

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего профессионального образования
**«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»**

Инженерная школа ядерных технологий

Направление подготовки 01.04.02 «Прикладная математика и информатика»

Отделение экспериментальной физики

МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ

Тема работы
Математические методы ординации многомерных объектов и признаков УДК 514.182(084.21)

Студент

Группа	ФИО	Подпись	Дата
ОВМ81	Абдуллоев Рустам Махмадсобирович		

Руководитель

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОЭФ ИЯТШ	Семенов М. Е.	Кандидат ф.-м. наук, доцент		

КОНСУЛЬТАНТЫ:

По разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОСГН ШБИП	Киселева Елена Станиславовна	к.э.н		

По разделу «Социальная ответственность»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Старший преподаватель отделения ОД ШБИП	Романова Светлана Владимировна			

ДОПУСТИТЬ К ЗАЩИТЕ:

Руководитель ООП	ФИО	Ученая степень, звание	Подпись	Дата
Профессор ОМИ ШБИП	Трифонов А.Ю.	Доктор ф.-м. наук, профессор		

Томск – 2020 г.

Планируемые результаты обучения.

Код результата	Результат обучения (выпускник должен быть готов)
<i>Профессиональные компетенции</i>	
ПК-1	К самостоятельной работе
ПК-2	Использовать современные прикладные программные средства и осваивать современные технологии программирования
ПК-3	Использовать стандартные пакеты прикладных программ для решения практических задач на ЭВМ, отлаживать, тестировать прикладное программное обеспечение
ПК-4	Настраивать, тестировать и осуществлять проверку вычислительной техники и программных средств
ПК-5	Демонстрировать знание современных языков программирования, операционных систем, офисных приложений, Интернета, способов и механизмов управления данными; принципов организации, состава и схемы работы операционных систем
ПК-6	Решать проблемы, брать на себя ответственность
ПК-7	Проводить организационно-управленческие расчеты, осуществлять организацию и техническое оснащение рабочих мест
ПК-8	Организовывать работу малых групп исполнителей
ПК-9	Определять экономическую целесообразность принимаемых технических и организационных решений
ПК-10	Владеть основными методами защиты производственного персонала и населения от возможных последствий аварий, катастроф, стихийных бедствий
ПК-11	Знать основные положения законы и методы естественных наук; выявлять естественнонаучную сущность проблем, возникающих в ходе профессиональной деятельности, использовать для их решения соответствующий естественнонаучный аппарат
ПК-12	Применять математический аппарат для решения поставленных задач, способен применять соответствующую процессу математическую модель и проверять ее адекватность
ПК-13	Применять знания и навыки управления информацией
ПК-14	Самостоятельно изучать новые разделы фундаментальных наук
<i>Универсальные компетенции</i>	
УК-1	Владеть культурой мышления, иметь способности к обобщению, анализу, восприятию информации, постановке цели и выбору путей ее достижения
УК-2	Логически, верно, аргументировано и ясно строить устную и письменную речь
УК-3	Уважительно и бережно относится к историческому наследию и культурным традициям, толерантно воспринимать социальные и культурные различия; понимать движущие силы и закономерности исторического процесса, место

	человека в историческом процессе, политической организации общества
УК-4	Понимать и анализировать мировоззренческие, социально и лично значимые философские проблемы
УК-5	Владеть одним из иностранных языков на уровне бытового общения, а также переводить профессиональные тексты с иностранного языка
УК-6	К кооперации с коллегами, работе в коллективе
УК-7	Находить организационно-управленческие решения в нестандартных ситуациях и готов нести за них ответственность
УК-8	Использовать нормативно-правовые документы в своей деятельности
УК-9	Стремиться к саморазвитию, повышению своей квалификации и мастерства
УК-10	Осознавать социальную значимость своей будущей профессии, обладать высокой мотивацией к выполнению профессиональной деятельности
УК-11	Использовать основные положения и методы социальных, гуманитарных и экономических наук при решении социальных и профессиональных задач
УК-12	Анализировать социально значимые проблемы и процессы
УК-13	Использовать основные законы естественнонаучных дисциплин в профессиональной деятельности, применять методы математического анализа и моделирования, теоретического и экспериментального исследования
УК-14	Понимать сущность и значение информации в развитии современного информационного общества, осознавать опасности и угрозы, возникающие в этом процессе, соблюдать основные требования информационной безопасности, в том числе защиты государственной тайны
УК-15	Оформлять, представлять и докладывать результаты выполненной работы
УК-16	Создавать и редактировать тексты профессионального назначения
УК-17	Использовать для решения коммуникативных задач современные технические средства и информационные технологии
УК-18	Эффективно работать индивидуально и в коллективе, демонстрировать ответственность за результаты работы и готовность следовать корпоративной культуре организации.

Министерство науки и высшего образования Российской Федерации
федеральное государственное автономное образовательное учреждение
высшего образования
**«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»**

Инженерная школа ядерных технологий
Направление подготовки 01.04.02 «Прикладная математика и информатика»
Отделение экспериментальной физики

УТВЕРЖДАЮ:
Руководитель ООП
_____ Трифонов А.Ю.
(Подпись) (Дата) (Ф.И.О.)

ЗАДАНИЕ
на выполнение выпускной квалификационной работы

В форме:

Магистерской диссертации

Студенту:

Группа	ФИО
0BM81	Абдуллоев Рустам Махмадсобирович

Тема работы:

Математические методы ординации многомерных объектов и признаков	
Утверждена приказом директора (дата, номер)	28.02.2020, № 59-75/с

Срок сдачи студентом выполненной работы:	31.05.2020
--	------------

ТЕХНИЧЕСКОЕ ЗАДАНИЕ:

Исходные данные к работе	<i>Анкетные данные соискателей вакансий, зарегистрированных на сайте «Инженерное волонтерство».</i>
---------------------------------	---

Перечень подлежащих исследованию, проектированию и разработке вопросов	<ol style="list-style-type: none"> 1. Провести сравнительный анализ методов ординации многомерных объектов и признаков. 2. Разработать веб-инструмент для сбора исходных данных. 3. Программно реализовать методы ординации многомерных данных на плоскость. 4. Построить ординационные диаграммы различными методами (PCA, RDA).
Перечень графического материала	<ol style="list-style-type: none"> 1. Ординационная диаграмма на плоскости. 2. Ординационная диаграмма, построенная методом PCA и RDA.
Консультанты по разделам выпускной квалификационной работы <i>(если необходимо, с указанием разделов)</i>	
Раздел	Консультант
Финансовый менеджмент, ресурс эффективность и ресурсосбережение	Киселева Елена Станиславовна, доцент ОСГН ШБИП
Социальная ответственность	Романова Светлана Владимировна, старший преподаватель отделения ОД ШБИП
Иностранный язык	Смирнова Ульяна Александровна, специалист по учебно-методической работе отделения иностранных языков.

Дата выдачи задания на выполнение выпускной квалификационной работы по линейному графику	15.03.2020
---	------------

Задание выдал руководитель:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОЭФ	Семенов Михаил Евгеньевич	Кандидат ф.-м. н., доцент		

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
0BM81	Абдуллоев Рустам Махмадсобирович		

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

Студенту:

Группа	ФИО
0BM81	Абдуллоеву Рустаму Махмадсобировичу

Школа	ИЯТШ	Отделение школы (НОЦ)	ОЭФ
Уровень образования	Магистратура	Направление/специальность	01.04.02 Прикладная математика и информатика

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

1. Стоимость ресурсов научного исследования (НИ): <i>материально-технических, энергетических, финансовых, информационных и человеческих</i>	1. Стоимость расходных материалов 2. Стоимость расхода электроэнергии 3. Норматив заработной платы
2. Нормы и нормативы расходования ресурсов	1. Тариф на электроэнергию 2. Коэффициенты для расчета заработной платы
3. Используемая система налогообложения, ставки налогов, отчислений, дисконтирования и кредитования	1. Отчисления во внебюджетные фонды (30,2%)

Перечень вопросов, подлежащих исследованию, проектированию и разработке:

1. Оценка коммерческого и инновационного потенциала НТИ	1. Анализ конкурентных технических решений; 2. SWOT – анализ.
2. Разработка устава научно-технического проекта	1. Цели и результаты научно-технического проекта 2. Организационная структура научно-технического проекта 3. Ограничения и допущения научно-технического проекта
3. Планирование процесса управления НТИ: <i>структура и график проведения, бюджет, риски и организация закупок</i>	1. Структура работ в рамках научного исследования; 2. Определение трудоемкости выполнения работ и разработка графика проведения научного исследования; 3. Бюджет научно - технического исследования (НТИ).
4. Определение ресурсной, финансовой, экономической эффективности	1. Расчет интегрального показателя ресурсоэффективности НТИ 2. Расчет сравнительной эффективности НТИ

Перечень графического материала (с точным указанием обязательных чертежей):

1. «Портрет» потребителя результатов НТИ 2. Сегментирование рынка 3. Оценка конкурентоспособности технических решений 4. Матрица SWOT 5. График проведения и бюджет НТИ 6. Оценка ресурсной, финансовой и экономической эффективности НТИ 7. Потенциальные риски
--

Дата выдачи задания для раздела по линейному графику

Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОСГН ШБИП	Киселева Елена Станиславовна	К.Э.Н.		15.03.2020

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
0BM81	Абдуллоев Рустам Махмадсобирович		15.03.2020

**ЗАДАНИЕ ДЛЯ РАЗДЕЛА
«СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ»**

Студенту:

Группа		ФИО	
ОВМ81		Абдуллоеву Рустаму Махмадсобировичу	
Школа	ИЯТШ	Отделение (НОЦ)	Экспериментальной физики
Уровень образования	Магистратура	Направление/специальность	01.04.02 Прикладная математика и информатика

Тема ВКР:

Математические методы ординации многомерных объектов и признаков	
Исходные данные к разделу «Социальная ответственность»:	
1. Характеристика объекта исследования (вещество, материал, прибор, алгоритм, методика, рабочая зона) и области его применения	<p><i>Объектом исследования является сравнительный анализ математических методов ординации многомерных объектов и признаков.</i></p> <p><i>Область применения: Проект можно использовать для быстрого ранжирования многомерных объектов из различных предметных областей.</i></p>
Перечень вопросов, подлежащих исследованию, проектированию и разработке:	
<p>1. Правовые и организационные вопросы обеспечения безопасности:</p> <p>специальные (характерные при эксплуатации объекта исследования, проектируемой рабочей зоны) правовые нормы трудового законодательства;</p> <p>организационные мероприятия при компоновке рабочей зоны.</p>	<p><i>ГОСТ 12.2.032-78 «Система стандартов безопасности труда (ССБТ). Рабочее место при выполнении работ сидя. Общие эргономические требования»</i></p> <p><i>СанПиН 2.2.2/2.4.1340-03 «Гигиенические требования к персональным электронно-вычислительным машинам и организации работы»</i></p>

<p>2. Производственная безопасность:</p> <p>2.1. Анализ выявленных вредных и опасных факторов</p> <p>2.2. Обоснование мероприятий по снижению воздействия</p>	<p><i>Недостаточная освещенность рабочей зоны;</i></p> <p><i>Повышенный уровень статического электричества;</i></p> <p><i>Повышенное значение напряжения в электрической цепи, замыкание которой может произойти через тело человека.</i></p> <p><i>Повышенная или пониженная температура воздуха рабочей зоны</i></p>
<p>3. Экологическая безопасность:</p>	<p><i>- анализ воздействия при работе на ПЭВМ на атмосферу, гидросферу, литосферу;</i></p> <p><i>- наличие отходов (бумага, картриджи, компьютеры и т. д.);</i></p> <p><i>- методы утилизации отходов.</i></p>
<p>4. Безопасность в чрезвычайных ситуациях:</p>	<p><i>Сильные морозы, диверсия и пожар</i></p> <p><i>Потери при авариях на электросетях</i></p> <p><i>Наиболее типичная ЧС: пожар</i></p>
<p>Дата выдачи задания для раздела по линейному графику</p>	<p>10.03.2020</p>

Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
<p>Старший преподаватель отделения ОД ШБИП</p>	<p>Романова Светлана Владимировна</p>			

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
<p>ОВМ81</p>	<p>Абдуллоев Рустам Махмадсобирович</p>		

РЕФЕРАТ

Выпускная квалификационная работа содержит 112 листов, 11 рисунков, 28 таблиц, 27 источника, 5 приложений.

ОРДИНАЦИЯ, МЕТОД ГЛАВНЫХ КОМПОНЕНТ, АНАЛИЗ ИЗБЫТОЧНОСТИ, БАЗА ДАННЫХ, РАНЖИРОВАНИЕ.

Объектом исследования являются метод главных компонент и его модификации.

Цель данной работы - проецирования многомерных анализируемых объектов на координатную плоскость.

Методы исследования: метод главных компонент (Principal component analysis, PCA) и анализ избыточности (Redundancy analysis, RDA).

Область применения - заключается в применении рассматриваемых методов ординации в различных предметных областях, таких как, биоинформатика, эконометрика, обработка изображений, для сжатия данных, а также во многих общественных науках.

Оглавление

Введение	12
1. Теоретическая часть.....	12
1.1. Вводные определения информация	13
1.2. Основные типы данных	13
1.3. Ординация многомерных объектов	16
1.4. Описание необходимых пакетов и функции в среде R.....	17
2. Практическая часть	19
2.1. Разработка веб инструмента для сбора исходных данных	20
2.1.1. Проектирование базы данных и СУБД.....	20
2.1.2. Построение инфологической модели данных	21
2.1.3. Проектирование целостности базы данных	24
2.1.4. Создание таблиц на структурированном языке SQL	24
2.1.5. Реализация проектируемой базы данных в СУБД MySQL	30
2.1.6. Разработка интерфейса пользователя	32
2.1.7. Реализация сайта.	34
2.2. Программная реализация методов ординации многомерных данных на плоскость	37
2.3. Построение ординационных диаграммы различными методами.....	45
3. Финансовый менеджмент, ресурсоэффективность и ресурсосбережение. 50	
3.1. Потенциальные потребители результатов исследования.	50
3.2. Анализ конкурентных решений.	50
3.3. SWOT-анализ.	52
3.4. Инициация проекта.	53
3.5. Цели и результаты проекта.....	53
3.6. Организационная структура проекта.	55
3.7. Ограничения и допущения проекта.	56
3.8. Планирование научно-исследовательской работы.	57
3.8.1. Организация и планирование работы.	57
3.8.2. Продолжительность этапов работ.	58
3.9. Расчет сметы затрат на выполнение проекта.	62
3.9.1. Расчет материальных затрат.....	62
3.9.2. Расчет заработной платы для исполнителей.	63

3.9.3.	Расчет затрат на социальный налог.	64
3.9.4.	Расчет затрат на электроэнергию.	64
3.9.5.	Расчет амортизационных расходов.	65
3.9.6.	Расчет прочих расходов.	66
3.9.7.	Расчет общей себестоимости разработки.	67
3.9.8.	Расчет прибыли.	67
3.9.9.	Расчет НДС.	67
3.10.	Оценка экономической эффективности проекта.	68
3.11.	Вывод по разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение:	68
4.	Социальная ответственность.	69
4.1.	Правовые и организационные вопросы обеспечения безопасности. ..	71
4.2.	Производственная безопасность.	74
4.3.	Анализ опасных и вредных производственных факторов и обоснование мероприятий по их устранению.	74
4.3.1.	Недостаточная освещенность рабочей зоны.	74
4.3.2.	Повышенный уровень статического электричества.	75
4.3.3.	Повышенное значение напряжения в электрической цепи, замыкание которой может произойти через тело человека.	76
4.3.4.	Повышенная или пониженная температура воздуха рабочей зоны. 77	
4.4.	Экологическая безопасность.	79
4.5.	Безопасность в чрезвычайных ситуациях.	79
4.6.	Выводы и рекомендации по разделу.	81
	Заключение.	82
	Приложения.	86
	Приложение 1.	86
	Приложение 2.	89
	Приложение 3.	93
	Приложение 4.	95
	Приложение 5.	96
	Приложение 6.	97

Введение

Ординация (приведение в порядок) — это собирательный термин для многомерных методов, которые при отображении многомерного объекта (набора данных) в низкоразмерное пространство позволяют обнаружить шаблон в исходных данных. Одним из таких методов является метод главных компонент (Principal Component Analysis, PCA).

В узком смысле ординация представляет собой нахождение таких координатных осей на плоскости, относительно которых можно выполнить оптимальное проецирование многомерных анализируемых объектов [13].

Цель данной работы - проецирование многомерных анализируемых объектов на координатную плоскость.

Для достижения поставленной цели необходимо последовательно решить следующие задачи.

1. Провести сравнительный анализ методов ординации многомерных объектов и признаков.
2. Разработать веб-инструмент для сбора исходных данных.
3. Программно реализовать методы ординации многомерных данных на плоскость.
4. Построить ординационные диаграммы различными методами (PCA, RDA).

Актуальности работы - заключается в применении рассматриваемых методов ординации во многих областях, таких как, биоинформатика, эконометрика, обработка изображений, для сжатия данных, а также во многих общественных науках.

Объект исследования - метод главных компонент и его модификации. Исходные данные – анкетные данные соискателей вакансий на сайте «Инженерное волонтерство».

1. Теоретическая часть

1.1. Вводные определения информация

В данной работе в качестве многомерного объекта будем понимать анкетные данные соискателей вакансий на сайте “Инженерное волонтерство” (engineervol.ru). Сайт был разработан в рамках производственной практики. При разработке был использован конструктор Тильда, который - позволяет создавать сайты различных типов. С помощью разработанного веб-инструмента мы собрали анкетные данные для проведения дальнейших исследований [15]. Пример регистрационной анкеты приведен на рисунке 1.

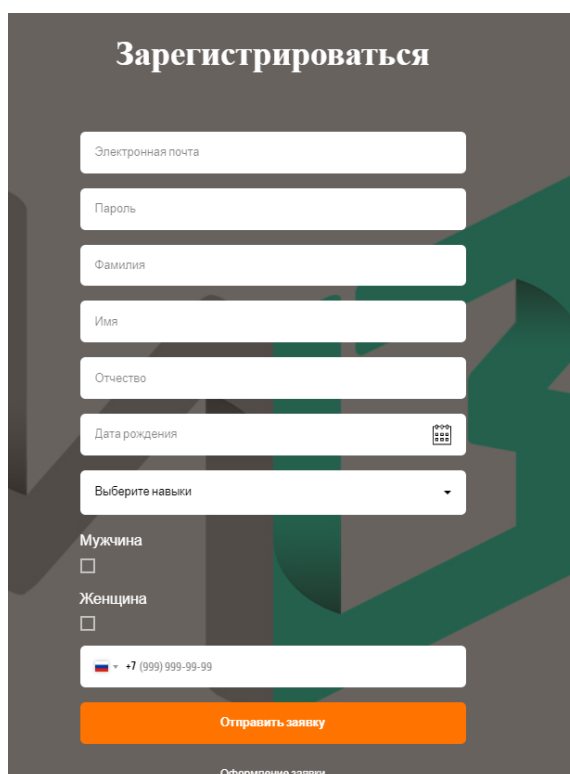


Рисунок 1 - Оформление заявки соискателей вакансий на сайте “Инженерное волонтерство” (engineervol.ru)

Как видно из рисунка 1, при заполнении анкеты пользователь может дать ответ с использованием различных типов данных - порядковые, интервальные и номинальные. Остановимся на этих типах подробнее.

1.2. Основные типы данных

Говоря о номинальных данных, следует отметить, что с ними можно сделать следующие операции: их можно посчитать, определить процент от

целого, при этом вычислить среднее значение не получится. Можно определить, сколько студентов были приглашены на собеседование, или сколько процентов соискателей являются более подходящими к требованиям.

Если возможны только две категории данных, то в таком случае мы имеем дело с дихотомическими данными. Например, к таким данным можно отнести ответы на вопросы, которые требуют ответа «да» или «нет». Если, проводя анализ, мы выявили, что соискатель подходит или нет, то это тоже будет дихотомическое решение [16].

Данные, которые естественным образом упорядочиваются по категориям, называются *порядковыми*. Вопросы в анкетах, ответами на которые могут быть такие фразы, как «полностью не согласен», «не согласен», «нейтрально отношусь», «согласен», «полностью согласен», предназначены для сбора порядковых данных. Ни одна из категорий порядковой шкалы не имеет фактической математической величины. Числовые значения зачастую присваиваются категориям для того, чтобы облегчить запись или анализ данных. Например: 1 = полностью не согласен, 5 = полностью согласен, но это распределение условно, и можно выбрать любую группу упорядоченных чисел для обозначения групп. Например, с такой же легкостью можно обозначить, что цифра 5 будет обозначать «полностью не согласен», а 1 – «полностью согласен» [16].

Цифры, которые мы присваиваем порядковым категориям, влияют на толкование конечного результата, но мы можем выбрать любой набор цифр, при условии соблюдения порядка нумерации.

Для порядковых данных, также как и в случае с номинальными данными, мы можем посчитать и определить процент от целого. При этом нет конкретной точки зрения по поводу того, можно ли для порядковых данных посчитать среднее значение. С одной стороны, невозможно определить среднее значение для категории «полностью согласен», например, и даже если вы определите их числовые значения, они не будут иметь фактической

математической величины. Каждое числовое значение представляет определенную категорию, а не количество чего бы то ни было.

В некоторых задачах настоятельно не рекомендуется использовать порядковые данные для проведения подобных расчетов, в то время как в других — это постоянная практика [16].

Интервальные данные – числовые данные, поэтому мы можем производить с ними математические операции, но для таких данных не существует «значимой» нулевой точки – другими словами, если интервальные данные принимают значение ноль, то, что мы измеряем, не отсутствует. В частности, 00:00 часов указывает на начало нового дня, а не отсутствие времени. К интервальным данным, с которыми мы сталкиваемся в повседневной жизни относится, например, календарный год и температура. Нулевое значение для годов не значит, что ранее времени не существовало, а нулевая температура (измеряемая в градусах Цельсия или Фаренгейта) отнюдь не показатель того, что тепла нет [16].

Нормативные данные – это тип числовых данных. Они также схожи с интервальными данными, но, помимо этого, в отличие от интервальных, имеют значимую нулевую точку. Для нормативных данных ноль обозначает несуществование того, что вы измеряете, – ноль минут, ноль людей в нашей плоскости, ноль подходящих людей. Во всех этих случаях ноль означает, что у вас нет того, что вы измеряете, и это отличается от того, что мы обсуждали в разделе интервальных данных. Другие часто встречаемые переменные, которые можно отнести к нормативным данным, – рост, возраст.

Нормативные и интервальные данные существуют двух типов: непрерывные и дискретные. В случае непрерывных данных, они принимают любое значение на шкале. Но это говорит о том, что такие данные принимают все возможные числовые значения. Они принимают только те значения, которые возможны в рамках границ шкалы. Дискретные же данные могут быть выражены только ограниченным набором значений (как правило, целыми

числами), величины между этими значениями невозможны. Например, мы часто округляем непрерывные данные до конкретного количества цифр. Однако все равно эти данные остаются непрерывными, а не дискретными [16].

1.3. Ординация многомерных объектов

Говоря о термине «ординация», следует отметить, что оно имеет отношение к поиску наиболее "естественных" системных закономерностей и особенно эффективен при наличии комплекса взаимно коррелированных влияющих факторов.

В данный момент в среде программирования R реализовано и разработано огромное множество алгоритмов сжатия информационного пространства. Например, можно выделить следующие [13]:

- метод главных компонент (principal component analysis, PCA) - оперирующий с корреляционной матрицей, представленный в пакете PCoA.
- метод главных координат (PCoA) и неметрическое многомерное шкалирование NMDS (Non-metric multidimensional scaling)-выполняют последовательную процедуру преобразования любой матрицы дистанций.
- анализ соответствий или корреспондентный анализ (correspondence analysis, CA) - основан на методе итерации встречного усреднения взвешивающих коэффициентов для переменных и объектов.
- функции многомерного факторного анализа (factor analysis, FA)-включая версии для смешанных и иерархических данных, они представлены в пакете FactoMineR.
- анализ скрытых классов (latent class analysis, LCA) — это статистический метод определения неизмеренного членства в классе среди субъектов с использованием категориальных и/или

непрерывных наблюдаемых переменных, представлен в пакете e1071.

Метод главных координат PCoA, или многомерное шкалирование (MDS, multidimensional scaling) схож по своей сути с методом главных компонент, однако может выполнять вычисление собственных значений и собственных векторов произвольной квадратной симметричной матрицы вместо матрицы корреляции [13].

1.4. Описание необходимых пакетов и функции в среде R

Пакеты R представляют собой набор функций R, выполняемый код и примеры наборов данных. Они хранятся в каталоге под названием «пакет». По умолчанию R устанавливает набор пакетов во время установки. Дополнительные пакеты добавляются позже, когда они необходимы для какой-то конкретной цели. Когда мы запускаем консоль R, по умолчанию доступны только пакеты по умолчанию. Другие пакеты, которые уже установлены, должны быть явно загружены для использования программой R, которая будет их использовать [17].

Пакетов, доступных на языке R, более тысячи. Перечислим основные пакеты и функции для нашей реализации следующим образом.

1. Ggplot2 – пакет служит для визуального построения ординационной диаграммы.
2. Vegan(), rda(), princomp(), prcomp(), PCA() – функции для расчета главных компонент.
3. FactoMineR(), Factoextra() - функции предназначены для визуализации результатов многомерного и кластерного анализа.

Ввод и сохранение данных

4. data(x) - для загрузки таблицы данных, ранее сохранённых командой save().

5. Функция *viz_pca()* - осуществляет построение диаграммы для разных объектов.
6. Функция *library()* - загружает дополнительный пакет.
7. Функция *read.table(file)* - считывает файл с данными и записывает его в таблицу данных.
8. *read.csv(filename, header=True)* – функция чтения данных из файла.
9. *print (a, ...)* - выводит текст на экран.

Создание векторов и таблиц с данными

10. *c(...)* функция общего назначения (от concatenation - объединение, слияние);
11. *data.frame(...)* – функция для создания таблицы данных.

Конвертация R-объект

12. *as.data.frame(x)* – создание таблицы данных.
13. *Unique(x)* – поиск уникальных значений среди элементов вектора *x*.
14. *apply (x, MARGIN, FUN = ...)* – возвращает вектор, массив или список значений, которые были получены путем применения функции *FUN* к некоторым элементам массива или матрицы *x*; те элементы *x*, которые подлежат обработке, указываются при помощи аргумента *MARGIN*; к примеру, если *x* – это матрица, то при значении *MARGIN* = 1 функция *FUN* будет применена к каждой строке матрицы, при *MARGIN* = 2 - к каждому столбцу матрицы.
15. *lapply (x, FUN = ...)* возвращает список той же длины, что и *x*; при этом значения в новом списке будут результатом применения функции *FUN* к элементам исходного объекта *x*.

Математические функции

16. *max(x)* – определяет максимальное значение в числовом векторе *x*.
17. *min(x)* – определяет минимальное значение в числовом векторе *x*.
18. *mean(x)* – вычисляет среднее арифметическое совокупности *x*.

19. `sd(x)` – возвращает стандартное отклонение совокупности x .
20. `cov(x)` – вычисляет корреляционную матрицу, если x является матрицей или таблицей данных (результатом будет 1, если x является вектором).

Матрицы

21. `t(x)` - транспонирование матрицы x .

Построение графиков

22. `plot(x)` – строит график значений вектора x , упорядоченных вдоль оси x .
23. `plot(x, y)` – отображает график зависимости y от x .
24. `xlab=`, `ylab=` задают названия соответствующих осей.
25. `polygon(x, y)` рисует полигон, чьи координаты хранятся в объектах x и y .
26. `paste (... , sep=...)` преобразует векторы в текстовые переменные и объединяет их в одном текстовом выражении; `sep` аргумент позволяет спросить Текстовое выражение, отделяя ценности слитого векторы (по умолчанию - это пробел).

Программирование

27. `function(список аргументов){ выражение }` – позволяет создать функцию пользователя.
28. `If(условие){ выражение }` – проверяет, выполняется ли условие в скобках; если условие выполняется, то эта функция возвращает значение выражения.
29. `If(условие){ выражение 1 }else{ выражение 2 }` – проверяет, выполняется ли условие в скобках; если условие выполняется, то эта функция возвращает выражение 1, иначе – выражение 2[18].

2. Практическая часть

2.1. Разработка веб инструмента для сбора исходных данных

Веб-сайт представляет собой набор документов и файлов, которые можно просмотреть в веб-браузере. Типичный веб-сайт состоит из страниц, называемых веб-страницами. Эти страницы могут содержать текстовые и мультимедийные элементы, такие как изображения, видео, аудио или анимацию [6].

Веб-страницы разрабатываются с использованием языка гипертекстовой разметки HTML (HyperText Markup Language), который состоит из тегов, обеспечивающих структуру веб-страницы, то есть выделяют заголовки, абзацы, списки и т.д. [7]. В нашем случае для разработки веб-сайта использовали конструктор сайтов Тильда [15].

2.1.1. Проектирование базы данных и СУБД

Система управления базами данных (СУБД) является программным обеспечением, которое взаимодействует с пользователем, другими приложениями и самой базой данных для сбора и анализа данных [12]. Общее назначение СУБД заключается в разработке и выполнении запросов, обновления и администрирования баз данных.

Применение СУБД для хранения и управления данными имеет ряд преимуществ и недостатков. Одно из главных преимуществ применения СУБД это получение доступа и использование одних и тех же данных с сохранением их целостности конечным пользователям и прикладным программистам. СУБД является центральным хранилищем данных. К нему могут обращаться несколько пользователей одновременно [12].

База данных – это место (контейнер) для хранения данных. В базе данных даже самая маленькая часть информации становится данными. Например, волонтер – это данные, сотрудники – это данные, рост, возраст– все это данные.

Современные системы баз данных могут быть использованы для работы несколькими пользователями. Такие СУБД дают возможность одновременно многим пользователям обращаться к одной и той же базе данных. Функции, с помощью которых возможно такое обращение с СУБД, называются стратегиями управления параллелизмом. Они гарантируют, что доступ к данным всегда корректен и целостность данных поддерживается.

СУБД должны обеспечивать возможность определения и применения определенных ограничений для обеспечения ввода пользователями достоверной информации и сохранения целостности данных. Ограничение базы данных – это ограничение или правило, которое определяет, что можно ввести или изменить в таблице, например, дата рождения, используя определенный формат ГГГГ.ММ.ДД [12].

По своей сущности СУБД позволяет многим пользователям иметь доступ к своей базе данных как индивидуально, так и одновременно. Для пользователей не имеет значение, где и как хранятся данные, к которым они обращаются [12].

2.1.2. Построение инфологической модели данных

База данных предназначена для хранения и добавления данных об объектах и отношениях между ними. В ходе проведения анализа предметной области были выделены следующие объекты и их атрибуты.

- ❖ Пользователи(волонтеры);
- ❖ Сотрудники организации;
- ❖ Город;
- ❖ Задачи;

У объекта «Пользователи (волонтеры)» можно выделить следующие атрибуты:

- ❖ Код пользователя (В);

- ❖ Имя;
- ❖ Фамилия;
- ❖ Телефон;
- ❖ Дата рождения;
- ❖ Город проживания;
- ❖ Стаж волонтера;
- ❖ Авторизация через социальные сети (vk.com);
- ❖ Электронная почта;

У объекта «Сотрудники организации» можно выделить следующие атрибуты:

- ❖ Код сотрудника организации (НС);
- ❖ Название организации;
- ❖ Сфера деятельности;
- ❖ Город;
- ❖ Об организации;
- ❖ Номер телефона;
- ❖ Должность;

У объекта «Город» можно выделить один атрибут:

- ❖ Код города;
- ❖ Название города;

У объекта «Задачи» можно выделить следующие атрибуты:

- ❖ Код задачи;
- ❖ Функционал;
- ❖ Сколько человек необходимо;
- ❖ Сроки (с – до);

Анализ определенных выше объектов и атрибутов позволяет выделить сущности проектируемой базы данных и построить ее инфологическую модель.

К основным сущностям можно отнести:

- 1) **Пользователи (волонтеры) (НП) (id_user)**, Имя, Фамилия, Телефон, Дата рождения, Город проживания, Стаж волонтера, Авторизация через социальные сети (vk.com), Электронная почта. Эта сущность предназначена для хранения информации о волонтерах.
- 2) **Номер сотрудник организации (НС) (id_organization)**, Название организации, Сфера деятельности, Город, об организации, Номер телефона, Должность. Сущность содержит информацию о сотрудниках.
- 3) **Город (id_city)**, Название города.
- 4) **Задачи (Код задачи, Функционал, Сколько человек необходимо, Сроки (с – до).**

Инфологическую модель на языке "Таблицы-связи" можно изобразить следующим образом (Рисунок 2):

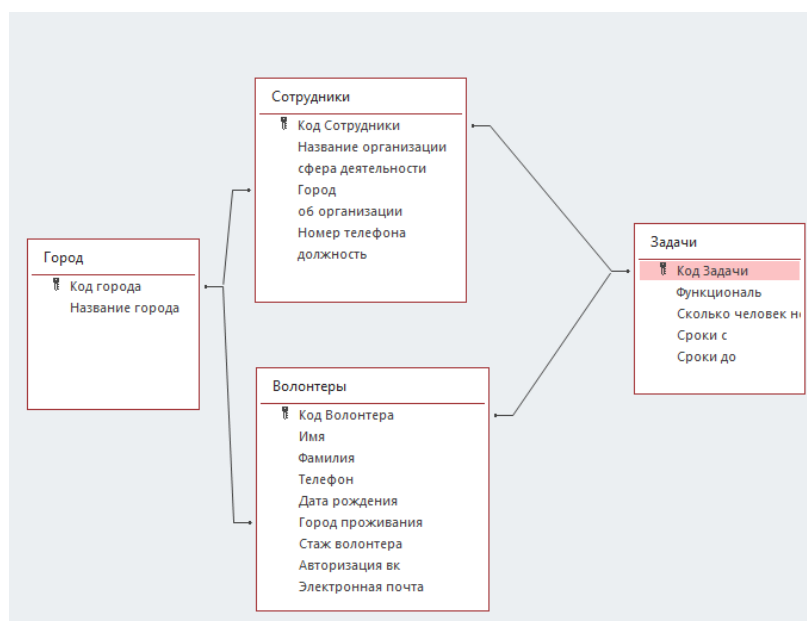


Рисунок 2 – Инфологическая модель базы данных "Инженерное волонтерство".

2.1.3. Проектирование целостности базы данных

В соответствии с процедурой проектирования баз данных каждая из полученных сущностей должна быть представлена базовой таблицей. Эти таблицы можно описать следующим образом:

СОЗДАТЬ ТАБЛИЦУ **Сотрудники организации** *(Главная сущность)

ПЕРВИЧНЫЙ КЛЮЧ (КС)

ПОЛЯ (Код сотрудники Целое Без знака Название организации Текст 20, Сфера деятельности Текст 20, Город Текст 15, об организации Текст 50, Номер телефона Целое 15, Должность Текст 15);

СОЗДАТЬ ТАБЛИЦУ **Задачи** *(Обозначающая сущность)

ПЕРВИЧНЫЙ КЛЮЧ (Код задачи)

ПОЛЯ (Код задачи Целое, Функционал Текст 25, Сколько человек необходимо Целое, Сроки с Целое, Сроки до Целое);

СОЗДАТЬ ТАБЛИЦУ **Волонтеры** *(Стержневая сущность)

ПЕРВИЧНЫЙ КЛЮЧ (Код Волонтера)

ПОЛЯ (КВ Целое, Имя Текст 15, Фамилия Текст 15, Телефон Целое, Дата рождения Дата и время 6, Город проживания Текст 15, Стаж волонтера Целое, Авторизация в вк Текст 20, Электронная почта Текст 20);

СОЗДАТЬ ТАБЛИЦУ **Город** *(Стержневая сущность)

ПЕРВИЧНЫЙ КЛЮЧ (Код Город)

ПОЛЯ (КГ Целое, Город Текст 15);

2.1.4. Создание таблиц на структурированном языке SQL

Проект был реализован с помощью среды разработки Phpstorm и PhpMyAdmin.

Код для создания сущности в Phpstorm приведен в приложении 1.

Создание БД в PhpMyAdmin.

Создаем базу данных в PhpMyAdmin MySQL:

```
CREATE DATABASE valanter  
CHARACTER SET utf8mb4  
COLLATE utf8mb4_unicode_ci;
```

Создаем таблицу Сотрудники организации.

```
CREATE TABLE `organisation` (  
  `id` int NOT NULL,  
  `name` varchar(20) CHARACTER SET utf8mb4 COLLATE utf8mb4_unicode_ci  
  NOT NULL COMMENT 'Название организации',  
  `area` varchar(20) CHARACTER SET utf8mb4 COLLATE utf8mb4_unicode_ci  
  NOT NULL COMMENT 'Сфера деятельности',  
  `city_id` int NOT NULL,  
  `about` varchar(50) CHARACTER SET utf8mb4 COLLATE utf8mb4_unicode_ci  
  DEFAULT NULL COMMENT 'Об организации'  
)  
ENGINE=InnoDB  
DEFAULT CHARSET=utf8mb4  
COLLATE=utf8mb4_unicode_ci;
```

Данные таблицы Сотрудники организации.

```
INSERT INTO `organisation` (`id`, `name`, `area`, `city_id`, `about`) VALUES  
(1, 'ТПУ', 'Доцент', 1, 'Образовательная организация '),  
(2, 'Кванториум', 'преподаватель', 1, 'Детский парк'),  
(3, 'Полиглот', 'Помощник ', 1, 'Утилизация токсичных отходов'),  
(4, 'ПАО Сбербанк', 'консультант', 1, 'российский государственный  
финансовый конгломерат');
```

Индексы таблицы Сотрудники организации.

```
ALTER TABLE `organisation`  
  ADD PRIMARY KEY (`id`),  
  ADD KEY `IDX_E6E132B48BAC62AF` (`city_id`);  
MODIFY `id` int NOT NULL AUTO_INCREMENT, AUTO_INCREMENT=6;
```

Ограничения внешнего ключа таблицы Сотрудники организации.

```
ALTER TABLE `organisation`  
  ADD CONSTRAINT `FK_E6E132B48BAC62AF` FOREIGN KEY (`city_id`)  
REFERENCES `city` (`id`);
```

Создаем таблицу Пользователи (Волонтеры).

```
CREATE TABLE `user` (  
  `id` int NOT NULL,  
  `Имя` varchar(15) CHARACTER SET utf8mb4 COLLATE utf8mb4_unicode_ci  
NOT NULL COMMENT 'Имя',  
  `lastname` varchar(15) CHARACTER SET utf8mb4 COLLATE  
utf8mb4_unicode_ci DEFAULT NULL COMMENT 'Фамилия',  
  `phone` varchar(15) CHARACTER SET utf8mb4 COLLATE  
utf8mb4_unicode_ci NOT NULL COMMENT 'Номер телефона',  
  `Возраст` int NOT NULL COMMENT 'Возраст',  
  `city_id` int NOT NULL COMMENT 'Город',  
  `description` varchar(255) COLLATE utf8mb4_unicode_ci DEFAULT NULL  
COMMENT 'О себе',  
  `auth_vk` varchar(15) CHARACTER SET utf8mb4 COLLATE  
utf8mb4_unicode_ci DEFAULT NULL COMMENT 'ссылка в соц. сеть vk',  
  `email` varchar(20) CHARACTER SET utf8mb4 COLLATE utf8mb4_unicode_ci  
NOT NULL COMMENT 'Email',
```

```
`experience` int DEFAULT NULL COMMENT 'Стаж волонтера (сколько лет)'
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4
COLLATE=utf8mb4_unicode_ci;
```

Дамп данных таблицы Пользователи (Волонтеры).

```
INSERT INTO `user` (`id`, `Имя`, `lastname`, `phone`, `Возраст`, `city_id`,
`description`, `auth_vk`, `email`, `experience`) VALUES
(1, 'Сохиб', 'Сохибназар', '89254946467', 24, 1, 'Студент', 'vk.com/dajgdga',
'sts4@tpu.ru', 2),
(2, 'Рустам ', 'Абдуллоев', '89254946467', 25, 1, 'Студент ', 'vk.com/ricco',
'rma8@tpu.ru', 1),
(3, 'Абдурахман', 'Ахроров', '+79138659187', 24, 1, 'спортсмен',
'vk.com/borec.21', 'abdu.95@gmail.com', 2),
(4, 'Григорий', 'Погосян', '8945715423', 25, 1, 'Бармен', 'vk.com/grisha95',
'grisha95@gmai.com', 1),
(5, ' Игорь', 'Блажевич', '79021100098', 22, 2, 'Студент', 'alukoin@ya.ru',
'margari1@yandex.ru', 2),
(6, ' Руфина', 'Валиева', '79021100099', 21, 2, 'Студентка', '@bk.ru', '@avangard-
dsl.ru', 3),
(7, 'Александр', 'Возвышаев', '79021100147', 20, 3, 'Студент', 'aertrhjp@bk.ru',
'ehunxuxjsr@bk.ru', 3),
(8, 'Алексей', 'Гриненко', '79032478845', 24, 1, 'Студент', 'oligarx_@bk.ru',
'jcrckyhadr@bk.ru', 1),
(9, 'Родион', 'Жигляев', '79032478917', 23, 2, 'Студент', 'jcrckuhr@bk.ru',
'oligarx_@bk.ru', 2),
(10, 'Анастасия ', 'Журавлева', '79032478979', 25, 1, 'Студентка', 'deb@e-
mails.ru', 'info@asc.com', 1),
(11, 'Кирилл', 'Зиборов', '79032480839', 26, 3, 'Студент', 'ptbyeady@bk.ru',
'@deks.ru', 3),
```

(12, 'Дмитрий', 'Колосов', '79032481196', 27, 1, 'Студент', '@corp.mail.ru', 'antioch@сmail.ru', 2),

(13, 'Алексей ', 'Красных', '79032511037', 21, 2, 'Студент', 'nbiyl@email.ru', 'vmxzgjwuyziy@c2v.com', 2),

(14, 'Иван', 'Кузнецов ', '79032511075', 22, 3, 'Студент', 'nbizm@email.ru', 'nbiylll@email.ru', 2),

(15, 'Александр', 'Куницын ', '79032511134', 22, 2, 'Студент', 'det@e-mails.ru', 'nbizulyb@email.ru', 1),

(16, 'Александр', 'Мазурин', '79032511175', 22, 1, 'Студент', 'sxsy@front.ru', 'iebkseryu@emails.ru', 1),

(17, 'Никита ', 'Мареев', '79032511178', 23, 1, 'Студент', 'sxsy@front.ru', 'deegez@gmail.com', 4),

(18, 'Кирилл', 'Мосиевич', '79037672163', 23, 2, 'Студент', 'ln01@gmail.com', 'mm27i25@gmail.com', 2),

(19, 'Павел', 'Носиков', '79037672179', 24, 3, 'Студент', 'a.wo@gmail.com', 'pier002@gmail.com', 2),

(20, 'Николай', 'Носков', '79037672527', 21, 2, 'Студент', 'visil@gmail.com', 'upds@gmail.com', 2),

(21, 'Никита', 'Патрашкин', '79037672339', 20, 3, 'Студент', 'pd3@gmail.com', 'savtrios@gmail.com', 4),

(22, 'Никита', 'Пивнев', '79037672682', 20, 1, 'Студент', 'goldmail.ru', 'boeaagau@goldmail.ru', 1),

(23, 'Екатерина ', 'Профорова', '79037672529', 19, 1, 'Студентка', 'boehy@gold', 'boelovx@goldmail.ru', 4),

(24, ' Егор ', 'Скударев', '79037672689', 18, 1, 'Студент', 'boeu@gold', 'buqlz@goldmail.ru', 3),

(25, ' Ксения ', 'Хазова', '79037672687', 19, 2, 'Студентка', 'ebwqfnt@gold', 'hotmail.com', 3),

(26, 'Феликс', 'Юсупов ', '79037672730', 24, 2, 'Студент', 'wocnqcdkguk', 'lojflajjemz@', 1),

(27, ' Артур', 'Хайруллин', '79037672780', 23, 1, 'Студент', 'mxcrkupzou', 'nbizulyb@email.ru', 2),

(28, 'Екатерина ', 'Юшкина', '79037672846', 22, 1, 'Студентка', 'ajtelkhuxf', 'aefebrxpxv@list.ru', 2);

Индексы таблицы Пользователи (Волонтеры).

```
ALTER TABLE `user`  
ADD PRIMARY KEY (`id`),  
ADD KEY `IDX_8D93D6498BAC62AF` (`city_id`);;  
MODIFY `id` int NOT NULL AUTO_INCREMENT, AUTO_INCREMENT=4;
```

Ограничения внешнего ключа таблицы Пользователи (Волонтеры).

```
ALTER TABLE `user`  
ADD CONSTRAINT `FK_8D93D6498BAC62AF` FOREIGN KEY (`city_id`)
```

Создаем таблицу Задания для волонтеров.

```
CREATE TABLE `task` (  
  `id` int NOT NULL,  
  `name` varchar(255) COLLATE utf8mb4_unicode_ci DEFAULT NULL  
  COMMENT 'Название задачи или мероприятия',  
  `task_funct` varchar(255) COLLATE utf8mb4_unicode_ci NOT NULL  
  COMMENT 'функционал',  
  `count` int DEFAULT NULL COMMENT 'сколько человек необходимо',  
  `date_start` date DEFAULT NULL COMMENT 'Дата с',  
  `date_end` date DEFAULT NULL COMMENT 'Дата до')  
ENGINE=InnoDB          DEFAULT          CHARSET=utf8mb4  
COLLATE=utf8mb4_unicode_ci;
```

Данные таблицы Задания для волонтеров.

```
INSERT INTO `task` (`id`, `name`, `task_func`, `count`, `date_start`, `date_end`)  
VALUES
```

```
(1, 'Благотворительный фонд - Мастер ОК ', 'Помоги людям, чьи возможности  
меньше твоих.', 11, '2020-08-28', '2021-01-31'),
```

```
(2, 'Благотворительный фонд - «Близкие Другие»', 'Наша цель - качественное  
улучшение жизни людей с тяжелыми нарушениями развития.', 10, '2020-08-25',  
'2021-03-12'),
```

```
(3, 'Благотворительная организация - Фонд \"Искорка\" - onco70.ru', 'Лечение  
онкобольных детей на уровне мировых стандартов', 7, '2021-02-01', '2021-05-  
25');
```

Индексы таблицы Задания для волонтеров.

```
ALTER TABLE `task`  
ADD PRIMARY KEY (`id`);
```

2.1.5. Реализация проектируемой базы данных в СУБД MySQL

Получение данных. Данные для таблиц **Сотрудники организации**, **Пользователи** и **Задания для волонтеров** были сгенерированы специально для данного проекта. В итоге для таблицы **Сотрудники организации** данные выглядят следующим образом:

Таблица 1. Сотрудники организации

№	Название орг-ции	Позиция	Город	Об организации
1	ТПУ	доцент	Томск	Образовательная организация
2	Кванториум	преподаватель	Томск	Детский парк
3	Полиглот	помощник	Томск	Утилизация токсичных отходов
4	ПАО Сбербанк	консультант	Томск	российский государственный финансовый конгломерат

Таблица 2. Пользователи (Волонтеры)

id	Имя	Фамилия	Телефон	Возраст	Город	О себе	Аккаунт	Электронная почта	Опыт
1	Сохиб	Сохибназар	89254946467	24	1	Студент	vk.com/dajgdga	sts4@.tpu.ru	2
2	Рустам	Абдуллоев	89254946467	25	1	Студент	vk.com/ricco	rma8@tpu.ru	1
3	Абдурахман	Ахроров	+79138659187	24	1	спортсмен	vk.com/borec.21	abdu.95@gmail.com	2
4	Григорий	Погосян	8945715423	25	1	Бармен	vk.com/grisha95	grisha95@gmai.com	1
5	Игорь	Блажевич	79021100098	22	2	Студент	alukoin@ya.ru	margaril@yandex.ru	2
6	Руфина	Валиева	79021100099	21	2	Студентка	@bk.ru	@avangard-dsl.ru	3
7	Александр	Возвышаев	79021100147	20	3	Студент	aertrhjp@bk.ru	ehunxuxjsr@bk.ru	3
8	Алексей	Гриненко	79032478845	24	1	Студент	oligarx_@bk.ru	jrckkyhadp@bk.ru	1
9	Родион	Жигляев	79032478917	23	2	Студент	jrckkyhp@bk.ru	oligarx_@bk.ru	2
10	Анастасия	Журавлева	79032478979	25	1	Студентка	deb@e-mails.ru	info@asc.com	1
11	Кирилл	Зиборов	79032480839	26	3	Студент	ptbyeady@bk.ru	@deks.ru	3
12	Дмитрий	Колосов	79032481196	27	1	Студент	@corp.mail.ru	antioch@cmail.ru	2
13	Алексей	Красных	79032511037	21	2	Студент	nbiyl@email.ru	vmxzgjuwuziy@c2v.com	2
14	Иван	Кузнецов	79032511075	22	3	Студент	nbizm@email.ru	nbiylll@email.ru	2
15	Александр	Куницын	79032511134	22	2	Студент	det@e-mails.ru	nbizulyb@email.ru	1
16	Александр	Мазурин	79032511175	22	1	Студент	sxsy@front.ru	iebksergy@emails.ru	1
17	Никита	Мареев	79032511178	23	1	Студент	sxsy@front.ru	deegez@gmail.com	4
18	Кирилл	Мосиевич	79037672163	23	2	Студент	ln01@gmail.com	mm27i25@gmail.com	2
19	Павел	Носиков	79037672179	24	3	Студент	a.wo@gmail.com	pier002@gmail.com	2
20	Николай	Носков	79037672527	21	2	Студент	visil@gmail.com	upds@gmail.com	2
21	Никита	Патрашкин	79037672339	20	3	Студент	pd3@gmail.com	savtrios@gmail.com	4
22	Никита	Пивнев	79037672682	20	1	Студент	goldmail.ru	boeaagau@goldmail.ru	1
23	Екатерина	Профорова	79037672529	19	1	Студентка	boehy@gold	boelovx@goldmail.ru	4
24	Егор	Скударев	79037672689	18	1	Студент	boeu@gold	buqlz@goldmail.ru	3
25	Ксения	Хазова	79037672687	19	2	Студентка	ebwqfnt@gold	hotmail.com	3
26	Феликс	Юсупов	79037672730	24	2	Студент	wocnqcdkguk	lojflajjemz@	1
27	Артур	Хайруллин	79037672780	23	1	Студент	mxcrcupzou	nbizulyb@email.ru	2
28	Екатерина	Юшкина	79037672846	22	1	Студентка	ajtelkhuxf	aefebxpxv@list.ru	2

Таблица 3. Задания для волонтеров

№	Название	Функционал	Кол-во	Дата с	Дата до
1	Благотворительный фонд - Мастер ОК	Помоги людям, чьи возможности меньше твоих.	24	2020-08-28	2021-01-31
2	Благотворительный фонд - «Близкие Другие»	Качественное улучшение жизни людей с тяжелыми нарушениями развития.	23	2020-08-25	2021-03-12
3	Благотворительная организация - Фонд "Искорка" - onco70.ru	Лечение онкобольных детей на уровне мировых стандартов.	13	2021-02-01	2021-05-25

В итоге для каждой таблицы получено следующее количество записей:

- 1) Сотрудники организации – 4 записи;
- 2) Пользователи (Волонтеры)– 60 записей;
- 3) Задания для волонтеров – 3 записи;

В свою очередь для выполнения каждого задания необходимо определенное количество людей. Например, для выполнения первого задания необходимо 11 - человек. Соискатели вакансий из таблицы «Волонтеры» смогут выбрать любое из трех заданий для выполнения.

2.1.6. Разработка интерфейса пользователя

Основная задача интерфейса предоставить пользователю возможность просматривать информацию о проекте, вызвать помощи и зарегистрироваться. Разработанный пользовательский интерфейс включает следующие блоки.

1. “О проекте” – содержит краткую и подробную информацию (в виде pdf – для скачивания) информацию о проекте.
2. “Мне нужна помощь” - в этом разделе размещены задания, предлагаемые различными организациями, в частности можно указать:
 1. Название задачи или мероприятия, на котором нужна помощь
 2. Функционал
 3. Сколько человек необходимо

4. Сроки (с – до)
3. “Я хочу помогать” – этот блок предназначен для волонтеры и здесь существует ячейки с заданиями по принципу сайта Procharity (<https://procharity.ru/>)
4. “Дополнительная информация” – содержит информацию о навыках с выбором: дизайн, 3D – моделирование, конструирование, программирование, электроника, схемотехника.
5. “Зарегистрироваться” – блок предназначен для регистрации, в нем нужно указать:
- ❖ Электронная почта
 - ❖ Пароль
 - ❖ Фамилия
 - ❖ Имя
 - ❖ Отчества
 - ❖ Дата рождения
 - ❖ Навык
 - ❖ Выбор пола – (“Мужчина”, “Женщина”)
 - ❖ Мобильный телефон
 - ❖ Отправка заявки (оформление заявки)

2.1.7. Реализация сайта.

Сайт «Инженерное волонтерство» был реализован с помощью конструктора Tilda.

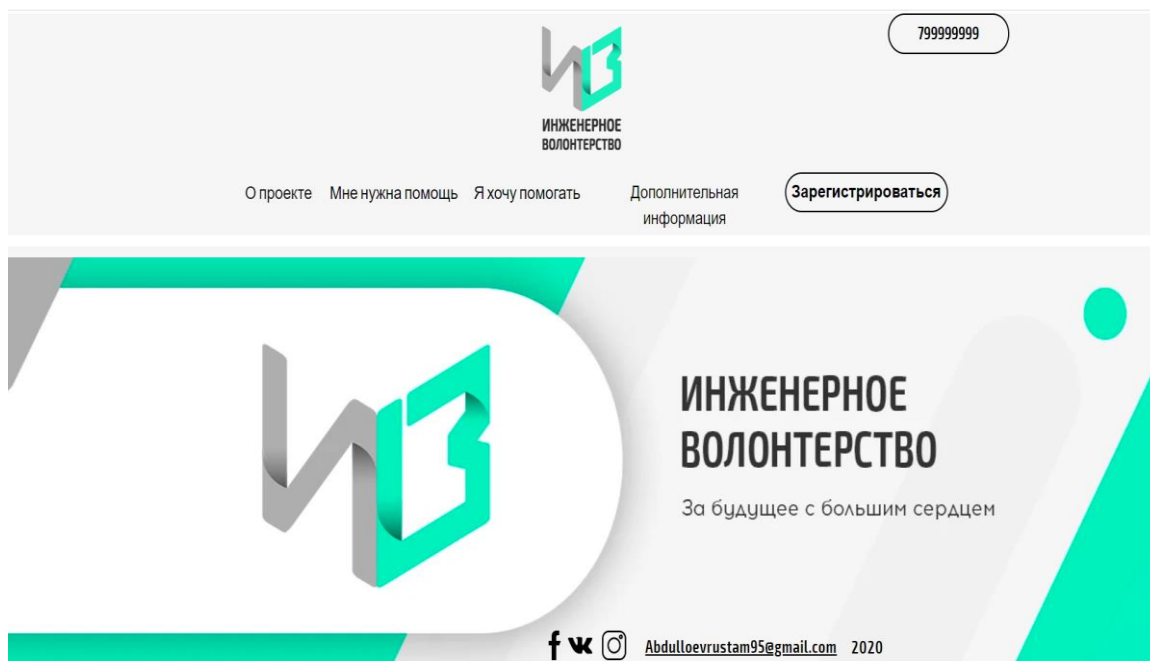


Рисунок 3 - Главная страница проекта.

Заходя в блок “ О проекте” мы увидим информацию в виде галереи. Здесь можно пролистать слайдер и прочитать информацию.

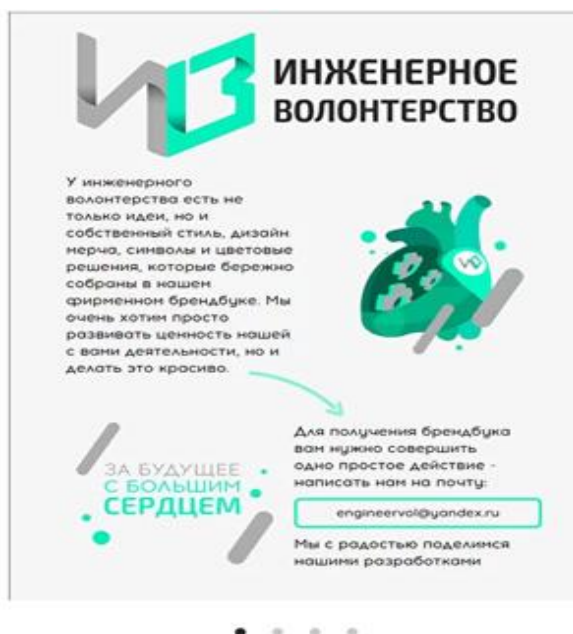
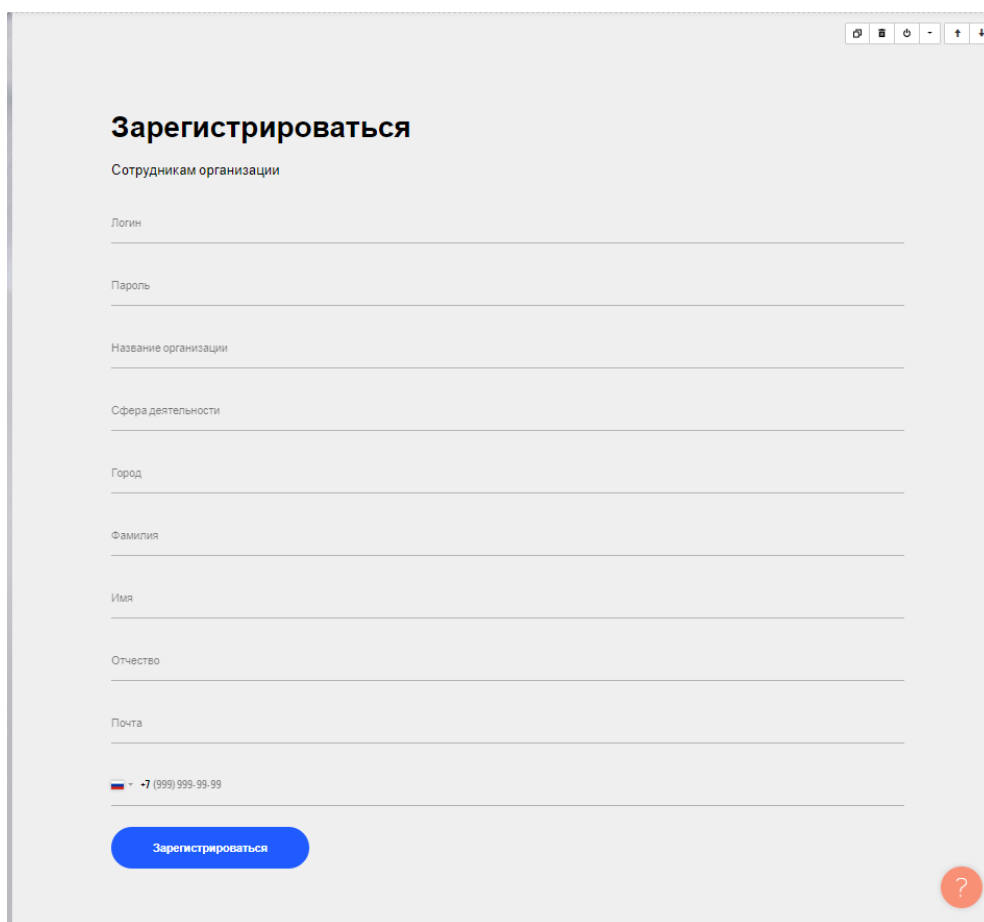


Рисунок 4 - О проекте.

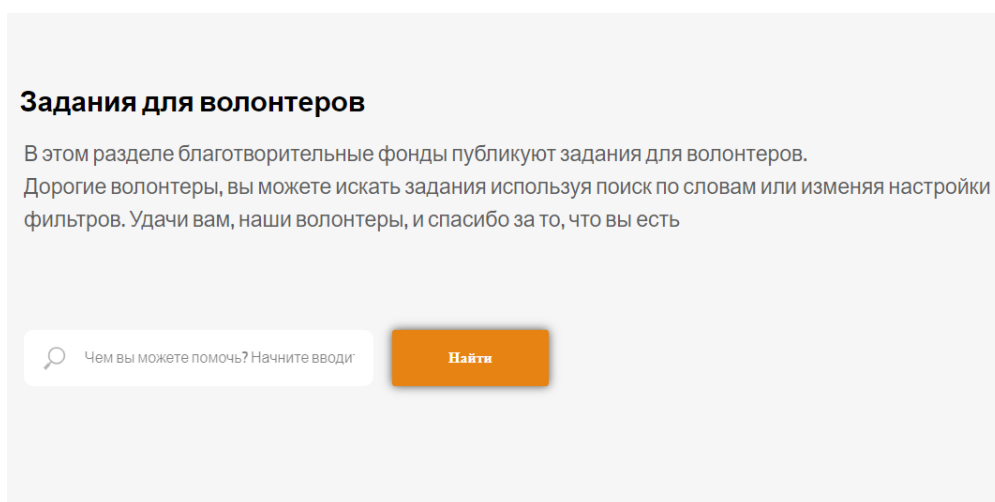
В следующем блоке “Мне нужна помощь” можно пройти регистрацию как сотрудник организации.



The screenshot shows a registration form titled "Зарегистрироваться" (Register) for "Сотрудникам организации" (Employees of the organization). The form contains several input fields: "Логин" (Login), "Пароль" (Password), "Название организации" (Organization name), "Сфера деятельности" (Area of activity), "Город" (City), "Фамилия" (Surname), "Имя" (Name), "Отчество" (Patronymic), and "Почта" (Email). Below the email field, there is a small flag icon and the text "+7 (999) 999-99-99". A blue button labeled "Зарегистрироваться" is positioned at the bottom left of the form, and a red question mark icon is at the bottom right.

Рисунок 5 - Регистрационный блок.

Блок “Я хочу помочь” предназначен для волонтеров, здесь располагаются задания.



The screenshot shows a section titled "Задания для волонтеров" (Tasks for volunteers). Below the title, there is a paragraph of text: "В этом разделе благотворительные фонды публикуют задания для волонтеров. Дорогие волонтеры, вы можете искать задания используя поиск по словам или изменяя настройки фильтров. Удачи вам, наши волонтеры, и спасибо за то, что вы есть". At the bottom, there is a search bar with a magnifying glass icon and the placeholder text "Чем вы можете помочь? Начните вводить". To the right of the search bar is an orange button labeled "Найти" (Find).

Рисунок 6 - Блок задания для волонтеров.

В блоке «Дополнительная информация» приведен информация о необходимых навыках.

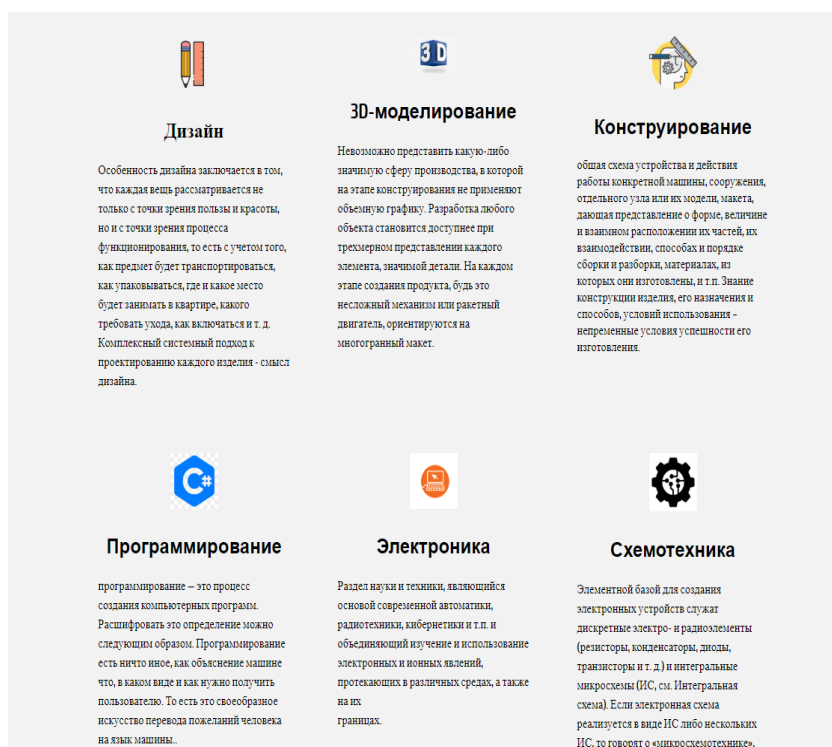


Рисунок 7 - Блок дополнительной информации.

Блок «Регистрация» служит для присоединения волонтеров.

The image shows a registration form titled "Зарегистрироваться" (Register) with the following fields and options:

- Электронная почта (Email)
- Пароль (Password)
- Фамилия (Family name)
- Имя (First name)
- Отчество (Surname)
- Дата рождения (Date of birth) with a calendar icon
- Выберите навыки (Select skills) with a dropdown arrow
- Мужчина (Male)
- Женщина (Female)
- Phone number field with a Russian flag icon and the number "+7 (999) 999-99-99"
- Отправить заявку (Send application) button
- Оформление заявки (Application form) text below the button

Рисунок 8 - Блок регистрации.

2.2. Программная реализация методов ординации многомерных данных на плоскость

Многомерный анализ является одним из разделов статистики. Он выполняет обработку и интерпретацию результатов, которые были получены на основе наблюдений одновременно нескольких взаимосвязанных случайных переменных, каждая из которых представляется одинаково важной, по крайней мере, первоначально. Если мы измеряем данные в метрических шкалах и предполагаем линейные отношения между переменными, то в анализе применяются многомерные статистические методы. Такие методы основаны на операциях матричной алгебры [13].

При решении некоторых задач обработки многомерных наблюдений исследователя интересует, прежде всего, можно ли снизить размерность, т.е. способ выделения небольшого набора исходных признаков или их линейных комбинаций, которые в наибольшей степени объясняют изменчивость наблюдаемых объектов. Это вызвано следующими факторами:

- а) возможностью наглядного представления (визуализация и ординация).
- б) стремлению к лаконизму исследуемых моделей.
- в) необходимостью сжатия объемов информации.

Принцип ординации наблюдений состоит в использовании различных методов оптимального целенаправленного проецирования (projecting pursuit) облака точек из многомерного пространства в пространство малой размерности (с 2 или 3 осями координат) [13].

Существуют различные методы или модификации принципа ординации. В целом сущность этого метода состоит в представлении исходной матрицы данных Y в виде совокупности p латентных переменных:

$$F: Y_1, Y_2, \dots, Y_m \Rightarrow F_1, F_2, \dots, F_p, \quad (1)$$

которые и являются осями ординационной диаграммы (двумерной при $p = 2$ или трехмерной в случае с тремя такими латентными переменными). При выборе осей ординации мы руководствуемся принципом оптимальности: а

именно мы стремимся достичь минимума потерь содержательной информации в исходных данных [13].

Чаще всего, первая ось F_1 совпадает с направлением наибольшей по длине оси эллипсоида рассеяния и проходит через центр сгущения значений y_{ij} . Вторая же ось F_2 проводится перпендикулярно к первой, имеет такое же направление, как и вторая из главных полуосей эллипсоида рассеяния и проходит через центр распределения. Такое решение позволяет формировать "картинку" данных с минимально возможными искажениями, а новые обобщающие переменные F_1 и F_2 , которые мы создали, становятся ортогональными (т.е. взаимно некоррелированными) [13].

Основная часть методов снижения размерности реализована на основе анализа собственных значений и собственных векторов. Данный анализ является важнейшим разделом линейной (матричной) алгебры. В общем суть преобразований $Y \Rightarrow F$ заключается в следующем:

1. Мы находим собственные значения и собственные векторы. Решения находим путем реализации операции линейного преобразования квадратной симметричной матрицы C размерностью m . В качестве матрицы C может использоваться любая матрица дистанций, однако в большинстве случаев используется именно дисперсионно-ковариационная матрица $C = \frac{Y^T Y}{(n-1)}$ стандартизованных исходных данных.

2. Собственные значения квадратной матрицы C – это такие значения λ_k , при которых система m уравнений вида $(C - \lambda_k * I) * u_k = 0$ имеет нетривиальное решение. Где u_k – собственные векторы матрицы C , соответствующие собственным значениям λ_k , $k = 1, 2, \dots, m$, I – единичная матрица.

3. Матрица из k собственных векторов U_k представляет собой веса для пересчета из исходного в редуцированное информационное пространство, т.е. матрицы переменных Y и F связаны соотношением

$$F_{n \times k} = Y_{n \times m} U_{m \times k} \quad (2)$$

4. Каждое собственное значение λ_k соответствует величине дисперсии, объясняемой на m -м уровне, т.е. сумма всех собственных значений будет равняться сумме дисперсий всех исходных переменных [13].

5. Ряд собственных значений чаще всего ранжируется от максимального до минимального: первое собственное значение λ_1 объясняет наибольшую долю вариации данных. Его часто называют "доминирующим" или "ведущим". Собственные значения λ_1 и λ_2 определяют меру значимости осей F_1 и F_2 ординационной диаграммы, т.е. концентрацию исходных точек вдоль каждой оси и, в итоге, ее выраженность.

6. Собственные значения λ_k находят методом итерации, и точность, с которой они будут вычислены, зависит от степени обусловленности исходной матрицы C .

Одним из методов, которым можно реализовать снижение размерности матрицы C , является метод главных компонент, который широко используется в различных областях науки и техники [13].

Используем в качестве примера анкетные данные соискателей вакансий.

id	Возраст	Стаж	Образование	Гражданство	Командировки	Место проживания	Семейное положение	Пол
1	27	3	1	1	1	2	2	0
2	19	4	3	2	1	3	2	1
3	19	4	3	3	0	1	1	1
4	30	2	1	4	1	1	3	0
5	30	3	1	1	1	1	2	1
6	27	2	1	2	1	2	1	0
7	18	2	3	3	0	1	2	1
8	18	5	3	4	1	1	1	1
9	29	4	1	4	1	4	1	1
10	18	2	3	3	1	4	3	1

Рисунок 9 - Анкетные данные.

Например, на вопрос "Ваше образование?" соискатель может выбрать из списка: "1 - Магистр, 2 - Бакалавр, 3 - Среднее". В данном случае мы имеем дело с порядковой шкалой. Она является монотонно убывающей. При помощи порядковой шкалы можно установить отношение больше/меньше между величинами, которые характеризуют данное свойство. Остальные вопросы и варианты ответов приведены в таблице 4.

Таблица 4. Таблица вопросов и ответов

Вопросы	Варианты ответов	Типы данных
Образование	1 - Магистратура 2 - Бакалавр 3 - Среднее	Порядковый
Гражданство	1. Гражданин РФ 2. Гражданин РФ и ИГ 3. Иност. гражданин 4. Лицо без гражданства	Номинальный
Командировки	0 - Нет, не готов 1 - Да, готов	Дихотомический
Место проживания	1. Томск 2. Томская область 3. РФ 4. Иностранное государство	Номинальный
Семейное - положение	1 – Холост/Незамужем 2– Разведен(а) 3 – Состою в браке	Номинальный
Пол	0 – Мужчина 1 - Женщина	Номинальный

Предварительно рассмотрим описательные статистики отдельных переменных:

```
DVol <- read.table(file = file.choose(), header = TRUE, sep = ",", row.names = 1)#
```

считывает таблицу данных и создаёт по ним data.frame

```
print(t(apply(DVol[1:2], 2, function(x) {c(Минимум = min(x), Максимум  
max(x), Среднее = mean(x), Отклонение = sd(x),
```

```
Корреляция = cor(x, DVol[3])) # признаков с Образование })), 3)
```


Таблица 5. Вывод нормированных данных на рабочем экране

id	Возраст	Стаж	Образование	Гражданство	Командировки	Место проживания	Семейное положение	Пол
1	0,800197445	-0,057813471	1	-1,33041347	1	-0,198655737	-0,196758509	0
2	-1,169519343	0,635948184	3	-0,443471157	1	0,718216896	-0,196758509	1
3	-1,169519343	0,635948184	3	0,443471157	0	-1,115528371	-1,377309561	1
4	1,538841241	-0,751575127	1	1,33041347	1	-1,115528371	0,983792544	0
5	1,538841241	-0,057813471	1	-1,33041347	1	-1,115528371	-0,196758509	1
6	0,800197445	-0,751575127	1	-0,443471157	1	-0,198655737	-1,377309561	0
7	-1,415733942	-0,751575127	3	0,443471157	0	-1,115528371	-0,196758509	1
8	-1,415733942	1,32970984	3	1,33041347	1	-1,115528371	-1,377309561	1
9	1,292626643	0,635948184	1	1,33041347	1	1,63508953	-1,377309561	1
10	-1,415733942	-0,751575127	3	0,443471157	1	1,63508953	0,983792544	1

Для данных “возраст”, “СтажВол” можно вычислить значения минимума, максимума, среднее, отклонение и корреляцию между возрастом и образованием и корреляция между стаж волонтером и образованием.

Таблица 6. Таблица статистики отдельных переменных.

Название признаков	Минимум	Максимум	Среднее	Отклонение	Корреляция
Возраст	-1,61	1,52	6.67e-11	1	0.954
СтажВол	-1,45	1,33	1.33e-10	1	-0.212

В среде R расчет главных компонент может быть осуществлен с использованием функций `princomp()` или `prcomp()`, но мы будем ориентироваться на многообразие возможностей пакета `vegan()` и его функцию `rda()`. Протокол анализа ограничим четырьмя главными компонентами, объясняющими 68% общей вариации данных:

```
library(vegan)
library(ggplot2)
Y <- as.data.frame(DVol[,-3])
mod.pca <- vegan::rda(Y ~ 1)
head(summary(mod.pca))
FAC <- as.factor(ifelse(DVol$Образование == 1, 1, 2))
pca.scores <- as.data.frame(summary(mod.pca)$sites[, 1:2])
```

```

pca.scores <- cbind(pca.scores, FAC)
# Составляем таблицу для "каркаса" точек на графике:
l <- lapply(unique(pca.scores$FAC), function(c)
{ f <- subset(pca.scores, FAC == c); f[chull(f), ]})
hull <- do.call(rbind, l)
# Включаем в названия осей доли объясненной дисперсии:
axX <- paste("PC1 (",
as.integer(100*mod.pca$CA$eig[1]/sum(mod.pca$CA$eig)), "%)")
axY <- paste("PC2 (",
as.integer(100*mod.pca$CA$eig[2]/sum(mod.pca$CA$eig)), "%)")
# Выводим ординационную диаграмму:
ggplot() +
  geom_polygon(data = hull,
              aes(x = PC1, y = PC2, fill = FAC), alpha = 0.4, linetype = 0) +
  geom_point(data = pca.scores,
            aes(x = PC1, y = PC2, shape = FAC, colour = FAC), size = 3) +
  scale_colour_manual(values = c('purple', 'blue')) +
  xlab(axX) + ylab(axY) + coord_equal() + theme_bw()

```

Таблица 7. Значения компонентов:

	PC1	PC2	PC3	PC4	PC5	PC6	PC7
Соб. значение	1.3466	1.1536	1.0112	0.7944	0.7545	0.24392	0.19921
Объясненная дисперсия	0.2444	0.2097	0.1838	0.1444	0.1371	0.04434	0.03621
Суммарная доля дисперсии	0.2444	0.4541	0.6379	0.7823	0.9195	0.96379	1.00000

Рассмотрим, какую конфигурацию имеет распределение наблюдений в пространстве первых двух главных компонент PC1 и PC2. Наибольший интерес для нас представляет образования, всех магистров относим к одной группе, всех остальных объединяем во вторую группу. Выполним построение ординационной диаграммы с использованием пакета ggplot2.

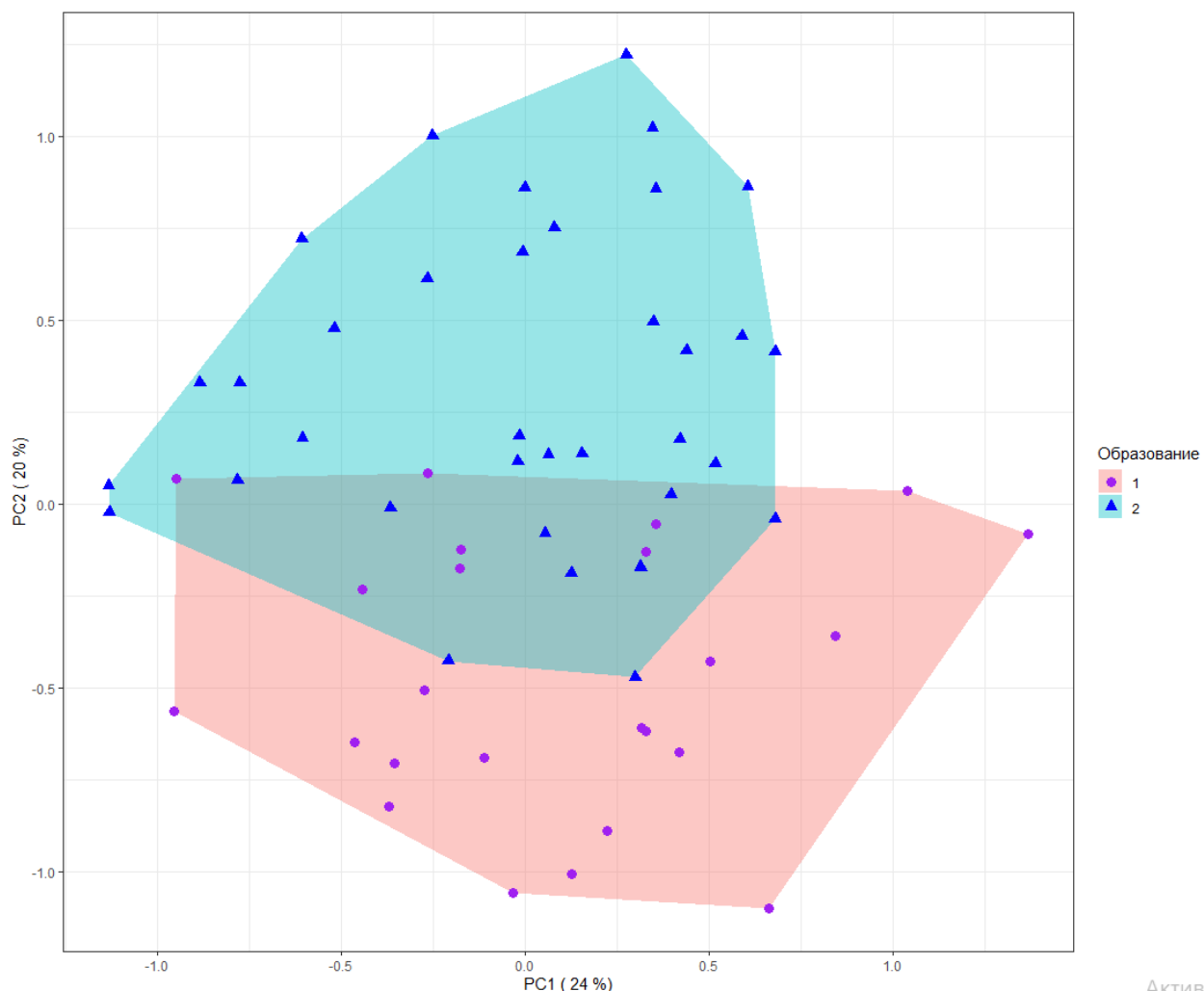


Рисунок 10 - Ординационная диаграмма на плоскости.

Заметим, что для выделения наблюдений каждого компонента на диаграмме обычно применяется вспомогательный графический объект “каркас” (hull) или контур, проводимый через крайние точки который состоит из двух факторов.

1) Маркер «Круг» — это магистранты, которые встречаются 24 раз, из них 8 маркеров образуют каркас. 2) Маркер «Треугольник» — это бакалавры

и школьники, которые встречаются 36 раз, из них 7 маркеров образуют каркас как видно на таблице 8.

Таблица 8. Факторные значения двух главных компонент.

	PC1	PC2	Образование
31	0.66364339	-1.10017959	1
4	-0.03259526	-1.05600926	1
12	-0.95476470	-0.56172184	1
54	-0.94954474	0.06859279	1
33	-0.26392002	0.08336578	1
22	1.04043057	0.03610504	1
40	1.36729372	-0.08205508	1
25	0.68042320	-0.04008515	2
32	0.29963614	-0.46984607	2
43	-0.20813645	-0.42668791	2
44	-1.13025591	-0.02331149	2
16	-1.13177085	0.05089217	2
29	-0.60721679	0.72108673	2
53	-0.25141423	1.00186874	2
49	0.27552481	1.22156527	2
35	0.60539695	0.86464524	2
59	0.68017331	0.41446991	2

По итогам расчетов можно отметить следующее:

- первая главная компонента объясняет 24% совокупной дисперсии в данных.
- вторая главная компонента объясняет 21% совокупной дисперсии в данных.
- третья главная компонента объясняет 18 % совокупной дисперсии в данных.
- четвертая главная компонента объясняет 14 % совокупной дисперсии в данных.

Совместно – 78%, что является хорошим результатом.

Как видно из таблицы собственных векторов можно выделить четыре компоненты.

Таблица 9. Собственные векторы.

	PC1	PC2	PC3	PC4	PC5	PC6
Возраст	0.0008938	-1.39490	-0.68803	0.660011	0.64404	-0.07109
Стаж	-1.1119558	-0.03842	1.12643	0.003148	0.86364	0.02742
Гражданство	0.8116309	-0.63887	1.21399	0.624915	-0.57855	-0.08280
Командировки	0.3646300	-0.15760	-0.11374	-0.017713	-0.13774	0.27129
Место. проживания	0.8674029	1.12784	-0.05199	0.874458	0.69334	-0.03948
Семейное. положение	-1.2738894	0.34868	-0.27846	0.985215	-0.68945	-0.03023
Пол	0.0064508	0.06312	-0.04878	-0.199010	-0.02148	-0.84265

- ось **PC1** в значительной мере определяет **семейное положение**.
- ось **PC2** определяет **возраст волонтера**.
- ось **PC3** определяет **гражданство волонтера**.
- ось **PC4** определяет **место проживания**.

Ординация показала, что облака точек всех компонентов являются разделяемыми.

2.3. Построение ординационных диаграммы различными методами

Для установления зависимости между двумя комплексами переменных используется канонический анализ. Говоря математическим языком, каноническая форма – это самая примитивная и универсальная форма, к которой приводятся определенные функции, отношения, или выражения без потери общности. Самой известной формой канонической ординации является анализ избыточности RDA (redundancy analysis) [13]. Он основан на приведении к собственным значениям и собственным векторам.

Анализ избыточности является диверсификацией метода множественной регрессии на моделировании данных с многомерным откликом. Метод RDA асимметричен, т.е., если $Y(n \times m)$ - таблица зависимых и $X(n \times q)$ - таблица объясняющих переменных, то $m \neq q$. Анализ избыточности выполняется следующим образом:

1. Оценка коэффициентов **В** модели множественной линейной регрессии **Y** на **X** и формирование матрицы прогнозируемых значений:

$$\mathbf{Y}^{\wedge} = \mathbf{X}\mathbf{B} = \mathbf{f}\mathbf{X}[\mathbf{X}'\mathbf{X}]^{-1} * \mathbf{X}'\mathbf{Y} \quad (3)$$

2. Далее вычисляется матрица канонических коэффициентов $\mathbf{S} = \mathbf{B}\mathbf{U}_c$ размерностью $(m * q)$, она рассчитывается на основе собственных векторов \mathbf{U}_c , взвешивающих влияние объясняющих переменных.
3. Рассчитывается ковариационная матрица $\mathbf{C}_{y|x}$ на основе значений \mathbf{Y}^{\wedge} и выполняется анализ главных: $(\mathbf{C}_{y|x} - \lambda_k \mathbf{I})\mathbf{u}_k = \mathbf{0}$, (4)

где

$\mathbf{C}_{y|x} = \mathbf{C}_{yX} * \mathbf{C}_{XX}^{-1} * \mathbf{C}_{yX}^T$ и обычным путем вычисляется матрица на основе k собственных векторов, не связанных с объясняющими переменными X [13].

Для примера, описанного в разделе 3, матрица \mathbf{X} сводится к вектору типа образование FAC, и анализ избыточности приводит к следующим результатам:

```
library(vegan) DVol <- read.table(file = file.choose(), header = TRUE,
sep = ",", row.names = 1) #считывает таблицу данных и создаёт по
ним data.frame.
```

Таблица 10. Вывод данных на рабочем экране

id	Возраст	Стаж	Образование	Гражданство	Командировки	Место проживания	Семейное положение	Пол
1	0,800197445	-0,057813471	1	-1,33041347	1	-0,198655737	-0,196758509	0
2	-1,169519343	0,635948184	3	-0,443471157	1	0,718216896	-0,196758509	1
3	-1,169519343	0,635948184	3	0,443471157	0	-1,115528371	-1,377309561	1
4	1,538841241	-0,751575127	1	1,33041347	1	-1,115528371	0,983792544	0
5	1,538841241	-0,057813471	1	-1,33041347	1	-1,115528371	-0,196758509	1
6	0,800197445	-0,751575127	1	-0,443471157	1	-0,198655737	-1,377309561	0
7	-1,415733942	-0,751575127	3	0,443471157	0	-1,115528371	-0,196758509	1
8	-1,415733942	1,32970984	3	1,33041347	1	-1,115528371	-1,377309561	1
9	1,292626643	0,635948184	1	1,33041347	1	1,63508953	-1,377309561	1
10	-1,415733942	-0,751575127	3	0,443471157	1	1,63508953	0,983792544	1

Далее идет выполнение проверки.

```
FAC <- as.factor(ifelse(DVol$Образование == 1, 1, 2))
```

```
Y <- as.data.frame(DVol[, -3])
```

```
mod.rda <- vegan::rda(Y ~ FAC)
```

```
head(summary(mod.rda))
```

Вызов:

```
rda (formula = Y ~ FAC)
```

Распределение дисперсии по компонентам:

Таблица 11. Значения компонентов

	RDA1	PC1	PC2	PC3	PC4	PC5	PC6	PC7
Соб. значение	0.8356	1.3403	1.0317	0.8887	0.7722	0.28562	0.20065	0.14666
Объясненная дисперсия	0.1519	0.2436	0.1875	0.1615	0.1404	0.05192	0.03647	0.02666
Суммарная доля дисперсии	0.1519	0.3955	0.5831	0.7446	0.8850	0.93687	0.97334	1.00

После включения в анализ объясняющего фактора “образование” появляется связанная с ним ось новой переменной **RDA1**, а к матрице на главные компоненты добавляется столбец соответствующих канонических коэффициентов. Собственное значение, определяющее RDA1, составляет только 15% от общей вариации данных.

Сформируем ординационную диаграмму в координатах **{RDA1, PC1}**, на которой обозначим координаты центроидов и выделим 95% доверительные области в виде эллипсов.

```
library(ggplot2)
```

```
rda.scores <- as.data.frame(vegan::scores(mod.rda, display = "sites", scales = 3))
```

```
rda.scores <- cbind(FAC, rda.scores) centroids <- aggregate(cbind(RDA1, PC1) ~
```

```
FAC, rda.scores, mean)
```

```
# функция для std.errf <- function(z)sd(z)/sqrt(length(z))
```

```
se <- aggregate(cbind(RDA1, PC1) ~ FAC,
```

```
rda.scores, f) names(se) <- c("FAC", "RDA1.se", "PC1.se")
```

```
# объединяем координаты центроидов и стандартные ошибки centroids <-  
merge (centroids, se, by = "FAC")
```

```

# формируем диаграмму ggplot() + geom_point(data = rda.scores, aes(x =
RDA1, y = PC1, shape = FAC, colour = FAC), size = 2) + xlim(c(-4, 4.5)) +
ylim(c(-2.5, 2)) + scale_colour_manual(values = c('purple', 'blue')) +
stat_ellipse(data = rda.scores,
aes(x = RDA1, y = PC1, fill = FAC), geom = "polygon", level = 0.8, alpha = 0.2)
+ geom_point(data = centroids,
aes(x = RDA1, y = PC1, colour = FAC), shape = 1, size = 4) + theme_bw()

```

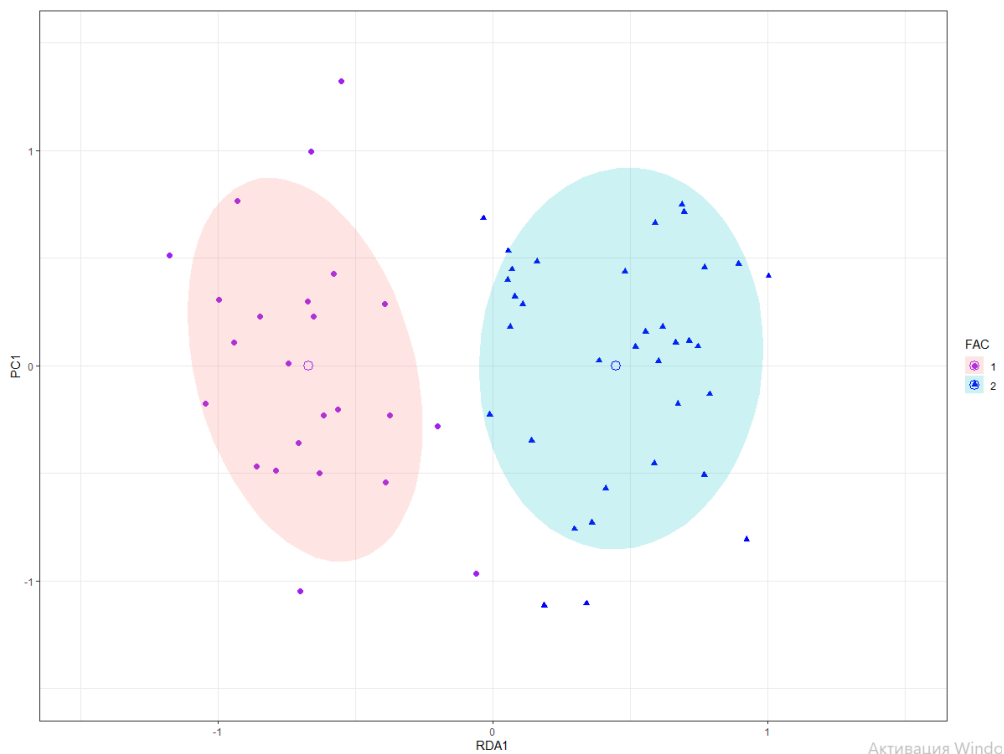


Рисунок 11 - Ординационная диаграмма, построенная методом RDA и PCA.

Несмотря на небольшую долю дисперсии, объясняемой фактором способа образования (F), координаты центроидов, определяющих центры тяжести сравниваемых групп, стали располагаться на некотором расстоянии друг от друга что позволяет мысленно провести разделяющую линию между ними. На данном графике существует 60 точек, которые разделены на два фактора.

1. Маркер «Круг» – встречается 18 раз, на интервале PC1[-2.5, 2] и RDA1[-4:4.5].
2. Маркер «Треугольник» – встречается 10 раз.

Таблица 13. Факторные значения двух переменных.

	FAC	RDA1	PC1
1	1	-0.37543298	-0.22816847
2	2	0.74533378	0.09005871
3	2	0.51815981	0.08793838
4	1	-1.04531194	-0.17460705
5	1	-0.85969660	-0.46720304
6	1	-0.58136555	0.42700979
7	2	0.71374944	0.11595110
8	2	0.55433561	0.15686444
9	1	-0.93013791	0.76744726
10	2	0.89364579	0.47264480

По итогам расчетов можно отметить следующее:

- первая каноническая переменная объясняет 15,1% совокупной дисперсии в данных.
- вторая неканоническая переменная объясняет 24,3% совокупной дисперсии в данных.
- третья неканоническая переменная объясняет 18,7% совокупной дисперсии в данных.
- четвертая неканоническая переменная объясняет 16,1% совокупной дисперсии в данных.

Таблица 12. Собственные векторы

	RDA1	PC1	PC2	PC3	PC4	PC5
Возраст	-1.60506	-0.23137	0.149072	-0.17611	-0.04090	0.49776
Стаж	0.08614	-1.09442	-0.908898	0.83543	0.72203	0.04976
Гражданство	-0.27533	0.74851	-1.363655	0.24159	-0.84210	-0.01972
Командировки	-0.11175	0.34094	0.007204	-0.21232	-0.11677	0.19420
Место проживания	0.12523	0.97331	0.662616	1.36428	0.06762	0.03881
Семейное положение	0.21988	-1.23039	0.477677	0.45632	-1.12432	-0.07453
Пол	-0.02483	0.01413	0.059233	-0.09144	0.11990	-0.79997

- ось **RDA** в значительной мере определяет **возраст**.
- ось **PC1** определяет **семейное положение**.
- ось **PC2** определяет **гражданство волонтера**.
- ось **PC3** определяет **место проживания**.

3. Финансовый менеджмент, ресурсоэффективность и ресурсосбережение.

Цель магистерской диссертации заключается в проецировании многомерных анализируемых объектов на координатную плоскость для соискателей вакантных позиций.

Объектом исследования является сравнительный анализ математических методов ординации многомерных объектов и признаков.

Целью данного раздела работы является экономическое обоснование магистерской диссертации. Для достижения поставленной цели необходимо решить следующие задачи:

- 1) Оценить конкурентоспособность разработки на основе методологии SWOT-анализа;
- 2) Провести организацию и планирование комплекса работ на создание проекта;
- 3) Провести стоимостную оценку разработки;
- 4) Оценить уровень научной новизны.

3.1. Потенциальные потребители результатов исследования.

Для анализа потребителей результатов исследования необходимо рассмотреть целевой рынок и провести его сегментирование.

Целевой рынок – сегменты рынка, на котором будет продаваться в будущем разработка. В свою очередь, сегмент рынка – это особым образом выделенная часть рынка, группы потребителей, обладающих определенными общими признаками.

Потенциальными потребителями результатов исследования будут являться волонтеры и сотрудники организации.

3.2. Анализ конкурентных решений.

Существуют различные языки программирования для реализации

методов ординации многомерных данных на плоскость. В данной работе мы используем язык R (язык программирования для статистической обработки данных и работы с графикой, а также свободная программная среда для вычислений с открытым исходным кодом в рамках проекта GNU). Также существует язык Python — это высокоуровневый язык программирования общего назначения, ориентированный на повышение производительности разработчика. Каждая из них имеет свои особые признаки, анализ которых представлен Таблице 1. Для оценочной карты были выбраны следующие критерии:

- Анализ данных
- Визуализация данных
- Синтаксис

Таблица 14. Оценочная карта для сравнения конкурентных решений

Критерии оценки	Вес критерия	Баллы		Конкурентоспособность	
		K1	Ф	K1	Ф
Экономические критерии оценки эффективности					
1. Анализ данных	0,5	2	4	1	2
2. Визуализация данных	0,3	3	5	0,9	1,5
3. Синтаксис	0,2	5	5	1	1
Итого	1	10	14	2,9	4,5

Позиция разработки оценивается по каждому показателю экспертным путем по пятибалльной шкале, где 1 – наиболее слабая позиция, а 5 – наиболее сильная. Анализ конкурентных решений определяется по формуле:

$$K = \sum B_i \cdot \Phi_i \quad (5)$$

где K – конкурентоспособность решения или конкурента, B_i – вес показателя (в долях единицы), Φ_i – балл i -го показателя.

Таким образом, можно сделать вывод, что язык R по многим показателям

является более предпочтительной, чем Python язык программирования (значение 5 является максимальным).

3.3. SWOT-анализ.

SWOT-анализ представляет собой сводную таблицу, иллюстрирующую связь между внутренними и внешними факторами компании. Целью SWOT-анализа является предоставление возможности оценки риска и конкурентоспособности компании или товара в данной отрасли производства.

Методика SWOT-анализа необходима для того, чтобы определить наиболее прозрачное на положение компании, продукции или услуги в данной отрасли.

Приведем матрицу SWOT - анализа для языка R.

Таблица 15. Матрица SWOT-анализа

	Сильные стороны	Слабые стороны
	С1. Экосистема С2. лингва-франка науки о данных	Сл1. Медленный язык Сл2. Сложен в изучении
Возможности вычисления и графического отображения В2. набор программных средств для манипулирования данными	В1С2 эффективная обработка и хранение данных; В1С1 набор операторов для вычислений на массивах, особенно матрицах	В1Сл2 потеря качества при увеличении: картинка рассыпается в цветные квадратики – пиксели; невозможно передать плавные цветовые переходы, как в растре.

Угрозы	У2С1С2	У1Сл1
У1. Требование	Это независимый от платформы язык, который может выполнять программы в Windows, Linux и Mac	R требует, чтобы его объекты были сохранены в физической памяти.
У2. Пакеты в R	Имеет более 10 000 пакетов в своем репозитории CRAN	медленнее, чем другие конкурирующие языки

3.4. Инициация проекта.

Группа процессов инициации состоит из процессов, которые выполняются для определения нового проекта или новой фазы существующего. В рамках процессов инициации определяются изначальные цели и содержание и фиксируются изначальные финансовые ресурсы. Определяются внутренние и внешние заинтересованные стороны проекта, которые будут взаимодействовать и влиять на общий результат.

Данная информация закрепляется в Уставе проекта.

Устав проекта документирует бизнес-потребности, текущее понимание потребностей заказчика проекта, а также новый продукт, услугу или результат, который планируется создать.

Устав проекта данной ВКР имеет следующую структуру:

- цели и результат проекта;
- организационная структура проекта;
- ограничения и допущения проекта.

3.5. Цели и результаты проекта.

Под заинтересованными сторонами проекта понимаются лица или

организации, которые активно участвуют в проекте или интересы которых могут быть затронуты как положительно, так и отрицательно в ходе исполнения или в результате завершения проекта. Информацию по заинтересованным сторонам приведём в таблице 3.

Таблица 16. Заинтересованные стороны проекта

Заинтересованные стороны проекта	Ожидания заинтересованных сторон
<ul style="list-style-type: none"> • Волонтеры; • Сотрудники 	Программа для бескорыстной помощи от профессионалов образовательной и социальной среды-уже не тренд, а успешная практика.

Под целями и результатом проекта необходимо приведем информацию о заинтересованных сторонах проекта, иерархии целей проекта и критериях достижения целей согласно таблице.

Таблица 17. Цели и результаты проекта

Цели проекта:	Реализация методы ординации многомерных данных на плоскость и сравнение ординационных диаграмм различными методами (PCA, RDA) с использованием метрик.
Ожидаемые результаты проекта:	<ul style="list-style-type: none"> • Веб-инструмент(сайт) для сбора исходных данных. • Реализация ординации многомерных данных на плоскость с помощью языков R/Phyton. • Проектирование база данных и СУБД с визуальным интерфейсом.
Критерии приемки результата проекта	
Требования к	Требование:

результату проекта:	<ul style="list-style-type: none"> • Подключение домена к веб-инструменту(сайт) • Тестирование разработанных методов и проинтерпретировать результаты.
---------------------	--

3.6. Организационная структура проекта.

В части «организационная структура проекта» необходимо принять решение о количестве человек, которые будут входить в рабочую группу, определить роль каждого участника в данном проекте, а также прописать функции, выполняемые каждым из участников и их трудозатраты в проекте. Эта информация представлена в таблице 5.

В ходе реализации научного проекта, помимо магистранта, задействован ряд специалистов:

Руководитель проекта – отвечает за реализацию проекта в пределах заданных ограничений по ресурсам, координирует деятельность участников проекта. В большинстве случаев эту роль выполняет руководитель магистерской диссертации.

Исполнитель по проекту – специалист, выполняющий отдельные работы по проекту. В случае, если магистерская работа является законченным научным исследованием – исполнителем проекта является магистрант. В случае, если магистерская работа является частью научного проекта, исполнителей может быть несколько.

Таблица 18. Рабочая группа проекта

№п/п	ФИО, основное место работы, должность	Роль в проекте	Функции	Трудозатраты, час.
------	---------------------------------------	----------------	---------	--------------------

1	Семенов Михаил Евгеньевич, ТПУ, доцент	Руководитель	Составление и утверждение научного задания, оценка эффективности полученных результатов	42
2	Абдуллоев Рустам Махмадсобиро вич, ТПУ, инженер	Исполнитель	Выполнение поставленных задач, составление пояснительной записки к ВКР	474
ИТОГО:				516

3.7. Ограничения и допущения проекта.

Ограничения проекта – это все факторы, которые могут послужить ограничением степени свободы участников команды проекта, а также «границы проекта» - параметры проекта или его продукта, которые не будут реализованных в рамках данного проекта.

Таблица 19. Ограничения проекта.

Фактор	Ограничения/ допущения
3.1. Бюджет проекта	180836,29 руб.
3.1.1.Источник финансирования	НИ ТПУ
3.2. Сроки проекта:	С 29.01.2019 до 27.05.2019
3.2.1. Дата утверждения плана управления проектом	06.02.2019
3.2.2. Дата завершения проекта	27.05.2019
3.3. Прочие ограничения и допущения	отсутствуют

3.8. Планирование научно-исследовательской работы.

3.8.1. Организация и планирование работы.

При организации процесса реализации данного исследования необходимо рационально планировать занятость каждого из его участников и сроки проведения отдельных работ.

В данном пункте составляется полный перечень проводимых работ, определяемых их исполнители и рациональная продолжительность. Так как число исполнителей редко превышает двух в большинстве случаев, то для наглядного результата чаще пользуются линейным графиком. Для построения такого графика приведем в таблице - 6 перечень работ и занятость исполнителей.

Таблица – 20. Перечень работ и продолжительность их выполнения.

№ Этапа	Этапы работы	Исполнители	Загрузка исполнителей
1	Постановка целей и задач, получение исходных данных	Научный руководитель	НР – 100%
2	Составление и утверждение ТЗ	Научный руководитель, студент	НР – 100% С – 10%
3	Подбор и изучение материалов по тематике	Научный руководитель, студент	НР – 50% С – 100%
4	Разработка календарного плана	Научный руководитель, студент	НР – 100% С – 10%
5	Обсуждение литературы	Научный руководитель, студент	НР – 30% С – 100%
6	Выбор СУБД и проектирование БД	студент	С – 100%
7	Разработка веб – сайта	студент	С – 100%
8	Оформление расчетно-пояснительной записки	студент	С – 100%
9	Оформление графического материала	студент	С – 100%

10	Анализ полученных результатов	Научный руководитель, студент	НР – 60% С – 100%
----	-------------------------------	-------------------------------	----------------------

3.8.2. Продолжительность этапов работ.

Трудовые затраты в большинстве случаев образуют основную часть стоимости разработки, поэтому важным моментом является определение трудоемкости работ каждого из участников научного исследования.

Трудоемкость выполнения проекта оценивается экспертным путем в человеко-днях и носит вероятностный характер, т.к. зависит от множества трудно учитываемых факторов. Для определения ожидаемого (среднего) значения трудоемкости $t_{ож}$ используется следующая формул:

$$t_{ож} = \frac{3*t_{min} + 2*t_{max}}{5} \quad (6)$$

где $t_{ож}$ – ожидаемая трудоемкость выполнения i -ой работы чел.-дн.;

t_{min} – минимальная продолжительность работы, дн.;

t_{max} – максимальная продолжительность работы, дн.

Для выполнения перечисленных в таблице 6 работ, требуется группа специалистов из следующего состава:

- Студент (С), соискатель степени магистра;
- Научный руководитель (НР).

Исходя из ожидаемой трудоемкости работ, определяется продолжительность каждой работы в рабочих днях $T_{рд}$, учитывающая параллельность выполнения работ несколькими исполнителями. Так, для построения линейного графика необходимо рассчитать длительность этапов в рабочих днях, а затем перевести ее в календарные дни. Расчет продолжительности выполнения каждого этапа в рабочих днях ($T_{рд}$ ведется по формуле:

$$T_{рд} = \frac{t_{ож}}{K_{ВН}} * K_{Д} \quad (7)$$

где $t_{ож}$ – продолжительность работы, дн.;

$K_{ВН}$ – коэффициент выполнения работ, учитывающий влияние внешних факторов на соблюдение предварительно определенных длительностей, в частности, возможно $K_{ВН} = 1$;

$K_{Д}$ – коэффициент, учитывающий дополнительное время на компенсацию непредвиденных задержек и согласование работ ($K_{Д} = 1-1,2$; в этих границах конкретное значение принимает сам исполнитель). Возьмем значение $K_{Д} = 1$. Продолжительность этапов работ и их трудоемкости по исполнителям, занятым на каждом этапе представлена в таблице – 7.

Таблица – 21. Временные показатели проведения научного исследования.

№ Этапа	Исполнители	Продолжительность работ, дни			Трудоемкость работ, дни			
					$T_{рд}$		$T_{кд}$	
		t_{min}	t_{max}	$t_{ож}$	НР	С	НР	С
1	Научный руководитель	1	2	1,6	1,6	-	1,92	-
2	Научный руководитель, студент	5	10	7	7	0,7	8,4	0,84
3	Научный руководитель, студент	10	15	12	6	12	7,2	14,4
4	Научный руководитель, студент	5	10	7	7	0,7	8,4	0,84
5	Научный руководитель, студент	1	2	1,6	1,6	0,48	1,92	0,58
6	Студент	15	20	17	-	17	-	20,4
7	Студент	3	5	3,8	-	3,8	-	4,56
8	Студент	10	20	14	-	14	-	16,8
9	Студент	1	2	1,6	-	1,6	-	1,92
10	Научный руководитель, студент	5	10	7	4,2	7	5,04	8,4
Итого:		46	96	72,6	27,4	57,28	32,88	68,74

Разработка графика проведения научного исследования

Выполнение ВКР является небольшим по объему исследованием, поэтому наиболее удобным и наглядным является построение ленточного графика проведения научных работ в форме диаграммы Гантта.

Так, построим ленточный график. Для удобства построения графика, длительность каждого из этапов работ из рабочих дней следует перевести в календарные дни. Для этого необходимо воспользоваться следующей формулой:

$$T_{\text{КД}} = T_{\text{РД}} * T_{\text{К}} \quad (8)$$

где $T_{\text{КД}}$ – продолжительность выполнения этапа в календарных днях;

$T_{\text{К}}$ – коэффициент календарности, позволяющий перейти от длительности работ в рабочих днях к их аналогам в календарных днях, который определяется по следующей формуле:

$$T_{\text{К}} = \frac{T_{\text{КАЛ}}}{T_{\text{КАЛ}} - T_{\text{ВД}} - T_{\text{ПД}}} \quad (9)$$

где $T_{\text{КАЛ}}$ – календарные дни ($T_{\text{КАЛ}} = 365$);

$T_{\text{ВД}}$ – выходные дни ($T_{\text{ВД}} = 52$ для при шестидневной рабочей неделе);

$T_{\text{ПД}}$ – праздничные дни ($T_{\text{ПД}} = 10$).

$$T_{\text{К}} = \frac{365}{365 - 52 - 10} = 1,20+$$

Таким образом, коэффициент календарности $T_{\text{К}}$ равен 1,20.



Величины трудоемкости этапов по исполнителям $T_{\text{КД}}$ (данные столбцов 8 и 9 кроме итогов) позволяют построить линейный график осуществления проекта.

Пример построения линейного графика приведен в таблице – 22.

Таблица 22. Линейная графика.

Этап	Вид работ	НР	С	Продолжительность выполнения работ										
				март			апрель			май				
				10	20	30	10	20	30	10	20	30		

1	Постановка целей и задач, получение исходных данных	1,92	-										
2	Составление и утверждение ТЗ	8,4	0,84										
3	Подбор и изучение материалов по тематике	7,2	14,4										
4	Разработка календарного плана	8,4	0,84										
5	Обсуждение литературы	1,92	0,58										
6	Выбор СУБД и проектирование БД	-	20,4										
7	Разработка веб – сайта	-	4,56										
8	Оформление расчетно- пояснительной записки	-	16,8										
9	Оформление графического материала	-	1,92										
10	Анализ полученных результатов	5,04	8,4										

НР –  ; С – 

3.9. Расчет сметы затрат на выполнение проекта.

В состав затрат на создание проекта включается величина всех расходов, необходимых для реализации комплекса работ, составляющих содержание данной разработки. Расчет сметной стоимости ее выполнения производится по следующим статьям затрат:

- Материалы и покупные изделия;
- Заработная плата;
- Социальный налог;
- Расходы на электроэнергию (без освещения);
- Амортизационные отчисления;
- Оплата услуг связи;
- Прочие (накладные расходы) расходы.

3.9.1. Расчет материальных затрат.

К данной статье расходов относится стоимость материалов, покупных изделий, расходуемых непосредственно в процессе выполнения работ над объектом исследования.

Покажем отражение стоимости всех материалов, используемых при работе над проектом, включая расходы на их приобретение и, при необходимости, доставку. Расчет затрат на материалы производится по форме, приведенной в таблице – 9.

Таблица – 23. Материальные затраты

Наименование	Единица	Кол-во	Цена за ед, руб.	Сумма, руб
Бумага	пачка	1	300	300
Канцелярские принадлежности	шт.	3	100	300
Картридж для принтера	шт.	1	3000	3000
Итого:				3600

Допустим, что ТЗР составляют 5 % от отпускной цены материалов, тогда расходы на материалы с учетом ТЗР равны:

$$C_{\text{мат}} = 3\,600 * 1,05 = 3\,780 \text{ руб}$$

3.9.2. Расчет заработной платы для исполнителей.

Данная статья расходов включает заработную плату научного руководителя и студента (в его роли выступает исполнитель проекта), а также премии, входящие в фонд заработной платы.

Расчет основной заработной платы выполняется на основе трудоемкости выполнения каждого этапа и величины месячного оклада исполнителя.

Среднедневная тарифная заработная плата ($ЗП_{\text{дн-т}}$) рассчитывается по формуле:

$$ЗП_{\text{дн-т}} = \frac{МО}{25,083} \quad (10)$$

Учитывающей, что в году 301 рабочий день и, следовательно, в месяце в среднем 25,083 рабочих дня (при шестидневной рабочей неделе).

Пример расчета затрат на полную заработную плату приведены в таблице 10. Затраты времени по каждому исполнителю в рабочих днях с округлением до целого взяты из таблицы – 7. Для учета в ее составе премий, дополнительной зарплаты и районной надбавки используется следующий ряд коэффициентов: $K_{\text{ПР}} = 1,1$; $K_{\text{доп.ЗП}} = 1,188$; $K_{\text{Р}} = 1,3$. Таким образом, для перехода от тарифной (базовой) суммы заработка исполнителя, связанной с участием в проекте, к соответствующему полному заработку (зарплатной части сметы) необходимо первую умножить на интегральный коэффициент $K_{\text{И}} = 1,1 * 1,188 * 1,3 = 1,699$. Вышеуказанное значение $K_{\text{доп.ЗП}}$ применяется при шестидневной рабочей неделе, при пятидневной оно равно 1,113, соответственно в этом случае $K_{\text{И}} = 1,62$.

Таблица – 24. Затраты на заработную плату.

Исполнитель	Оклад, руб./мес.	Среднедневная ставка, руб./раб. день	Затраты времени, раб. дни	Коэффициен	Фонд з/платы, руб.
НР	35120	1400,15	28	1,699	66 607,94
С	12000	478,41	58	1,62	44951,40
Итого:					111559,34

3.9.3. Расчет затрат на социальный налог.

Затраты на единый социальный налог (ЕСН), включающий в себя отчисления в пенсионный фонд, на социальное и медицинское страхование, составляют 30,2 % от полной заработной платы по проекту, т.е. $C_{соц} = Сзп * 0,302$.

Итак, в нашем случае:

$$C_{соц} = 111559,34 * 0,302 = 33690,92 \text{ руб.}$$

3.9.4. Расчет затрат на электроэнергию.

Данный вид расходов включает в себя затраты на электроэнергию, потраченную в ходе выполнения проекта на работу используемого оборудования, рассчитываемые по формуле:

$$C_{эл.об} = P_{об} * t_{об} * ЦЭ, \quad (11)$$

где $P_{об}$ – мощность, потребляемая оборудованием, кВт;

ЦЭ – тариф на 1 кВт·час;

$t_{об}$ – время работы оборудования, час.

Для ТПУ ЦЭ = 5,748 руб./квт·час (с НДС).

Время работы оборудования вычисляется на основе итоговых данных таблицы – 7 для студента (ТРД) из расчета, что продолжительность рабочего дня равна 8 часов.

$$t_{об} = T_{рд} * K_t, \quad (12)$$

где $K_t \leq 1$ – коэффициент использования оборудования по времени. Возьмем его равным 1.

Мощность, потребляемая оборудованием, определяется по формуле:

$$P_{об} = P_{ном} * K_C, \quad (13)$$

где $P_{ном}$ – номинальная мощность оборудования, кВт;

$K_C \leq 1$ – коэффициент загрузки, зависящий от средней степени использования номинальной мощности. Для технологического оборудования малой мощности $K_C = 1$.

Пример расчета затраты на электроэнергию для технологических целей приведен в таблице 11

Таблица – 25. Затраты на электроэнергию технологическую.

Наименование оборудования	Время работы оборудования $t_{об}$, час	Потребляемая мощность $P_{об}$, кВт	Затраты $\mathcal{E}_{об}$, руб.
Персональный компьютер	464	0,3	800,12
Струйный принтер	2	0,1	1,15
Итого:			801,27

3.9.5. Расчет амортизационных расходов.

В статье «Амортизационные отчисления» рассчитывается амортизация используемого оборудования за время выполнения проекта.

Используется формула:

$$C_{ам} = \frac{N_A * C_{об} * t_{рф} * n}{F_d}, \quad (14)$$

где N_A – годовая норма амортизации единицы оборудования;

$C_{об}$ – балансовая стоимость единицы оборудования с учетом ТЗР;

$t_{рф}$ – фактическое время работы оборудования в ходе выполнения проекта, учитывается исполнителем проекта;

n – число задействованных однотипных единиц оборудования.

Например, для ПК в 2020 г. (299 рабочих дней при шестидневной рабочей неделе) можно принять $FД = 299 * 8 = 2392$ часа.

Для принтера из справочника $FД = 500$ часов.

При использовании нескольких типов оборудования расчет по формуле делается соответствующее число раз, затем результаты суммируются.

Для ПК найдем $N_A = 0,4$. Для принтера $N_A = 0,5$.

Стоимость ПК= 25 000 рублей. Время использования 304 часа, тогда для него:

$$C_{AM(ПК)} = \frac{0,4 * 25\ 000 * 304 * 1}{2392} = 1270,90 \text{ руб.}$$

Стоимость принтера 6000 руб. Время использования 2 часа, тогда для него:

$$C_{AM(ПР)} = \frac{0,5 * 6\ 000 * 2}{1\ 500} = 12 \text{ руб.}$$

Итого начислено амортизации 1 282,9 руб.

3.9.6. Расчет прочих расходов.

В статье «Прочие расходы» отражены расходы на выполнение проекта, которые не учтены в предыдущих статьях, их следует принять равными 10% от суммы всех предыдущих расходов, т.е.

$$C_{\text{проч}} = (C_{\text{мат}} + C_{\text{зд}}) + C_{\text{соц}} + C_{\text{эл.об}} + C_{\text{ам}}) * 0,1$$

(15)

Для нашего примера это:

$$C_{\text{проч}} = (3\ 990 + 111\ 559,34 + 33\ 690,92 + 801,27 + 1270,9) * 0,1 = 15\ 131,24 \text{ руб.}$$

3.9.7. Расчет общей себестоимости разработки.

Проведя расчет по всем статьям сметы затрат на разработку, можно определить общую себестоимость проекта. Данные результаты можно посмотреть в таблице – 12

Таблица – 26. Смета затрат на разработку проекта.

Статья затрат	Условное обозначение	Сумма, руб.
Материалы и покупные изделия	$C_{\text{мат}}$	3 990
Основная заработная плата	$C_{\text{зп}}$	111 559,34
Отчисления в социальные фонды	$C_{\text{соц}}$	33 690,92
Расходы на электроэнергию	$C_{\text{эл.}}$	801,27
Амортизационные отчисления	$C_{\text{ам}}$	1270,90
Прочие расходы	$C_{\text{проч}}$	16439,67
Итого:		167 752,10

Таким образом, затраты на разработку составили $C = 167\,752,10$ руб.

3.9.8. Расчет прибыли.

Прибыль примем в размере 10 % от полной себестоимости проекта. В нашем примере она составляет 16 775,21 руб. (10 %) от расходов на разработку проекта.

3.9.9. Расчет НДС.

НДС составляет 20% от суммы затрат на разработку и прибыли. В нашем случае:

$\text{НДС} = (167\,752,10 + 16\,775,21) * 0,2 = 36\,905,46$ руб. Цена разработки НИР
Цена равна сумме полной себестоимости, прибыли и НДС:

$$C_{\text{НИР(КР)}} = 167\,752,10 + 16\,775,21 + 36\,905,46 = 221\,432,77 \text{руб.}$$

3.10. Оценка экономической эффективности проекта.

В ходе выполнения части работы по финансовому менеджменту, ресурсоэффективности и ресурсосбережению был проведен анализ разрабатываемого исследования.

Проведено планирование НИР, а именно: определена структура и календарный план работы, трудоемкость, составлена ленточная диаграмма Ганта, и определен бюджет научно-исследовательской работы. Результаты соответствуют требованиям ВКР по срокам и иным параметрам.

Эффективность данной исследовательской работы состоит в том, что нахождение вакантных соискателей облегчает работу сотрудников.

3.11. Вывод по разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение:

1. Оценочная карта сравнения технических решений конкурентов показала, что, язык R по многим показателям является более подходящий, чем язык программирования Python для статистических и графических вычислений

2. В ходе планирования научно-исследовательской работы определён перечень работ, выполняемый рабочей группой. В данном случае рабочая группа состоит из двух человек: научный руководитель и студент. Согласно составленному плану работ длительность трудовой занятости сотрудников исследовательского проекта составила 102 дней (69 дня – занятость студента, 33 дня – длительность работы научного руководителя). На основе временных показателей по каждой из произведенных работ был построен календарный план-график, построенный на основе диаграммы Ганта, по которому можно

увидеть, что самая продолжительная по времени работа – Выбор СУБД и проектирование БД.

3. Бюджет научно-технического исследования составил 167 752,10 руб. Бюджет НТИ состоит из материальных затрат (3 990 рублей), затрат на оплаты труда (111 559,34 рублей), отчислений в социальные фонды (33 690,92 рубль), расходов на электроэнергию (801,27 рубль), амортизационных отчислений (1270,90) и прочих расходов (16 439,67рубля).

4. Проведена оценка результатов ресурсоэффективности, результаты соответствуют требованиям ВКР по срокам и иным параметрам.

4. Социальная ответственность.

В настоящее время большое внимание уделяется вопросам безопасности трудящихся на рабочем месте. Одной из основных задач является охрана здоровья сотрудников предприятий, сведение к минимуму или ликвидация различных видов производственных травм и снижение риска заболеваний. Такие факторы рассматривает социальная ответственность, а именно: состояние рабочего места, помещения, режим работы, обеспечение мероприятий по защите трудящихся в случае возникновения чрезвычайной ситуации. Существует международный стандарт ИСО 26000 «Руководство по социальной ответственности», разработанный в 2012 г., в соответствии с которым осуществляется разработка системы социальной ответственности на предприятии. При применении организацией данного стандарта для разработки системы социальной ответственности, она решает ряд проблем, таких как: проблемы, связанные с потребителями, трудовые практики, участие в жизни сообществ и их развитие, права человека, окружающая среда, добросовестные деловые практики.

В современной жизни компьютер широко применяется в повседневную жизнь людей и их использование постоянно увеличивается. Несоблюдение требований безопасности персональных компьютеров приводит к тому, что

при работе за компьютером пользователь может ощущать дискомфорт: возникают головные боли и резь в глазах, появляются усталость и раздражительность.

В данном разделе рассмотрены вопросы, связанные с организацией рабочего места в соответствии с нормами производственной санитарии, техники производственной безопасности и охраны окружающей среды.

В данной работе рассмотрена проектировка рабочего места и помещения, в котором оно находится.

Под проектированием рабочего места понимается целесообразное пространственное размещение в горизонтальной и вертикальной плоскостях функционально взаимосвязанных средств производства (оборудования, оснастки, предметов труда и др.), необходимых для осуществления трудового процесса.

При проектировании рабочих мест должны быть учтены освещенность, температура, влажность, давление, шум, наличие вредных веществ, электромагнитных полей и другие санитарно-гигиенические требования к организации рабочих мест. Работа выполняется преимущественно за компьютером, поэтому в соответствии СанПиН 2.2.2/2.4.1340-03 в помещении на одного работника, работающего за ПК с ЖК мониторами предусмотрено 4,5 кв.м.

При проектировании рабочей зоны необходимо уделить внимание охране окружающей среды, а в частности, организации безотходного производства.

Также необходимо учитывать возможность чрезвычайных ситуаций. Так как рабочая зона находится в городе Томске, наиболее типичной ЧС является мороз. Так же, в связи с неспокойной ситуацией в мире, одной из возможных ЧС может быть диверсия.

В данной части магистерской работе рассмотрен анализ вредных и опасных факторов производства, методы их предупреждения, организация

мероприятий защиты в случае возникновения чрезвычайной ситуации. Для эффективного ведения разработки обозначенной системы необходимо учитывать социальные, правовые и экологические вопросы охраны здоровья и безопасности жизнедеятельности.

В рамках данного раздела необходимо:

- определить и изучить возможные вредные факторы, влияющие на исполнителей при разработке программного продукта;
- разработать решения для минимизации их влияния;
- произвести анализ разработанных решений, с точки зрения социальной ответственности за моральные, общественные, экономические и экологические возможные негативные последствия, и ущерб здоровью человека в результате их внедрения.

4.1. Правовые и организационные вопросы обеспечения безопасности.

Для обеспечения высокопроизводительного труда необходимо создать для работника благоприятные условия труда.

Условия труда – это совокупность факторов производственной среды и трудового процесса, оказывающих влияние на здоровье и работоспособность человека в процессе труда.

Регулирование отношений между работником и работодателем, касающихся оплаты труда, трудового распорядка, особенности регулирования труда женщин, детей, людей с ограниченными способностями и проч., осуществляется законодательством РФ, а именно трудовым кодексом РФ [19].

Нормальная продолжительность рабочего времени не может превышать 40 часов в неделю. Порядок исчисления нормы рабочего времени на определенные календарные периоды (месяц, квартал, год) в зависимости от установленной продолжительности рабочего времени в неделю определяется федеральным органом исполнительной власти, осуществляющим функции по

выработке государственной политики и нормативно-правовому регулированию в сфере труда.

В течение рабочего дня (смены) работнику должен быть предоставлен перерыв для отдыха и питания. Время предоставления перерыва и его конкретная продолжительность устанавливаются правилами внутреннего трудового распорядка или по соглашению между работником и работодателем.

Всем работникам предоставляются выходные дни (еженедельный непрерывный отдых).

Организация-работодатель выплачивает заработную плату работникам. Возможно удержание заработной платы только в случаях, установленных ТК РФ ст. 137. В случае задержки заработной платы более чем на 15 дней, работник имеет право приостановить работу, письменно уведомив работодателя.

Законодательством РФ запрещена дискриминация по любым признакам и принудительный труд.

Если пользователь постоянно загружен работой с ЭВМ, приемлемой является поза сидя. В положении сидя основная нагрузка падает на мышцы, поддерживающие позвоночный столб и голову. В связи с этим при длительном сидении время от времени необходимо сменять фиксированные рабочие позы.

Исходя из общих принципов организации рабочего места, в нормативно-методических документах сформулированы требования к конструкции рабочего места. Основными элементами рабочего места программиста являются: рабочий стол, рабочий стул (кресло), монитор, клавиатура, мышь; вспомогательными - пюпитр, подставка для ног. Взаимное расположение элементов рабочего места должно обеспечивать возможность осуществления всех необходимых движений и перемещений для эксплуатации и технического обслуживания оборудования. Рабочее место при выполнении работ в положении сидя должно соответствовать требованиям ГОСТ 12.2.032-78.

Рабочие места с ЭВМ должны располагаться на расстоянии не менее 1,5 метра от стены с оконными проемами, от других стен – на расстоянии 1 м, между собой – на расстоянии не менее 1,5 м. При размещении рабочих мест необходимо исключить возможность прямой засветки экрана источником естественного освещения.

Располагать ЭВМ на рабочем месте необходимо так, чтобы поверхность экрана находилась на расстоянии 400 – 700 мм от глаз пользователя. Конструкция рабочего места и взаимное расположение всех его элементов (сиденье, органы управления, средства отображения информации и т.д.) должны соответствовать антропометрическим, физиологическим и психологическим требованиям, а также характеру работы. Рабочее кресло обеспечивает поддержание рабочей позы в положении сидя, и чем длительнее это положение в течение рабочего дня, тем жестче должны быть требования к созданию удобных и правильных рабочих сидений. Высота поверхности сиденья должна регулироваться в пределах 400 – 550 мм. При организации рабочего пространства необходимо учитывать индивидуальные антропометрические параметры пользователя с соответствующими допусками на возможные изменения рабочих поз и потребность в перемещениях.

Рациональной рабочей позой может считаться такое расположение тела, при котором ступни работника расположены на плоскости пола или на подставке для ног, бедра сориентированы в горизонтальной плоскости, верхние части рук – вертикальный угол локтевого сустава колеблется в пределах 70 – 90, запястья согнуты под углом не более чем 20, наклон головы – в пределах 15 – 20, а также исключены частые ее повороты.

Продолжительность непрерывной работы с компьютером без регламентированного перерыва не должна превышать 2-х часов. Это связано с тем, что на человека при работе с компьютером оказывают влияние опасные и вредные производственные факторы, а также наступает общее утомление, что негативно сказывается на здоровье и самочувствии человека.

4.2. Производственная безопасность.

Были определены вредные факторы, влияющие на условия труда разработчика, а также произведен анализ источников вредных факторов. Полученные данные представлены в виде сводной таблицы 27.

Таблица 27 - Возможные опасные и вредные факторы

Факторы (ГОСТ 12.0.003-2015)	Этапы работ			Нормативные документы
	Разработка	Исследования	Работа за персональным компьютером	
1. Недостаточная освещенность рабочей зоны;		+	+	Параметры освещения определены в СНиП 23-05-95 [1]; Предельно допустимые уровни напряжений прикосновения и токов установлены в ГОСТ 12.1.038-82 [2]; Мероприятия по устранению статического электричества перечислены в ГОСТ 12.1.030-81 [3]; Правила защиты от поражения электрическим током приведены в ГОСТ Р 50571.3-94 [4]; СанПиН 2.2.2.542-96 Гигиенические требования к персональным электронно-вычислительным машинам и организации работ [5].
2. Повышенный уровень статического электричества;		+	+	
3. Повышенное значение напряжения в электрической цепи, замыкание которой может произойти через тело человека.	+	+	+	
4. Повышенная или пониженная температура воздуха рабочей зоны	+	+	+	

4.3. Анализ опасных и вредных производственных факторов и обоснование мероприятий по их устранению.

4.3.1. Недостаточная освещенность рабочей зоны.

Работа с компьютером подразумевает постоянный зрительный контакт с дисплеем ПЭВМ и занимает от 80 % рабочего времени. Недостаточность освещения снижает производительность труда, увеличивает утомляемость и

количество допускаемых ошибок, а также может привести к появлению профессиональных болезней зрения.

Достаточная освещенность рабочего места важна для эффективной работы. Она достигается использованием как естественного, так и искусственного света вместе.

Согласно СНиП 23-05-93 проводимые помещения работы по разработке программного продукта относятся к V разряду зрительной работы. Согласно пункту 4.2 СНиП 23-05-95 [5] уровень общей освещенности должен быть не меньше 200 лк. Величина нормируемой освещенности помещений составляет 300 лк, что является допустимым. Минимальный размер объекта различения 3-5 мм – работа малой точности. Нормы коэффициента пульсации для обозначенного вида работ составляют не более 20%. В помещениях присутствует как естественное, так и искусственное освещение. Последнее обеспечивается люминесцентными источниками света в потолочных светильниках и светильниками, установленными непосредственно на рабочем месте трудящегося.

4.3.2. Повышенный уровень статического электричества.

Статическое электричество является одним из опасных факторов. Оно возникает в результате накопления заряда на поверхности или объеме диэлектриков или на изолированных проводниках. Опасность возникает при разряде этого заряда, поскольку заряд может быть велик, что может вызвать поломку оборудования, угрозу для жизни оператора в результате прохождения заряда через тело человека. Накопление заряда может быть вызвано изоляцией оборудования или оператора.

Для устранения необходимо обеспечить заземление компьютера, чтобы заряд мог уходить на землю, не накапливаясь на корпусе компьютера. Правила заземления и другие мероприятия по устранению статического электричества перечислены в ГОСТ 12.1.030-81 [22].

4.3.3. Повышенное значение напряжения в электрической цепи, замыкание которой может произойти через тело человека.

Нормативное регулирование охраны труда при осуществлении трудовой деятельности за компьютерами осуществляется посредством следующих документов: Типовая инструкция СанПиН 2.2.2.; СанПиН 2.2.2. / 2.4. 1340-03 (далее – СанПиН); ТК РФ; Приказ Минздравсоцразвития РФ № 302н; 426-ФЗ.

Возможность поражения электрическим током на рабочем месте существует, т.к. ЭВМ работает от источника тока. При прикосновении к электроустановкам, по которым протекает ток возможно прохождение тока через тело человека, так как земля имеет нулевой потенциал. Сила поражения зависит от времени воздействия, характера тока, состояния человека, места соприкосновения, пути прохождения тока. Поражающими факторами являются ожоги, сбои в работе сердечной мышцы, нервной системы. Причиной прикосновения является как правило нарушение изоляции изделия и нарушение техники безопасности. Для питания персональных компьютеров в России как правило используется сетевое напряжение 230В переменного тока при частоте 50 Гц. После прохождения через блок питания ток трансформируется в постоянный ток напряжения 5 и 12 В. Согласно ГОСТ 12.1.038-82 допустимо прохождение через тело человека переменного тока с напряжением 2В, силой 0.3мА, постоянного тока 8В, силой 1мА.

Для того чтобы исключить опасность поражения электрическим током необходимо соблюдать правила электрической безопасности:

1) Перед включением ЭВМ в сеть должна быть визуально проверена ее электропроводка на отсутствие возможных видимых нарушений изоляции, а также на отсутствие замыкания токопроводящих частей на корпусе компьютера;

2) Запрещается при включенной ЭВМ одновременно прикасаться к приборам, имеющим естественное заземление;

- 3) При появлении признаков замыкания необходимо немедленно отключить от электрической сети ЭВМ и устранить неисправность;
- 4) Применения напряжений 42 В и ниже переменного тока частотой 50 Гц и 110 В и ниже постоянного тока;
- 5) Заземления или зануления корпусов электрооборудования и элементов электроустановок, которые могут оказаться под напряжением вследствие повреждения изоляции.

Для защиты от прохождения тока по телу необходимо соблюдать технику безопасности использования персонального компьютера, использовать ГОСТ Р 50571.3-94 [23] по защите от поражения электрического тока. При обнаружении нарушений изоляции необходимо в первую очередь отключить прибор от источника питания, использовать изоляционные материалы.

4.3.4. Повышенная или пониженная температура воздуха рабочей зоны.

Микроклимат помещений – это климат внутренней среды помещений, который определяется действующими на организм сочетаниями температуры, влажности и скорости движения воздуха, а также температуры окружающих поверхностей. Показатели микроклимата должны обеспечивать сохранение теплового баланса человека с окружающей средой и поддержание оптимального или допустимого теплового состояния организма человека [25]. Нормы оптимальных и допустимых метеорологических условий устанавливает СанПиН 2.2.4.548–96. Все категории работ разграничиваются на основе интенсивности энергозатрат организма в ккал/ч (Вт). Работа, производимая сидя и сопровождающаяся незначительным физическим напряжением, относится к категории Ia – работа с интенсивностью энергозатрат до 120 ккал/ч (до 139 Вт). Для данной категории допустимые нормы микроклимата представлены в таблице 28.

Таблица 28 – Допустимые нормы микроклимата в рабочей зоне

производственных помещений.

Период года	Категория тяжести выполняемых работ	Температура, °С		Относительная влажность, %		Скорость движения воздуха, м/с	
		Фактическое значение	Допустимое значение	Фактическое значение	Допустимое значение	Фактическое значение	Допустимое значение
Холодный	Ia	(20÷24)	(15÷28)	55	(15÷75)	0.1	≤ 0.1
Теплый	Ia	(23÷25)	(15÷28)	55	(15÷75)	0.1	≤ 0.2

Анализируя таблицу 27, можно сделать вывод, что в рассматриваемом помещении параметры микроклимата соответствуют нормам СанПиН. Допустимый уровень микроклимата помещения обеспечивается системой водяного центрального отопления и естественной вентиляцией.

В производственных помещениях, где допустимые нормативные величины микроклимата поддерживать не представляется возможным, необходимо проводить мероприятия по защите работников от возможного перегревания и охлаждения. Это достигается различными средствами:

- применением систем местного кондиционирования воздуха;
- использованием индивидуальных средств защиты от повышенной или пониженной температуры;
- регламентацией периодов работы в неблагоприятном микроклимате и отдыха в помещении с микроклиматом, нормализующим тепловое состояние;
- сокращением рабочей смены и др.

Профилактика перегревания работников в нагревающем микроклимате включает следующие мероприятия: нормирование верхней границы внешней термической нагрузки на допустимом уровне применительно к 8-часовой рабочей смене; регламентация продолжительности воздействия нагревающей среды (непрерывно и за рабочую смену) для поддержания среднесменного

теплового состояния на оптимальном или допустимом уровне.

4.4. Экологическая безопасность.

На сегодняшние дни во всем мире присутствуют проблемы рационального использования природных ресурсов, проблемы охраны окружающей среды. При этом особое внимание уделяется организации воздействия вредных факторов на природу.

С точки зрения потребления ресурсов компьютер потребляет сравнительно небольшое количество электроэнергии, что положительным образом сказывается на общей экономии потребления электроэнергии в целом, соответственно меньше загрязняется окружающая среда.

При выполнении работы на персональных компьютерах, некоторые из частей компьютера могут прийти в негодность или устареть, тогда необходимо их утилизировать. В качестве отходов в результате деятельности программирования может выступать бумага, отходы электроники, ртутьсодержащие лампы. Эти твердые бытовые отходы загрязняют окружающую среду и их необходимо сдавать на утилизацию в специальные организации, чтобы максимально снизить их вредное воздействие. Для складирования мусора перед утилизацией предлагается вывоз на территории, предназначенные для этого.

При написании ВКР вредных выбросов в атмосферу, почву и водные источники не производилось, радиационного заражения не произошло, чрезвычайные ситуации не наблюдались, поэтому не оказывались существенные воздействия на окружающую среду.

4.5. Безопасность в чрезвычайных ситуациях.

Процесс выполнения работ и написания данной выпускной квалификационной работы, была выполнена в городе Томске, где преобладает континентально-циклонический климат.

Возможными ЧС могут быть сильные морозы, диверсия и пожар. Для Томска в зимнее время года характерны морозы. Достижение критически

низких температур приведет к авариям систем теплоснабжения и жизнеобеспечения, приостановке работы и т.д. В случае переморозки труб должны быть предусмотрены запасные обогреватели. Их количества и мощности должно хватать для того, чтобы работа на производстве не прекратилась. Чтобы избежать потерь при авариях на электросетях у каждого компьютера должен быть блок бесперебойного питания, а также должен быть аварийный генератор, который обеспечит дальнейшую бесперебойную работу персонала.

В процессе работы на персональном компьютере возможной чрезвычайной ситуацией, которую создаёт компьютер может стать пожар. Пожар в течении работы персонального компьютера может возникнуть в результате замыкания электропроводки, приводящее к возгоранию аппаратуры. Далее огонь может распространиться на другие объекты рабочего помещения, создавая техногенную чрезвычайную ситуацию.

Федеральный закон №123 "Технический регламент о требованиях пожарной безопасности" [19] устанавливает основные принципы установления пожарной безопасности. Для предупреждения возникновения пожаров необходимо прежде всего следовать технике безопасности пользования персональным компьютером.

Исключение условий образования горючей среды в сложившихся рабочих условиях включает в себя: применение минимального количества горючих веществ и материалов и наиболее безопасного их размещения, изоляция источников зажигания от горючих веществ, понижение концентрации окислителя в горючей среде в защищаемом объеме, удаление из помещений пожароопасных отходов производства, отложений пыли и пуха.

Для предупреждения вероятности осуществления диверсии предприятие необходимо оборудовать системой видеонаблюдения, круглосуточной охраной, пропускной системой, надежной системой связи, а также исключения распространения информации о системе охраны объекта, расположении

помещений и оборудования в помещениях, системах охраны, сигнализаторах, их местах установки и количестве.

При возникновении пожара необходимо обеспечить технические и организационные мероприятия по оповещению, эвакуации людей, тушению 75 пожара. При обнаружении пожара необходимо оповестить пожарную службу посредством пожарной сигнализации и провести эвакуацию через эвакуационные пути и выходы. Ликвидация последствий пожара обеспечивается автоматическими установками пожаротушения и пожарными службами.

4.6. Выводы и рекомендации по разделу

Проанализировав условия труда на рабочем месте, где была разработана бакалаврская работа, можно сделать вывод, что помещение удовлетворяет необходимым нормам и в случае соблюдения техники безопасности и правил пользования компьютером работа в данном помещении не приведет к ухудшению здоровья пользователя.

Помещение, в котором находится рабочее место, относится к помещениям без повышенной опасности поражения электрическим током. Температурный режим, влажность воздуха не способствуют разрушению изоляции электрооборудования. По состоянию пожаробезопасности помещение соответствует нормам.

Монитор компьютера служит источником электромагнитное поле – вредного фактора, который отрицательно влияет на зрение и в целом на здоровье работника при непрерывной работе более 4 часов. Рекомендуется: во избежание негативного влияния на здоровье необходимо делать перерывы при работе с ЭВМ и проводить специализированные комплексы упражнений для глаз. При рассмотрении вопроса об охране окружающей среды можно делать вывод, что деятельность помещения не является экологически опасной. По состоянию пожаробезопасности помещение соответствует нормам.

Заключение

В ходе выполнения магистерской диссертации была разработана платформа, предназначенная для сбора исходных данных с целью дальнейшего проведения ординации многомерных объектов.

Для эффективной реализации платформы был проведен анализ предметной области, выполнено проектирование базы данных, которая была реализована в среде Phpstorm с использованием PhpMyAdmin, а также разработан сайт «Инженерное волонтерство», доступный по ссылке **engineervol.ru**.

В ходе расчетов методом главных компонент были отобраны четыре главные компоненты, общая доля вариации которых равна 78%.

В результате проведенного анализ избыточности можно отметить следующее:

- первая каноническая переменная объясняет 15,1% совокупной дисперсии в данных.
- Вторая, третья, четвертая неканонические переменные объясняют 24,3%, 18,7%, 16,1% совокупной дисперсии данных соответственно.

Список публикации студента.

Абдуллоев, Р.М. Математические методы ординации многомерных объектов и признаков/Р.М. Абдуллоев//Точная наука. – Кемерово, 2020. – № 86. – с. 56 – 68.

Список использованных источников

- 1) Лабораторные работы по курсу «Базы данных». [Электронный ресурс] – URL: http://orloff.am.tpu.ru/lab_data_base2/. (Дата обращения: 05.05.2020).
- 2) Все о PHP, MySQL и не только. [Электронный ресурс] – URL: <http://www.php.su/>. (Дата обращения: 05.05.2020).
- 3) Документация по MySQL. [Электронный ресурс] – URL: <http://www.mysql.ru/docs/>. (Дата обращения: 06.05.2020).
- 4) HTML 4: Энциклопедия пользователя / Р. Дарнелл, Д. Бэсори-Киц, Дж. Г. Брайан, Б. Кемпбелл. —Киев: DiaSoft, 1998.—688 с.
- 5) MySQL. Справочное руководство по MySQL (последние версии на сайте разработчиков) [Электронный ресурс] – URL: <http://www.mysql.com/documentation/>. (Дата обращения: 06.05.2020).
- 6) Понятия веб-сайт. [Электронный ресурс] - URL: <https://www.webdevelopersnotes.com/>. (Дата обращения: 07.06.2020).
- 7) PHP: настольная книга программиста / Мазуркевич А., Дмитрий, Еловой Д. — Мн.: Новое знание, 2003. — 480 с.
- 8) Разница между сайтом и веб-сайтом. [Электронный ресурс] - URL: <https://writingexplained.org/>. (Дата обращения: 07.06.2020).
- 9) Пользовательский интерфейс. [Электронный ресурс] - URL: https://en.wikipedia.org/wiki/User_interface/. (Дата обращения: 07.06.2020).
- 10) Понятия PHP. [Электронный ресурс] – URL: <https://ru.wikipedia.org/wiki/PHP/>. (Дата обращения: 07.06.2020).
- 11) Самоучитель PHP 7 / М. В. Кузнецов, И. В. Симдянов. - СПб.: БХВПетербург, 2018. -448 с.
- 12) Системы управления базами данных. [Электронный ресурс] - URL: https://studopedia.ru/20_7663_sistemi-upravleniya-bazami-dannih.html/. (Дата обращения: 05.06.2020).

- 13) Шитиков В.К., Мастицкий С.Э. Классификация, регрессия и другие алгоритмы Data Mining с использованием R. - Тольятти, Лондон: - 2017. - 351 с.
- 14) Джонгман Р.Г.Г., тер Браак С.Дж.Ф., ван Тонгерен О.Ф.Р. Анализ данных в экологии сообществ и ландшафтов. М.: РАСХН, 1999. 306 с.
- 15) Обзор о конструкторе сайтов Tilda. [Электронный ресурс] – URL: <https://tilda.education/constructor-saitov-tilda/>. (Дата обращения: 20.02.2020).
- 16) Основные типы данных [Электронный ресурс] – URL: <https://books.irrp.org.ua/data-design/osnovnye-tipy-dannyh/>. (Дата обращения: 16.02.2020).
- 17) Уроки по программированию [Электронный ресурс] – URL:<https://coderlessons.com/>. (Дата обращения: 16.02.2020).
- 18) Анализ и визуализация данных [Электронный ресурс] – URL:https://r-analytics.blogspot.com/p/blog-page_06.html/.(Дата обращения: 23.03.2020).
- 19) Трудовой кодекс Российской Федерации от 30.12.2001 N 197-ФЗ (ред. от 27.12.2018)
- 20) ГОСТ 12.2.032-78 ССБТ. Рабочее место при выполнении работ сидя. Общие эргономические требования.
- 21) ГОСТ 12.2.033-78 ССБТ. Рабочее место при выполнении работ стоя.
- 22) Общие эргономические требования.
- 23) ГОСТ 12.2.049-80 ССБТ. Оборудование производственное. Общие эргономические требования.
- 24) ГОСТ 12.1.030-81 ССБТ. Электробезопасность. Защитное заземление, зануление.
- 25) ГОСТ 12.1.019-2017 ССБТ. Электробезопасность. Общие требования и номенклатура видов защиты.
- 26) СанПиН 2.2.2.542-96 Гигиенические требования к видеодисплейным терминалам, персональным электронно-вычислительным машинам и организации работ. URL: <http://docs.cntd.ru/document/5200235/>.

- 27) Безопасность жизнедеятельности: Учебник для вузов/ С.В. Белов, А.В. Ильницкая, А.Ф. Козьяков и др.; Под общ. ред. С.В. Белова. 7-е изд., стер. – М.: Высш.шк., 2007. – 616 с.

Приложения

Приложение 1

Создание сущности Сотрудники организации в Phpstorm

```
<?php
namespace App\Entity;
use ApiPlatform\Core\Annotation\ApiResource;
use Doctrine\Common\Collections\ArrayCollection;
use Doctrine\Common\Collections\Collection;
use Doctrine\ORM\Mapping as ORM;
/**
 * @ApiResource()
 * @ORM\Entity(repositoryClass="App\Repository\OrganisationRepository")
 */
class Organisation
{
    /**
     * @ORM\Id()
     * @ORM\GeneratedValue()
     * @ORM\Column(type="integer")
     */
    private $id;
    /**
     * @ORM\Column(type="string", length=255, options={"comment":"Название
организации"})
     */
    private $name;
    /**
     * @ORM\Column(type="string", length=255, options={"comment":"Сфера
деятельности"})
     */
    private $area;
    /**
     * @ORM\Column(type="string", length=255, nullable=true,
options={"comment":"Об организации"})
     */

```

```

*/
private $about;
/**
 * staff organisation - Сотрудники организации
 *
 * @ORM\ManyToOne(targetEntity="App\Entity\User",
mappedBy="organisation")
*/
private $users;
/**
 * @ORM\ManyToOne(targetEntity="App\Entity\City")
 * @ORM\JoinColumn(nullable=false)
 */
private $city;
public function __construct()
{
    $this->users = new ArrayCollection();
}
public function getId(): ?int
{
    return $this->id;
}
public function getName(): ?string
{
    return $this->name;
}
public function setName(string $name): self
{
    $this->name = $name;
    return $this;
}
public function getArea(): ?string
{
    return $this->area;
}
public function setArea(string $area): self
{
    $this->area = $area;
    return $this;
}
public function getAbout(): ?string
{
    return $this->about;
}

```

```

}
public function setAbout(string $about): self
{
    $this->about = $about;
    return $this;
}
/**
 * @return Collection|User[]
 */
public function getUsers(): Collection
{
    return $this->users;
}
public function addUser(User $user): self
{
    if (!$this->users->contains($user)) {
        $this->users[] = $user;
        $user->addOrganisation($this);
    }
    return $this;
}
public function removeUser(User $user): self
{
    if ($this->users->contains($user)) {
        $this->users->removeElement($user);
        $user->removeOrganisation($this);
    }
    return $this;
}
public function getCity(): ?City
{
    return $this->city;
}
public function setCity(?City $city): self
{
    $this->city = $city;
    return $this;}}

```


Приложение 2

Создание сущности Пользователи (Волонтеры) в Phpstorm

```
<?php
namespace App\Entity;
use ApiPlatform\Core\Annotation\ApiResource;
use Doctrine\Common\Collections\ArrayCollection;
use Doctrine\Common\Collections\Collection;
use Doctrine\ORM\Mapping as ORM;/**
 * @ApiResource()
 * @ORM\Entity(repositoryClass="App\Repository\UserRepository") */
class User
{
    /**
     * @ORM\Id()
     * @ORM\GeneratedValue()
     * @ORM\Column(type="integer") */
    private $id; /**
     * @ORM\Column(type="string", length=255, options={"comment":"Имя"}) */
    private $name; /**
     * @ORM\Column(type="string", length=255, nullable=true,
options={"comment":"Фамилия"}) */
    private $lastname;
    /** * @ORM\Column(type="string", length=255, options={"comment":"Номер
телефона"}) */ private $phone;
    /**
     * @ORM\Column(type="date", options={"comment":"Дата рождения"})
     */
    private $date_birth;
    /**
     * @ORM\Column(type="string", length=255, nullable=true,
options={"comment":"О себе"})
     */
    private $description;
    /**
     * @ORM\Column(type="string", length=255, nullable=true,
options={"comment":"ссылка в соц. сеть vk"})
     */
    private $auth_vk;
    /**
     * @ORM\Column(type="string", length=255, options={"comment":"Email"})
     */
    private $email;
```

```

/**
 * Контактное лицо организации
 *
 * @ORM\ManyToOne(targetEntity="App\Entity\Organisation",
inversedBy="users")
 */
private $organisation;
/**
 * @ORM\Column(type="integer", nullable=true, options={"comment":"Стаж
волонтера (сколько лет)"})
 */
private $experience;
/**
 * @ORM\ManyToOne(targetEntity="App\Entity\City")
 * @ORM\JoinColumn(nullable=false)
 */
private $city;
/**
 * @ORM\Column(type="string", length=255)
 */
private $password;
/** * @ORM\Column(type="json")
 */ private $roles = [];

public function __construct()
{
    $this->organisation = new ArrayCollection();
}
public function getId(): ?int
{ return $this->id; }
public function getName(): ?string
{
    return $this->name;
}
public function setName(string $name): self
{
    $this->name = $name;
    return $this;
}
public function getLastname(): ?string
{
    return $this->lastname;
}

```

```

public function setLastname(?string $lastname): self
{
    $this->lastname = $lastname;
    return $this;
}
public function getPhone(): ?string
{
    return $this->phone;
}
public function setPhone(string $phone): self
{
    $this->phone = $phone;
    return $this;
}
public function getDateBirth(): ?\DateTimeInterface
{
    return $this->date_birth;
}
public function setDateBirth(\DateTimeInterface $date_birth): self
{
    $this->date_birth = $date_birth;
    return $this;
}
public function getDescription(): ?string
{
    return $this->description;
}
public function setDescription(string $description): self
{
    $this->description = $description;
    return $this;
}
public function getAuthVk(): ?string
{
    return $this->auth_vk;
}
public function setAuthVk(string $auth_vk): self
{
    $this->auth_vk = $auth_vk;
    return $this;
}
public function getEmail(): ?string
{

```

```

    return $this->email;
}
public function setEmail(string $email): self
{
    $this->email = $email;

    return $this;
}
/**
 * @return Collection|Organisation[]
 */
public function getOrganisation(): Collection
{
    return $this->organisation;
}
public function addOrganisation(Organisation $organisation): self
{
    if (!$this->organisation->contains($organisation)) {
        $this->organisation[] = $organisation;
    }
    return $this;
}
public function removeOrganisation(Organisation $organisation): self
{
    if ($this->organisation->contains($organisation)) {
        $this->organisation->removeElement($organisation);
    }
    return $this;
}
public function getExperience(): ?int
{
    return $this->experience;
}
public function setExperience(?int $experience): self
{
    $this->experience = $experience;
    return $this;
}

public function getCity(): ?City
{
    return $this->city;
}

```

```

public function setCity(?City $city): self
{
    $this->city = $city;
    return $this;
}
public function getPassword(): ?string
{
    return $this->password;
}
public function setPassword(string $password): self
{
    $this->password = $password;
    return $this;
}
public function getRoles(): ?array
{
    $roles = $this->roles;
    // guarantee every user at least has ROLE_USER
    $roles[] = 'ROLE_USER';
    return array_unique($roles);
}
public function setRoles(array $roles): self
{
    $this->roles = $roles;
    return $this;
}
}

```

Приложение 3

Создание сущности Задачи в Phpstorm

```

<?php namespace App\Entity;
use ApiPlatform\Core\Annotation\ApiResource;
use Doctrine\ORM\Mapping as ORM;/**
 * @ApiResource()
 * @ORM\Entity(repositoryClass="App\Repository\TaskRepository")
 *class Task{ /**
 * @ORM\Id()
 * @ORM\GeneratedValue()
 * @ORM\Column(type="integer")
 */
private $id;

```

```

/**
 *      @ORM\Column(type="string",          length=255,          nullable=true,
options={"comment":"Нание задачи или мероприятия"})
 */
private $name;
/**
 *              @ORM\Column(type="string",          length=255,
options={"comment":"функционал"})
 */private $task_func;
/**
 * @ORM\Column (type="integer", nullable=true, options= {"comment":"сколько
человек необходимо"})
 */private $count;/**
 * @ORM\Column (type="date", nullable=true, options={"comment":"Дата с"})
 */private $date_start;/**
 * @ORM\Column(type="date", nullable=true, options={"comment":"Дата до" })
 */
private $date_end;
public function getId(): ?int
{
return $this->id;
}
public function getName(): ?string
{
return $this->name;
}
public function setName(?string $name): self
{
$this->task_func = $task_func;

return $this;
}

public function getCount(): ?int
{
return $this->count;
}
public function setCount(?int $count): self

```

```

{
$this->count = $count;
return $this;
}
public function getDateStart(): ?\DateTimeInterface
{
return $this->date_start;
}
public function setDateStart(?\DateTimeInterface $date_start): self
{
$this->date_start = $date_start;
return $this;
}
public function getDateEnd(): ?\DateTimeInterface
{
return $this->date_end;
}
public function setDateEnd(?\DateTimeInterface $date_end): self
{
$this->date_end = $date_end;
return $this;}}

```

Приложение 4

Создание сущности Города в Phpstorm

```

<?php
namespace App\Entity;
use ApiPlatform\Core\Annotation\ApiResource;
use Doctrine\ORM\Mapping as ORM;
/**
 * @ApiResource()
 * @ORM\Entity(repositoryClass="App\Repository\CityRepository")
 */
class City
{
/**

```

```

* @ORM\Id()
* @ORM\GeneratedValue()
* @ORM\Column(type="integer")
*/
private $id;

/**
 * @ORM\Column(type="string", length=255, options={"comment":"Название
города"})
 */
private $name;
public function getId(): ?int
{
    return $this->id;
}
public function getName(): ?string
{
    return $this->name;}
public function setName(string $name): self
{
    $this->name = $name;
    return $this;  }}

```

Приложение 5

Построение даталогической модели данных

CREATE [TEMPORARY] TABLE [IF NOT EXISTS] Имя_таблицы

[(Create-определение,...)]

[Опции_таблицы]

[SELECT ... (любое корректное выражение SELECT)]

Create-определение может быть одним из вариантов:

Определение поля:

Имя_столбца type [NOT NULL | NULL] [DEFAULT
Значение_по_умолчанию] [AUTO_INCREMENT]

Определение первичного ключа:

PRIMARY KEY (Имя_столбца,...)

Определение индекса (ключа):

{INDEX | KEY} [index_name] (Имя_столбца,...)

Определение внешнего ключа:

[CONSTRAINT имя] FOREIGN KEY [index_name] (Имя_столбца,...)

[Определение_ссылки]

Определение_ссылки:

REFERENCES Имя_таблицы [(Имя_столбца,...)]

[ON DELETE Определение_стратегии]

[ON UPDATE Определение_стратегии]

Определение_стратегии:

RESTRICT | CASCADE | SET NULL | NO ACTION | SET DEFAULT

Опции_таблицы:

ENGINE = {InnoDB | MYISAM | ...} и множество других опций.

Приложение 6

(справочное)

Mathematical methods of ordination of multidimensional objects and features

Студент

Группа	ФИО	Подпись	Дата
0BM81	Абдуллоев Рустам Махмадсобирович		

Руководитель ВКР

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОЭФ ИЯТШ	Семенов Михаил Евгеньевич	к.ф.-м.н., доцент		

Консультант-лингвист Отделения иностранных языков ШБИП:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Старший преподаватель	Смирнова Ульяна Александрова			

Introduction

Ordination is a collective term for multidimensional methods that, when mapping a multidimensional object (data set) to a low-dimensional space, allow you to detect a pattern in the source data. This method includes the Principal Component Analysis (PCA) method.

In a narrow sense, ordination is the finding of such coordinate axes on a plane relative to which optimal projection of multidimensional analyzed objects can be performed [13].

The objective of this work is to project multidimensional analyzed objects onto the coordinate plane for applicants for vacant positions.

To achieve this purpose, it is necessary to consistently solve the following tasks.

1. To conduct a comparative analysis of the methods of ordination of multidimensional objects and signs.
2. To develop a web-based tool for collecting source data.
3. To programmatically implement methods for ordinating multidimensional data onto a plane.
4. To build ordination diagrams using various methods (PCA, RDA), compare using various metrics
5. To test the developed methods, draw conclusions and interpret the results.

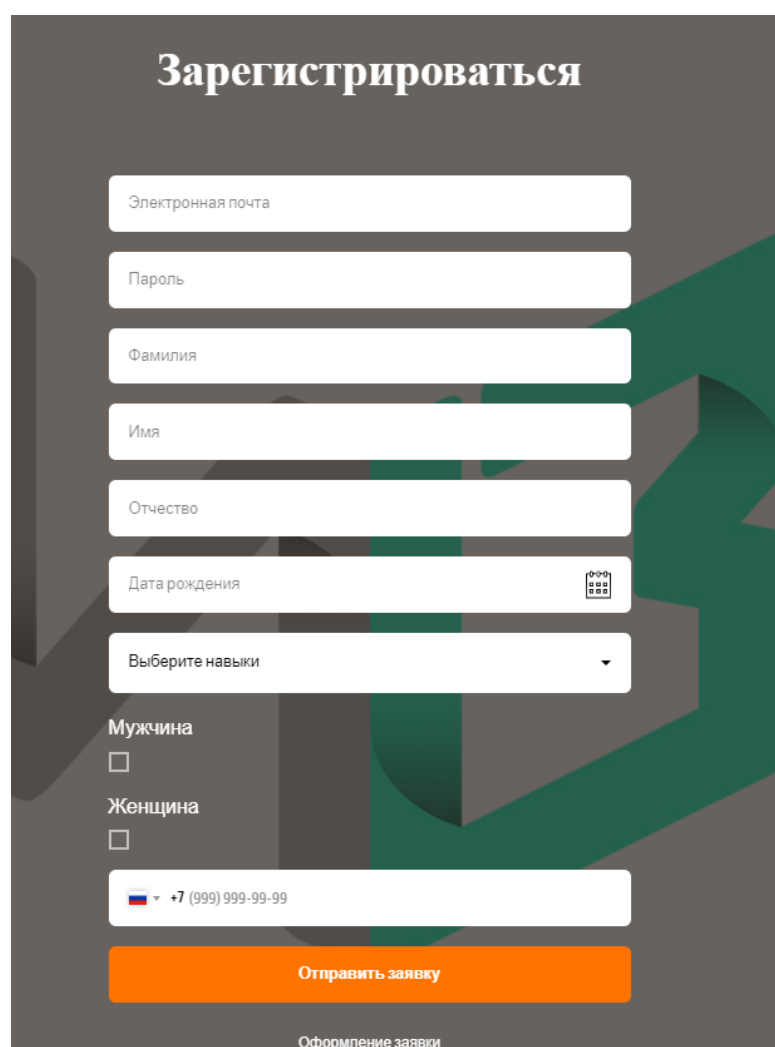
The relevance of the work is the application of the methods of ordination under consideration in many areas, such as econometrics, bioinformatics, image processing, for data compression, in the social sciences.

The object of the study is the method of the main components and its modification.

Theoretical part

1.1. Introductory information

In this work, we will understand the personal data of job applicants on the site "Engineering volunteering" (IV engineervol.ru). The site was developed as a part of a production practice. During development, we used the Tilde constructor, which allows you to create a large number of site types. Using the developed web tool, we collected personal data for further research [15]. An example of the questionnaire is shown in figure 1.



The image shows a registration form titled "Зарегистрироваться" (Register) on a dark grey background. The form consists of several white input fields and a dropdown menu, followed by gender selection options and a phone number field. At the bottom, there is an orange button and a small text label.

Зарегистрироваться


Электронная почта


Пароль

Фамилия

Имя


Отчество

Дата рождения 

Выберите навыки 

Мужчина

Женщина

 +7 (999) 999-99-99

Отправить заявку

Оформление заявки

Figure 1-application Form for job seekers on the site "Engineering volunteering" (engineervol.ru)

As can be seen from figure 1, when filling out the questionnaire, the user can give an answer using various types of data - ordinal, interval, nominal, and normative. Let's look at these types in more detail.

1.2. Basic data types

The nominal data can be calculated, and the percentage of the whole can be determined, but the average value cannot be calculated. You can determine how many students were invited to an interview, or how many percent of applicants are more suitable for the requirements.

If only two categories are available, the data is classified as dichotomous. Answers to questions that require a Yes-no answer are dichotomous data. If the analysis reveals that the applicant is suitable for us or not, then this will also be a dichotomous solution [16].

Data that is naturally ordered by category is called ordinal data. Questions in questionnaires that can be answered with phrases such as "completely disagree", "disagree", "neutral", "agree", "completely agree" are intended for collecting ordinal data. None of the categories on the ordinal scale has an actual mathematical value. Numeric values are often assigned to categories to make it easier to record or analyze data. For example: 1 = completely disagree, 5 = completely agree, but this distribution is conditional, and you can choose any group of ordered numbers to denote groups. For example, you can just as easily decide that the number 5 will mean "completely disagree", and 1 - "completely agree" [16].

The numbers that we assign to ordinal categories affect the interpretation of the final analysis, but we can choose any set of numbers, provided that the numbering order is followed.

In the same way as nominal data, ordinal data can be calculated and a percentage of the whole can be determined, but there is no consensus on whether the average value can be calculated for ordinal data. On the one hand, it is impossible to

determine the average value for the "fully agree" category, for example, and even if you define their numerical values, they will not have an actual mathematical value. Each numeric value represents a specific category, not a quantity of anything.

In some tasks, it is strongly discouraged to use ordinal data for such calculations, while in others it is a constant practice [16].

Interval data is numerical data, so we can perform mathematical operations with it, but such data does not have a "significant" zero point – that is, when the value is zero, what we measure is not missing. For example, 0:00 hours does not mean no time, but the beginning of a new day. Other interval data that we encounter in everyday life is the calendar year and temperature. A zero value for years does not mean that time did not exist before, and a zero temperature (measured in degrees Celsius or Fahrenheit) is not an indicator that there is no heat [16].

Normative data – numeric data that has a lot in common with interval data, except that normative data, unlike interval data, has a significant zero point. In regulatory data, zero means that what you are measuring is missing-zero minutes, zero people in our plane, zero suitable people. In all these cases, zero means that you don't have what you are measuring, and this is different from what we discussed in the interval data section. Other common variables that can be attributed to normative data are height and age.

Interval and normative data can be either discrete or continuous. Discrete data is expressed by a limited set of values (usually integers), and values between these values are not possible. Continuous data can take any value on the scale. This does not mean that the data can take all possible numeric values – only all values within the scale boundaries. For ease of explanation, we often round up continuous data to a certain number of digits. But in this case, these data remain continuous, not discrete [16].

1.2. Ordination of multidimensional objects

Ordination is associated with the search for the most "natural" system patterns and is particularly effective in the presence of a complex of mutually correlated influencing factors.

To date, a significant number of information space compression algorithms have been developed and implemented in the R programming environment. For example, we can distinguish the following (Shitikov and Mastitsky, 2017) [13].

- principal component analysis (PCA) - operating with a correlation matrix, presented in the PCoA package.
- principal coordinate method (PCoA) and non-metric multidimensional scaling (NMDS) - performing a sequential procedure for converting any distance matrix.
- correspondence analysis (CA) - based on an iterative procedure for counter-averaging weighting coefficients for objects and variables.
- multi-dimensional factor analysis (FA) functions-including versions for hierarchical and mixed data provided in the FactoMineR package.
- latent class analysis (LCA) is a statistical method for determining unmeasured class membership among subjects using categorical and/or continuous observable variables, presented in the e1071 package.

The PCoA principal coordinate method, or multidimensional scaling (MDS), is very similar to the principal component method, but instead of a correlation matrix, it calculates eigenvalues and eigenvectors of an arbitrary square symmetric matrix [13].

1.3. Description of required packages in the R environment

R packages are a set of R functions, observed code, and sample data. They are stored in a folder called "library" in the R environment. By default, R installs a set of packages during installation. Additional packages are added later when they are needed for a specific purpose. When we launch the R console, only the default

packages are available by default. Other packages that are already installed must be explicitly loaded for use by the R program that will use them [17].

All packages available in R are over a thousand. The necessary packages and functions for our implementation are as follows.

1. `Vegan()`, `rda()`, `princomp()`, `prcomp()`, `PCA()` – functions for calculating the main components.

2. the `Ggplot2` package is used for visual construction of the ordination diagram.

3. `FactoMineR()`, `Factoextra()` - functions are intended for visualization of results of multidimensional and cluster analysis.

Entering and saving data

4. `Data(x)` - to load a data table that was previously saved with the `save ()` command.

5. Function `viz_pca()` - performs a diagram for different objects.

6. Function `Library()` - loads an additional package.

7. Function `read.table(file)` - reads a data file and creates a data table from it.

8. `Read.csv (filename, header=True)` – function for reading data from a file.

9. `Print(a, ...)` - displays text on the screen.

Creating vectors and tables with data

10. `c(...)` general purpose function (from concatenation-merge, merge);

11. `data.frame (...)` – function for creating a data table.

Converting an R-object

12. `as.data.frame(x)` - creating a data table.

13. Unique(x) – search for non-repeating values.

14. apply (x, MARGIN, FUN = ...) returns a vector, array, or list of values obtained by applying the FUN function to certain elements of the array or matrix x; the elements to be processed x are specified using the MARGIN argument; for example, if x is a matrix, then for MARGIN = 1, the FUN function will be applied to each row of the matrix, for MARGIN = 2 - to each column of the matrix.

15. lapply (x, FUN = ...) returns a list of the same length as x;

the values in the new list will be the result of applying the FUN function to the elements of the source object.

Mathematical function

16. max (x) the maximum value in the numeric vector x.

17. min (x) is the minimum value in the numeric vector x.

18. mean (x) arithmetic mean of the set of x.

19. sd (x) standard deviation of the population x.

20. cor (x) returns a correlation matrix if x is a matrix or data table (the result is 1 if x is a vector).

Matrixes

21. t(x) is the transposition of the matrix x.

Plotting

22. plot (x) graph of the values of the vector x ordered along the x-axis.

23. plot (x, y) graph of the dependence of y on x.

24. Xlab=, ylab=set the names of the corresponding axes.

25. `polygon (x, y)` draws a polygon whose coordinates are stored in the `x` and `y` objects.

26. `Paste (... , sep=...)` converts vectors to text variables and combines them into a single text expression; the `sep` argument allows you to query a Text expression by separating the values of the merged vectors (by default, it is a space).

Programming

27. `function (list of arguments){expression}` allows you to create a custom function.

28. `If (condition){the}` expression checks whether the condition is met; if it is met, it executes the expression specified in curly brackets.

29. `If (condition){expression 1}else{expression 2}` checks whether the condition is met; if it is met, it executes expression 1, otherwise it executes expression 2 [18].

2.1. Development of a web tool for collecting source data.

A website is a collection of documents and files that can be viewed in a web browser program. A typical website consists of pages called web pages. These pages may contain text and multimedia elements, such as images, videos, audio, or animations [6].

A collection of web pages together with multimedia files will be called a website. Web pages are developed using the hypertext Markup language HTML (HyperText Markup Language), which consists of tags that provide the structure of the web page, that is, highlight titles, paragraphs, lists, etc. [7]. In our case, you will need a website Builder to develop a website. Tilda allows you to create an online store without programming [15].

2.2. Database and DBMS design

A database management system (DBMS) is a software that interacts with the user, other applications, and the database itself to collect and analyze data [12]. The General purpose of a DBMS is to develop and execute queries, update and administer databases. Thus, DBMS are tools that are used to connect the administrator and the database itself.

Using a DBMS to store and manage data has advantages, but also overhead. One of the biggest advantages of using a DBMS is that it allows end users and application programmers to access and use the same data while maintaining its integrity. The DBMS provides a Central data store that can be accessed by multiple users in a managed manner [12].

A database is a place (container) for storing data. In a database, even the smallest part of information becomes data. For example, a volunteer is data, employees are data, height, age are all data.

When using a DBMS, the information we collect and add to the database is no longer subject to accidental disorganization. It becomes more accessible and integrated with the rest of our work. Managing information through a database allows us to become strategic users of the data we have.

A database system is called self-describing because it contains not only the database itself, but also metadata that defines and describes data and relationships between tables in the database. If necessary, this information is used by the DBMS software or database users. This separation of data and data information makes the database system completely different from the traditional file system, where data definition is part of application programs [12].

Modern database systems are designed for multiple users. In other words, they allow many users to access the same database at the same time. This access is achieved through functions called concurrency management strategies. These strategies ensure that data access is always correct and data integrity is maintained.

DBMS should provide the ability to define and apply certain restrictions to ensure that users enter reliable information and maintain data integrity. A database constraint is a restriction or rule that defines what you can enter or change in a table, such as date of birth, using a specific format YYYY. MM. DD [12].

By its very nature, a DBMS allows many users to access their database both individually and simultaneously. It is not important for users to know how and where the data they access is stored [12].

2.3. User interface development

The main task of the interface is to allow the user to view information about the project, help, and register.

This structure consists of the following functions

1. The first block – "About the project" - contains brief information about the project and in the same block there will be a link to download a pdf file about the project.
2. The second block "I need help" - in this section, organizations will place their tasks, i.e. there should be several fields where you can enter:
 - 2.1 Name of the task or event where you need help
 - 2.2 Functionality
 - 2.3 How many people are needed
 - 2.4 Terms (from-to)
3. The third block – "I want to help" - this block will be visited by volunteers, and there are cells with tasks like on Procharity(<https://procharity.ru/>)
4. The fourth block - "Additional information" – contains information about skills with a choice: design, 3D modeling, construction, programming, electronics, circuit design.
 - 4.1 The fifth block is "Register", where you can enter:
 - 4.2 Email address

- 4.3 Password
- 4.4 Last name
- 4.5 Name
- 4.6 Patronymics
- 4.7 Date of birth
- 4.8 Selection skills
- 4.9 The selection of sex - “Male”, “Female”
- 4.10 Mobile phone number
- 4.11 Sending an application (making an application)

3. Software implementation of methods for ordination of multidimensional data on a plane.

Multidimensional analysis is a part of statistics that processes and interprets results obtained from observations of several interrelated random variables at the same time, each of which is considered equally important, at least initially. If the data is measured in metric scales and linear relations between variables are assumed, then multidimensional statistical methods based on matrix algebra operations are used in the analysis [13].

In many problems of processing multidimensional observations, the researcher is primarily interested in the possibility of reducing the dimension, i.e., how to select a small set of initial features or their linear combinations, which most explain the variability of the observed objects. This is due to the following three reasons:

- a) the possibility of visual representation (visualization and ordination);
- b) the desire for laconism of the studied models
- C) the need to compress the volume of information.

The principle of ordination of observations consists in using various methods of optimal goal-oriented projection (projecting pursuit) of a point cloud from a multidimensional space to a small-dimensional space (with 2 or 3 coordinate axes) [13].

Various methods or modifications are used for this purpose, but in General the essence of ordination is to represent the original data matrix Y as a set of p latent variables

$$F: Y_1, Y_2, \dots, Y_m \Rightarrow F_1, F_2, \dots, F_p,$$

which are the axes of the ordination diagram (two-dimensional for $p = 2$ or three-dimensional in the case of three such latent variables). The choice of ordination axes is made using the principle of optimality: because the desire to achieve a minimum loss of meaningful information available in the source data [13].

As a rule, the first F_1 axis is drawn through the center of the y_{ij} values thickening and coincides with the direction of the longest axis of the scattering ellipsoid. The second axis F_2 also passes through the center of the distribution, but is drawn perpendicular to the first and coincides in the direction of the second of the main semi-axes of the scattering ellipsoid. This operation creates a " picture " of the data with minimal possible distortions, and the new generalizing variables F_1 and F_2 become orthogonal (i.e. mutually uncorrelated) [13].

Most methods of dimension reduction are based on the analysis of eigenvalues and eigenvectors, which is the most important section of linear (matrix) algebra. In summary, the essence of the $Y \Rightarrow F$ transformations is as follows:

1. Finding eigenvalues and eigenvectors is performed during the mathematical operation of linear transformation of a square symmetric matrix with dimension m . This can be any distance matrix, but most often the variance-covariance matrix

$$C = \frac{Y'Y}{(n-1)}$$

of standardized source data is used.

2. Eigenvalues of a square matrix \mathbf{C} are called such values of λ_k that the system m of equations of the form $(\mathbf{C} - \lambda_k * \mathbf{I}) * u_k = 0$ has a non-trivial solution. Here u_k is the eigenvectors of the matrix \mathbf{C} corresponding $\lambda_k, k = 1, 2, \dots, m, \mathbf{I}$ is the unit matrix.

3. The Matrix of k eigenvectors U_k represents the weights for conversion from the original to the reduced information space, i.e. the matrices of variables \mathbf{Y} and \mathbf{F} are connected by the relation

$$F_{n \times k} = Y_{n \times m} U_{m \times k}.$$

4. Each eigenvalue λ_k corresponds to the value of the variance explained at the k -th level, i.e. the sum of all eigenvalues will be equal to the sum of the variances of all the original variables [13].

5. a Number of eigenvalues are usually ranked from the largest to the lowest: the first eigenvalue λ_1 , which explains the largest proportion of data variation, is often referred to as " dominant "or «leading". The values λ_1 and λ_2 determine the measure of significance of the axes \mathbf{F}_1 and \mathbf{F}_2 of the ordination diagram, i.e. the concentration of starting points along each axis and, as a result, its severity.

6. The eigenvalues of λ_k are found during the iterative procedure, and the accuracy of their calculation depends on the degree of conditionality of the original matrix \mathbf{C} .

The classical method of reducing the data dimension is principal component analysis, which is widely used in various fields of science and technology [13].

	A	B	C	D	E	F	G	H	I
1	id	Age	Education	Citizenship	teer Expe	business Trips	Place of residence	marital status	Sex
2	1	25	3	2	2	0	3	4	1
3	2	21	3	1	1	1	2	4	0
4	3	26	1	3	4	0	3	3	0
5	4	22	2	2	2	0	2	2	0
6	5	39	2	2	4	1	1	4	1

Figure 2-Personal data

For example, the question " Your education?" the applicant can choose from the list: "1-Master's Degree, 2-bachelor's Degree, 3-Secondary". In this case, we are dealing with an ordinal scale. It is monotonically decreasing and allows you to set the ratio more/less between the values that characterize the specified property. Other questions and answers are listed in the table.

Questions	Answer choice	Data type
Education	1 - Magistracy 2 - Bachelor 3 - Average	Ordinal
Citizenship	5. 1. Russian Citizen 6. Russian Citizen and Foreign citizens 7. Foreign citizens 8. Stateless person	Nominal
business trip	0 - No, I'm not ready 1-Yes, ready	Dichotomous
Place of residence	1. Tomsk 2. Tomsk region 3. RF 4. Foreign country	Nominal
Marital status	1- Single/Unmarried 2 - Divorced 3 - Married	Nominal
Sex	0 – Man 1 - Woman	Nominal

Figure 3-question and answer table

4. Construction of ordination diagrams using various methods

The relationship between two sets of variables can be established in the course of canonical analysis. In mathematics, a canonical form is the simplest and most universal form to which certain functions, relations, or expressions can be reduced without loss of generality. The most common forms of canonical ordination based on reduction to eigenvalues and eigenvectors are RDA redundancy analysis [13].

Redundancy analysis is a direct extension of the ideas of multiple regression to the modeling of data with a multidimensional response. The analysis is asymmetric, i.e. if $\mathbf{Y}(n \times m)$ is a table of dependent variables and $\mathbf{X}(n \times q)$ is a table of explanatory variables, i.e. $m \neq q$. RDA is performed in three stages:

1. The coefficients \mathbf{B} of the multiple linear regression model \mathbf{Y} on \mathbf{X} are estimated and a matrix of predicted values is formed:

$$\mathbf{Y}^{\wedge} = \mathbf{XB} = \mathbf{fX}[\mathbf{X}'\mathbf{X}]^{-1} * \mathbf{X}'\mathbf{Y}$$

2. A matrix of canonical coefficients $\mathbf{S} = \mathbf{BU}_c$ with dimension $(m * q)$ is calculated based on the eigenvectors \mathbf{U}_c that weigh the influence of explanatory variables.

3. The covariance matrix $\mathbf{C}_{y|x}$ is calculated based on the values of \mathbf{Y}^{\wedge} , and the main ones are analyzed: $(\mathbf{C}_{y|x} - \lambda_k \mathbf{I})\mathbf{u}_k = \mathbf{0}$, where

$\mathbf{C}_{y|x} = \mathbf{C}_{\mathbf{YX}} * \mathbf{C}_{\mathbf{XX}}^{-1} * \mathbf{C}_{\mathbf{YX}}^T$ and the matrix is calculated in the usual way based on k eigenvectors that are not related to the explanatory variables \mathbf{X} [13].

Conclusion

In the course of the master's thesis, a platform was developed for collecting source data for further ordination of multidimensional objects.

For effective implementation of the platform, the subject area was analyzed, the database was designed, which was implemented in the Phpstorm environment using PhpMyAdmin, and the site "Engineering volunteering" was developed, available at the link engineervol.ru.

In the course of calculations using the principal component method, four principal components were selected, the total percentage of variation of which is 78%.

As a result of the redundancy analysis, the following can be noted:

- the first canonical variable explains 15.1% of the total variance in the data.
- The second, third, and fourth non-canonical variables explain 24.3%, 18.7%, and 16.1% of the total data variance, respectively.