**Приложение А**

**МИНИСТЕРСТВО ТРАНСПОРТА РОССИЙСКОЙ ФЕДЕРАЦИИ**

**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ   
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ**

**«РОССИЙСКИЙ УНИВЕРСИТЕТ ТРАНСПОРТА»**

**(РУТ (МИИТ)**

ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ДОПОЛНИТЕЛЬНОЙ ПРОФЕССИОНАЛЬНОЙ ПРОГРАММЫ –   
ПРОГРАММЫПОВЫШЕНИЯ КВАЛИФИКАЦИИ

«Инструменты и библиотеки Python для обработки, анализа и визуализации больших данных транспортных систем»

**Содержание**

# 1 Исходные данные

## 1.1 Перечень учебно-методической документации, нормативных правовых актов, нормативной технической документации, иной документации, учебной литературы и иных изданий, информационных ресурсов, использованных при подготовке оценочных материалов

Таблица 1 – Учебно-методическая документация, нормативные правовые акты, нормативная техническая документация, иная документация, учебная литература и иные издания, информационные ресурсы

|  |
| --- |
| **Вид информационного и учебно-методического обеспечения** |
| **1 Учебно-методическая документация** |
| 1.1 Конспект лекций |
| 1.2 Методические указания к организации и проведению практических занятий |
| **2 Список используемых источников** |
| 2.1 Ильке Кандан Бенги. PyCharm против Anaconda: какой инструмент Python подходит для вашего проекта? / Ильке Кандан Бенги. — Текст : электронный // HashDork : [сайт]. — 2023. — 19 нояб.— URL: https://hashdork.com/ru/pycharm-vs-anaconda (дата обращения: 27.11.2023). |
| 2.2 Генерация признаков из временных рядов / NewTechAudit : [профессиональное сообщество]. — Текст : электронный // Хабр : [сайт] — 2021. — 22 окт. — URL: https://habr.com/ru/articles/584896 (дата обращения: 20.11.2023). |
| 2.3 Python для анализа данных : операции в pandas. — Текст : электронный // Физтех.Статистика : [сайт] — URL: https://mipt-stats.gitlab.io/courses/python/10\_pandas2.html (дата обращения: 22.11.2023). |
| 2.4 Задойный Алексей. Шпаргалка по визуализации данных в Python с помощью Plotly / А. Задойный. — Текст : электронный // Хабр : [сайт]. — 2020. — 25 мая. — URL: https://habr.com/ru/articles/502958 (дата обращения: 13.10.2023). |
| 2.5 Как выполнить тест тренда Манна-Кендалла в Python / Редакция Кодкампа. — Текст: электронный // Кодкамп : онлайн-институт цифровых навыков : [сайт]. — 2022. — 17 авг.— URL: https://www.codecamp.ru/blog/mann-kendall-test-python (дата обращения: 10.10.2023). |
| 2.6 Роман Котюбеев. Диаграмма рассеяния и столбчатая диаграмма в Plotly Express / Р. Котюбеев. — Текст : электронный // python-school : [сайт]. — 2023. — 21 окт. — URL:https://python-school.ru/blog/визуализация-данных/plotly-express (дата обращения: 30.10.2023). |
| 2.7 С.В. Рындина Базовые возможности языка Python для анализа данных. — Пенза : Издательство ПГУ, 2022. — 72 с. |
| 2.8 Бизли Д. Python. Подробный справочник. – пер. с англ.. — СПб : Символ-Плюс, 2010. — 864 с. |
| 2.9 Уэс Маккинни Python и анализ данных: Первичная обработка данных с применением pandas, NumPy и Jupiter / пер. с англ. А. А. Слинкина. 3-е изд. . — М : МК Пресс, 2023. — 536 с. |
| 2.10 ГрасД. Data Science. Наука о данных с нуля: пер. с англ. - 2-е изд., перераб. и доп. — СПб : БХВ-Петербурr, 2021. — 416 с. |
| 2.11 Плас, Д. В. Python для сложных задач: наука о данных и машинное обучение : [перевод с английского] / Дж. Вандер Плас. — Санкт-Петербург : Питер, 2021. — 572 с. |
| **3 Информационное обеспечение** |
| 3.1 http://library.miit.ru/ |

## 1.2 Планируемые результаты освоения, соотнесенные с результатами обучения по дополнительной профессиональной программе – программе повышения квалификации (далее – программа)

Таблица 2 – Планируемые результаты освоения, соотнесенные с результатами обучения

| Планируемые результаты освоения | Планируемые результаты обучения |
| --- | --- |
| Способен анализировать большие данные при помощи инструментов Python, полученных из корпоративных АСУ | **Знания:**  Знать способы группировки данных по нескольким столбцам и применение агрегирующих функций, Знать возможности сортировки агрегированных данных, Знать последовательность выявления аномалий (ошибок) в данных, Знать последовательность приведения агрегированных данных к табличному виду, Знать методику применения агрегационной функции diff() к данным формата «дата», Знать порядок формирования таблицы с новыми аналитическими признаками, Знать начало работы в jupyter notebook, Знать импорт библиотек для загрузки данных, Знать применение функции read\_csv для загрузки файлов, Знать методы обработки пропущенных значений, Знать определение типов данных, Знать последовательность разработки стратегий обработки пропусков, Знать способы обработки полей с датами и извлечение из даты новых признаков, Знать последовательность проверки и удаление дубликатов, Знать порядок инсталлирования библиотеки plotly, Знать последовательность формирования среза данных одного временного ряда, Знать алгоритм визуализации множества временных рядов на одном графике, Знать инструмент для формирования набора изображения с графиками временных рядов, Знать инструмент проверки гипотезы о наличии тренда во временном ряду, Знать инструменты для выполнения циклов расчетов, Знать алгоритм проверки гипотезы о наличии тренда во временном ряду для групп объектов, Знать алгоритм проверки гипотезы об однородности связанных выборок для оценки значимости изменения количества событий, Знать инструмент для объединения двух датафреймов в итоговую таблицу, Знать структуру данных в Python, Знать основные библиотеки Python для аналитики данных, Знать среду для работы с Python, Знать основные возможности библиотеки pandas.  **Умения:**  Уметь подготавливать данные для решения бизнес-задач при помощи инструментов группировки, агрегации и визуализации данных в условиях некачественных данных, Уметь загружать файлы большого объема для дальнейшей обработки при помощи библиотек Pandas и Numpy в различных форматах, Уметь проводить разведочный анализ данных при помощи базовых библиотек Python в условиях некачественных данных, Уметь формировать статистические гипотезы при помощи представления данных в виде временных рядов в условиях больших данных, характеризующихся множеством параметров. |

# 2 Спецификация заданий для проверки знаний

Таблица 3 – Спецификация заданий для проверки знаний

| **Предмет оценки (знание)** | **Критерии оценки** | **Шкала оценки** | **Тип и  № задания** |
| --- | --- | --- | --- |
| Знать структуру данных в Python | Модельный ответ | 1 балл – за правильный ответ; 0 баллов – за неверный ответ | Задания с выбором ответа: 1, 2, 3, 4, 5, 6 Задания с открытым ответом: 7, 8, 9 Задания на установление соответствия: 10 |
| Знать методы обработки пропущенных значений | Модельный ответ | 1 балл – за правильный ответ; 0 баллов – за неверный ответ | Задания с выбором ответа: 11, 12 Задания с открытым ответом: 13 Задания на установление соответствия: 14, 15 |
| Знать начало работы в jupyter notebook | Модельный ответ | 1 балл – за правильный ответ; 0 баллов – за неверный ответ | Задания с выбором ответа: 16, 17, 18 Задания с открытым ответом: 19, 20 Задания на установление соответствия: 21 |
| Знать импорт библиотек для загрузки данных | Модельный ответ | 1 балл – за правильный ответ; 0 баллов – за неверный ответ | Задания с выбором ответа: 22, 23, 24, 25, 26, 27, 28, 29 |
| Знать применение функции read\_csv для загрузки файлов | Модельный ответ | 1 балл – за правильный ответ; 0 баллов – за неверный ответ | Задания с выбором ответа: 30, 31, 32, 33, 34, 35, 36, 37 Задания с открытым ответом: 38 |
| Знать основные библиотеки Python для аналитики данных | Модельный ответ | 1 балл – за правильный ответ; 0 баллов – за неверный ответ | Задания с выбором ответа: 39, 40, 41, 42, 43, 44, 45 Задания с открытым ответом: 46, 47, 48, 49 Задания на установление соответствия: 50 |
| Знать среду для работы с Python | Модельный ответ | 1 балл – за правильный ответ; 0 баллов – за неверный ответ | Задания с выбором ответа: 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67 Задания с открытым ответом: 68, 69, 70 Задания на установление соответствия: 71 |
| Знать определение типов данных | Модельный ответ | 1 балл – за правильный ответ; 0 баллов – за неверный ответ | Задания с выбором ответа: 72, 73, 74, 75, 76, 77, 78, 79 Задания с открытым ответом: 80 |
| Знать последовательность разработки стратегий обработки пропусков | Модельный ответ | 1 балл – за правильный ответ; 0 баллов – за неверный ответ | Задания с выбором ответа: 81 Задания с открытым ответом: 82 Задания на установление соответствия: 83, 84, 85 |
| Знать способы обработки полей с датами и извлечение из даты новых признаков | Модельный ответ | 1 балл – за правильный ответ; 0 баллов – за неверный ответ | Задания с выбором ответа: 86, 87, 88 Задания с открытым ответом: 89 Задания на установление соответствия: 90 |
| Знать последовательность проверки и удаление дубликатов | Модельный ответ | 1 балл – за правильный ответ; 0 баллов – за неверный ответ | Задания с выбором ответа: 91, 92, 93 Задания с открытым ответом: 94, 95 |
| Знать способы группировки данных по нескольким столбцам и применение агрегирующих функций | Модельный ответ | 1 балл – за правильный ответ; 0 баллов – за неверный ответ | Задания с выбором ответа: 96, 97 Задания с открытым ответом: 98, 99 Задания на установление соответствия: 100 |
| Знать возможности сортировки агрегированных данных | Модельный ответ | 1 балл – за правильный ответ; 0 баллов – за неверный ответ | Задания с выбором ответа: 101, 102 Задания с открытым ответом: 103, 104 Задания на установление соответствия: 105, 106 |
| Знать последовательность выявления аномалий (ошибок) в данных | Модельный ответ | 1 балл – за правильный ответ; 0 баллов – за неверный ответ | Задания с выбором ответа: 107, 108, 109, 110, 111 |
| Знать последовательность приведения агрегированных данных к табличному виду | Модельный ответ | 1 балл – за правильный ответ; 0 баллов – за неверный ответ | Задания с выбором ответа: 112, 113, 114, 115, 116 |
| Знать методику применения агрегационной функции diff() к данным формата «дата» | Модельный ответ | 1 балл – за правильный ответ; 0 баллов – за неверный ответ | Задания с выбором ответа: 117, 118, 119, 120, 121, 122, 123, 124 |
| Знать порядок формирования таблицы с новыми аналитическими признаками | Модельный ответ | 1 балл – за правильный ответ; 0 баллов – за неверный ответ | Задания с выбором ответа: 125, 126, 127, 128, 129, 130 |
| Знать порядок инсталлирования библиотеки plotly | Модельный ответ | 1 балл – за правильный ответ; 0 баллов – за неверный ответ | Задания с выбором ответа: 131, 132, 133 Задания с открытым ответом: 134 Задания на установление соответствия: 135 |
| Знать последовательность формирования среза данных одного временного ряда | Модельный ответ | 1 балл – за правильный ответ; 0 баллов – за неверный ответ | Задания с выбором ответа: 136, 137, 138, 139, 140, 141, 142 Задания с открытым ответом: 143, 144 |
| Знать алгоритм визуализации множества временных рядов на одном графике | Модельный ответ | 1 балл – за правильный ответ; 0 баллов – за неверный ответ | Задания с выбором ответа: 145, 146, 147, 148, 149, 150, 151, 152 Задания с открытым ответом: 153 |
| Знать основные возможности библиотеки pandas | Модельный ответ | 1 балл – за правильный ответ; 0 баллов – за неверный ответ | Задания с выбором ответа: 154, 155 Задания с открытым ответом: 156, 157 Задания на установление соответствия: 158 |
| Знать инструмент для формирования набора изображения с графиками временных рядов | Модельный ответ | 1 балл – за правильный ответ; 0 баллов – за неверный ответ | Задания с выбором ответа: 159, 160, 161, 162, 163, 164, 165 Задания с открытым ответом: 166 |
| Знать инструмент проверки гипотезы о наличии тренда во временном ряду | Модельный ответ | 1 балл – за правильный ответ; 0 баллов – за неверный ответ | Задания с выбором ответа: 167, 168, 169 Задания с открытым ответом: 170 Задания на установление соответствия: 171, 172 |
| Знать инструменты для выполнения циклов расчетов | Модельный ответ | 1 балл – за правильный ответ; 0 баллов – за неверный ответ | Задания с выбором ответа: 173, 174, 175, 176, 177, 178, 179 |
| Знать алгоритм проверки гипотезы о наличии тренда во временном ряду для групп объектов | Модельный ответ | 1 балл – за правильный ответ; 0 баллов – за неверный ответ | Задания с выбором ответа: 180, 181, 182, 183, 184, 185, 186 |
| Знать алгоритм проверки гипотезы об однородности связанных выборок для оценки значимости изменения количества событий | Модельный ответ | 1 балл – за правильный ответ; 0 баллов – за неверный ответ | Задания с выбором ответа: 187, 188, 189, 190, 191, 192, 193 Задания с открытым ответом: 194, 195 |
| Знать инструмент для объединения двух датафреймов в итоговую таблицу | Модельный ответ | 1 балл – за правильный ответ; 0 баллов – за неверный ответ | Задания с выбором ответа: 196, 197, 198 Задания на установление соответствия: 199, 200 |

Общая информация по структуре заданий для проверки знаний:

* количество заданий с выбором ответа: 148;
* количество заданий на установление последовательности: 0;
* количество заданий на установление соответствия: 19;
* количество заданий с открытым ответом: 33;
* время выполнения заданий для проверки знаний: 1 ак. ч.

# 3 Спецификация заданий для проверки умений

Таблица 4 – Спецификация заданий для проверки умений

| Предмет оценки (умение) | Критерии оценки | Шкала оценки | Тип и  № задания |
| --- | --- | --- | --- |
| Уметь подготавливать данные для решения бизнес-задач при помощи инструментов группировки, агрегации и визуализации данных в условиях некачественных данных | Модельный ответ | 1 балл – за правильный ответ;  0 баллов – за неверный ответ | задание на применение умений в модельных условиях № 4, 5 |
| Уметь загружать файлы большого объема для дальнейшей обработки при помощи библиотек Pandas и Numpy в различных форматах | Модельный ответ | 1 балл – за правильный ответ;  0 баллов – за неверный ответ | задание на применение умений в модельных условиях № 1 |
| Уметь проводить разведочный анализ данных при помощи базовых библиотек Python в условиях некачественных данных | Модельный ответ | 1 балл – за правильный ответ;  0 баллов – за неверный ответ | задание на применение умений в модельных условиях № 2, 3 |
| Уметь формировать статистические гипотезы при помощи представления данных в виде временных рядов в условиях больших данных, характеризующихся множеством параметров | Модельный ответ | 1 балл – за правильный ответ;  0 баллов – за неверный ответ | задание на применение умений в модельных условиях № 6 |

Время выполнения практических всех практических заданий: 6 ак. ч.

Время выполнения практических заданий, необходимых для прохождения итоговой аттестации: 6 ак. ч.

# 4 Требования безопасности к проведению оценочных мероприятий

Стандартные требования безопасности при проведении работ за компьютером.

# 5 Задания для проверки знаний

## 5.1 Материально-техническое обеспечение (далее – МТО) для проведения итоговой аттестации на проверку знаний

Таблица 5 – Состав МТО

| **Наименование** | **Кол-во** | **Ед. изм.** | **Примечание** |
| --- | --- | --- | --- |
| **1 Помещения** | | | |
| 1.1 Для лекционных занятий | | | |
| 1.1.1 Лекционная аудитория | 1 | шт. | доска, средства отображения данных на большой экран, доступ в интернет |
| **2 Мебель** | | | |
| 2.1 Учебных аудиторий | | | |
| 2.1.1 Стол | 30 | шт. | посадочные места по количеству обучающихся |
| 2.1.2 Стул | 30 | шт. | посадочные места по количеству обучающихся |
| **3 Оборудование** | | | |
| 3.1 Учебной аудитории для проведения занятий лекционного и семинарского типа | | | |
| 3.1.1 Персональный компьютер преподавателя с веб-камерой, доступ к сети Интернет | 1 | шт. | компьютер с лицензионным программным обеспечением, должен удовлетворять минимальным системным требованиям специализированного ПО, обеспечивать возможность отображения информации на большой экран |
| 3.1.2 Персональные компьютеры для обучающихся с веб-камерой, доступ к сети Интернет | 30 | шт. | компьютер с лицензионным программным обеспечением, должен удовлетворять минимальным системным требованиям специализированного ПО. Количество компьютеров по количеству обучающихся |
| 3.1.3 Мультимедиа-комплекс | 1 | шт. | возможность отображения информации на большой экран |
| 3.1.4 Периферийное оборудование для ПК (принтер, сканер, сетевое оборудование, интерактивная доска) | 1 | шт. |  |
| **4 Расходные материалы** | | | |
| 4.1 Бумага | 1 | уп. |  |
| 4.2 Ручки | 1 | уп. |  |
| **5 Программное обеспечение** | | | |
| 5.1 Офисное | | | |
| 5.1.1 Лицензионное программное обеспечение (Microsoft Office) | 1 | шт. | программное обеспечение по количеству персональных компьютеров |
| 5.2 Специализированное | | | |
| 5.2.1 Тестирующий программный комплекс системы | 1 | шт. | создание библиотеки контрольных вопросов различных типов; формирование тестов на основе библиотеки вопросов (с возможностью случайной выборки, ограничениями по времени и другими параметрами); включение тестов в состав электронных курсов; назначение тестов в качестве самостоятельных оценочных процедур; детальная аналитика по итогам тестирования |
| **6 Иные** | | | |
| 6.1 Информационно-телекоммуникационные сети | 1 |  | обеспечивают передачу по линиям связи учебной информации и обратную связь между обучающимся и средством обучения |
| 6.2 Библиотека электронных образовательных ресурсов | 1 |  | доступ к электронным образовательным ресурсам, контроль знаний обучающихся (тестирование); персональные компьютеры, программа для создания интерактивных и мультимедийных электронных образовательных ресурсов |

## 5.2 Тестовые задания

**1 С помощью каких символов можно обращаться к элементам кортежа:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) квадратных скобок; |
| б) круглых скобок; |
| в) двоеточия; |
| г) точки с запятой. |

**2 Укажите метод кортежа, возвращающий количество вхождений значения:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) count; |
| б) rest; |
| в) tuple; |
| г) list. |

**3 Какой метод используется для добавления элемента в конец списка:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) append; |
| б) insert; |
| в) remove; |
| г) in. |

**4 Какой метод удаляет из списка элемент, находившийся в указанной позиции:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) pop; |
| б) insert; |
| в) remove; |
| г) in. |

**5 Какие методы возвращают список ключей и список значений:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) keys и values; |
| б) get и pop; |
| в) remove и in. |

**6 Какие теоретико-множественные операции поддерживают множества:**

\_

Выбор нескольких правильных ответов

|  |
| --- |
| a) объединение; |
| б) хеширование; |
| в) пересечение; |
| г) разность; |
| д) симметрическую разность. |

**7 Метод, добавляющий элемент в конец списка называется ...**

\_

Вопрос с открытым ответом

**8 Функция, преобразовывающая любую последовательность или итератор в кортеж, называется ...**

\_

Вопрос с открытым ответом

**9 Функция, перебирающая элементы последовательности в обратном порядке, называется ...**

\_

Вопрос с открытым ответом

**10 Установите соответствие между названием структур данных в Python и их определением:**

\_

Установление соответствия

| **Колонка 1** | **Колонка 2** |
| --- | --- |
| 1) кортежи | a) одномерная неизменяемая последовательность объектов Python фиксированной длины, которую нельзя изменить после первоначального присваивания |
| 2) списки | б) упорядоченный набор элементов, каждый из которых имеет свой номер, или индекс, позволяющий быстро получить к нему доступ |
| 3) словари | в) коллекция пар ключ-значение, в которой и ключ, и значение – объекты Python |
| 4) множества | г) неупорядоченная коллекция уникальных элементов |

**11 Как можно оценить ситуацию наличия отсутствующих данных:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) типичное явление в большинстве аналитических приложений; |
| б) не типичное явление в большинстве аналитических приложений; |
| в) аналитическим приложениям это не свойственно. |

**12 Укажите методы, которые применяются для фильтрации отсутствующих данных:**

\_

Выбор нескольких правильных ответов

|  |
| --- |
| a) pandas.isna; |
| б) dropna; |
| в) tuple; |
| г) list. |

**13 Для восполнения отсутствующих данных служит метод ...**

\_

Вопрос с открытым ответом

**14 Установите соответствие между названием метода и его описанием:**

\_

Установление соответствия

| **Колонка 1** | **Колонка 2** |
| --- | --- |
| 1) «dropna» | a) фильтрует метки оси в зависимости от того, существуют ли для метки отсутствующие данные, причем есть возможность указать различные пороги, определяющие, какое количество отсутствующих данных считать допустимым |
| 2) «fillna» | б) восполняет отсутствующие данные указанным значением или использует какой-нибудь метод интерполяции, например «ffill» или «bfill» |
| 3) «isna» | в) возвращает объект, содержащий булевы значения, которые показывают, какие значения отсутствуют |
| 4) «notna» | г) логическое отрицание isna; возвращает True для присутствующих и False для отсутствующих значений |

**15 Установите соответствие между названием аргументами метода fillna и их описанием:**

\_

Установление соответствия

| **Колонка 1** | **Колонка 2** |
| --- | --- |
| 1) «value» | a) скалярное значение или похожий на словарь объект для восполнения отсутствующих значений |
| 2) «method» | б) метод интерполяции: "bfill" (обратное восполнение) или "ffill" (прямое восполнение). По умолчанию None |
| 3) «axis» | в) ось, по которой производится восполнение ("index" или "columns"); по умолчанию axis="index" |
| 4) «limit» | г) для прямого и обратного восполнений максимальное количество непрерывно следующих друг за другом промежутков, подлежащих восполнению |

**16 Что является одним из основных компонентов Jupyter-проекта:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) блокнот (notebook) – интерактивный документ, содержащий код, текст (простой или размеченный), визуализацию и другие результаты выполнения кода; |
| б) среда разработки программного кода на языке python; |
| в) среда тестирования программного кода на языке python. |

**17 С каким расширением будет создан файл после сохранения блокнота (команда «Save and Checkpoint» в меню File):**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) .ipynb; |
| б) .ipyn; |
| в) .ipy; |
| г) .ipynby. |

**18 Как можно использовать Jupyter:**

\_

Выбор нескольких правильных ответов

|  |
| --- |
| a) в качестве локальной среды вычислений; |
| б) его можно развернуть на сервере и обращаться удаленно; |
| в) в качестве среды для масштабных проектов; |
| г) в качестве универсальной среды, из которой можно перенести код куда угодно. |

**19 Для создания нового блокнота необходимо нажать кнопку ...**

\_

Вопрос с открытым ответом

**20 С расширением ... будет создан файл после сохранения блокнота (команда «Save and Checkpoint» в меню File):**

\_

Вопрос с открытым ответом

**21 Установите соответствие между командой и ее назначением:**

\_

Установление соответствия

| **Колонка 1** | **Колонка 2** |
| --- | --- |
| 1) «Save and Checkpoint» | a) сохранение блокнота |
| 2) «Close and Halt» | б) закрытие блокнота |
| 3) Shift-Enter | в) выполнение кода |

**22 Что такое модуль в языке Python:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) файл с расширением .py, содержащий определения и другие инструкции языка Python; |
| б) структура данных для хранения информации; |
| в) код, используемый для построения графиков; |
| г) математические операции. |

**23 Какое ключевое слово используется для импортирования библиотеки:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) load; |
| б) import; |
| в) include; |
| г) from. |

**24 Какая библиотека используется для тестирования статистических гипотез в Python:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) pandas; |
| б) matplotlib; |
| в) seaborn; |
| г) scipy. |

**25 Каким образом можно импортировать конкретные функции из модуля в Python:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) import module; |
| б) from module import function; |
| в) module.function(); |
| г) import function from module. |

**26 Что делает функция datasets.load\_iris():**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) загружает набор данных о лилиях из CSV-файла; |
| б) печатает на экране информацию о наборе данных; |
| в) загружает встроенный набор данных о лилиях; |
| г) отображает график распределения данных о лилиях. |

**27 Какое ключевое слово используется для создания сокращенного имени при импортировании библиотеки:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) as; |
| б) to; |
| в) with; |
| г) as\_short. |

**28 Какая функция используется для чтения данных из CSV-файла с использованием библиотеки pandas:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) read\_csv(); |
| б) load\_csv(); |
| в) open\_csv(); |
| г) import\_csv(). |

**29 Какие библиотеки не используются для тестирования статистических гипотез в Python**

\_

Выбор нескольких правильных ответов

|  |
| --- |
| a) pandas; |
| б) matplotlib; |
| в) seaborn; |
| г) scipy. |

**30 Какой параметр функции read\_csv() используется для указания имени файла или URL-адреса, из которого будут прочитаны данные:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) filepath\_or\_buffer; |
| б) delimiter; |
| в) header; |
| г) encoding. |

**31 Какая кодировка используется Python по умолчанию для работы с данными:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) utf-16; |
| б) utf-8; |
| в) cp1251; |
| г) ASCII. |

**32 Какой метод используется для сортировки строк таблицы по значениям столбцов в DataFrame:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) sort(); |
| б) sort\_index(); |
| в) sort\_values(); |
| г) order(). |

**33 Какой метод используется для удаления строк и столбцов в DataFrame:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) drop(); |
| б) remove(); |
| в) delete(); |
| г) discard(). |

**34 Какой метод используется для удаления дубликатов строк из DataFrame:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) remove\_duplicates(); |
| б) delete\_duplicates(); |
| в) drop\_duplicates(); |
| г) clear\_duplicates(). |

**35 Каким способом можно загрузить файл с локального диска в облачное окружение Google Colab:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) используя библиотеку pandas; |
| б) используя библиотеку NumPy; |
| в) используя библиотеку seaborn; |
| г) используя библиотеку google.colab. |

**36 Как можно применить функцию к каждой ячейке столбца DataFrame, используя метод apply():**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) df.apply(custom\_function, axis=0); |
| б) df.applymap(custom\_function); |
| в) df.map(custom\_function, axis=0); |
| г) df.apply\_column(custom\_function). |

**37 Какие кодировки не используются Python по умолчанию для работы с данными:**

\_

Выбор нескольких правильных ответов

|  |
| --- |
| a) utf-16; |
| б) utf-8; |
| в) cp1251. |

**38 Какая функция позволяет вывести информацию о типах данных в DataFrame ...**

\_

Вопрос с открытым ответом

**39 Что представляет собой библиотека NumPy:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) основной пакет для выполнения научных расчетов на Python; |
| б) библиотека для обработки структурированных данных; |
| в) самый популярный в Python инструмент для создания графиков и других способов визуализации двумерных данных. |

**40 Что представляет собой библиотека Pandas:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) основной пакет для выполнения научных расчетов на Python; |
| б) библиотека для обработки структурированных данных; |
| в) самый популярный в Python инструмент для создания графиков и других способов визуализации двумерных данных. |

**41 Что представляет собой библиотека Matplotlib:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) основной пакет для выполнения научных расчетов на Python; |
| б) библиотека для обработки структурированных данных; |
| в) самый популярный в Python инструмент для создания графиков и других способов визуализации двумерных данных. |

**42 Что представляет собой библиотека SciPy:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) собрание пакетов, предназначенных для решения различных стандартных вычислительных задач; |
| б) основной пакет для выполнения научных расчетов на Python; |
| в) библиотека для обработки структурированных данных. |

**43 Что представляет собой библиотека Scikit-learn:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) основной инструмент машинного обучения для программистов на Python; |
| б) основной пакет для выполнения научных расчетов на Python; |
| в) библиотека для обработки структурированных данных. |

**44 Что представляет собой библиотека Statsmodels:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) пакет для статистического анализа, содержащий алгоритмы классической (прежде всего частотной) статистики и эконометрики; |
| б) основной пакет для выполнения научных расчетов на Python; |
| в) библиотека для обработки структурированных данных. |

**45 Что представляет собой библиотека Jupyter:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) продуктивная среда для интерактивных исследовательских вычислений; |
| б) основной пакет для выполнения научных расчетов на Python; |
| в) библиотека для обработки структурированных данных. |

**46 Основной пакет для выполнения научных расчетов на Python называется ...**

\_

Вопрос с открытым ответом

**47 Самый популярный в Python инструмент для создания графиков и других способов визуализации двумерных данных называется ...**

\_

Вопрос с открытым ответом

**48 Библиотека для обработки структурированных данных – это библиотека ...**

\_

Вопрос с открытым ответом

**49 Основной пакет для выполнения научных расчетов на Python – это библиотека ...**

\_

Вопрос с открытым ответом

**50 Установите соответствие между названием пакетов SciPy, предназначенных для решения различных стандартных вычислительных задач и их особенностями:**

\_

Установление соответствия

| **Колонка 1** | **Колонка 2** |
| --- | --- |
| 1) scipy.integrate | a) подпрограммы численного интегрирования и решения дифференциальных уравнений |
| 2) scipy.linalg | б) подпрограммы линейной алгебры и разложения матриц |
| 3) scipy.optimize | в) алгоритмы оптимизации функций (нахождения минимумов) и поиска корней |
| 4) scipy.signal | г) средства обработки сигналов |
| 5) scipy.sparse | д) обертка вокруг SPECFUN, написанной на Fortran-библиотеке, содержащей реализации многих стандартных математических функций, в том числе гамма-функции |

**51 Что представляет собой IDE (Integrated Development Environment) или «интегрированная среда разработки»:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) готовый комплекс средств, необходимых для разработки и создания ПО; |
| б) готовый комплекс средств, необходимых для тестирования ПО; |
| в) готовый комплекс средств, необходимых для интеграции ПО; |
| г) готовый комплекс средств, необходимых для отладки ПО. |

**52 Для чего программистам нужно использовать среду разработки:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) для формирования более качественного и читаемого кода; |
| б) для экономии времени; |
| в) для упрощения работы; |
| г) для возможности совместной работы. |

**53 Для чего в среде разработки нужен режим «поддержка отладки»:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) для возможности поэтапного выполнения программы с целью поиска ошибок; |
| б) для возможности сохранения и перезагрузки файлов с кодом; |
| в) для подсветки синтаксиса; |
| г) для автоматического форматирования кода. |

**54 Для чего в среде разработки нужен режим «Сохранение и перезагрузка файлов с кодом»:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) для возможности поэтапного выполнения программы с целью поиска ошибок; |
| б) для возможности сохранить код, закрыть файл и вновь запустить его через некоторое время; |
| в) для подсветки синтаксиса; |
| г) для автоматического форматирования кода. |

**55 Для чего в среде разработки нужен режим «Подсветка синтаксиса»:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) для упрощения чтения кода и нахождения отдельных переменных; |
| б) для автоматического форматирования кода; |
| в) для возможности поэтапного выполнения программы с целью поиска ошибок; |
| г) для возможности сохранения и перезагрузки файлов с кодом. |

**56 Для чего в среде разработки нужен режим «Автоматическое форматирование кода»:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) для правильного написания основных операторов языка; |
| б) для возможности поэтапного выполнения программы с целью поиска ошибок; |
| в) для возможности сохранения и перезагрузки файлов с кодом; |
| г) для упрощения чтения кода и нахождения отдельных переменных. |

**57 Для чего в среде разработки присутствует большое число плагинов:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) для расширения возможностей среды разработки; |
| б) для автоматического форматирования кода; |
| в) для возможности поэтапного выполнения программы с целью поиска ошибок; |
| г) для возможности сохранения и перезагрузки файлов с кодом. |

**58 Что представляет собой среда Eclipse:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) интегрированная среда разработки, которая особенно популярна для программирования на Java, но отлично встраивается и в инструментарий Python; |
| б) интегрированная среда разработки, которая особенно популярна для программирования на С++, но отлично встраивается и в инструментарий Python; |
| в) интегрированная среда разработки, которая особенно популярна для программирования на Go, но отлично встраивается и в инструментарий Python; |
| г) интегрированная среда разработки, которая особенно популярна для программирования на PHP, но отлично встраивается и в инструментарий Python. |

**59 Что представляет собой среда IDLE (Integrated Development and Learning Environment):**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) базовая среда для работы с кодом Python; |
| б) продвинутая среда для работы с кодом Python; |
| в) комплексная среда разработки на Python. |

**60 Что представляет собой среда PyCharm:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) базовая среда для работы с кодом Python; |
| б) продвинутая среда для работы с кодом Python; |
| в) комплексная среда разработки на Python. |

**61 Что представляет собой среда Thonny:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) IDE с простым и интуитивно понятным интерфейсом; |
| б) веб-приложение с открытым исходным кодом, поддерживающее рабочие среды для несколько языков программирования; |
| в) комплексная среда разработки на Python. |

**62 Что представляет собой среда Jupyter Notebook:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) веб-приложение с открытым исходным кодом, поддерживающее рабочие среды для несколько языков программирования; |
| б) IDE с простым и интуитивно понятным интерфейсом; |
| в) комплексная среда разработки на Python. |

**63 Какая главная особенность и преимущество Jupyter Notebook:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) способность соединять в одном документе исполняемый код и фрагменты текста на основе разметочного языка Markdown; |
| б) в среду разработки уже встроена последняя версия Python; |
| в) отлично работает с MicroPython и автоматически подсвечивает неверный синтаксис. |

**64 Что представляет собой среда Spyder:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) бесплатная IDE с открытым исходным кодом, сочетающая расширенный анализ, отладку, редактирование, интерактивное выполнение, углубленную проверку кода и возможности визуализации; |
| б) веб-приложение с открытым исходным кодом, поддерживающее рабочие среды для несколько языков программирования; |
| в) интегрированная среда разработки, которая особенно популярна для программирования на Java, но отлично встраивается и в инструментарий Python. |

**65 Что представляет собой среда Wing:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) IDE поставляется с мощным отладчиком и интеллектуальным редактором, которые делают интерактивную разработку Python быстрой, точной и плавной.; |
| б) веб-приложение с открытым исходным кодом, поддерживающее рабочие среды для несколько языков программирования; |
| в) интегрированная среда разработки, которая особенно популярна для программирования на Java, но отлично встраивается и в инструментарий Python. |

**66 Что представляет собой среда Eric Python:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) мощная и многофункциональная IDE и редактор Python, разработанная непосредственно на самом языке Python; |
| б) IDE поставляется с мощным отладчиком и интеллектуальным редактором, которые делают интерактивную разработку Python быстрой, точной и плавной.; |
| в) веб-приложение с открытым исходным кодом, поддерживающее рабочие среды для несколько языков программирования. |

**67 Как может запускаться рабочая среда Jupyter Notebook:**

\_

Выбор нескольких правильных ответов

|  |
| --- |
| a) на компьютере; |
| б) на популярных облачных сервисах в виде платформ для программирования; |
| в) на сервере; |
| г) на системе хранения данных (СХД). |

**68 Готовый комплекс средств, необходимых для разработки и создания ПО называется ...**

\_

Вопрос с открытым ответом

**69 Для возможности поэтапного выполнения программы с целью поиска ошибок в среде разработки используется режим ...**

\_

Вопрос с открытым ответом

**70 Для правильного написания основных операторов языка в среде разработки нужен режим ...**

\_

Вопрос с открытым ответом

**71 Установите соответствие между названием среды разработки и его особенностями:**

\_

Установление соответствия

| **Колонка 1** | **Колонка 2** |
| --- | --- |
| 1) ERIC | a) обеспечивает интегрированную поддержку unittest, CORBA и Google Protocol Buffers (protobuf) |
| 2) Wing | б) возможность эмулировать другие IDE — vi, emacs, Eclipse, Visual Studio и Xcod |
| 3) Spyder | в) включен как IDE по умолчанию в популярный дистрибутив Anaconda Python |
| 4) Jupyter Notebook | г) возможность быстрого старта разработки в режиме онлайн, без развертывания среды на своем устройстве |

**72 Какой тип данных имеет целое число:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) str; |
| б) float; |
| в) int; |
| г) bool. |

**73 Какой тип данных имеет число 4.00:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) str; |
| б) float; |
| в) int; |
| г) bool. |

**74 Допустимо ли использование функции replace к строке:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) да; |
| б) нет. |

**75 Если в переменной а хранится строка "сегодня хороший день" (а = "сегодня хароший день") допустимо ли исправить грамматическую ошибку в слове «хароший» следующим образом a[9] = 'о':**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) да; |
| б) нет. |

**76 В f-строке выражения, которые будут заменены заданным значением обозначаются:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) двойными кавычками; |
| б) фигурными скобками; |
| в) круглыми скобками; |
| г) одинарными кавычками; |
| д) квадратными скобками. |

**77 Укажите операцию, которая позволит привести строку s = "5.0" к типу float:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) float(s); |
| б) type(s); |
| в) нельзя такую строку привести к float. |

**78 Какой гиперпараметр позволит игнорировать несоответствие формата данных дате в столбце DataFrame:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) format; |
| б) errors; |
| в) ignore; |
| г) exact. |

**79 Укажите правильный способ задания строки:**

\_

Выбор нескольких правильных ответов

|  |
| --- |
| a) "сегодня хороший день; |
| б) "сегодня хороший день"; |
| в) сегодня хороший день"; |
| г) сегодня хороший день; |
| д) сегодня хороший день. |

**80 Тип данных ... имеет целое число**

\_

Вопрос с открытым ответом

**81 Какой самый простой способ работы с пропусками:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) удаление записей их содержащих; |
| б) добавление записей без пропусков; |
| в) заполнение пропусков данными. |

**82 Метод, выводящий сводку по параметрам описательной статистики для всех переменных набора называется ...**

\_

Вопрос с открытым ответом

**83 Установите соответствие между названием метода и его описанием:**

\_

Установление соответствия

| **Колонка 1** | **Колонка 2** |
| --- | --- |
| 1) count | a) Количество не-NA объектов |
| 2) sum | б) Сумма |
| 3) mean | в) Среднее значение |
| 4) mad | г) Среднее абсолютное отклонение |
| 5) median | д) Медиана |
| 6) min | е) Минимум |

**84 Установите соответствие между названием метода и его описанием:**

\_

Установление соответствия

| **Колонка 1** | **Колонка 2** |
| --- | --- |
| 1) max | a) Максимум |
| 2) mode | б) Мода |
| 3) std | в) Стандартное отклонение |
| 4) var | г) Несмещенная дисперсия |
| 5) sem | д) Стандартная ошибка среднего |
| 6) skew | е) Скошенность (момент 3-го порядка) |

**85 Установите соответствие между названием метода и его описанием:**

\_

Установление соответствия

| **Колонка 1** | **Колонка 2** |
| --- | --- |
| 1) kurt | a) Эксцесс (момент 4-го порядка) |
| 2) quantile | б) Квантиль (%) |
| 3) cumsum | в) Кумулятивная сумма |
| 4) cumprod | г) Кумулятивное произведение |
| 5) cummax | д) Кумулятивный максимум |
| 6) cummin | е) Кумулятивный минимум |

**86 Как можно характеризовать библиотеку Pandas:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) эффективный инструмент для обработки данных, представляющих собой временные ряды; |
| б) эффективный инструмент для обработки данных, представляющих собой таблицы; |
| в) эффективный инструмент для обработки неструктурированных данных. |

**87 Закончите фразу «Для того чтобы была бы возможность пользоваться функциями pandas, предназначенными для работы с датой и временем ...»:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) даты в DataFrame должны быть представлены значениями подходящего типа данных; |
| б) даты в DataFrame должны быть представлены значениями любого типа данных; |
| в) даты в DataFrame должны быть. |

**88 С помощью какой функции можно прибавлять к датам временные интервалы или вычитать их из дат:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) dateOffset; |
| б) dropna; |
| в) tuple; |
| г) list. |

**89 Для получения «месяца» служит метод ...**

\_

Вопрос с открытым ответом

**90 Установите соответствие между названием функции и ее назначением:**

\_

Установление соответствия

| **Колонка 1** | **Колонка 2** |
| --- | --- |
| 1) date\_range | a) выдаёт даты, находящиеся между начальной и конечной датами |
| 2) apply | б) преобразовывает объекты |
| 3) explode | в) позволяет создавать новую строку для каждой записи в этих списках |

**91 Как определить общее число дубликатов по всем набору столбцов в DataFrame data:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) data.drop\_duplicates(); |
| б) data.drop\_duplicates().sum(); |
| в) data.drop\_duplicates(axis = 1). |

**92 Метод uplicated объекта DataFrame возвращает:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) список с булевыми значениями; |
| б) количества True; |
| в) DataFrame с булевыми значениями. |

**93 Укажите что из следующего является гиперпараметров функции drop\_duplicates():**

\_

Выбор нескольких правильных ответов

|  |
| --- |
| a) keep; |
| б) dropna; |
| в) subset; |
| г) list; |
| д) axis. |

**94 Сколько строчек останется в DataFrame data (содержание на картинке) после выполнения следующего кода: data.drop\_duplicates(subset=[‘ loco\_2.fail\_type’]) ...**



Вопрос с открытым ответом

**95 Для того чтобы был оставлен последний из дубликатов необходимо передать значение ... гиперпараметру keep (укажите ответ в двойных кавычках)**

\_

Вопрос с открытым ответом

**96 Закончите фразу «Для групповых операций существует термин …»**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) разделение – применение – объединение; |
| б) объединение – применение – разделение; |
| в) применение – объединение – разделение. |

**97 Закончите фразу «Совместное использование функций, массивов, словарей и объектов Series ...»**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) вполне допустимо, потому что внутри все преобразуется в массивы; |
| б) вполне допустимо, потому что внутри все преобразуется в кортежи; |
| в) не допустимо. |

**98 Для вычисления количества отличных от null значений в каждой группе служит функция ...**

\_

Вопрос с открытым ответом

**99 Для возврата объекта Series, содержащего размеры групп, служит метод ...**

\_

Вопрос с открытым ответом

**100 Установите соответствие между названием функции и ее назначением:**

\_

Установление соответствия

| **Колонка 1** | **Колонка 2** |
| --- | --- |
| 1) count | a) вычисляет количество отличных от null значений в каждой группе |
| 2) size | б) возвращает количество элементов в массиве |
| 3) GroupBy | в) разделяет и выделяет некоторую часть данных из всего набора данных на основе определенных предопределенных условий или параметров |

**101 Что понимают под агрегированием:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) любое преобразование данных, которое порождает скалярные значения из массива; |
| б) любое преобразование данных; |
| в) любое преобразование массива. |

**102 Чем объясняется то, что пользовательские функции агрегирования работают гораздо медленнее оптимизированных функций:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) большими ресурсными затратами (на вызовы функций и реорганизацию данных) при построении промежуточных блоков данных; |
| б) малыми ресурсными затратами при построении промежуточных блоков данных; |
| в) пользовательские функции работают с той же скоростью, что и оптимизированные. |

**103 Для выбора из данных запрошенного количества наименьших значений служит метод:**

\_

Вопрос с открытым ответом

**104 Любое преобразование данных, которое порождает скалярные значения из массива, называется ...**

\_

Вопрос с открытым ответом

**105 Установите соответствие между названием функции и его описанием:**

\_

Установление соответствия

| **Колонка 1** | **Колонка 2** |
| --- | --- |
| 1) any, all | a) возвращают True, если хотя бы одно или все значения, отличные от NA, «похожи на истину» |
| 2) count | б) количество отличных от NA значений в группе |
| 3) cumprod | в) накопительное произведение отличных от NA значений |
| 4) first, last | г) первое и последнее отличные от NA значения |
| 5) mean | д) среднее отличных от NA значений |

**106 Установите соответствие между названием функции и его описанием:**

\_

Установление соответствия

| **Колонка 1** | **Колонка 2** |
| --- | --- |
| 1) median | a) Медиана отличных от NA значений |
| 2) nth | б) Значение, которое занимало бы позицию n, если бы данные были отсортированы |
| 3) ohlc | в) Статистика типа «начальное–наибольшее–наименьшее–конечное» для данных типа временных рядов |
| 4) prod | г) Произведение отличных от NA значений |
| 5) quantile | д) Выборочный квантиль |

**107 Какой основном вопрос применения операций с массивами**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) фильтрация или преобразование выбросов; |
| б) подготовка данных; |
| в) агрегирование данных. |

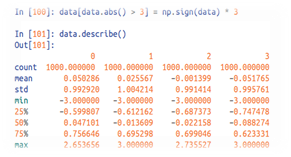
**108 Каким методом нужно воспользоваться для того, чтобы выбрать все строки, в которых встречаются значения, по абсолютной величине превышающие определенную величину**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) any; |
| б) np.sign; |
| в) data. |

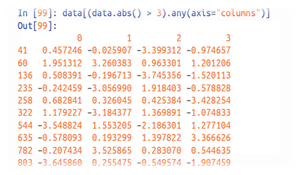
**109 Поясните, что выполняет следующий программный код**



Выбор одного правильного ответа

|  |
| --- |
| a) «срезает» значения, выходящие за границы интервала от –3 до 3; |
| б) «срезает» значения, выходящие за границы интервала от 0 до 3; |
| в) «срезает» значения, выходящие за границы интервала от –3 до 0. |

**110 Поясните, зачем необходимы скобки вокруг data.abs() > 3 в коде**



Выбор одного правильного ответа

|  |
| --- |
| a) они означают, что метод any применяется к результату операции сравнения; |
| б) они означают, что метод any применяется к результату операции агрегации; |
| в) они означают, что метод any применяется для всех колонок массива. |

**111 Поясните, что мы хотим реализовать данных кодом**



Выбор одного правильного ответа

|  |
| --- |
| a) хотим найти в одном из столбцов значения, превышающие 3 по абсолютной величине; |
| б) хотим найти во всех столбцах значения, превышающие 3 по абсолютной величине. |

**112 Какая операция применяется при объединении данных массивов, имеющие общие столбцы (ключи):**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) merge(); |
| б) pd.concat(); |
| в) groupby(). |

**113 Функция, с помощью которой можно изменить названия столбцов:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) rename(); |
| б) reset\_index(); |
| в) rename\_columns(). |

**114 Гиперпараметр функции merge, определяющий тип объединения:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) by; |
| б) how; |
| в) left\_index; |
| г) on. |

**115 Groupby\_data – результат применения агрегационной функции к сгруппированным данным. Привести к формату DataFrame Groupby\_data можно путем применения функции:**

\_

Выбор нескольких правильных ответов

|  |
| --- |
| a) reset\_index(); |
| б) fillna(); |
| в) pd.DataFrame(); |
| г) count(). |

**116 Какие значения нужно передать параметру on для того, чтобы объединить таблицы по результатам одной группировки, но с применением разных агрегационных функций:**

\_

Выбор нескольких правильных ответов

|  |
| --- |
| a) никакие; |
| б) список столбцов, по которым происходит группировка; |
| в) список столбцов, одинаковых для обеих таблиц; |
| г) индексы. |

**117 Какой метод используется для повторной выборки данных из временных рядов в Pandas:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) resample(); |
| б) reshape(); |
| в) reformat(); |
| г) resize(). |

**118 Какой метод используется для создания скользящего окна в Pandas:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) slide(); |
| б) shift(); |
| в) smooth(); |
| г) rolling(). |

**119 Какой параметр задает размер окна при использовании метода для создания скользящего окна:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) size; |
| б) window\_size; |
| в) window; |
| г) dimensions. |

**120 Какой метод используется для вычисления разности временного ряда в Pandas:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) difference(); |
| б) delta(); |
| в) diff(); |
| г) change(). |

**121 Какой параметр в методе shift() задает количество шагов сдвига временного ряда:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) time; |
| б) steps; |
| в) periods; |
| г) shift. |

**122 Какой метод используется для удаления строк с пропущенными значениями в pandas DataFrame:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) remove\_nan(); |
| б) drop\_nan(); |
| в) dropna(); |
| г) remove\_missing(). |

**123 Какой метод используется для создания новых признаков из даты в pandas DataFrame:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) create\_features(); |
| б) generate\_features(); |
| в) extract\_features(); |
| г) add\_features(). |

**124 Какой метод рекомендуется использовать для уменьшения размерности значений даты в pandas DataFrame:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) reduce\_dimension(); |
| б) shrink\_date(); |
| в) compress\_date(); |
| г) subtract\_date(). |

**125 Укажите функцию с помощью которой можно объединить более двух DataFrame:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) merge; |
| б) join; |
| в) concat. |

**126 Что является результатом группировки после применения функции groupby:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) объект, состоящий из пар (имя группы, подтаблица); |
| б) объект groupby из групп объектов. |

**127 Какое максимальное количество колонок можно передать параметру by функции groupby:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) 2; |
| б) 100; |
| в) 2 в 16 степени; |
| г) равно максимальному количеству столбцов в DataFrame. |

**128 Сводная таблица получается путем применения функции:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) pivot\_table; |
| б) crosstab. |

**129 В функции concat() для добавления строк необходимо параметру axis передать значение:**

\_

Выбор нескольких правильных ответов

|  |
| --- |
| a) 0; |
| б) 1; |
| в) index; |
| г) columns. |

**130 Выберите функции, с помощью которых можно объединить несколько DataFrame:**

\_

Выбор нескольких правильных ответов

|  |
| --- |
| a) merge; |
| б) concat; |
| в) corr; |
| г) lambda; |
| д) join. |

**131 Что такое Plotly**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) библиотека для визуализации данных; |
| б) библиотека для подготовки данных; |
| в) библиотека для обработки данных. |

**132 Как называется специальная библиотека для создания дашбордов, созданная на базе plotly и веб-сервера Flask:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) Dash; |
| б) Mash; |
| в) Pandas. |

**133 Выберите составные части библиотеки Plotly:**

\_

Выбор нескольких правильных ответов

|  |
| --- |
| a) Front-End на JS; |
| б) Back-End на Python (за основу взята библиотека Seaborn); |
| в) Back-End на R; |
| г) Front-End на C#; |
| д) Back-End на Java. |

**134 Библиотека, позволяющая строить интерактивные визуализации, называется ...**

\_

Вопрос с открытым ответом

**135 Установите соответствие между названием среды разработки и кодом для установки библиотеки в ней:**

\_

Установление соответствия

| **Колонка 1** | **Колонка 2** |
| --- | --- |
| 1) Anaconda | a) pip install plotly |
| 2) Jupyter Notebook | б) pip install plotly |

**136 Какое правило соответствует фильтру «равно»:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) +; |
| б) ==; |
| в) =; |
| г) =. |

**137 Какое правило соответствует фильтру «не равно»:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) +; |
| б) ==; |
| в) =; |
| г) =. |

**138 Можно ли функции line() модулю plotly.express передать гиперпараметру x или y значение в формате Datetime:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) да; |
| б) нет. |

**139 Можно ли функции line() модулю plotly.express передать гиперпараметру x или y значение в формате str:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) да; |
| б) нет. |

**140 Можно ли функции line() модулю plotly.express передать гиперпараметру x или y значение в формате int:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) да; |
| б) нет. |

**141 С помощью какой функции можно сгенерировать последовательность дат:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) groupby(); |
| б) reindex(); |
| в) count(); |
| г) date\_range(). |

**142 Выберите верные утверждения. Срез данных это:**

\_

Выбор нескольких правильных ответов

|  |
| --- |
| a) Подмножество многомерного куба; |
| б) Часть таблицы, сформированная с применения фильтров; |
| в) Случайно сформированная таблица; |
| г) Исходная таблица, в которой все столбцы приведены к формату str. |

**143 Часть таблицы, сформированная с применения фильтров, называется …**

\_

Вопрос с открытым ответом

**144 Правило ... соответствует фильтру «равно»**

\_

Вопрос с открытым ответом

**145 Какая функция используется для создания графика по точкам с помощью библиотеки Plotly Express:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) px.plot(); |
| б) go.plot(); |
| в) px.scatter(); |
| г) go.scatter(). |

**146 Какой метод используется для отображения созданного графика с помощью Plotly Express:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) show(); |
| б) display(); |
| в) plot(); |
| г) view(). |

**147 Какие основные шаги необходимо выполнить для создания графика с помощью объектов в Plotly:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) создание фигуры, добавление объектов, отображение фигуры; |
| б) загрузка данных, анализ данных, построение графика; |
| в) создание объектов, установка осей, добавление точек; |
| г) подключение библиотеки, определение функций, вывод результатов. |

**148 Каким образом можно добавить график по точкам на фигуру в Plotly:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) go.add\_trace(); |
| б) fig.add\_trace(); |
| в) fig.add\_scatter(); |
| г) go.plot(). |

**149 Какой объект используется для более гибкого создания графиков в Plotly:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) go.Scatter(); |
| б) plotly.express.scatter(); |
| в) go.Figure(); |
| г) plotly.subplots(). |

**150 Что позволяет сделать атрибут name при создании графика в Plotly:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) установить цвет линии; |
| б) изменить тип маркера; |
| в) указать тип графика; |
| г) отобразить имя графика в легенде. |

**151 Какой атрибут используется для изменения размеров полей графика в Plotly:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) fig.resize\_layout(); |
| б) fig.update\_layout(margin=dict(); |
| в) fig.layout.set\_margin(); |
| г) fig.set\_layout\_margin(). |

**152 Какие методы не используются для отображения созданного графика с помощью Plotly Express:**

\_

Выбор нескольких правильных ответов

|  |
| --- |
| a) show(); |
| б) display(); |
| в) plot(); |
| г) view(). |

**153 Функция ... используется для создания графика по точкам с помощью библиотеки Plotly Express**

\_

Вопрос с открытым ответом

**154 В чем смысл ранжирования:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) в присваивании рангов – от единицы до числа присутствующих в массиве элементов, начиная с наименьшего значения; |
| б) в сортировке по нескольким столбцам; |
| в) сортировать по значениям в одном или нескольких столбцах. |

**155 Укажите метод, который применяется для ранжирования:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) rank; |
| б) rest; |
| в) tuple; |
| г) list. |

**156 Для сортировки Series по значениям служит метод ...**

\_

Вопрос с открытым ответом

**157 Для лексикографической сортировки по индексу служит метод ...**

\_

Вопрос с открытым ответом

**158 Установите соответствие между названием способов обработки равных рангов и их описанием:**

\_

Установление соответствия

| **Колонка 1** | **Колонка 2** |
| --- | --- |
| 1) «average» | a) по умолчанию: всем элементам группы присвоить средний ранг |
| 2) «min» | б) всем элементам группы присвоить минимальный ранг |
| 3) «max» | в) всем элементам группы присвоить максимальный ранг |
| 4) «first» | г) присваивать ранги в порядке появления значений в наборе данных |
| 5) «dense» | д) как method="min", но при переходе к следующей группе элементов с одинаковым рангом ранг всегда увеличивается на 1, а не на количество элементов в группе |

**159 Выберете верное утверждение. Фасеточное построение графиков это:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) выведение на одном рисунке несколько графиков; |
| б) отображение нескольких графиков в табличном виде; |
| в) отображение графика с легендой. |

**160 Какие значения передаются аргументу facet\_row:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) название столбца (атрибута), по которому нужно проводить разделение; |
| б) количество строк, которые нужно отображать на одном графике; |
| в) выбранное количество столбцов (атрибутов), по которым нужно проводить разделение. |

**161 Какой аргумент отвечает за высоту рисунка:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) height; |
| б) Height; |
| в) category\_orders; |
| г) barmode. |

**162 Нужно ли передавать значения по осям х и у при фасеточном построении графика:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) да; |
| б) нет; |
| в) на усмотрения разработчика. |

**163 Какой аргумент выполняет туже функцию, что и метод groupby библиотеки Pandas:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) Height; |
| б) category\_orders; |
| в) barmode. |

**164 Какой аргумент позволяет располагать графики в таблице в заданном пользователе порядком:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) height; |
| б) Height; |
| в) category\_orders; |
| г) barmode. |

**165 При фасеточном построении для каждого элемента таблицы рисунков будет отображаться диапазоны изменений по осям х и у:**

\_

Выбор нескольких правильных ответов

|  |
| --- |
| a) нет, по умолчанию; |
| б) да, по умолчанию; |
| в) в зависимости от настроек пользователя. |

**166 Фасеточное построение графиков это - ...**

\_

Вопрос с открытым ответом

**167 Для чего используется тест тренда Манна-Кендалла:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) для определения тренда в данных временного ряда; |
| б) это метрический алгоритм для автоматической классификации объектов или регрессии; |
| в) это алгоритм обучения с учителем, использующийся для задач классификации и регрессионного анализа. |

**168 Какой пакет нужно установить в Python, чтобы выполнить тест тренда Манна-Кендалла:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) Pymannkendall; |
| б) Pykendallmann; |
| в) любой из предложенных. |

**169 С помощью какой библиотеки наряду с выполнением теста тренда Манна-Кендалла можно создать быстрый линейный график:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) Matplotlib; |
| б) Pandas; |
| в) Numpy. |

**170 Для определения тренда в данных временного ряда используется ...**

\_

Вопрос с открытым ответом

**171 Установите соответствие между наименованием выходного параметра теста тренда Манна-Кендалла и его интерпретацией:**

\_

Установление соответствия

| **Колонка 1** | **Колонка 2** |
| --- | --- |
| 1) trend | a) говорит о тренде (возможные выходные данные включают увеличение, уменьшение или отсутствие тренда) |
| 2) h | б) True, если тренд присутствует. False, если тренд отсутствует |
| 3) p | в) p-значение теста |
| 4) z | г) статистика нормализации теста |
| 5) Тау | д) Кендалл Тау |
| 6) s: | е) оценка Манна-Кендала |
| 7) var\_s: | ж) дисперсия S |
| 8) slope | з) оценка Тейла-Сена/наклон |
| 9) Intercept | и) перехват надежной линии Кендалла-Тейла |

**172 Установите соответствие между названием гипотезы для теста и ее назначением:**

\_

Установление соответствия

| **Колонка 1** | **Колонка 2** |
| --- | --- |
| 1) H 0 | a) в данных отсутствует тренд |
| 2) H A | б) в данных присутствует тенденция (это может быть как положительная, так и отрицательная тенденция) |
| 3) если p-значение теста ниже некоторого уровня значимости (обычно выбираются значения 0,10, 0,05 и 0,01) | в) имеется статистически значимое свидетельство того, что в данных временного ряда присутствует тенденция |

**173 Инструкция while выполняет инструкции, пока условное выражение не вернет значение:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) True; |
| б) False; |
| в) 1; |
| г) 0; |
| д) Finished; |
| е) End. |

**174 Как называется i в инструкции for i in s:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) итератор; |
| б) инструкция; |
| в) переменная; |
| г) константа. |

**175 Выберете верное утверждение о функции zip(s,t):**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) объединит последовательности s и t; |
| б) выполняет инструкции, пока условное выражение не вернет значение False; |
| в) прерывает выполнение инструкции For после исчерпания элементов в самой короткой из; |
| г) последовательностей, если их длины не равны. |

**176 С помощью какой инструкции можно перейти к следующей итерации, не прерывая цикл:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) break; |
| б) zip(); |
| в) for; |
| г) continue. |

**177 Основное назначение инструкции else в циклах:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) избежать преждевременного прерывания цикла; |
| б) установить флаг прерывания цикла; |
| в) остановить выполнение цикла. |

**178 С каким объекта может работать инструкция for:**

\_

Выбор нескольких правильных ответов

|  |
| --- |
| a) единичные числа; |
| б) строки; |
| в) объекты, поддерживающие итерации; |
| г) списки. |

**179 Функция enumerate() позволяет получить, помимо элементов последовательности доступ к:**

\_

Выбор нескольких правильных ответов

|  |
| --- |
| a) порядковому номеру итерации; |
| б) размеру последовательности; |
| в) кортежам вида (0, s[0]); |
| г) количеству символов последовательности. |

**180 По каким столбцам будет проходить объединение, если явно не указаны ключи (столбцы, по которым производить соединение):**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) по номерам строк; |
| б) по первым столбцам из всех объединяемых массивов; |
| в) по столбцам с одинаковыми именами. |

**181 Какое значение имеет параметр ‘how’ в функции pd.merge() по умолчанию:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) left’; |
| б) outer’; |
| в) inner’; |
| г) ничего из вышеперечисленного; |
| д) right. |

**182 При значение параметра ‘how’ = ‘outer’ в функции pd.merge(how = ‘outer’) в случае если в одном из массивов не нашлось равного ключа, какое значение будет в итоговой табличке:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) NaN или <NA>; |
| б) 0; |
| в) &#2013266073;; |
| г) (пробел). |

**183 Возможно ли в функции pd.merge() объединение по нескольким ключам:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) да; |
| б) нет. |

**184 Если в функции pd.merge() ключи для объединения переданы явно (указано значение аргумента on), что произойдет со столбцами, имеющими одинаковое название, но не перееденных аргументу on:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) они будут добавлены в ключи автоматически; |
| б) такие столбцы будут исключены из обоих массивов; |
| в) такие столбцы останутся только в первом массиве, остальные будут исключены; |
| г) они не будут добавлены в ключи, а новом массиве будут переименованы. |

**185 Выберите функции и инструкции, с помощью которых можно комбинировать объекты pandas:**

\_

Выбор нескольких правильных ответов

|  |
| --- |
| a) merge(); |
| б) concat(); |
| в) zip(); |
| г) For. |

**186 Выберите правильные/правильные утверждения. Операция слияния или соединения, комбинирует наборы данных, соединяя строки:**

\_

Выбор нескольких правильных ответов

|  |
| --- |
| a) по одному ключу; |
| б) по нескольким ключам; |
| в) по одному или нескольким ключам; |
| г) без ключей; |
| д) по номера строк; |
| е) по номерам столбцов. |

**187 К каком типу относится критерий знакового ранга Уилкоксона:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) параметрический; |
| б) непараметрический; |
| в) критерий среднего; |
| г) метрический критерий. |

**188 Что проверяется с помощью критерий знакового ранга Уилкоксона:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) наличие значительной разницы между средними значениями генеральной совокупности; |
| б) наличие значительной между средними значениями генеральной совокупности; |
| в) наличие значительной разницы между среднеквадратичным отклонением генеральной совокупности. |

**189 Критерий знакового ранга Уилкоксона можно использовать, если распределение различий между двумя выборки не относится к нормальному распределению:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) нет; |
| б) да. |

**190 Как звучит нулевая гипотеза в критерии знакового ранга Уилкоксона:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) двухсторонний; |
| б) больше; |
| в) меньше; |
| г) равны между собой; |
| д) неодинаково; |
| е) заметно неодинаково. |

**191 Какое значение сравнивается с параметром α:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) pvalue; |
| б) теоретическое значением статистики Уилкоксона; |
| в) выборочное значение статистики по критерию Уилкоксона. |

**192 Какие значение передаются функции wilcoxon библиотеки scipy:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) два средних значений по двум выборкам; |
| б) два массива значений одной длины; |
| в) два массива значений различной длины. |

**193 Сколько значений возвращает функция wilcoxon библиотеки scipy:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) ни одного; |
| б) одно; |
| в) два; |
| г) по запросы пользователя; |
| д) три. |

**194 Критерий знакового ранга Уилкоксона относится к ... типу**

\_

Вопрос с открытым ответом

**195 Наличие значительной разницы между средними значениями генеральной совокупности проверяется с помощью ...**

\_

Вопрос с открытым ответом

**196 Продолжите фразу «По сравнению с pandas.merge, метод join объекта DataFrame по умолчанию выполняет…»:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) левое внешнее соединение по ключам соединения; |
| б) правое внешнее соединение по ключам соединения; |
| в) левое внешнее разъединение по ключам соединения. |

**197 Для чего можно использовать метод экземпляра join:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) для комбинирования нескольких объектов DataFrame, обладающих одинаковыми или похожими индексами, но непересекающимися столбцами; |
| б) для комбинирования нескольких объектов DataFrame, обладающих разными индексами, но непересекающимися столбцами; |
| в) для комбинирования нескольких объектов DataFrame, обладающих одинаковыми или похожими индексами и пересекающимися столбцами. |

**198 Выберите правильное описание функции CONCAT:**

\_

Выбор одного правильного ответа

|  |
| --- |
| a) функция, соединяющая несколько текстовых значений в одну строку, другими словами, выполняет конкатенацию строк; |
| б) функция, создающая строки путем замены некоторых частей другой строки; |
| в) функция, возвращающая новую строку, в которой каждый символ в строке заменяется с использованием определенной таблицы перевода. |

**199 Установите соответствие между наименованием типа слияния и его интерпретацией:**

\_

Установление соответствия

| **Колонка 1** | **Колонка 2** |
| --- | --- |
| 1) Один-к-одному | a) одна единица данных в одной таблице связана только с одной единицей данных в другой таблице |
| 2) Один-ко-многим | б) в этом типе данных одна строка с информацией связана со многими строками в другой таблице |
| 3) Многие-к-одному | в) построение этого типа связи похоже на предыдущий тип, но работает в обратном направлении — есть первичная и второстепенная таблица |
| 4) Многие-ко-многим | г) в этом типе связей много строк в одной таблице имеет связь со многими строками в другой таблице |

**200 Установите соответствие между названием vtnjlf Pandas и еuj назначением:**

\_

Установление соответствия

| **Колонка 1** | **Колонка 2** |
| --- | --- |
| 1) Inner join | a) возвращает только строки, в которых ключи из левой таблицы совпадают с ключами из правой таблицы |
| 2) Full join | б) вернет все строки из всех таблиц, участвующих в соединении, соединив между собой те, которые подошли под условие |
| 3) Left join | в) выбираются все записи первой (левой) таблицы, даже если они не соответствуют записям во второй (правой) таблице |
| 4) Right join | г) выбираются все записи второй (правой) таблицы, даже если они не соответствуют записям в первой (левой) таблице |

## 5.3 Критерии и шкала оценки (ключи к заданиям), правила обработки результатов теста

Таблица 6 – Критерии и шкала оценки (ключи к заданиям)

| № задания | Правильные варианты ответа, модельные ответы | Шкала оценки |
| --- | --- | --- |
| 1 | a | 1 балл – за правильный ответ |
| 2 | a | 1 балл – за правильный ответ |
| 3 | a | 1 балл – за правильный ответ |
| 4 | a | 1 балл – за правильный ответ |
| 5 | a | 1 балл – за правильный ответ |
| 6 | a,в,г,д | 1 балл – за правильный ответ |
| 7 | append | 1 балл – за правильный ответ |
| 8 | tuple | 1 балл – за правильный ответ |
| 9 | reversed | 1 балл – за правильный ответ |
| 10 | 1-a,2-б,3-в,4-г | 1 балл – за правильный ответ |
| 11 | a | 1 балл – за правильный ответ |
| 12 | a,б | 1 балл – за правильный ответ |
| 13 | fillna | 1 балл – за правильный ответ |
| 14 | 1-a,2-б,3-в,4-г | 1 балл – за правильный ответ |
| 15 | 1-a,2-б,3-в,4-г | 1 балл – за правильный ответ |
| 16 | a | 1 балл – за правильный ответ |
| 17 | a | 1 балл – за правильный ответ |
| 18 | a,б | 1 балл – за правильный ответ |
| 19 | new | 1 балл – за правильный ответ |
| 20 | .ipynb | 1 балл – за правильный ответ |
| 21 | 1-a,2-б,3-в | 1 балл – за правильный ответ |
| 22 | a | 1 балл – за правильный ответ |
| 23 | б | 1 балл – за правильный ответ |
| 24 | г | 1 балл – за правильный ответ |
| 25 | б | 1 балл – за правильный ответ |
| 26 | в | 1 балл – за правильный ответ |
| 27 | a | 1 балл – за правильный ответ |
| 28 | a | 1 балл – за правильный ответ |
| 29 | a,б,в | 1 балл – за правильный ответ |
| 30 | a | 1 балл – за правильный ответ |
| 31 | б | 1 балл – за правильный ответ |
| 32 | в | 1 балл – за правильный ответ |
| 33 | a | 1 балл – за правильный ответ |
| 34 | в | 1 балл – за правильный ответ |
| 35 | г | 1 балл – за правильный ответ |
| 36 | г | 1 балл – за правильный ответ |
| 37 | a,в | 1 балл – за правильный ответ |
| 38 | dtypes() | 1 балл – за правильный ответ |
| 39 | a | 1 балл – за правильный ответ |
| 40 | б | 1 балл – за правильный ответ |
| 41 | в | 1 балл – за правильный ответ |
| 42 | a | 1 балл – за правильный ответ |
| 43 | a | 1 балл – за правильный ответ |
| 44 | a | 1 балл – за правильный ответ |
| 45 | a | 1 балл – за правильный ответ |
| 46 | NumPy | 1 балл – за правильный ответ |
| 47 | Matplotlib | 1 балл – за правильный ответ |
| 48 | Pandas | 1 балл – за правильный ответ |
| 49 | NumPy | 1 балл – за правильный ответ |
| 50 | 1-a,2-б,3-в,4-г,5-д | 1 балл – за правильный ответ |
| 51 | a | 1 балл – за правильный ответ |
| 52 | a | 1 балл – за правильный ответ |
| 53 | a | 1 балл – за правильный ответ |
| 54 | б | 1 балл – за правильный ответ |
| 55 | a | 1 балл – за правильный ответ |
| 56 | a | 1 балл – за правильный ответ |
| 57 | a | 1 балл – за правильный ответ |
| 58 | a | 1 балл – за правильный ответ |
| 59 | a | 1 балл – за правильный ответ |
| 60 | в | 1 балл – за правильный ответ |
| 61 | a | 1 балл – за правильный ответ |
| 62 | a | 1 балл – за правильный ответ |
| 63 | a | 1 балл – за правильный ответ |
| 64 | a | 1 балл – за правильный ответ |
| 65 | a | 1 балл – за правильный ответ |
| 66 | a | 1 балл – за правильный ответ |
| 67 | a,б | 1 балл – за правильный ответ |
| 68 | IDE | 1 балл – за правильный ответ |
| 69 | «поддержка отладки» | 1 балл – за правильный ответ |
| 70 | «Автоматическое форматирование кода» | 1 балл – за правильный ответ |
| 71 | 1-a,2-б,3-в,4-г | 1 балл – за правильный ответ |
| 72 | в | 1 балл – за правильный ответ |
| 73 | б | 1 балл – за правильный ответ |
| 74 | a | 1 балл – за правильный ответ |
| 75 | б | 1 балл – за правильный ответ |
| 76 | б | 1 балл – за правильный ответ |
| 77 | a | 1 балл – за правильный ответ |
| 78 | в | 1 балл – за правильный ответ |
| 79 | б,г | 1 балл – за правильный ответ |
| 80 | int | 1 балл – за правильный ответ |
| 81 | a | 1 балл – за правильный ответ |
| 82 | .describe() | 1 балл – за правильный ответ |
| 83 | 1-a,2-б,3-в,4-г,5-д,6-е | 1 балл – за правильный ответ |
| 84 | 1-a,2-б,3-в,4-г,5-д,6-е | 1 балл – за правильный ответ |
| 85 | 1-a,2-б,3-в,4-г,5-д,6-е | 1 балл – за правильный ответ |
| 86 | a | 1 балл – за правильный ответ |
| 87 | a | 1 балл – за правильный ответ |
| 88 | a | 1 балл – за правильный ответ |
| 89 | month | 1 балл – за правильный ответ |
| 90 | 1-a,2-б,3-в | 1 балл – за правильный ответ |
| 91 | б | 1 балл – за правильный ответ |
| 92 | б | 1 балл – за правильный ответ |
| 93 | a,в | 1 балл – за правильный ответ |
| 94 | 0 | 1 балл – за правильный ответ |
| 95 | "last" | 1 балл – за правильный ответ |
| 96 | a | 1 балл – за правильный ответ |
| 97 | a | 1 балл – за правильный ответ |
| 98 | Count | 1 балл – за правильный ответ |
| 99 | size | 1 балл – за правильный ответ |
| 100 | 1-a,2-б,3-в | 1 балл – за правильный ответ |
| 101 | a | 1 балл – за правильный ответ |
| 102 | a | 1 балл – за правильный ответ |
| 103 | nsmallest | 1 балл – за правильный ответ |
| 104 | агрегированием | 1 балл – за правильный ответ |
| 105 | 1-a,2-б,3-в,4-г,5-д | 1 балл – за правильный ответ |
| 106 | 1-a,2-б,3-в,4-г,5-д | 1 балл – за правильный ответ |
| 107 | a | 1 балл – за правильный ответ |
| 108 | a | 1 балл – за правильный ответ |
| 109 | a | 1 балл – за правильный ответ |
| 110 | a | 1 балл – за правильный ответ |
| 111 | a | 1 балл – за правильный ответ |
| 112 | a | 1 балл – за правильный ответ |
| 113 | a | 1 балл – за правильный ответ |
| 114 | б | 1 балл – за правильный ответ |
| 115 | a,в | 1 балл – за правильный ответ |
| 116 | б,в | 1 балл – за правильный ответ |
| 117 | a | 1 балл – за правильный ответ |
| 118 | г | 1 балл – за правильный ответ |
| 119 | в | 1 балл – за правильный ответ |
| 120 | в | 1 балл – за правильный ответ |
| 121 | в | 1 балл – за правильный ответ |
| 122 | в | 1 балл – за правильный ответ |
| 123 | в | 1 балл – за правильный ответ |
| 124 | г | 1 балл – за правильный ответ |
| 125 | в | 1 балл – за правильный ответ |
| 126 | a | 1 балл – за правильный ответ |
| 127 | г | 1 балл – за правильный ответ |
| 128 | a | 1 балл – за правильный ответ |
| 129 | a,в | 1 балл – за правильный ответ |
| 130 | a,б,д | 1 балл – за правильный ответ |
| 131 | a | 1 балл – за правильный ответ |
| 132 | a | 1 балл – за правильный ответ |
| 133 | a,б,в | 1 балл – за правильный ответ |
| 134 | Plotly | 1 балл – за правильный ответ |
| 135 | 1-a,2-б | 1 балл – за правильный ответ |
| 136 | б | 1 балл – за правильный ответ |
| 137 | г | 1 балл – за правильный ответ |
| 138 | б | 1 балл – за правильный ответ |
| 139 | a | 1 балл – за правильный ответ |
| 140 | a | 1 балл – за правильный ответ |
| 141 | г | 1 балл – за правильный ответ |
| 142 | a,б | 1 балл – за правильный ответ |
| 143 | срезом данных | 1 балл – за правильный ответ |
| 144 | == | 1 балл – за правильный ответ |
| 145 | в | 1 балл – за правильный ответ |
| 146 | a | 1 балл – за правильный ответ |
| 147 | a | 1 балл – за правильный ответ |
| 148 | б | 1 балл – за правильный ответ |
| 149 | в | 1 балл – за правильный ответ |
| 150 | г | 1 балл – за правильный ответ |
| 151 | a | 1 балл – за правильный ответ |
| 152 | б,в,г | 1 балл – за правильный ответ |
| 153 | show() | 1 балл – за правильный ответ |
| 154 | a | 1 балл – за правильный ответ |
| 155 | a | 1 балл – за правильный ответ |
| 156 | sort\_values | 1 балл – за правильный ответ |
| 157 | sort\_index | 1 балл – за правильный ответ |
| 158 | 1-a,2-б,3-в,4-г,5-д | 1 балл – за правильный ответ |
| 159 | б | 1 балл – за правильный ответ |
| 160 | a | 1 балл – за правильный ответ |
| 161 | a | 1 балл – за правильный ответ |
| 162 | a | 1 балл – за правильный ответ |
| 163 | в | 1 балл – за правильный ответ |
| 164 | в | 1 балл – за правильный ответ |
| 165 | a,в | 1 балл – за правильный ответ |
| 166 | отображение нескольких графиков в табличном виде | 1 балл – за правильный ответ |
| 167 | a | 1 балл – за правильный ответ |
| 168 | a | 1 балл – за правильный ответ |
| 169 | a | 1 балл – за правильный ответ |
| 170 | тест тренда Манна-Кендалла | 1 балл – за правильный ответ |
| 171 | 1-a,2-б,3-в,4-г,5-д,6-е,7-ж,8-з,9-и | 1 балл – за правильный ответ |
| 172 | 1-a,2-б,3-в | 1 балл – за правильный ответ |
| 173 | б | 1 балл – за правильный ответ |
| 174 | в | 1 балл – за правильный ответ |
| 175 | a | 1 балл – за правильный ответ |
| 176 | г | 1 балл – за правильный ответ |
| 177 | a | 1 балл – за правильный ответ |
| 178 | б,в,г | 1 балл – за правильный ответ |
| 179 | a,в | 1 балл – за правильный ответ |
| 180 | в | 1 балл – за правильный ответ |
| 181 | в | 1 балл – за правильный ответ |
| 182 | a | 1 балл – за правильный ответ |
| 183 | a | 1 балл – за правильный ответ |
| 184 | г | 1 балл – за правильный ответ |
| 185 | a,б | 1 балл – за правильный ответ |
| 186 | a,б,в | 1 балл – за правильный ответ |
| 187 | б | 1 балл – за правильный ответ |
| 188 | a | 1 балл – за правильный ответ |
| 189 | б | 1 балл – за правильный ответ |
| 190 | г | 1 балл – за правильный ответ |
| 191 | a | 1 балл – за правильный ответ |
| 192 | б | 1 балл – за правильный ответ |
| 193 | в | 1 балл – за правильный ответ |
| 194 | непараметрическому | 1 балл – за правильный ответ |
| 195 | критерия знакового ранга Уилкоксона | 1 балл – за правильный ответ |
| 196 | a | 1 балл – за правильный ответ |
| 197 | a | 1 балл – за правильный ответ |
| 198 | a | 1 балл – за правильный ответ |
| 199 | 1-a,2-б,3-в,4-г | 1 балл – за правильный ответ |
| 200 | 1-a,2-б,3-в,4-г | 1 балл – за правильный ответ |

Правила обработки результатов теста: тест считается выполненным при правильном выполнении обучающимся не менее 70 % заданий.

# 6 Задания для проверки умений

**Задание № 1**

**(обязательное)**

Предмет оценки (умение):

Уметь загружать файлы большого объема для дальнейшей обработки при помощи библиотек Pandas и Numpy в различных форматах

Описание ситуации и постановка задачи:

**Необходимо:**

1) Запустить jupyter notebook.

2) Импортировать библиотеки, необходимые для загрузки и анализа данных.

3) Определить расширение (тип/формат) и размер файла loco\_2\_corr.tsv.

4) Выбрать способ загрузки файла (функцию библиотеки) loco\_2\_corr.tsv, получите DataFrame.

5) Рассчитать общее количество строк в полученном DataFrame.

6) Рассчитать общее количество колонок в рассчитанном DataFrame.

7) Рассчитать количество строк для серии 135 (поле loco\_2.ser\_loc).

Место выполнения: Учебный портал/учебная аудитория.

Источники информации для выполнения:

Таблица 7 – Источники информации для выполнения задания:

|  |
| --- |
| **1 Учебно-методическая документация** |
| 1.1 Конспект лекций |
| 1.2 Методические указания к организации и проведению практических занятий |
| **2 Список используемых источников** |
| 2.1 Учебники, монографии |
| 2.1.1 С.В. Рындина Базовые возможности языка Python для анализа данных. — Пенза : Издательство ПГУ, 2022. — 72 с. |
| **3 Информационное обеспечение** |
| 3.1 http://library.miit.ru/ |

Дополнительные материалы:  
1. Описание полей файла loco\_2\_corr.tsv,  
URL: https://constructor-api.emiit.ru/tasks/287/additional\_files/156/download  
2. Датасет loco\_2\_corr,  
URL: https://constructor-api.emiit.ru/tasks/287/additional\_files/157/download

Максимальное время выполнения: 45 минут.

МТО для выполнения задания:

Таблица 8 – Состав МТО

|  |  |  |  |
| --- | --- | --- | --- |
| **1 Помещения** | | | |
| 1.1 Для практических занятий | | | |
| 1.1.1 Аудитория для практических занятий | 1 | шт. | доска, средства отображения данных на большой экран, доступ в интернет |
| **2 Мебель** | | | |
| 2.1 Учебных аудиторий | | | |
| 2.1.1 Стол | 30 | шт. | посадочные места по количеству обучающихся |
| 2.1.2 Стул | 30 | шт. | посадочные места по количеству обучающихся |
| **3 Оборудование** | | | |
| 3.1 Учебной аудитории для проведения занятий лекционного и семинарского типа | | | |
| 3.1.1 Персональный компьютер преподавателя с веб-камерой, доступ к сети Интернет | 1 | шт. | компьютер с лицензионным программным обеспечением, должен удовлетворять минимальным системным требованиям специализированного ПО, обеспечивать возможность отображения информации на большой экран |
| 3.1.2 Персональные компьютеры для обучающихся с веб-камерой, доступ к сети Интернет | 30 | шт. | компьютер с лицензионным программным обеспечением, должен удовлетворять минимальным системным требованиям специализированного ПО. Количество компьютеров по количеству обучающихся |
| 3.1.3 Мультимедиа-комплекс | 1 | шт. | возможность отображения информации на большой экран |
| 3.1.4 Периферийное оборудование для ПК (принтер, сканер, сетевое оборудование, интерактивная доска) | 1 | шт. |  |
| **4 Расходные материалы** | | | |
| 4.1 Бумага | 1 | уп. |  |
| 4.2 Ручки | 1 | уп. |  |
| **5 Программное обеспечение** | | | |
| 5.1 Офисное | | | |
| 5.1.1 Лицензионное программное обеспечение (Microsoft Office) | 1 | шт. | программное обеспечение по количеству персональных компьютеров |
| 5.2 Специализированное | | | |
| 5.2.1 Среда разработки Anaconda Python | 1 | шт. | Anaconda Python включает в себя среду разработки Jupyter Notebook, библиотеки NumPy, SciPy, Pandas, Matplotlib, Scikit-Learn и многие другие популярные пакеты Python. |
| 5.2.2 Тестирующий программный комплекс системы | 1 | шт. | создание библиотеки контрольных вопросов различных типов; формирование тестов на основе библиотеки вопросов (с возможностью случайной выборки, ограничениями по времени и другими параметрами); включение тестов в состав электронных курсов; назначение тестов в качестве самостоятельных оценочных процедур; детальная аналитика по итогам тестирования |
| **6 Иные** | | | |
| 6.1 Информационно-телекоммуникационные сети | 1 |  | обеспечивают передачу по линиям связи учебной информации и обратную связь между обучающимся и средством обучения |
| 6.2 Библиотека электронных образовательных ресурсов | 1 |  | доступ к электронным образовательным ресурсам, контроль знаний обучающихся (тестирование); персональные компьютеры, программа для создания интерактивных и мультимедийных электронных образовательных ресурсов |

**Критерии оценки**

Таблица 9 – Критерии оценки

| **Предмет оценки** | **Объект оценки** | **Критерий оценки** |
| --- | --- | --- |
| Уметь загружать файлы большого объема для дальнейшей обработки при помощи библиотек Pandas и Numpy в различных форматах | 1. Файл loco\_2.tsv 2. Описание полей файла loco\_2.tsv | Соответствие модельному ответу |

**Модельный ответ**

Таблица 10 – Модельный ответ

| **Объект оценки** | **Модельный ответ (индикатор)** |
| --- | --- |
| 1. Файл loco\_2.tsv 2. Описание полей файла loco\_2.tsv | Формат файла: tsv Общее количество строк: 11390506 Общее количество столбцов: 16 Общее количество строк для локомотива серии 135: 542390 |

**Задание № 2**

**(обязательное)**

Предмет оценки (умение):

Уметь проводить разведочный анализ данных при помощи базовых библиотек Python в условиях некачественных данных

Описание ситуации и постановка задачи:

Необходимо:

1. Сформировать новый DataFrame: срез данных по критерию «тип отказа» (loco\_2.fail\_type) равен ‘производственный при ремонте’.
2. Определить количество пропущенных значений по столбцам в новом DataFrame.
3. Рассчитать общее количество строчек в массиве данных.
4. Определить столбец (столбцы), доля пропусков в которых самая большая.
5. Провести логический анализ и определить столбец, пропуски в котором делают невозможным провести анализ последствий отказов в различных местах возникновения отказа.
6. Удалить строки, в которых есть пропуски в столбце, выделенном по п.5.
7. Определить количество строк, в которых есть пропуски. Определить в записях с каким индексом из следующего списка содержится 4 пропуска (индекс строки): 6275915, 11390499, 11390500, 1.
8. Разработать стратегию обработки пропусков для различных столбцов, используя стратегию удаления строк только для поля loco\_2.zns\_loc (после выполнения п.6).
9. Восстановить пропущенные значения, при условии, что удалять строки можно только в том случае, если есть пропуск по полю loco\_2.zns\_loc (после выполнения п.6).
10. Определить количество итоговых строк в массиве, количество удаленных строк, долю удаленных строк (округление до второго знака после запятой).

Место выполнения: Учебный портал/учебная аудитория.

Источники информации для выполнения:

Таблица 11 – Источники информации для выполнения задания:

|  |
| --- |
| **1 Учебно-методическая документация** |
| 1.1 Конспект лекций |
| 1.2 Методические указания к организации и проведению практических занятий |
| **2 Список используемых источников** |
| 2.1 Учебники, монографии |
| 2.1.1 С.В. Рындина Базовые возможности языка Python для анализа данных. — Пенза : Издательство ПГУ, 2022. — 72 с. |
| 2.1.2 Уэс Маккинни Python и анализ данных: Первичная обработка данных с применением pandas, NumPy и Jupiter / пер. с англ. А. А. Слинкина. 3-е изд. . — М : МК Пресс, 2023. — 536 с. |
| **3 Информационное обеспечение** |
| 3.1 http://library.miit.ru/ |

Максимальное время выполнения: 24 минут.

МТО для выполнения задания:

Таблица 12 – Состав МТО

|  |  |  |  |
| --- | --- | --- | --- |
| **1 Помещения** | | | |
| 1.1 Для практических занятий | | | |
| 1.1.1 Аудитория для практических занятий | 1 | шт. | доска, средства отображения данных на большой экран, доступ в интернет |
| **2 Мебель** | | | |
| 2.1 Учебных аудиторий | | | |
| 2.1.1 Стол | 30 | шт. | посадочные места по количеству обучающихся |
| 2.1.2 Стул | 30 | шт. | посадочные места по количеству обучающихся |
| **3 Оборудование** | | | |
| 3.1 Учебной аудитории для проведения занятий лекционного и семинарского типа | | | |
| 3.1.1 Персональный компьютер преподавателя с веб-камерой, доступ к сети Интернет | 1 | шт. | компьютер с лицензионным программным обеспечением, должен удовлетворять минимальным системным требованиям специализированного ПО, обеспечивать возможность отображения информации на большой экран |
| 3.1.2 Персональные компьютеры для обучающихся с веб-камерой, доступ к сети Интернет | 30 | шт. | компьютер с лицензионным программным обеспечением, должен удовлетворять минимальным системным требованиям специализированного ПО. Количество компьютеров по количеству обучающихся |
| 3.1.3 Мультимедиа-комплекс | 1 | шт. | возможность отображения информации на большой экран |
| 3.1.4 Периферийное оборудование для ПК (принтер, сканер, сетевое оборудование, интерактивная доска) | 1 | шт. |  |
| **4 Расходные материалы** | | | |
| 4.1 Бумага | 1 | уп. |  |
| 4.2 Ручки | 1 | уп. |  |
| **5 Программное обеспечение** | | | |
| 5.1 Офисное | | | |
| 5.1.1 Лицензионное программное обеспечение (Microsoft Office) | 1 | шт. | программное обеспечение по количеству персональных компьютеров |
| 5.2 Специализированное | | | |
| 5.2.1 Среда разработки Anaconda Python | 1 | шт. | Anaconda Python включает в себя среду разработки Jupyter Notebook, библиотеки NumPy, SciPy, Pandas, Matplotlib, Scikit-Learn и многие другие популярные пакеты Python. |
| 5.2.2 Тестирующий программный комплекс системы | 1 | шт. | создание библиотеки контрольных вопросов различных типов; формирование тестов на основе библиотеки вопросов (с возможностью случайной выборки, ограничениями по времени и другими параметрами); включение тестов в состав электронных курсов; назначение тестов в качестве самостоятельных оценочных процедур; детальная аналитика по итогам тестирования |
| **6 Иные** | | | |
| 6.1 Информационно-телекоммуникационные сети | 1 |  | обеспечивают передачу по линиям связи учебной информации и обратную связь между обучающимся и средством обучения |
| 6.2 Библиотека электронных образовательных ресурсов | 1 |  | доступ к электронным образовательным ресурсам, контроль знаний обучающихся (тестирование); персональные компьютеры, программа для создания интерактивных и мультимедийных электронных образовательных ресурсов |

**Критерии оценки**

Таблица 13 – Критерии оценки

| **Предмет оценки** | **Объект оценки** | **Критерий оценки** |
| --- | --- | --- |
| Уметь проводить разведочный анализ данных при помощи базовых библиотек Python в условиях некачественных данных | DataFrame, полученный по итогам задания 1. | Соответствие модельному ответу |

**Модельный ответ**

Таблица 14 – Модельный ответ

| **Объект оценки** | **Модельный ответ (индикатор)** |
| --- | --- |
| DataFrame, полученный по итогам задания 1. | Количество пропущенных значений по столбцу loco2.znsloc: 2507262 Доля пропущенных значений по столбцу loco2.znsloc (до второго знака после запятой): 0,43 Название столбцов с наибольшей долей пропусков (через запятую с проблем): loco2.znsloc, loco2.prsloc Название столбца, который делает невозможным анализ последствий отказов: loco2.failconsequence Индекс строки, в которой содержится 4 пропуска: 6275915 количество строк в массиве: 5772009 количество удаленных строк: 108054 доля удаленных строк: 0.018 |

**Задание № 3**

**(обязательное)**

Предмет оценки (умение):

Уметь проводить разведочный анализ данных при помощи базовых библиотек Python в условиях некачественных данных

Описание ситуации и постановка задачи:

**Необходимо:**

1. Привести значение в поле date\_end к формату даты и создать новые признаки: месяц в году, год, день, день в году (новые признаки окончания даты расследования отказа).
2. Определить в каком месяце чаще всего завершается расследование отказов.
3. Осуществить выбор локомотива и просмотр количества дней датами окончания расследования случаев отказа:
4. Сделайте срез данных для следующих характеристик:

* (loco\_2\_new['loco\_2.ser\_loc'] == 135)&(loco\_2\_new['loco\_2.prs\_loc'] == 2)
* &(loco\_2\_new['loco\_2.zns\_8zn'] == 739)
* &(loco\_2\_new["год"] ==2019)

1. Сбросьте индексы.
2. Удалите неинформативные столбцы (всего 3 неинформативных столбца два из них: 'index', 'loco\_2.ctime').
3. Отсортируйте массив по номеру дня расследования в году.
4. Рассчитайте количество дней между окончаниями расследования.
5. Определите в какие даты разница в количестве дней между расследования была больше 1.

Место выполнения: Учебный портал/учебная аудитория.

Источники информации для выполнения:

Таблица 15 – Источники информации для выполнения задания:

|  |
| --- |
| **1 Учебно-методическая документация** |
| 1.1 Конспект лекций |
| 1.2 Методические указания к организации и проведению практических занятий |
| **2 Список используемых источников** |
| 2.1 Учебники, монографии |
| 2.1.1 С.В. Рындина Базовые возможности языка Python для анализа данных. — Пенза : Издательство ПГУ, 2022. — 72 с. |
| 2.1.2 Уэс Маккинни Python и анализ данных: Первичная обработка данных с применением pandas, NumPy и Jupiter / пер. с англ. А. А. Слинкина. 3-е изд. . — М : МК Пресс, 2023. — 536 с. |
| **3 Информационное обеспечение** |
| 3.1 http://library.miit.ru/ |

Максимальное время выполнения: 45 минут.

МТО для выполнения задания:

Таблица 16 – Состав МТО

|  |  |  |  |
| --- | --- | --- | --- |
| **1 Помещения** | | | |
| 1.1 Для практических занятий | | | |
| 1.1.1 Аудитория для практических занятий | 1 | шт. | доска, средства отображения данных на большой экран, доступ в интернет |
| **2 Мебель** | | | |
| 2.1 Учебных аудиторий | | | |
| 2.1.1 Стол | 30 | шт. | посадочные места по количеству обучающихся |
| 2.1.2 Стул | 30 | шт. | посадочные места по количеству обучающихся |
| **3 Оборудование** | | | |
| 3.1 Учебной аудитории для проведения занятий лекционного и семинарского типа | | | |
| 3.1.1 Персональный компьютер преподавателя с веб-камерой, доступ к сети Интернет | 1 | шт. | компьютер с лицензионным программным обеспечением, должен удовлетворять минимальным системным требованиям специализированного ПО, обеспечивать возможность отображения информации на большой экран |
| 3.1.2 Персональные компьютеры для обучающихся с веб-камерой, доступ к сети Интернет | 30 | шт. | компьютер с лицензионным программным обеспечением, должен удовлетворять минимальным системным требованиям специализированного ПО. Количество компьютеров по количеству обучающихся |
| 3.1.3 Мультимедиа-комплекс | 1 | шт. | возможность отображения информации на большой экран |
| 3.1.4 Периферийное оборудование для ПК (принтер, сканер, сетевое оборудование, интерактивная доска) | 1 | шт. |  |
| **4 Расходные материалы** | | | |
| 4.1 Бумага | 1 | уп. |  |
| 4.2 Ручки | 1 | уп. |  |
| **5 Программное обеспечение** | | | |
| 5.1 Офисное | | | |
| 5.1.1 Лицензионное программное обеспечение (Microsoft Office) | 1 | шт. | программное обеспечение по количеству персональных компьютеров |
| 5.2 Специализированное | | | |
| 5.2.1 Среда разработки Anaconda Python | 1 | шт. | Anaconda Python включает в себя среду разработки Jupyter Notebook, библиотеки NumPy, SciPy, Pandas, Matplotlib, Scikit-Learn и многие другие популярные пакеты Python. |
| 5.2.2 Тестирующий программный комплекс системы | 1 | шт. | создание библиотеки контрольных вопросов различных типов; формирование тестов на основе библиотеки вопросов (с возможностью случайной выборки, ограничениями по времени и другими параметрами); включение тестов в состав электронных курсов; назначение тестов в качестве самостоятельных оценочных процедур; детальная аналитика по итогам тестирования |
| **6 Иные** | | | |
| 6.1 Информационно-телекоммуникационные сети | 1 |  | обеспечивают передачу по линиям связи учебной информации и обратную связь между обучающимся и средством обучения |
| 6.2 Библиотека электронных образовательных ресурсов | 1 |  | доступ к электронным образовательным ресурсам, контроль знаний обучающихся (тестирование); персональные компьютеры, программа для создания интерактивных и мультимедийных электронных образовательных ресурсов |

**Критерии оценки**

Таблица 17 – Критерии оценки

| **Предмет оценки** | **Объект оценки** | **Критерий оценки** |
| --- | --- | --- |
| Уметь проводить разведочный анализ данных при помощи базовых библиотек Python в условиях некачественных данных | DataFrame, полученный по итогам задания 2. | Соответствие модельному ответу |

**Модельный ответ**

Таблица 18 – Модельный ответ

| **Объект оценки** | **Модельный ответ (индикатор)** |
| --- | --- |
| DataFrame, полученный по итогам задания 2. | Номер месяца, в котором чаще всего завершается расследование отказа: 3 Количество строк по результату выполнения пункта 3.f: 6. |

**Задание № 4**

**(обязательное)**

Предмет оценки (умение):

Уметь подготавливать данные для решения бизнес-задач при помощи инструментов группировки, агрегации и визуализации данных в условиях некачественных данных

Описание ситуации и постановка задачи:

**Необходимо:**

1. Подготовить данные:
   1. Подсчитать число записей до удаления дубликатов.
   2. Удалить неинформативные колонки (две) и удалите дубликаты.
   3. Подсчитать количество записей после удаления дубликатов.
2. Сгруппировать данные по набору столбцов ['год', 'loco\_2.ser\_loc', 'loco\_2.prs\_loc', 'loco\_2.zns\_8zn'] и рассчитать количество расследований отказов.
3. Результаты, полученные по п.2 перевести в формат DataFrame, до сброса индексов. Задать новые имена колонок: ['количество расследований'].
4. Определить, наблюдается ли проблема со значениями по полю loco\_2.zns\_8zn.
5. Сделать срез данных, убрав локомотивы с zns\_8zn равным 0 и отсортировать полученный массив по полям ['loco\_2.ser\_loc','loco\_2.prs\_loc', 'loco\_2.zns\_8zn', 'loco\_2.fail\_place', 'loco\_2.date\_end']
6. Сгруппировать данные так, чтобы рассчитать количество дней между окончаниями расследований случаев отказов в разрезе места отказа ('loco\_2.fail\_place') и рассчитать это количество. Привести данные к таком формату, чтобы в столбце было записано целое число (количество дней) без дополнительных символов.
7. Добавить к срезу данных, полученных по п.5. Новому признаку дайте имя diff.
8. На основе DataFrame, полученного по результатам п.7 сформировать новый DataFrame и рассчитать среднее число дней между расследованиями.
9. Определить среднее число дней между отказами электрооборудования силовой цепи, вспомогательных цепей и цепей управления локомотива, МВПС для локомотивов серии 694 в марте 2021 года: 71.
10. Определить количество месяцев для локомотива серии 694, когда среднее число дней между расследованиями отказов дизеля и дизельного оборудование локомотива, МВПС.

Место выполнения: Учебный портал/учебная аудитория.

Источники информации для выполнения:

Таблица 19 – Источники информации для выполнения задания:

|  |
| --- |
| **1 Учебно-методическая документация** |
| 1.1 Конспект лекций |
| 1.2 Методические указания к организации и проведению практических занятий |
| **2 Список используемых источников** |
| 2.1 Нормативные правовые акты, нормативная техническая документация, иная документация |
| 2.1.1 Генерация признаков из временных рядов / NewTechAudit : [профессиональное сообщество]. — Текст : электронный // Хабр : [сайт] — 2021. — 22 окт. — URL: https://habr.com/ru/articles/584896 (дата обращения: 20.11.2023). |
| 2.1.2 Python для анализа данных : операции в pandas. — Текст : электронный // Физтех.Статистика : [сайт] — URL: https://mipt-stats.gitlab.io/courses/python/10\_pandas2.html (дата обращения: 22.11.2023). |
| 2.2 Учебники, монографии |
| 2.2.1 Уэс Маккинни Python и анализ данных: Первичная обработка данных с применением pandas, NumPy и Jupiter / пер. с англ. А. А. Слинкина. 3-е изд. . — М : МК Пресс, 2023. — 536 с. |
| **3 Информационное обеспечение** |
| 3.1 http://library.miit.ru/ |

Максимальное время выполнения: 45 минут.

МТО для выполнения задания:

Таблица 20 – Состав МТО

|  |  |  |  |
| --- | --- | --- | --- |
| **1 Помещения** | | | |
| 1.1 Для практических занятий | | | |
| 1.1.1 Аудитория для практических занятий | 1 | шт. | доска, средства отображения данных на большой экран, доступ в интернет |
| **2 Мебель** | | | |
| 2.1 Учебных аудиторий | | | |
| 2.1.1 Стол | 30 | шт. | посадочные места по количеству обучающихся |
| 2.1.2 Стул | 30 | шт. | посадочные места по количеству обучающихся |
| **3 Оборудование** | | | |
| 3.1 Учебной аудитории для проведения занятий лекционного и семинарского типа | | | |
| 3.1.1 Персональный компьютер преподавателя с веб-камерой, доступ к сети Интернет | 1 | шт. | компьютер с лицензионным программным обеспечением, должен удовлетворять минимальным системным требованиям специализированного ПО, обеспечивать возможность отображения информации на большой экран |
| 3.1.2 Персональные компьютеры для обучающихся с веб-камерой, доступ к сети Интернет | 30 | шт. | компьютер с лицензионным программным обеспечением, должен удовлетворять минимальным системным требованиям специализированного ПО. Количество компьютеров по количеству обучающихся |
| 3.1.3 Мультимедиа-комплекс | 1 | шт. | возможность отображения информации на большой экран |
| 3.1.4 Периферийное оборудование для ПК (принтер, сканер, сетевое оборудование, интерактивная доска) | 1 | шт. |  |
| **4 Расходные материалы** | | | |
| 4.1 Бумага | 1 | уп. |  |
| 4.2 Ручки | 1 | уп. |  |
| **5 Программное обеспечение** | | | |
| 5.1 Офисное | | | |
| 5.1.1 Лицензионное программное обеспечение (Microsoft Office) | 1 | шт. | программное обеспечение по количеству персональных компьютеров |
| 5.2 Специализированное | | | |
| 5.2.1 Среда разработки Anaconda Python | 1 | шт. | Anaconda Python включает в себя среду разработки Jupyter Notebook, библиотеки NumPy, SciPy, Pandas, Matplotlib, Scikit-Learn и многие другие популярные пакеты Python. |
| 5.2.2 Тестирующий программный комплекс системы | 1 | шт. | создание библиотеки контрольных вопросов различных типов; формирование тестов на основе библиотеки вопросов (с возможностью случайной выборки, ограничениями по времени и другими параметрами); включение тестов в состав электронных курсов; назначение тестов в качестве самостоятельных оценочных процедур; детальная аналитика по итогам тестирования |
| **6 Иные** | | | |
| 6.1 Информационно-телекоммуникационные сети | 1 |  | обеспечивают передачу по линиям связи учебной информации и обратную связь между обучающимся и средством обучения |
| 6.2 Библиотека электронных образовательных ресурсов | 1 |  | доступ к электронным образовательным ресурсам, контроль знаний обучающихся (тестирование); персональные компьютеры, программа для создания интерактивных и мультимедийных электронных образовательных ресурсов |

**Критерии оценки**

Таблица 21 – Критерии оценки

| **Предмет оценки** | **Объект оценки** | **Критерий оценки** |
| --- | --- | --- |
| Уметь подготавливать данные для решения бизнес-задач при помощи инструментов группировки, агрегации и визуализации данных в условиях некачественных данных | DataFrame, полученный по итогам задания 3. | Соответствие модельному ответу |

**Модельный ответ**

Таблица 22 – Модельный ответ

| **Объект оценки** | **Модельный ответ (индикатор)** |
| --- | --- |
| DataFrame, полученный по итогам задания 3. | 1. Записей до удаления дубликатов: 5663955 2. Записей после дубликатов: 38172 3. Число дней между отказами электрооборудования силовой цепи, вспомогательных цепей и цепей управления локомотива, МВПС для локомотивов серии 694 в марте 2021 года: 71. 4. Количество месяцев для локомотива серии 694, когда среднее число дней между расследованиями отказов Дизеля и дизельного оборудование локомотива, МВПС: 2 |

**Задание № 5**

**(обязательное)**

Предмет оценки (умение):

Уметь подготавливать данные для решения бизнес-задач при помощи инструментов группировки, агрегации и визуализации данных в условиях некачественных данных

Описание ситуации и постановка задачи:

Необходимо:

1. Построить график среднего числа дней между расследования по месяцам для локомотива серии 108, определить по графику количество пиков, превышающих значение 400.
2. Передать гиперпараметру facet\_col значение ["loco\_2.fail\_place"], определить сколько получится рисунков. Для корректного отображения графиков передать функции px.line() следующие параметры: facet\_col\_wrap=2, width=2000, height=1500.
3. Сформировать DataFrame с количеством отказов по сериям по месяцам.
4. Построить графики количества отказов (используя DataFrame, полученный по результатам выполнения п. 4.1). Определить, на какой серии локомотива в течении всего периода наблюдения отказов было больше, чем на других сериях.

Место выполнения: Учебный портал/учебная аудитория.

Источники информации для выполнения:

Таблица 23 – Источники информации для выполнения задания:

|  |
| --- |
| **1 Учебно-методическая документация** |
| 1.1 Конспект лекций |
| 1.2 Методические указания к организации и проведению практических занятий |
| **2 Список используемых источников** |
| 2.1 Нормативные правовые акты, нормативная техническая документация, иная документация |
| 2.1.1 Генерация признаков из временных рядов / NewTechAudit : [профессиональное сообщество]. — Текст : электронный // Хабр : [сайт] — 2021. — 22 окт. — URL: https://habr.com/ru/articles/584896 (дата обращения: 20.11.2023). |
| 2.1.2 Python для анализа данных : операции в pandas. — Текст : электронный // Физтех.Статистика : [сайт] — URL: https://mipt-stats.gitlab.io/courses/python/10\_pandas2.html (дата обращения: 22.11.2023). |
| 2.2 Учебники, монографии |
| 2.2.1 Уэс Маккинни Python и анализ данных: Первичная обработка данных с применением pandas, NumPy и Jupiter / пер. с англ. А. А. Слинкина. 3-е изд. . — М : МК Пресс, 2023. — 536 с. |
| **3 Информационное обеспечение** |
| 3.1 http://library.miit.ru/ |

Максимальное время выполнения: 45 минут.

МТО для выполнения задания:

Таблица 24 – Состав МТО

|  |  |  |  |
| --- | --- | --- | --- |
| **1 Помещения** | | | |
| 1.1 Для практических занятий | | | |
| 1.1.1 Аудитория для практических занятий | 1 | шт. | доска, средства отображения данных на большой экран, доступ в интернет |
| **2 Мебель** | | | |
| 2.1 Учебных аудиторий | | | |
| 2.1.1 Стол | 30 | шт. | посадочные места по количеству обучающихся |
| 2.1.2 Стул | 30 | шт. | посадочные места по количеству обучающихся |
| **3 Оборудование** | | | |
| 3.1 Учебной аудитории для проведения занятий лекционного и семинарского типа | | | |
| 3.1.1 Персональный компьютер преподавателя с веб-камерой, доступ к сети Интернет | 1 | шт. | компьютер с лицензионным программным обеспечением, должен удовлетворять минимальным системным требованиям специализированного ПО, обеспечивать возможность отображения информации на большой экран |
| 3.1.2 Персональные компьютеры для обучающихся с веб-камерой, доступ к сети Интернет | 30 | шт. | компьютер с лицензионным программным обеспечением, должен удовлетворять минимальным системным требованиям специализированного ПО. Количество компьютеров по количеству обучающихся |
| 3.1.3 Мультимедиа-комплекс | 1 | шт. | возможность отображения информации на большой экран |
| 3.1.4 Периферийное оборудование для ПК (принтер, сканер, сетевое оборудование, интерактивная доска) | 1 | шт. |  |
| **4 Расходные материалы** | | | |
| 4.1 Бумага | 1 | уп. |  |
| 4.2 Ручки | 1 | уп. |  |
| **5 Программное обеспечение** | | | |
| 5.1 Офисное | | | |
| 5.1.1 Лицензионное программное обеспечение (Microsoft Office) | 1 | шт. | программное обеспечение по количеству персональных компьютеров |
| 5.2 Специализированное | | | |
| 5.2.1 Среда разработки Anaconda Python | 1 | шт. | Anaconda Python включает в себя среду разработки Jupyter Notebook, библиотеки NumPy, SciPy, Pandas, Matplotlib, Scikit-Learn и многие другие популярные пакеты Python. |
| 5.2.2 Тестирующий программный комплекс системы | 1 | шт. | создание библиотеки контрольных вопросов различных типов; формирование тестов на основе библиотеки вопросов (с возможностью случайной выборки, ограничениями по времени и другими параметрами); включение тестов в состав электронных курсов; назначение тестов в качестве самостоятельных оценочных процедур; детальная аналитика по итогам тестирования |
| **6 Иные** | | | |
| 6.1 Информационно-телекоммуникационные сети | 1 |  | обеспечивают передачу по линиям связи учебной информации и обратную связь между обучающимся и средством обучения |
| 6.2 Библиотека электронных образовательных ресурсов | 1 |  | доступ к электронным образовательным ресурсам, контроль знаний обучающихся (тестирование); персональные компьютеры, программа для создания интерактивных и мультимедийных электронных образовательных ресурсов |

**Критерии оценки**

Таблица 25 – Критерии оценки

| **Предмет оценки** | **Объект оценки** | **Критерий оценки** |
| --- | --- | --- |
| Уметь подготавливать данные для решения бизнес-задач при помощи инструментов группировки, агрегации и визуализации данных в условиях некачественных данных | DataFrame, полученный по итогам задания 4. | Соответствие модельному ответу |

**Модельный ответ**

Таблица 26 – Модельный ответ

| **Объект оценки** | **Модельный ответ (индикатор)** |
| --- | --- |
| DataFrame, полученный по итогам задания 4. | 1. Количество пиков, превышающих значение 400: 3 2. Количество рисунков: 8 3. Серия локомотива в течении всего периода наблюдения отказов было больше, чем на других сериях: 241 |

**Задание № 6**

**(обязательное)**

Предмет оценки (умение):

Уметь формировать статистические гипотезы при помощи представления данных в виде временных рядов в условиях больших данных, характеризующихся множеством параметров

Описание ситуации и постановка задачи:

Необходимо:

1. Инсталлировать библиотеку pymannkendall.
2. Импортировать библиотеку pymannkendall.
3. Исключить из DataFrame записи с loco\_2.zns\_8zn равным 0.
4. Сформировать (после п.3) два DataFrame:
   1. с количеством отказов по месяцам, возникающих в Электрооборудовании силовой цепи, вспомогательных цепей и цепей управления локомотива, МВПС (см. полную формулировку по полю loco\_2.fail\_place).
   2. с количеством отказов по месяцам, возникающих в Тяговые электрические машины локомотива, МВПС (см. полную формулировку по полю loco\_2.fail\_place).
5. Для каждого из полученных по п. 4 DataFrame определить наличие трендов в разрезе серий локомотивов. Результаты записать в DataFrame со столбцами ‘серия’, ‘статус динамики отказа’.
6. Для каждого из полученных по п. 4 DataFrame определить наличие среднего числа отказов и добавить эту информацию к DataFrame, полученным по результатам п. 5.
7. Инсталлировать библиотеку scipy.
8. Из библиотеки scipy импортировать модуль wilcoxon.
9. Для DataFrame, полученного по результатам п. 4.а рассчитать значение p-value по критерию Виклонсана при сравнении временных рядов за 2020 и 2021 года для серии 135.
10. По Критерию wilcoxon сравнить значимость изменения числа отказов за 2019 и 2020 год для каждого из DataFrame, полученного по п.4.
11. Объединить результаты пунктов 6 и 10.
12. Определить перечень серий, для которых наблюдается рост числа отказов, а мероприятия не эффективны как для Электрооборудования так и для тяговых электрических машин.
13. Провести анализ для серии 234 и определить, на основе какого критерия (wilcoxon, pymannkendall или среднее число отказов) можно установить, по которому можно ранжировать место возникновения отказа по результатам проведенных расчетов.

Место выполнения: Учебный портал/учебная аудитория..

Источники информации для выполнения:

Таблица 27 – Источники информации для выполнения задания:

|  |
| --- |
| **1 Учебно-методическая документация** |
| 1.1 Конспект лекций |
| 1.2 Методические указания к организации и проведению практических занятий |
| **2 Список используемых источников** |
| 2.1 Нормативные правовые акты, нормативная техническая документация, иная документация |
| 2.1.1 Задойный Алексей. Шпаргалка по визуализации данных в Python с помощью Plotly / А. Задойный. — Текст : электронный // Хабр : [сайт]. — 2020. — 25 мая. — URL: https://habr.com/ru/articles/502958 (дата обращения: 13.10.2023). |
| 2.1.2 Как выполнить тест тренда Манна-Кендалла в Python / Редакция Кодкампа. — Текст: электронный // Кодкамп : онлайн-институт цифровых навыков : [сайт]. — 2022. — 17 авг.— URL: https://www.codecamp.ru/blog/mann-kendall-test-python (дата обращения: 10.10.2023). |
| 2.1.3 Роман Котюбеев. Диаграмма рассеяния и столбчатая диаграмма в Plotly Express / Р. Котюбеев. — Текст : электронный // python-school : [сайт]. — 2023. — 21 окт. — URL:https://python-school.ru/blog/визуализация-данных/plotly-express (дата обращения: 30.10.2023). |
| 2.2 Учебники, монографии |
| 2.2.1 Бизли Д. Python. Подробный справочник. – пер. с англ.. — СПб : Символ-Плюс, 2010. — 864 с. |
| 2.2.2 Уэс Маккинни Python и анализ данных: Первичная обработка данных с применением pandas, NumPy и Jupiter / пер. с англ. А. А. Слинкина. 3-е изд. . — М : МК Пресс, 2023. — 536 с. |
| 2.2.3 ГрасД. Data Science. Наука о данных с нуля: пер. с англ. - 2-е изд., перераб. и доп. — СПб : БХВ-Петербурr, 2021. — 416 с. |
| **3 Информационное обеспечение** |
| 3.1 http://library.miit.ru/ |

Максимальное время выполнения: 66 минут.

МТО для выполнения задания:

Таблица 28 – Состав МТО

|  |  |  |  |
| --- | --- | --- | --- |
| **1 Помещения** | | | |
| 1.1 Для практических занятий | | | |
| 1.1.1 Аудитория для практических занятий | 1 | шт. | доска, средства отображения данных на большой экран, доступ в интернет |
| **2 Мебель** | | | |
| 2.1 Учебных аудиторий | | | |
| 2.1.1 Стол | 30 | шт. | посадочные места по количеству обучающихся |
| 2.1.2 Стул | 30 | шт. | посадочные места по количеству обучающихся |
| **3 Оборудование** | | | |
| 3.1 Учебной аудитории для проведения занятий лекционного и семинарского типа | | | |
| 3.1.1 Персональный компьютер преподавателя с веб-камерой, доступ к сети Интернет | 1 | шт. | компьютер с лицензионным программным обеспечением, должен удовлетворять минимальным системным требованиям специализированного ПО, обеспечивать возможность отображения информации на большой экран |
| 3.1.2 Персональные компьютеры для обучающихся с веб-камерой, доступ к сети Интернет | 30 | шт. | компьютер с лицензионным программным обеспечением, должен удовлетворять минимальным системным требованиям специализированного ПО. Количество компьютеров по количеству обучающихся |
| 3.1.3 Мультимедиа-комплекс | 1 | шт. | возможность отображения информации на большой экран |
| 3.1.4 Периферийное оборудование для ПК (принтер, сканер, сетевое оборудование, интерактивная доска) | 1 | шт. |  |
| **4 Расходные материалы** | | | |
| 4.1 Бумага | 1 | уп. |  |
| 4.2 Ручки | 1 | уп. |  |
| **5 Программное обеспечение** | | | |
| 5.1 Офисное | | | |
| 5.1.1 Лицензионное программное обеспечение (Microsoft Office) | 1 | шт. | программное обеспечение по количеству персональных компьютеров |
| 5.2 Специализированное | | | |
| 5.2.1 Среда разработки Anaconda Python | 1 | шт. | Anaconda Python включает в себя среду разработки Jupyter Notebook, библиотеки NumPy, SciPy, Pandas, Matplotlib, Scikit-Learn и многие другие популярные пакеты Python. |
| 5.2.2 Тестирующий программный комплекс системы | 1 | шт. | создание библиотеки контрольных вопросов различных типов; формирование тестов на основе библиотеки вопросов (с возможностью случайной выборки, ограничениями по времени и другими параметрами); включение тестов в состав электронных курсов; назначение тестов в качестве самостоятельных оценочных процедур; детальная аналитика по итогам тестирования |
| **6 Иные** | | | |
| 6.1 Информационно-телекоммуникационные сети | 1 |  | обеспечивают передачу по линиям связи учебной информации и обратную связь между обучающимся и средством обучения |
| 6.2 Библиотека электронных образовательных ресурсов | 1 |  | доступ к электронным образовательным ресурсам, контроль знаний обучающихся (тестирование); персональные компьютеры, программа для создания интерактивных и мультимедийных электронных образовательных ресурсов |

**Критерии оценки**

Таблица 29 – Критерии оценки

| **Предмет оценки** | **Объект оценки** | **Критерий оценки** |
| --- | --- | --- |
| Уметь формировать статистические гипотезы при помощи представления данных в виде временных рядов в условиях больших данных, характеризующихся множеством параметров | DataFrame, полученный по результатам пункта 4. | Соответствие модельному ответу |

**Модельный ответ**

Таблица 30 – Модельный ответ

| **Объект оценки** | **Модельный ответ (индикатор)** |
| --- | --- |
| DataFrame, полученный по результатам пункта 4. | 1) Значение p-value по критерию Виклонсана при сравнении временных рядов за 2020 и 2021 года для серии 135 (до 5-го знака после запятой 0,000ХХ): 0,00048. 2) Серии, для которых наблюдается рост числа отказов, а мероприятия не эффективны как для электрооборудования, так и для тяговых электрических машин: 234. 3) Критерий ранжирования: среднее число отказов. |

Правила обработки результатов итоговой аттестации на проверку умений: аттестация на проверку умений включает решение практических заданий и считается пройденной при правильном выполнении обучающимся 6 (из 6) практических заданий.