

№ <имя функции>

- Назначение функции
- Параметры входа
- Параметры выхода

1. **check_duplicate_events**

- Поиск дубликатов событий с одинаковыми временными метками и типами событий.
 - ts, event, counter_id, randPAS_session_id, ip.
 - Словарь с дубликатами или сообщением об их отсутствии.
-

2. **check_session_start**

- Проверка корректности начала сессий.
 - randPAS_session_id, page_view_order_number, event_order_number, ip.
 - Словарь с ошибками или подтверждением корректности.
-

3. **check_order_relation**

- Проверка соответствия между номерами событий и просмотров страниц.
 - randPAS_session_id, event_order_number, page_view_order_number, ip, ts.
 - Словарь с ошибками или подтверждением корректности.
-

4. **check_numbering_sequence**

- Проверка последовательности нумерации событий.
 - randPAS_session_id, event_order_number, ip, ts, event.
 - Словарь с ошибками или подтверждением корректности.
-

5. **detect_location_changes**

- Анализ смены местоположения пользователей.
 - randPAS_user_passport_id, ts, ip, geo_city_id.
 - DataFrame с информацией о сменах местоположения.
-

6. **analyze_city_activity**

- Выявление аномальной активности по городам.

- geo_city_id, ts.
 - Кортеж из DataFrame с аномалиями и общей активностью.
-

7. detect_suspicious_ips

- Поиск IP с аномальным числом пользователей.
 - ip, randPAS_user_passport_id, ts.
 - DataFrame с подозрительными IP.
-

8. detect_user_activity_spikes

- Обнаружение пользователей с аномально высокой активностью.
 - randPAS_user_passport_id, ts, ip.
 - DataFrame с подозрительными пользователями.
-

9. detect_anomalous_time_windows

- Поиск аномальных временных окон активности.
 - randPAS_user_passport_id, ts.
 - DataFrame с аномальными окнами.
-

10. detect_anomalous_device_shares

- Анализ аномальных изменений в распределении устройств.
 - ts, ua_device_family.
 - DataFrame с аномалиями.
-

11. detect_anomalous_page_views

- Выявление аномальных изменений в просмотрах страниц.
 - ts, url, randPAS_user_passport_id.
 - DataFrame с аномалиями.
-

12. detect_anomalous_users

- Поиск аномальных пользователей по времени на страницах.
- randPAS_user_passport_id, url, ts.

- DataFrame с аномальными пользователями.
-

13. zscore_detector

- Обнаружение аномалий методом Z-Score.
 - Одномерный массив значений.
 - Бинарный массив.
-

14. iqr_detector

- Обнаружение аномалий методом IQR.
 - Одномерный массив значений.
 - Бинарный массив.
-

15. modified_zscore_detector

- Обнаружение аномалий устойчивым методом Z-Score.
 - Одномерный массив значений.
 - Бинарный массив.
-

16. isolation_forest_detector

- Обнаружение аномалий через Isolation Forest.
 - Одномерный массив значений.
 - Бинарный массив.
-

17. lof_detector

- Обнаружение аномалий через LOF.
 - Одномерный массив значений.
 - Бинарный массив (1 — аномалия).
-

18. percentile_detector

- Обнаружение аномалий по перцентилям.
- Одномерный массив значений.
- Бинарный массив.

19. **majority_anomaly_vote**

- Объединение результатов разных методов.
- Произвольное число бинарных массивов.
- Бинарный массив.

20. **plot_anomalies_comparison**

- Визуализация сравнения методов обнаружения аномалий.
- Исходные данные и бинарные массивы аномалий.
- График.

Классификация по методам анализа:

1. Статистические методы

№	Функция	Используемый метод	Назначение
1	check_duplicate_events	Подсчет совпадений	Поиск дубликатов
2	check_session_start	Проверка начальных значений	Валидация сессий
3	check_order_relation	Сравнение порядковых номеров	Проверка последовательности
4	check_numbering_sequence	Анализ последовательности	Проверка нумерации
5	detect_location_changes	Анализ изменений во времени	Трекинг местоположений
6	analyze_city_activity	Z-score, скользящие окна	Детекция аномалий
13	zscore_detector	Z-score	Детекция выбросов
14	iqr_detector	Межквартильный размах	Детекция выбросов
15	modified_zscore_detector	Медианный Z-score	Устойчивая детекция
18	percentile_detector	Перцентильный анализ	Детекция выбросов

2. Методы машинного обучения

12	detect_anomalous_users	Isolation Forest	Аномалии
16	isolation_forest_detector	Isolation Forest	Аномалии
17	lof_detector	Local Outlier Factor (LOF)	Аномалии

3. Временные ряды

6	analyze_city_activity	Скользящие окна (rolling windows)	Анализ активности
---	-----------------------	-----------------------------------	-------------------

9	detect_anomalous_time_windows	Анализ временных интервалов	Детекция аномалий
10	detect_anomalous_device_shares	Анализ временных трендов	Изменение долей
11	detect_anomalous_page_views	Анализ изменений во времени	Детекция аномалий

4. Сетевой анализ

7	detect_suspicious_ips	Анализ распределения по IP	Детекция подозрительных IP
8	detect_user_activity_spikes	Анализ временных интервалов	Детекция ботов

5. Ансамблевые методы

19	majority_anomaly_vote	Голосование нескольких детекторов	Комбинированная детекция
----	-----------------------	-----------------------------------	--------------------------

6. Визуализация

20	plot_anomalies_comparison	Построение графиков	Сравнение методов
----	---------------------------	---------------------	-------------------