МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ

ФЕДЕРАЦИИ

Федеральное государственное автономное

образовательное учреждение высшего образования

«Самарский национальный исследовательский

университет имени академика С.П. Королёва»

(Самарский университет)

Институт информатики и кибернетики

Кафедра технической кибернетики

Лабораторная работа № 3

«STREAM PROCESSING WITH APACHE FLINK»

Выполнил студент: Ермаков Д.Д.

Гр. 6132 - 010204D

Проверил преподаватель: Попов С.Б.

Самара 2022

**ЗАДАНИЕ**

1. Выполнить следующие задания из набора заданий репозитория <https://github.com/ververica/flink-training-exercises>:

* RideCleanisingExercise
* RidesAndFaresExercise
* HourlyTipsExerxise
* ExpiringStateExercise

1. Сформировать отчет о ходе выполнения лабораторной работы.

**ХОД РАБОТЫ**

Подготовим необходимые файлы для лабораторной работы, для этого загрузим файлы с репозитория https://github.com/ververica/flink-training. Для каждого из 4 заданий, напишем свой код на языке программирования Scala.

Для каждого задания определена своя цель. Поэтому исходя из цели задания произведем необходимые вычисления и определим то, что нас просят.

1. RideCleanisingExercise — необходимо очистить поток событий поездки на такси, удалив события, которые начинаются или заканчиваются за пределами Нью-Йорка.
2. RidesAndFaresExercise — необходимо объединить записи для каждой поездки такси.
3. HourlyTipsExerxise — необходимо определить за каждый час водителя, получающего наибольшее количество чаевых.
4. ExpiringStateExercise — необходимо устранить ошибку в задании 2.

Ссылка на репозиторий: https://github.com/ShadowOfDen/BigData/tree/main/Lab3.

Внутри репозитория в файле README.md содержится краткая сводка по каждому заданию.

**ЗАКЛЮЧЕНИЕ**

В результате лабораторной работы мы дополнили код и получили необходимые данные для каждого из задания. А также по каждому из заданию был сформирован свой мини отчет.