)] = data.groupby('ОКАТО').apply(lambda df:
len(df[df['Год ликвидации'] == year]))
    temp['Действующие ' + str(year)] = data.groupby('ОКАТО').apply(lambda df:
len(df[(df['Год регистрации'] <= year) & (df['Год ликвидации'] == 0)]) + \
len(df[(df['Год регистрации'] <= year) & (df['Год ликвидации'] > year))))

getTownMainOKPO = lambda df:
OKATO[OKATO.index.str.contains(str(df.iloc[0]))].iloc[0, 4]
getTownMainOKED = lambda df:
OKATO[OKATO.index.str.contains(str(df.iloc[0]))].iloc[0, 5]

# Доля моноотрасли по ОС
for year in years:
    column = 'Основные средства ' + str(year)
    result_column = 'Доля моноотрасли ОС ' + str(year)
    temp[result_column] = data.groupby('ОКАТО').apply(lambda df:
df[df['ОКПО'].isin(getTownMainOKPO(df['ОКАТО'])) |
df['ОКВЭД*'].isin(getTownMainOKED(df['ОКАТО']))][column].sum())/temp[column]
    temp[result_column] = temp[result_column].fillna(0)

# Доля моноотрасли по Выручке
for year in years:
    column = 'Выручка ' + str(year)
    result_column = 'Доля моноотрасли Выручка ' + str(year)
    temp[result_column] = data.groupby('ОКАТО').apply(lambda df:
df[df['ОКПО'].isin(getTownMainOKPO(df['ОКАТО'])) |
df['ОКВЭД*'].isin(getTownMainOKED(df['ОКАТО']))][column].sum())/temp[column]
    temp[result_column] = temp[result_column].fillna(0)

# Доля моноотрасли по Оплате труда
for year in years:
    column = 'Оплата труда ' + str(year)
    result_column = 'Доля моноотрасли Оплата труда ' + str(year)
    temp[result_column] = data.groupby('ОКАТО').apply(lambda df:
df[df['ОКПО'].isin(getTownMainOKPO(df['ОКАТО'])) |
df['ОКВЭД*'].isin(getTownMainOKED(df['ОКАТО']))][column].sum())/temp[column]
    temp[result_column] = temp[result_column].fillna(0)

# НИИ
for year in years:
    temp['НИИ ОС ' + str(year)] = (data.groupby('ОКАТО').apply(lambda df:
(df['Основные средства ' + str(year)]**2).sum()/(temp['Основные средства ' +
str(year)]**2)).fillna(0)
    temp['НИИ Выручка ' + str(year)] = (data.groupby('ОКАТО').apply(lambda
df: (df['Выручка ' + str(year)]**2).sum()/(temp['Выручка ' +
str(year)]**2)).fillna(0)
    temp['НИИ Оплата труда ' + str(year)] =
(data.groupby('ОКАТО').apply(lambda df: (df['Оплата труда ' +
str(year)]**2).sum()/(temp['Оплата труда ' + str(year)]**2)).fillna(0)

def countHHImod(df, column):
    notMainCompanies = df[~(df['ОКПО'].isin(getTownMainOKPO(df['ОКАТО']))) |

```

```

df['ОКВЭД*'].isin(getTownMainOKED(df['ОКАТО']))]
    return
(notMainCompanies[column]**2).sum()/(notMainCompanies[column].sum()**2)

# НИИ модифицированный
for year in years:
    temp['`НИИ ОС ' + str(year)] = (data.groupby('ОКАТО').apply(countHHImod,
'Основные средства ' + str(year))).fillna(0)
    temp['`НИИ Выручка ' + str(year)] =
(data.groupby('ОКАТО').apply(countHHImod, 'Выручка ' + str(year))).fillna(0)
    temp['`НИИ Оплата труда ' + str(year)] =
(data.groupby('ОКАТО').apply(countHHImod, 'Оплата труда ' +
str(year))).fillna(0)

# E
def CountE(df, column):
    _sum = df[column].sum()
    return (np.log(_sum/df[column])*df[column]/_sum).sum()

for year in years:
    temp['E ОС ' + str(year)] = data.groupby('ОКАТО').apply(CountE, 'Основные
средства ' + str(year)).fillna(0)
    temp['E Выручка ' + str(year)] = data.groupby('ОКАТО').apply(CountE,
'Выручка ' + str(year)).fillna(0)
    temp['E Оплата труда ' + str(year)] = data.groupby('ОКАТО').apply(CountE,
'Оплата труда ' + str(year)).fillna(0)

del temp['ОКПО']
del temp['ОКВЭД']
temp.to_csv("Results.csv", index=False)

```

Приложение В

Листинг кода для визуализации моногородов по параметрам фильтрации и построения регрессионных моделей.

```
import folium
import math
from folium.plugins import MarkerCluster, Fullscreen
import pandas as pd
from django.shortcuts import render
from django.http import JsonResponse
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn.metrics import r2_score, mean_absolute_error, mean_squared_error
from sklearn.ensemble import RandomForestRegressor
from sklearn.linear_model import LinearRegression
import numpy as np

# Исправление проблем с визуализацией, связанных с искривлением поверхности
# Земли и соответствующих размеров
def reduce_latitude_scaling(x):
    x = abs(x)
    return 164**2/(164 + 6.35591933e-01*x - 1.31263708e-01*(x**2) +\
        1.26938922e-02*(x**3) - 4.7772578e-04*(x**4) +\
        9.44459319e-06*(x**5) - 9.44479840e-08*(x**6) +\
        3.96846223e-10*(x**7))

# Округления дробных чисел с разной точностью для разных целей
def round_number(number):
    if number < 1:
        i = 4
        if round(number, i) == 0 and number != 0:
            while round(number, i) == 0:
                i += 1
            return round(number, i)
    elif number > 1000:
        return round(number)
    else:
        return round(number, 2)

# Разделение строкового представления числа пробелом через три символа
def split_big_numbers(number):
    if math.isnan(float(number)):
        return '---'
    number = float(number)
    if number < 1000:
        return number
    result = ''
    number = str(round(number))
    c = len(number)%3
    for n in number:
        if c == 0:
            c = 3
            result += ' '
        c -= 1
        result += n
    return result.strip()

# Задание стартовой позиции карты в зависимости от количества городов
def get_start_position(towns):
    if len(towns) == 0:
```

```

        return (56.476947, 84.965567, False)
    return (towns["Широта"].iloc[0], towns["Долгота"].iloc[0], True)

# Удаление ненужного HTML текста, выводимого перед загрузкой картой
def replace_text(html):
    return html.replace('<span style="color:#565656">Make this Notebook
Trusted to load map: File -> Trust Notebook</span>', '')

# Визуализация карты с маркерами
def map(request):
    def textForPopup(i, year):
        return ('<center><strong>' + towns['Моногород'].iloc[i] +
'</strong></center>' +
'<br>ОКАТО - {}'.format(towns['ОКАТО'].iloc[i]) +
'<br>Численность населения - {}
ч'.format(split_big_numbers(towns['Численность населения'].iloc[i])) +
'<br>Выручка за {} год - {} Р'.format(year,
split_big_numbers(str(towns['Выручка ' + str(year)].iloc[i]))) +
'<br>Основные средства на {} год - {} Р'.format(year,
split_big_numbers(str(towns['Основные средства ' + str(year)].iloc[i]))) +
'<br>Оплата труда за {} год - {} Р'.format(year,
split_big_numbers(str(towns['Оплата труда ' + str(year)].iloc[i]))) +
'<br>Доля моноотрасли по ОС на {} год - {}'.format(year,
round_number(towns['Доля моноотрасли ОС ' + str(year)].iloc[i])) +
'<br>Доля моноотрасли по выручке на {} год - {}'.format(year,
round_number(towns['Доля моноотрасли Выручка ' + str(year)].iloc[i])) +
'<br>Доля моноотрасли по ОТ на {} год - {}'.format(year,
round_number(towns['Доля моноотрасли Оплата труда ' + str(year)].iloc[i])) +
'<br>ННИ по ОС на {} год - {} '.format(year,
round_number(towns['ННИ ОС ' + str(year)].iloc[i])) +
'<br>ННИ по выручке на {} год - {} '.format(year,
round_number(towns['ННИ Выручка ' + str(year)].iloc[i])) +
'<br>ННИ по ОТ на {} год - {} '.format(year,
round_number(towns['ННИ Оплата труда ' + str(year)].iloc[i])) +
'<br>Е по ОС на {} год - {} '.format(year,
round_number(towns['Е ОС ' + str(year)].iloc[i])) +
'<br>Е по выручке на {} год - {} '.format(year,
round_number(towns['Е Выручка ' + str(year)].iloc[i])) +
'<br>Е по ОТ на {} год - {} '.format(year,
round_number(towns['Е Оплата труда ' + str(year)].iloc[i])) +
'<br><br><center><input type="button"
onclick="window.open(\'/details/{}/{}\', \'_blank\')"
value="Подробнее" /></center>'.\
format(towns['Моногород'].iloc[i], year))

    version = request.GET.get('version', None)
    if version is not None and version == 'new':
        towns =
pd.read_csv("~/DataVisualizationLab/monotowns/api/Results_NEW.csv",
low_memory=False)
    else:
        towns =
pd.read_csv("~/DataVisualizationLab/monotowns/api/Results.csv",
low_memory=False)

    town_name = request.GET.get('town_name', None)
    OKATO = request.GET.get('OKATO', None)
    is_clustering_enabled = request.GET.get('is_clustering_enabled', False)
    year = request.GET.get('year', 2013)
    region = request.GET.get('region')

```

```

is_showing_only_hidden = request.GET.get('hidden')
parameter = request.GET.get('search_parameter')

if is_showing_only_hidden is not None:
    towns = towns[towns['hidden']]

if town_name is not None:
    towns =
towns[towns['Моногород'].str.lower().str.contains(town_name.lower())]
    if OKATO is not None:
        towns['OKATO'] = towns['OKATO'].astype('str')
        towns = towns[towns['OKATO'].str.startswith(OKATO)]
    if region is not None:
        towns = towns[(towns['Федеральный округ'].str.lower() ==
region.lower()) | (towns['Регион'].str.lower() == region.lower())]
    if parameter is not None:
        comparators = sorted(request.GET.get('comparators',
'>').split(','))[::-1]
        comparator_values = sorted([float(i) for i in
request.GET.get('comparator_values', '0').split(',')])
        for index in range(len(comparators)):
            if comparators[index] == '=':
                towns = towns[towns[parameter] == comparator_values[index]]
            elif comparators[index] == '>':
                towns = towns[towns[parameter] > comparator_values[index]]
            elif comparators[index] == '<':
                towns = towns[towns[parameter] < comparator_values[index]]

start_position = get_start_position(towns)
map = folium.Map(location=[start_position[0], start_position[1]],
zoom_start=6, max_zoom=12, min_zoom=3, width='70%', height='70%',
position='absolute')
Fullscreen().add_to(map)
if is_clustering_enabled:
    marker_cluster = MarkerCluster().add_to(map)

for i in range(len(towns)):
    popup = folium.Popup(folium.Html(textForPopup(i, year), script =
True), max_width = 300, min_width = 300)
    icon = folium.Icon(color = 'lightred', icon = 'info-sign')
    folium.Marker(location = [towns["Широта"].iloc[i],
towns["Долгота"].iloc[i]],
                    popup = popup, icon = icon,
                    tooltip=str(towns['Моногород'].iloc[i])).add_to(marker_cluster if
is_clustering_enabled else map)

    context = {'map': replace_text(map._repr_html_()), 'town_number':
str(len(towns)) }

return render(request, 'map.html', context)

# Визуализации карты с распределением моногородов по выбранному параметру
def parameter_distribution(request):
    version = request.GET.get('version', None)
    if version is not None and version == 'new':
        towns =
pd.read_csv("~/DataVisualizationLab/monotowns/api/Results_NEW.csv",
low_memory=False)
    else:
        towns =

```

```

pd.read_csv("~/DataVisualizationLab/monotowns/api/Results.csv",
low_memory=False)
regions = pd.read_csv("~/DataVisualizationLab/monotowns/api/regions.csv",
low_memory=False)

parameter = request.GET.get('parameter', 'Численность населения')
is_showing_only_hidden = request.GET.get('hidden')
is_showing_region_centers = request.GET.get('is_connection_enabled')

value_suffix = ''
simple_parameter = ' '.join(parameter.split(' ')[:-1])
if parameter in ['Численность населения']:
    value_suffix = ' ч'
elif simple_parameter in ['Выручка', 'Основные средства', 'Оплата
труда']:
    value_suffix = ' Р'
elif simple_parameter in ['Удельный показатель выручка',
'Удельный показатель основные средства',
'Удельный показатель оплата труда']:
    value_suffix = ' Р/ч'

if is_showing_only_hidden is not None:
    towns = towns[towns['hidden']]

empty_values = towns[towns[parameter] == 0]
not_empty_values = towns[towns[parameter] != 0].sort_values(parameter,
ascending=False)
min_value = not_empty_values.min()[parameter]
values_range = not_empty_values.max()[parameter] - min_value

start_position = get_start_position(not_empty_values)
if not start_position[-1]:
    start_position = get_start_position(empty_values)
map = folium.Map(location=[start_position[0], start_position[1]],
zoom_start=6, max_zoom=12, min_zoom=3, width='70%',
height='70%')
Fullscreen().add_to(map)

if is_showing_region_centers is None:
    for i in range(len(not_empty_values)):
        town = not_empty_values.iloc[i]
        folium.Circle(
            location=[town['Широта'], town['Долгота']],
            radius=((900*(town[parameter] - min_value)/values_range) +
100)*\
                reduce_latitude_scaling(town['Широта']),
            tooltip='{}: {} - {}'.format(town['Моногород'],
parameter,
str(split_big_numbers(round_number(town[parameter])) + value_suffix),
color="#e1898f",
fill=True,
fill_color="#e1898f"
).add_to(map)
    for i in range(len(empty_values)):
        town = empty_values.iloc[i]
        folium.Circle(
            location=[town['Широта'], town['Долгота']],
            radius=50*reduce_latitude_scaling(town['Широта']),
            tooltip='{}: {} - {}'.format(town['Моногород'],
parameter,

```

```

        '0' + value_suffix),
        color="#000000",
        fill=True,
        fill_color="#000000"
    ).add_to(map)

else:
    for _, region in regions.iterrows():
        c_lat, c_lon = region['center_lat'], region['center_lon']
        region_towns = towns[towns['Регион'] == region['region']]
        region_has_any_town = len(region_towns) > 0
        if region_has_any_town:
            folium.Circle(
                location=[c_lat, c_lon],
                radius=(250*reduce_latitude_scaling(c_lat)),
                tooltip=region['center'],
                color="#c99e6b",
                fill=True,
                fill_color="#c99e6b"
            ).add_to(map)

        for _, region in regions.iterrows():
            c_lat, c_lon = region['center_lat'], region['center_lon']
            region_towns = not_empty_values[not_empty_values['Регион'] ==
region['region']]
            for i in range(len(region_towns)):
                town = region_towns.iloc[i]
                folium.PolyLine([[town['Широта'], town['Долгота']], [c_lat,
c_lon]], color='#9c95c8aa').add_to(map)
                folium.Circle(
                    location=[town['Широта'], town['Долгота']],
                    radius=((900*(town[parameter] - min_value)/values_range
+ 100)*\
                        reduce_latitude_scaling(town['Широта']),
                    tooltip='{}: {} - {}'.format(town['Моногород'],
parameter,
str(split_big_numbers(round_number(town[parameter]))) + value_suffix),
                    color="#e1898f",
                    fill=True,
                    fill_color="#e1898f"
                ).add_to(map)

            for _, region in regions.iterrows():
                c_lat, c_lon = region['center_lat'], region['center_lon']
                region_towns = empty_values[empty_values['Регион'] ==
region['region']]
                for i in range(len(region_towns)):
                    town = region_towns.iloc[i]
                    folium.PolyLine([[town['Широта'], town['Долгота']], [c_lat,
c_lon]], color='#9c95c8aa').add_to(map)
                    folium.Circle(
                        location=[town['Широта'], town['Долгота']],
                        radius=50*reduce_latitude_scaling(town['Широта']),
                        tooltip='{}: {} - {}'.format(town['Моногород'],
parameter,
'0' + value_suffix),
                        color="#000000",
                        fill=True,
                        fill_color="#000000"
                    ).add_to(map)

```



```

context = {'map': replace_text(map._repr_html_())}

return render(request, 'empty_map.html', context)

# Получение списка доступных моногородов
def get_towns_list(request):
    version = request.GET.get('version', None)
    if version is not None and version == 'new':
        towns =
pd.read_csv("~/DataVisualizationLab/monotowns/api/Results_NEW.csv",
low_memory=False)
    else:
        towns =
pd.read_csv("~/DataVisualizationLab/monotowns/api/Results.csv",
low_memory=False)
        towns = towns['Моногород'].sort_values().to_list()
    return JsonResponse({'towns': towns})

# Получение детальной информации по моногороду
def get_details(request):
    town = request.GET.get('town')
    year = int(request.GET.get('year'))
    if town is None or year is None:
        return JsonResponse({'error': 'Параметра не указаны'}, status=400)

    version = request.GET.get('version', None)
    if version is not None and version == 'new':
        towns =
pd.read_csv("~/DataVisualizationLab/monotowns/api/Results_NEW.csv",
low_memory=False)
    else:
        towns =
pd.read_csv("~/DataVisualizationLab/monotowns/api/Results.csv",
low_memory=False)

    town = towns[towns['Моногород'].str.lower() == town.lower()].iloc[0]
    object = {
        'Предприятия': str(town['Предприятия']),
        'Численность населения': str(split_big_numbers(town['Численность
населения'])).replace('.0', ''),
        'Градообразующие предприятия': town['Градообразующие предприятия'],
        'Выручка': '{} Р'.format(split_big_numbers(town['Выручка ' +
str(year)])),
        'Удельная выручка': '{}
Р/ч'.format(split_big_numbers(round_number(town['Удельный показатель выручка
' + str(year)]))),
        'Основные средства': '{} Р'.format(split_big_numbers(town['Основные
средства ' + str(year)])),
        'Удельные основные средства': '{}
Р/ч'.format(split_big_numbers(round_number(town['Удельный показатель основные
средства ' + str(year)]))),
        'Оплата труда': '{} Р'.format(split_big_numbers(town['Оплата труда '
+ str(year)])),
        'Удельная оплата труда': '{}
Р/ч'.format(split_big_numbers(round_number(town['Удельный показатель оплата
труда ' + str(year)]))),
        'Доля моноотрасли ОС': str(round_number(town['Доля моноотрасли ОС ' +
str(year)])),
        'Доля моноотрасли Выручка': str(round_number(town['Доля моноотрасли
Выручка ' + str(year)])),

```

```

        'Доля моноотрасли Оплата труда': str(round_number(town['Доля
моноотрасли Оплата труда ' + str(year)])),
        'ННИ ОС': str(round_number(town['ННИ ОС ' + str(year)])),
        'ННИ Выручка': str(round_number(town['ННИ Выручка ' + str(year)])),
        'ННИ Оплата труда': str(round_number(town['ННИ Оплата труда ' +
str(year)])),

    'Е ОС': str(round_number(town['Е ОС ' + str(year)])),
    'Е Выручка': str(round_number(town['Е Выручка ' + str(year)])),
    'Е Оплата труда': str(round_number(town['Е Оплата труда ' +
str(year)]))
    }
    return JsonResponse({'details': object})

# Построение корреляционной матрицы
def get_corrilation_matrix(request):
    data =
pd.read_csv("~/DataVisualizationLab/monotowns/api/Results_NEW.csv",
low_memory=False)
    parameters_list = ['Выручка *', 'Удельный показатель выручка *',
'Основные средства *', 'Удельный показатель основные
средства *',
'Оплата труда *', 'Удельный показатель оплата труда
*',
'Широта', 'Долгота', 'Численность населения',
'Открытые *', 'Закрытые *', 'Действующие *',
'ННИ Выручка *', 'ННИ ОС *', 'ННИ Оплата труда *',
'`ННИ Выручка *', '`ННИ ОС *', '`ННИ Оплата труда *',
'Е Выручка *', 'Е ОС *', 'Е Оплата труда *',
'Доля моноотрасли Выручка *', 'Доля моноотрасли ОС *',
'Доля моноотрасли Оплата труда *']
    names = [u'Выручка', u'Удельная выручка',
u'Основные средства (ОС)', u'Удельные ОС',
u'Оплата труда (ОТ)', u'Удельная ОТ',
u'Широта', u'Долгота', u'Численность населения',
u'Открытые орг-ции', u'Закрытые орг-ции', u'Действующие орг-
ции',
u'ННИ по выручке', u'ННИ по ОС', u'ННИ по ОТ',
u'ННИ мод по выручке', u'ННИ мод по ОС', u'ННИ мод по ОТ',
u'Е по выручке', u'Е по ОС', u'Е по ОТ',
u'Доля моноотрасли по выручке', u'Доля моноотрасли по ОС',
u'Доля моноотрасли по ОТ']
    years = [2013, 2014, 2015, 2016, 2017]

    data_frame_correlation = dict()
    for parameter in parameters_list:
        data_frame_correlation[parameter] = []
    data_frame_correlation = pd.DataFrame(data_frame_correlation)
    for year in years:
        data_frame_for_concatination = dict()
        correct_data = data[((data['Доля моноотрасли Выручка ' + str(year)] >
0.1) & (data['Доля моноотрасли ОС ' + str(year)] > 0.1) & (data['Доля
моноотрасли Оплата труда ' + str(year)] > 0.1))]
        for parameter in parameters_list:
            data_frame_for_concatination[parameter] =
correct_data[parameter.replace('*', str(year))].to_list()
        data frame correlation = pd.concat([data frame correlation,
pd.DataFrame(data_frame_for_concatination)])

    correlation = np.array(data_frame_correlation.corr().values.tolist())

```

```

fig, ax = plt.subplots(figsize=(20, 20))
ax.imshow(correlation, cmap="YlGn")
ax.set_xticks(np.arange(len(parameters_list)))
ax.set_yticks(np.arange(len(parameters_list)))
ax.set_xticklabels(names)
ax.set_yticklabels(names)
plt.setp(ax.get_yticklabels(), rotation=20, ha="right",
          rotation_mode="anchor")

plt.setp(ax.get_xticklabels(), rotation=20, ha="right",
          rotation_mode="anchor")
for i in range(len(parameters_list)):
    for j in range(len(parameters_list)):
        color = '#000000'
        value = format(correlation[i, j], '.2f')
        if float(value) > 0.4:
            color = '#ffffff'
        ax.text(j, i, value, ha="center", va="center", color=color)
ax.set_title(u"Матрица корреляции")

fig.savefig('/home/dvk48/DataVisualizationLab/static/correlation_new.png',
            bbox_inches='tight')
return None

# Построение регрессионной модели и построение предсказаний
def get_predictions(request):
    year_result = request.GET.get('year', 2013)
    prediction_model = request.GET.get('prediction_model')
    region = request.GET.get('region')
    target_parameter = request.GET.get('parameter', 'OC')
    is_showing_only_hidden = request.GET.get('hidden')
    is_using_empty_for_training =
request.GET.get('is_using_empty_for_training')
    indep_vars_list = request.GET.get('indep_vars', 'Действующие *, `ННИ $
*').split(',')
    indep_vars_list = [var.replace('$', target_parameter) for var in
indep_vars_list]

    version = request.GET.get('version', None)
    if version is not None and version == 'new':
        data =
pd.read_csv("~/DataVisualizationLab/monotowns/api/Results_NEW.csv",
low_memory=False)
    else:
        data =
pd.read_csv("~/DataVisualizationLab/monotowns/api/Results.csv",
low_memory=False)
        data = data.sort_values(['Федеральный округ', 'Регион', 'Моногород'])
        parameters_list = ['Выручка *', 'Удельный показатель выручка *',
'Основные средства *', 'Удельный показатель основные
средства *',
'Оплата труда *', 'Удельный показатель оплата труда
*',
'Широта', 'Долгота', 'Численность населения',
'Открытые *', 'Закрытые *', 'Действующие *',
'ННИ Выручка *', 'ННИ ОС *', 'ННИ Оплата труда *',
'`ННИ Выручка *', '`ННИ ОС *', '`ННИ Оплата труда *',
'Е Выручка *', 'Е ОС *', 'Е Оплата труда *',
'Доля моноотрасли Выручка *', 'Доля моноотрасли ОС
*', 'Доля моноотрасли Оплата труда *']

```

```

years = [2013, 2014, 2015, 2016, 2017]
data_frame = dict()
data_frame['Моногород'] = []
data_frame['Федеральный округ'] = []
data_frame['Регион'] = []
for parameter in parameters_list:
    data_frame[parameter] = []
data_frame = pd.DataFrame(data_frame)

for year in years:
    data_frame_for_concatination = dict()
    data_frame_for_concatination['Моногород'] =
data['Моногород'].apply(lambda x: x + '*' + str(year)).to_list()
    data_frame_for_concatination['Федеральный округ'] = data['Федеральный
округ']
    data_frame_for_concatination['Регион'] = data['Регион']
    data_frame_for_concatination['hidden'] = data['hidden']
    for parameter in parameters_list:
        data_frame_for_concatination[parameter] =
data[parameter.replace('*', str(year))].to_list()
    data_frame = pd.concat([data_frame,
pd.DataFrame(data_frame_for_concatination)])

indep_vars = indep_vars_list
dep_var = 'Доля моноотрасли ' + target_parameter + '*'

if is_using_empty_for_training:
    data_frame_regression = data_frame
else:
    data_frame_regression = data_frame[data_frame[dep_var] != 0]

X = data_frame_regression.loc[:, indep_vars]
Y = data_frame_regression[dep_var]

X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size=0.3,
random_state=51)
if prediction_model == 'random_forest':
    model = RandomForestRegressor(random_state=51, max_depth=50,
n_estimators=20)
else:
    model = LinearRegression()
model.fit(X_train, Y_train)
Y_pred = model.predict(X_test)
r2 = r2_score(Y_test, Y_pred)
mae_score = mean_absolute_error(Y_test, Y_pred)
mse_score = mean_squared_error(Y_test, Y_pred)

data_result =
data_frame[data_frame['Моногород'].str.contains(str(year_result))]

if is_showing_only_hidden is not None:
    data_result = data_result[data_result['hidden']]
if region is not None:
    data_result = data_result[(data_result['Федеральный
округ'].str.lower() == region.lower()) |
    (data_result['Регион'].str.lower() == region.lower())]
predicted_result = model.predict(data_result.loc[:, indep_vars])

results = {
    'r2_score': r2,
    'mae_score': mae_score,

```

```

        'mse_score': mse_score,
        'data': []
    }
    i = 0
    for index, row in data_result.iterrows():
        results['data'].append({
            'label': row['Моногород'].split('*')[0],
            'x': i,

            'y': round_number(predicted_result[i]),
            'color': '#f5b169',
            'text': 'предсказанное значение'
        })
        results['data'].append({
            'label': row['Моногород'].split('*')[0],
            'x': i,
            'y': round_number(row['Доля моноотрасли ' + target_parameter + '
*']),
            'color': '#ff0000' if row['Доля моноотрасли ' + target_parameter + '
*'] == 0 else '#6395ff',
            'text': 'расчитанное значение'
        })
        i += 1

    return JsonResponse({'data': results})

```

Приложение Г

Матрица корреляции

Выручка	1.00	0.56	0.88	0.27	0.92	0.31	0.86	0.84	0.84	0.89	-0.26	-0.30	-0.37	-0.19	-0.16	-0.30	0.43	0.43	0.52	-0.12	-0.17	-0.27
Удельная выручка	0.56	1.00	0.57	0.58	0.51	0.70	0.21	0.19	0.22	0.26	0.10	-0.01	-0.17	0.02	-0.10	-0.18	-0.05	0.04	0.23	0.19	0.09	-0.09
Основные средства (ОС)	0.88	0.57	1.00	0.57	0.91	0.43	0.76	0.64	0.70	0.72	-0.25	-0.19	-0.39	-0.17	-0.14	-0.32	0.38	0.28	0.53	-0.07	-0.02	-0.24
Удельные ОС	0.27	0.58	0.57	1.00	0.32	0.53	0.09	0.05	0.08	0.09	0.12	0.25	-0.07	0.03	-0.08	-0.06	-0.10	-0.23	0.09	0.19	0.31	0.01
Оплата труда (ОТ)	0.92	0.51	0.91	0.32	1.00	0.51	0.85	0.75	0.81	0.83	-0.27	-0.24	-0.35	-0.22	-0.20	-0.37	0.44	0.38	0.51	-0.11	-0.09	-0.22
Удельная ОТ	0.31	0.70	0.43	0.53	0.51	1.00	0.10	0.05	0.11	0.12	0.15	0.17	0.04	-0.06	-0.12	-0.26	-0.11	-0.14	0.03	0.20	0.24	0.10
Численность населения	0.86	0.21	0.76	0.09	0.85	0.10	1.00	0.88	0.91	0.92	-0.38	-0.35	-0.47	-0.29	-0.27	-0.37	0.60	0.52	0.64	-0.23	-0.19	-0.36
Открытые орг-ции	0.84	0.19	0.64	0.05	0.75	0.05	0.88	1.00	0.90	0.96	-0.27	-0.27	-0.27	-0.17	-0.17	-0.20	0.48	0.46	0.42	-0.22	-0.23	-0.25
Закрытые орг-ции	0.84	0.22	0.70	0.08	0.81	0.11	0.91	0.90	1.00	0.96	-0.32	-0.29	-0.35	-0.21	-0.22	-0.26	0.54	0.47	0.51	-0.24	-0.19	-0.29
Действующие орг-ции	0.89	0.26	0.72	0.09	0.83	0.12	0.92	0.96	0.96	1.00	-0.32	-0.31	-0.35	-0.21	-0.22	-0.26	0.54	0.49	0.50	-0.24	-0.22	-0.29
НИИ по выручке	-0.26	0.10	-0.25	0.12	-0.27	0.15	-0.38	-0.27	-0.32	-0.32	1.00	0.72	0.76	0.12	0.28	0.20	-0.92	-0.72	-0.74	0.85	0.52	0.64
НИИ по ОС	-0.30	-0.01	-0.19	0.25	-0.24	0.17	-0.35	-0.27	-0.29	-0.31	0.72	1.00	0.70	0.14	0.15	0.25	-0.70	-0.95	-0.67	0.55	0.80	0.51
НИИ по ОТ	-0.37	-0.17	-0.39	-0.07	-0.35	0.04	-0.47	-0.27	-0.35	-0.35	0.76	0.70	1.00	0.22	0.37	0.31	-0.77	-0.71	-0.95	0.51	0.38	0.84
НИИ мод по выручке	-0.19	0.02	-0.17	0.03	-0.22	-0.06	-0.29	-0.17	-0.21	-0.21	0.12	0.14	0.22	1.00	0.49	0.40	-0.30	-0.22	-0.28	-0.01	0.03	0.15
НИИ мод по ОС	-0.16	-0.10	-0.14	-0.08	-0.20	-0.12	-0.27	-0.17	-0.22	-0.22	0.28	0.15	0.37	0.49	1.00	0.29	-0.39	-0.31	-0.40	0.26	-0.02	0.38
НИИ мод по ОТ	-0.30	-0.18	-0.32	-0.06	-0.37	-0.26	-0.37	-0.20	-0.26	-0.26	0.20	0.25	0.31	0.40	0.29	1.00	-0.34	-0.33	-0.45	0.12	0.09	0.17
Е по выручке	0.43	-0.05	0.38	-0.10	0.44	-0.11	0.60	0.48	0.54	0.54	-0.92	-0.70	-0.77	-0.30	-0.39	-0.34	1.00	0.78	0.82	-0.79	-0.49	-0.67
Е по ОС	0.43	0.04	0.28	-0.23	0.38	-0.14	0.52	0.46	0.47	0.49	-0.72	-0.95	-0.71	-0.22	-0.31	-0.33	0.78	1.00	0.73	-0.57	-0.76	-0.55
Е по ОТ	0.52	0.23	0.53	0.09	0.51	0.03	0.64	0.42	0.51	0.50	-0.74	-0.67	-0.95	-0.28	-0.40	-0.45	0.82	0.73	1.00	-0.50	-0.35	-0.77
Доля моноотрасли по выручке	-0.12	0.19	-0.07	0.19	-0.11	0.20	-0.23	-0.22	-0.24	-0.24	0.85	0.55	0.51	-0.01	0.26	0.12	-0.79	-0.57	-0.50	1.00	0.59	0.64
Доля моноотрасли по ОС	-0.17	0.09	-0.02	0.31	-0.09	0.24	-0.19	-0.23	-0.19	-0.22	0.52	0.80	0.38	0.03	-0.02	0.09	-0.49	-0.76	-0.35	0.59	1.00	0.44
Доля моноотрасли по ОТ	-0.27	-0.09	-0.24	0.01	-0.22	0.10	-0.36	-0.25	-0.29	-0.29	0.64	0.51	0.84	0.15	0.38	0.17	-0.67	-0.55	-0.77	0.64	0.44	1.00