|  |  |
| --- | --- |
| **Изображение выглядит как текст, Шрифт, логотип, Графика  Автоматически созданное описание** |  |

КОНКУРСНОЕ ЗАДАНИЕ КОМПЕТЕНЦИИ

«СПЕЦИАЛИСТ ПО АНАЛИЗУ ДАННЫХ (BI-АНАЛИТИК)»

Регионального этапа чемпионата по профессиональному мастерству «Профессионалы» в 2024г.

Конкурсное задание разработано экспертным сообществом и утверждено Менеджером компетенции, в котором установлены нижеследующие правила и необходимые требования владения профессиональными навыками для участия в соревнованиях по профессиональному мастерству.

**Конкурсное задание включает в себя следующие разделы:**

[1. ОСНОВНЫЕ ТРЕБОВАНИЯ КОМПЕТЕНЦИИ 3](#_Toc158539583)

[1.1. ОБЩИЕ СВЕДЕНИЯ О ТРЕБОВАНИЯХ КОМПЕТЕНЦИИ 3](#_Toc158539584)

[1.2. ПЕРЕЧЕНЬ ПРОФЕССИОНАЛЬНЫХ ЗАДАЧ СПЕЦИАЛИСТА ПО КОМПЕТЕНЦИИ «СПЕЦИАЛИСТ ПО АНАЛИЗУ ДАННЫХ (BI-АНАЛИТИК)» 4](#_Toc158539585)

[1.3. ТРЕБОВАНИЯ К СХЕМЕ ОЦЕНКИ 9](#_Toc158539586)

[1.4. СПЕЦИФИКАЦИЯ ОЦЕНКИ КОМПЕТЕНЦИИ 10](#_Toc158539587)

[1.5. КОНКУРСНОЕ ЗАДАНИЕ 11](#_Toc158539588)

[1.5.1. Разработка/выбор конкурсного задания 11](#_Toc158539589)

[1.5.2. Структура модулей конкурсного задания (инвариант/вариатив) 12](#_Toc158539590)

[2. СПЕЦИАЛЬНЫЕ ПРАВИЛА КОМПЕТЕНЦИ 19](#_Toc158539591)

[2.1. Личный инструмент конкурсанта 20](#_Toc158539592)

[2.2.Материалы, оборудование и инструменты, запрещенные на площадке 20](#_Toc158539593)

[3. Приложения 22](#_Toc158539594)

1. ОСНОВНЫЕ ТРЕБОВАНИЯ КОМПЕТЕНЦИИ

1.1. ОБЩИЕ СВЕДЕНИЯ О ТРЕБОВАНИЯХ КОМПЕТЕНЦИИ

Требования компетенции «Специалист по анализу данных (BI-аналитик)» определяют знания, умения, навыки и трудовые функции, которые лежат в основе наиболее актуальных требований работодателей отрасли.

Целью соревнований по компетенции является демонстрация лучших практик и высокого уровня выполнения работы по соответствующей рабочей специальности или профессии.

Требования компетенции являются руководством для подготовки конкурентоспособных, высококвалифицированных специалистов / рабочих и участия их в конкурсах профессионального мастерства.

В соревнованиях по компетенции проверка знаний, умений, навыков и трудовых функций осуществляется посредством оценки выполнения практической работы.

Требования компетенции разделены на четкие разделы с номерами и заголовками, каждому разделу назначен процент относительной важности, сумма которых составляет 100.

1.2. ПЕРЕЧЕНЬ ПРОФЕССИОНАЛЬНЫХ ЗАДАЧ СПЕЦИАЛИСТА ПО КОМПЕТЕНЦИИ «СПЕЦИАЛИСТ ПО АНАЛИЗУ ДАННЫХ (BI-АНАЛИТИК)»

**Таблица №1 - Перечень профессиональных задач специалиста**

|  |  |  |
| --- | --- | --- |
| **№** | **Раздел** | **Важность в %** |
| 1 | **Разработка модулей программного обеспечения для компьютерных систем** | 10 |
| Специалист должен знать и понимать:- основные этапы разработки программного обеспечения;- основные принципы технологии структурного и объектно-ориентированного программирования;- способы оптимизации и приемы рефакторинга;- основные принципы отладки и тестирования программных продуктов. |
| Специалист должен уметь:- осуществлять разработку кода программного модуля на языках низкого и высокого уровней;- создавать программу по разработанному алгоритму как отдельный модуль; - выполнять отладку и тестирование программы на уровне модуля;- осуществлять разработку кода программного модуля на современных языках программирования;- уметь выполнять оптимизацию и рефакторинг программного кода;- оформлять документацию на программные средства. |

|  |  |  |
| --- | --- | --- |
| **№** | **Раздел** | **Важность в %** |
| 2 | **Разработка и отладка программного кода** | 20 |
| Специалист должен знать и понимать:- методы и приемы формализации поставленных задач;- методы и приемы алгоритмизации поставленных задач;- синтаксис выбранного языка программирования, особенности программирования на этом языке, стандартные библиотеки языка программирования;- методологии разработки компьютерного программного обеспечения- методы повышения читаемости программного кода;- основные стандарты оформления технической документации на компьютерное программное обеспечение;- методы и приемы отладки программного кода.Специалист должен уметь:- применять алгоритмы решения типовых задач в области разработки;- использовать методы и приемы формализации поставленных задач;- использовать методы и приемы алгоритмизации поставленных задач;- применять выбранные языки программирования для написания программного кода;- использовать выбранную среду программирования;- использовать возможности имеющейся технической и/или программной архитектуры для написания программного кода;- применять нормативно-технические документы, определяющие требования к оформлению программного кода;- выявлять ошибки в программном коде;- применять методы и приемы отладки программного кода;- интерпретировать сообщения об ошибках, предупреждения, записи технологических журналов; |

|  |  |  |
| --- | --- | --- |
| **№** | **Раздел** | **Важность в %** |
| 3 | **Управление этапами жизненного цикла методологической и технологической инфраструктуры анализа больших данных в организации** | 25 |
| Специалист должен знать и понимать:Теория вероятностей и математическая статистикаСовременные методы и инструментальные средства анализа больших данныхСтандарты проведения анализа данныхМетоды оценки временных и стоимостных характеристик технологий больших данныхСовременная технологическая инфраструктура высокопроизводительных и распределенных вычисленийМетоды интерпретации и визуализации больших данныхСпециалист должен уметь:Проводить презентации при консультировании заказчика, согласовании и утверждении требований к результатам аналитических работ с использованием технологий больших данныхПодготавливать документы, регламентирующие требования к результатам аналитического исследования с использованием технологий больших данных в соответствии с существующими регламентами организации |

|  |  |  |
| --- | --- | --- |
| **№** | **Раздел** | **Важность в %** |
| 4 | **Анализ больших данных с использованием существующей в организации методологической и технологической инфраструктуры** | 35 |
| Специалист должен знать и понимать:Предметная область анализа больших данных в соответствии с требованиями заказчикаВозможности имеющейся у исполнителя методологической и технологической инфраструктуры анализа больших данныхСовременный опыт использования анализа больших данныхТеоретическая и прикладная информатикаТеоретические и прикладные основы анализа данныхОсновы бизнес-интеллекта, типы систем бизнес-интеллектаТеория принятия решенийМатематическое моделированиеТипы анализа больших данных, виды аналитики |
| Специалист должен уметь:Использовать имеющуюся у исполнителя методологическую и технологическую инфраструктуру анализа больших данных для выполнения аналитических работПроводить сравнительный анализ методов и инструментальных средств анализа больших данныхПроводить анализ больших данных в соответствии с утвержденными требованиями к результатам аналитического исследования |

|  |  |  |
| --- | --- | --- |
| **№** | **Раздел** | **Важность в %** |
| 5 | **Решение задач автоматизации информационно-аналитической деятельности с использованием информационно-аналитических систем в защищенном исполнении** | 5 |
| Специалист должен знать и понимать:Типовые средства и методы защиты информации в локальных и глобальных вычислительных сетяхБазовая конфигурация системы защиты информации ИАСОсобенности применения программных и программно-аппаратных средств защиты информации в ИАС |
| Специалист должен уметь:Конфигурировать параметры системы защиты информации ИАС в соответствии с ее эксплуатационной документациейОбнаруживать и устранять неисправности системы защиты информации ИАС согласно эксплуатационной документации |
| 6 | **Проектирование информационно-аналитических систем в защищенном исполнении** | 5 |
| Специалист должен знать и понимать:Типовые средства, методы и протоколы идентификации, аутентификации и авторизацииНормативные правовые акты Российской Федерации в области защиты информацииОрганизационные меры по защите информации |
| Специалист должен уметь:Производить монтаж и диагностику компонентов ИАСИспользовать типовые криптографические средства защиты информации, в том числе средства электронной подписи |

1.3. ТРЕБОВАНИЯ К СХЕМЕ ОЦЕНКИ

Сумма баллов, присуждаемых по каждому аспекту, должна попадать в диапазон баллов, определенных для каждого раздела компетенции, обозначенных в требованиях и указанных в таблице №2.

**Таблица №2** *-* **Матрица пересчета требований компетенции в критерии оценки**

|  |  |
| --- | --- |
| **Критерий/Модуль** | **Итого баллов за раздел ТРЕБОВАНИЙ КОМПЕТЕНЦИИ** |
| **Разделы ТРЕБОВАНИЙ КОМПЕТЕНЦИИ** |   | **A** | **Б** | **В** | **Г** | **Д** |  |
| **1** | 6  | 0 | 0 | 0 | 4 | 10 |
| **2** | 2,5 | 5 | 11,5 | 0 | 1 | 20 |
| **3** | 10,5 | 2,5 | 1,5 | 7,5 | 3 | 25 |
| **4** | 4 | 12 | 15 | 4 | 0 | 35 |
| **5** | 0 | 2 | 3  | 0 | 0 | 5 |
| **6** | 0 | 3 | 2 | 0 | 0 | 5 |
| **Итого баллов за критерий/модуль** | 23 | 24,5 | 33 | 11,5 | 8 | **100** |

1.4. СПЕЦИФИКАЦИЯ ОЦЕНКИ КОМПЕТЕНЦИИ

Оценка Конкурсного задания будет основываться на критериях, указанных в таблице №3:

**Таблица №3 -****Оценка конкурсного задания**

|  |  |
| --- | --- |
| **Критерий** | **Методика проверки навыков в критерии** |
| **А** | **Предобработка количественных данных** |  Критерии:1. Выбор источников данных2. Проверка типов данных3. Оценка объема данных4. Оценка погрешности |
| **Б** | **Анализ количественных данных** |  Критерии:1. Точность расчетов описательных статистик и корреляций2. Корректность работы дашборда (отсутствие ошибок при отображении информации) / корректность построения графиков и диаграмм |
| **В** | **Обработка и анализ текстовых данных** |  Критерии:1. Содержательная оценка точности распределения навыков по классам2. Оценка доли неклассифицируемых навыков3. Оценка показателей полноты и точности базы |
| **Г** | **Сетевой анализ данных** |  Критерии:1. Оценка правильности построения графа2. Оценка удобства прочтения графа (выбор укладки)3. Оценка описания метрик графа4. Оценка предложенных рекомендаций |
| **Д** | **Презентация решения** | Эксперты выслушивают презентации, оценивают содержание и выступление конкурсантов с помощью объективных и субъективных критериев. |

1.5. КОНКУРСНОЕ ЗАДАНИЕ

Общая продолжительность Конкурсного задания: 15 ч.

Количество конкурсных дней: 3 дней

Вне зависимости от количества модулей, КЗ должно включать оценку по каждому из разделов требований компетенции.

Оценка знаний участника должна проводиться через практическое выполнение Конкурсного задания. В дополнение могут учитываться требования работодателей для проверки теоретических знаний / оценки квалификации.

1.5.1. Разработка/выбор конкурсного задания

Конкурсное задание состоит из 5 модулей, включает обязательную к выполнению часть (инвариант) – 4 модуля: предобработка количественных данных, анализ количественных данных, обработка и анализ текстовых данных, презентация решения; и вариативная часть – 1 модуль: сетевой анализ данных. Общее количество баллов конкурсного задания составляет 100.

Обязательная к выполнению часть (инвариант) выполняется всеми регионами без исключения на всех уровнях чемпионатов.

Количество модулей из вариативной части, выбирается регионом самостоятельно в зависимости от материальных возможностей площадки соревнований и потребностей работодателей региона в соответствующих специалистах. В случае если ни один из модулей вариативной части не подходит под запрос работодателя конкретного региона, то вариативный модуль формируется регионом самостоятельно под запрос работодателя. При этом время на выполнение модуля и количество баллов в критериях оценки по аспектам не меняются.

1.5.2. Структура модулей конкурсного задания (инвариант/вариатив)

**Введение**

В рамках чемпионата конкурсанты решают задачу аналитики данных вакансий для выявления отраслевой специфики рынка труда. Конкурсанты определяют, какие вакансии отражают отраслевую специфику, какие требования предъявляют работодатели к соискателям, а также проводят сравнительный анализ по различным критериям. Эти данные могут быть использованы при разработке или обновлении образовательных программ, а также для помощи в профориентации граждан.

**Важно:** Предметная область может быть вариабельной исходя из потребностей конкретного региона.

**Модуль А. Предобработка количественных данных (инвариант)**

**Время на выполнения модуля:**2 часа

**Краткое описание задания:** написать запросы на выборку данных из базы данных либо работать с предоставленными данными (csv). Необходимо произвести очистку от ошибок и поиск аномальных значений, выполнить работу с пропущенными значениями, преобразовать типы данных (при необходимости).

**Задание:**

Необходимо собрать данные о вакансиях BI-аналитиков или о вакансиях экскурсоводов.

Первый способ:

Воспользоваться данными из открытых источников (сайтов-агрегаторов вакансий). Примеры источников: Работа России, Работа.ру, Авито. Также можно пользоваться другими открытыми источниками (например, сайты фриланса fl.ru и другие).

Второй способ:

Воспользоваться данными открытого источника - веб-сервиса РосНавык (<https://rosnavyk.ru/>). Для получения данных с этого источника можно воспользоваться вкладкой “Данные”, сделать запрос нужных вакансий и получить готовую таблицу в формате csv или xlsx.

В зависимости от выбранного типа сбора данных может быть начислено разное количество баллов.

Поля для выгрузки:

Выгружать можно все доступные данные. Среди полей для выгрузки могут быть: название вакансии, описание вакансии, предлагаемый уровень заработной платы, дата публикации вакансии, регион публикации, работодатель, ключевые навыки (требования), опыт работы, график работы.

Период выгрузки:

Последние 3 месяца

Регион:

Российская Федерация

В базе необходимо:

* провести предобработку данных (привести все к нижнему регистру, задать одинаковый формат данных),
* необходимо определить, какие пропущенные значения являются критическими и провести с ними работу, оценить погрешность при допустимом количестве пропущенных значений
* преобразовать типы данных (при необходимости)

Результат модуля: файл в формате csv или xlsx с готовой к обработке базой данных.

**Модуль Б. Анализ количественных данных (инвариант)**

**Время на выполнения модуля:**4 часа

**Краткое описание задания:** необходимо провести анализ основных свойств данных, нахождение в них общих закономерностей, распределений и аномалий, формулирование выводов, в том числе с использованием инструментов визуализации (построение дашборда).

**Задание:**

Необходимо проанализировать данные по следующим критериям:

* уровень предлагаемых заработных плат
* предлагаемый график работы
* требуемый опыт работы
* распределение по федеральным округам
* построение графика динамики вакансий по неделям/месяцам

Конкурсантам предоставляется на выбор два варианта представления результатов: интерактивный дашборд или представление результатов в виде аналитической записки с диаграммами и графиками.

1. Необходимо построить интерактивный дашборд, в котором отображались бы данные параметры и при выборе любого показателя изменялись бы другие.

Основные элементы дашборда:

* название вакансии
* регион
* заработная плата (выводить медианный уровень заработной платы)
* опыт работы
* график работы
* дата публикации вакансии (график динамики количества вакансий по месяцам)
* работодатель

Необходимо также подготовить аналитическую записку с описанием основных показателей базы (описательная статистика, корреляционный анализ). Рекомендуется также поиск зависимостей одних показателей от других.

1. Необходимо подготовить подробную аналитическую записку с графиками и диаграммами, отражающими основные связи между параметрами (можно построить до 20 графиков и диаграмм). Основные параметры, которые можно использовать в визуальных материалах:
* название вакансии
* регион
* заработная плата (выводить медианный уровень заработной платы)
* опыт работы
* график работы
* дата публикации вакансии (график динамики количества вакансий по месяцам)
* работодатель

Рекомендуется поиск зависимостей, использование корреляционного или других видов анализа.

Результат модуля: интерактивный дашборд с требуемыми параметрами и краткая аналитическая записка с основными показателями. Либо подробная аналитическая записка с основными показателями, графиками и диаграммами.

**Модуль В. Обработка и анализ текстовых данных (инвариант)**

**Время на выполнения модуля:**4 часа

**Краткое описание задания:**

Необходимо выполнить парсинг данных для сбора информации с указанных веб-ресурсов. Сформировать структуру набора данных. Провести предварительную обработку данных. Выполнить построение и отбор признаков. Выполнить семантический анализ текста. Провести визуальный анализ статистики текстовых признаков. Сделать выводы, в том числе с использованием инструментов визуализации.

**Задание:**

Необходимо из описаний вакансий вычленить ключевые навыки. Навыки - конкретные умения и требования, которые предъявляет работодатель к соискателю.

С помощью алгоритмов анализа естественного языка необходимо провести предварительную обработку данных, провести кластеризацию и классификацию выделенных навыков в близкие по смыслу категории.

Провести анализ полученных классов навыков: определить наиболее востребованные и наименее востребованные навыки для выбранной группы вакансий; определить наиболее высокооплачиваемые навыки; определить региональную специфику востребованных навыков.

Результат модуля:

1. Таблица в формате xlsx с классом навыков, его содержанием и частотой встречаемости.

2. Аналитическая записка в формате docx с основными выводами

**Модуль Г. Сетевой анализ данных (вариатив)**

**Время на выполнения модуля:**2 часа

**Краткое описание задания:** выявление связей и взаимодействия между элементами, в том числе с использованием инструментов визуализации. Определение ключевых узлов и групп навыков. Предложение рекомендаций на основе анализа данных.

**Задание:**

Необходимо построить граф связей навыков по выбранной группе вакансий. Граф должен давать возможность понять, как часто один навык упоминается в вакансии вместе с другим навыков, то есть, насколько связаны навыки между собой (насколько часто они упоминаются вместе в одной вакансии).

Необходимо описать основные метрики полученного графа и их интерпретацию.

Необходимо ответить на вопрос: какие рекомендации образовательным организациям можно дать на основании полученного графа?

Результат модуля: визуализированный граф (в формате jpeg или другом визуальном формате, позволяющем интерактивно взаимодействовать с графом) с указанием основных метрик и их описанием, а также с рекомендациями (в формате docx)

**Модуль Д. Презентация решения (инвариант)**

**Время на выполнения модуля:**3 часа

**Краткое описание задания:** необходимо создать презентацию, охватывающую все этапы выполнения задания. В ней должно быть указано ёмкое описание результатов работы с обоснованием выбора того или иного решения. Также в презентации необходимо отразить скриншоты результатов своей работы. Разрабатывать презентацию рекомендуется в Power Point или аналогичной среде. Подготовить устный доклад по результатам своей работы, включающие основные результаты по каждому модулю и выводы (не более 5 минут).

**Задание:**

Необходимо представить полученные результаты: подготовить презентацию в формате pptx или pdf и устный доклад (не более 5 минут).

Основные элементы презентации:

1. Техническое задание.
2. Описание готовой базы с основными метриками.
3. Аналитический отчет с основными результатами: описательные статистики, корреляционный анализ, анализ навыков.

2. СПЕЦИАЛЬНЫЕ ПРАВИЛА КОМПЕТЕНЦИ

В ходе проведения чемпионата есть вероятность возникновения спорных ситуаций, дающих преимущества некоторым участникам над другими. В таблице 4 представлены наиболее частые проблемы, а также алгоритм действия, при выявлении подобных ситуаций:

**Таблица №4 -****Решение спорных ситуаций**

|  |  |
| --- | --- |
| **Вопрос** | **Специальные правила**  |
| Сохранение решения конкурсантами | Решение должно быть сохранено по указанному в Конкурсном задании пути, в случае нарушения инструкций решение не проверяется. |
| Временной регламент | 1. В случае выполнения участником модуля Конкурсного задания в непредназначенное для этого время, в том числе во время выполнения другого модуля, такое решение не будет оценено.2. За не прекращение выполнения Конкурсного задания по команде «СТОП» Главного эксперта или ответственных за хронометраж Экспертов в соответствии с временем выполнения Модуля или в иных ситуациях к Конкурсантам применяются штрафные санкции. |
| Программное обеспечение | 1. Конкурсанты могут создавать программные продукты, оформлять инструкции или делать заметки, находясь на рабочей площадке, однако их никогда нельзя забирать с рабочей площадки.2. За использование материалов, файлов, подготовленных вне конкурсного времени и за пределами конкурсной площадки, в том числе шпаргалок, материалов, полученных в сети Интернет (если иное не указано в Конкурсном задании), выполняемый модуль Конкурсного задания не оценивается.  |
| Отказ оборудования | Если имеется явное доказательство того, что конкурсанты сами причинили ущерб оборудованию, им не будет предоставляться замена и дополнительное время. |
| Поведение конкурсантов | 1. За использование ненормативной лексики устно во время выполнения Конкурсного задания или во время защиты своих работ, а также письменно в представленных к проверке файлах к Конкурсанту применяются штрафные санкции в виде снижения общей суммы баллов на 5 баллов. 2. Общение участников во время выполнения Конкурсного задания запрещено, в случае неоднократного нарушения запрета участники дисквалифицируются на оставшееся время конкурсного дня. |

Помимо описанных выше спорных ситуаций ниже приведен перечень указаний к организации работы на площадке проведения чемпионата:

1) при наличии технической возможности на площадке необходимо обеспечить видеофиксацию мониторов конкурсантов (видеозахват рабочих столов на обоих мониторах) с выгрузкой видеозаписей на выделенный сервер;

2) при наличии технической возможности на площадке для сохранения и выгрузки работ конкурсантов необходимо использовать систему контроля версий. При отсутствии технической возможности на площадке допускается выгрузка работ конкурсантов на флеш-накопитель техническим экспертом, с корректировкой соответствующих критериев в Критериях оценки;

3) любая фото-, видеосъемка СМИ допускается только после согласования с Главным экспертом и Индустриальным экспертом.

2.1. Личный инструмент конкурсанта

Список оборудования неопределенный:

Конкурсант может привезти индивидуальное периферийное оборудование по списку: мышь, клавиатура, проводная гарнитура.

Ко всей периферии применяется следующее требование: оборудование не должно иметь возможности беспроводного подключения, а также заранее программируемых команд (макросов).

2.2.Материалы, оборудование и инструменты, запрещенные на площадке

Для участников соревнований: телефоны и иные средства связи, ноутбуки, смарт-часы, Bluetooth-гарнитуры, средства фото-, аудио-, видеозаписи, средства электронного переноса информации (USB-накопители).

Для экспертного сообщества: запрещено использование телефонов или иных средств связи, ноутбуков, Bluetooth-гарнитур, средств фото-, аудио-, видеозаписи во время формирования итоговой версии конкурсного задания и критериев оценивания (во время внесения 30% изменений) и во время оценивания работ конкурсантов. В остальное время ограничений нет.

3. Приложения

Приложение №1 Инструкция по заполнению матрицы конкурсного задания

Приложение №2 Матрица конкурсного задания

Приложение №3 Критерии оценки

Приложение №4 Инструкция по охране труда и технике безопасности по компетенции «Специалист по анализу данных (BI-аналитик)».