**Домашнее задание по теме «Построение гипотез. A/B тестирование. Автотестирование гипотез»**

**Формулировка задания**

Подготовка к проведению A/B тестирования гипотезы

Для выполнения задания выполнить следующие шаги:

1. Найти данные для анализа (таблица csv, xlsx в открытом доступе)
2. Изучить данные (тема, типы столбцов, есть ли зависимость между столбцами)
3. Составить список гипотез по данным. Предположения :

* Зависимость столбцов по датам, времени и другим критериям;
* Численный показатель или критерий оценки;
* Насколько есть разница в данных по выбранному числовому критериям оценки

1. Выбрать одну из гипотез и подтвердить её или опровергнуть.
2. (Дополнительно) В качестве критериев выбрать статистические критерии. Например, критерий Стьюдента и др.

**Ожидаемый результат**

1. Файл с исходными данными;
2. Создан список гипотез по данным (примеры гипотез ниже);
3. Проверка гипотезы
4. Результат по гипотезе - подтверждение, опровержение и вывод
5. Расчеты в MS Excel, Colab/Jupiter notebook или python файл с проверкой выбранной гипотезы
6. (Дополнительно) Указан критерий проверки выбранной гипотезы.

**Примеры гипотез в области промышленности для проведения A/B тестирования**

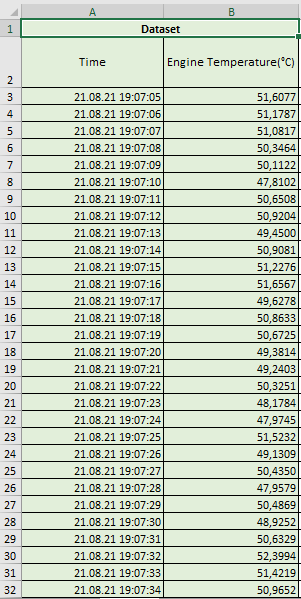
1. Добавление функции автоматической проверки качества на стадии производства увеличит производительность на 10%.
2. Оптимизация процесса производства с использованием роботизированных систем снизит затраты на 10%.
3. Обновление системы управления производством увеличит качество продукции на 20%.
4. Использование интеллектуальной системы принятия решений повысит производительность на 15%.
5. Увеличение доли устройств для автоматизации процессов производства повысит производительность на 20%.
6. Увеличение инвестиций в инструменты контроля качества продукции увеличит производительность на 25%.
7. Отказ от традиционных способов производства и переход на современные технологии увеличит качество продукции на 30%.
8. Использование машинного обучения и искусственного интеллекта для анализа данных повысит производительность на 15%.
9. Создание интегрированной системы управления снизит затраты на производство на 10%.
10. Использование программно-аппаратных средств для автоматической обработки данных увеличит производительность на 20%.

**Перечень инструментов, необходимых для реализации деятельности:**

1. MS Excel
2. Colab/Jupiter notebook
3. PyCharm

**ИСХОДНЫЕ ДАННЫЕ**

В качестве исходных данных взят датасет, содержащий в себе выгрузку данных с датчика температуры двигателя производственной установки по перекачке оборотной воды с привязкой ко времени.



**АНАЛИЗИРУЕМ ДАННЫЕ**

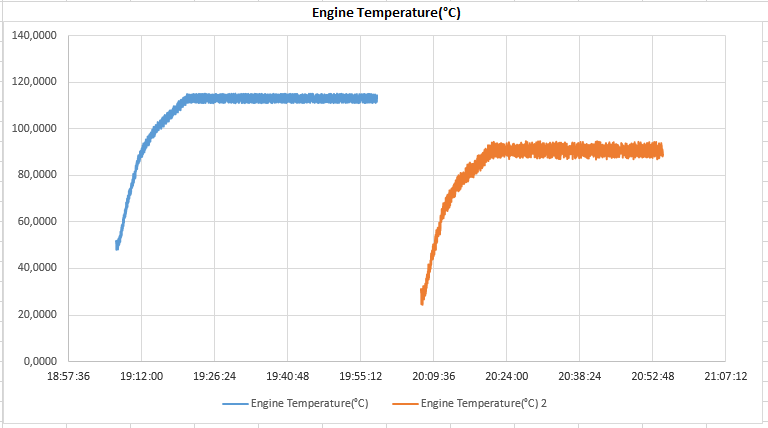
1. Датасет содержит в себе две выборки (данные за два промежутка времени): «21.08.2021 19:07:05 - 21.08.2021 19:58:25» и «21.08.2021 20:07:05 - 21.08.2021 20:54:48».
2. Датасет содержит в себе данные по временным отрезкам – выход двигателя в рабочий режим, т.к. с увеличением времени температурный показатель растёт, а затем принимает значение близкое к рабочей температуре и изменяется в пределе погрешности.
3. Между первой и второй выборкой с рабочим агрегатом были проведены манипуляции.

**ГИПОТЕЗЫ**

1. Принимаем численный показатель равный 90°C – типовая температура работы двигателя.
2. Выдвигаем гипотезу, что после проведения технического обслуживания двигателя агрегат работает в диапазоне температур близком к типовой рабочей температуре 90°C, как следствие прослужит регламентированный срок без поломок.

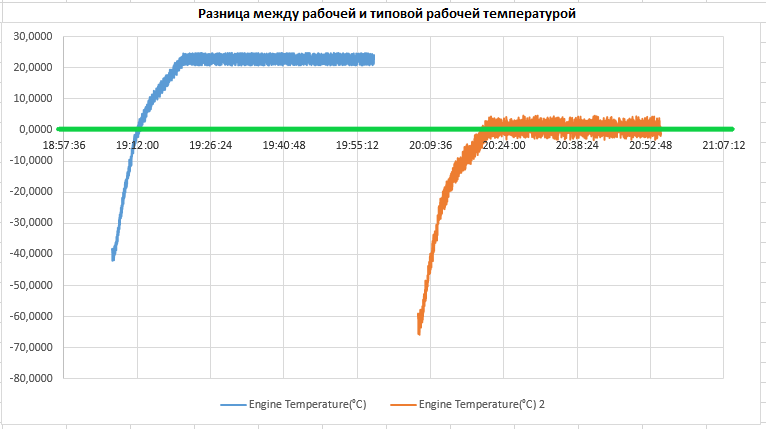
**ПРОВЕРКА ГИПОТЕЗЫ**

Строим графики

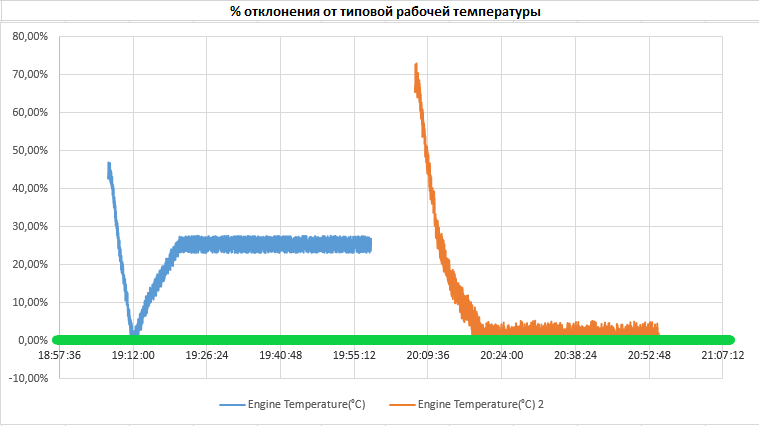


Исходя из графиков можно сделать вывод, что во второй выборке рабочая температура агрегата находится ближе к типовой рабочей температуре 90°C.

Сравниваем значения с эталонным - 90°C:



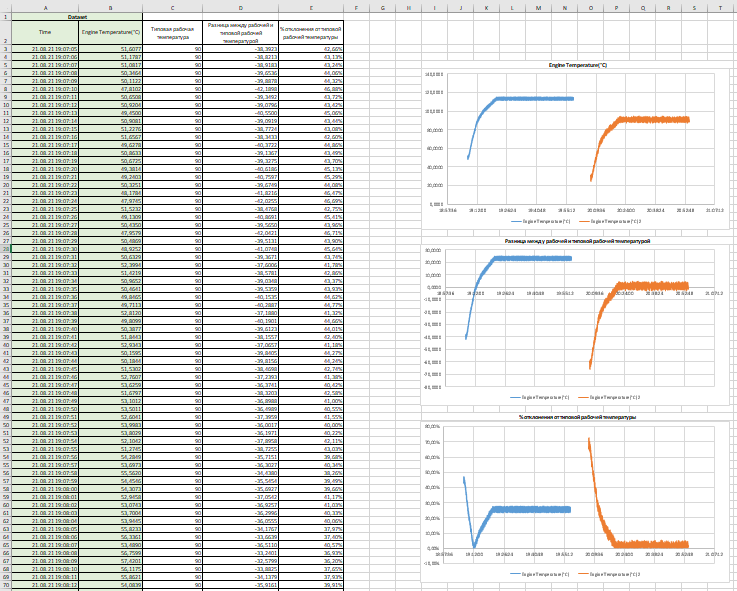
Отрисовываем график процентного отклонения от 90°C:



**Результат проверки гипотезы**

Гипотеза подтверждена. После проведения технического обслуживания двигателя агрегат работает в диапазоне температур близком к типовой рабочей температуре 90°C, как следствие прослужит регламентированный срок без поломок.

**Расчеты в MS Excel с проверкой выбранной гипотезы (скриншот и файл .xls)**



****