**Федеральное государственное автономное образовательное**

**учреждение высшего образования**

**«Национальный исследовательский университет**

**«Высшая школа экономики»**

**Факультет компьютерных наук**

**Основная образовательная программа**

**Прикладная математика и информатика**

**Домашнее задание**

**по курсу *“Основные методы анализа данных”***

**Выполнил:**

Студент 3 курса

Григорьев Дмитрий

**Преподаватель:**

Миркин Б. Г.

Москва, 2019

**Актуальность выбранных данных:**

Безопасность на дорогах является одной из самых важных мировых и демографических задач.

Аварии на автомобильном транспорте наносит огромный материальный и моральный ущерб как обществу в целом, так и отдельным гражданам.

Дорожно-транспортный травматизм приводит к исключению из сферы производства людей трудоспособного возраста. Гибнут и становятся инвалидами дети.

Мне стало интересно рассмотреть именно эту сферу данные, поэтому я выбрал для анализа данные аварий дорожного движения за 2018 год Data from the UK Department for Transport -RoadSafetyData\_Accidents\_2018, которые можно найти на сайте <https://data.gov.uk/dataset/cb7ae6f0-4be6-4935-9277-47e5ce24a11f/road-safety-data>.

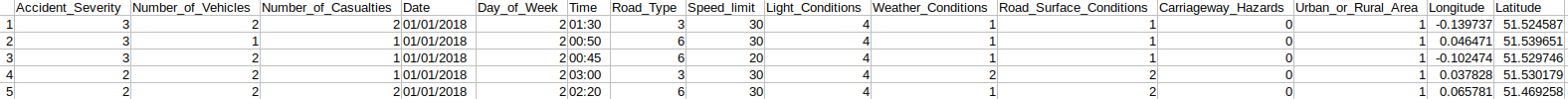
**Описание набора данных:**

В наборе выбранных мной данных присутствует 13 признаков:

1. Accident\_Severity(Степень Тяжести Несчастного Случая )
2. Number\_of\_Vehicles(Количество транспортных средств)
3. Number\_of\_Casualties(Число жертв)
4. Date(Дата)
5. Day\_of\_Week(День недели)
6. Time(Время)
7. Road\_Type(Тип дороги)
8. Speed\_limit(Ограничение скорости, миль в час)
9. Light\_Conditions(Условия освещения)
10. Weather\_Conditions(Метеорологические условия)
11. Road\_Surface\_Conditions(Состояние дорожного покрытия)
12. Carriageway\_Hazards(Опасность на проезжей части)
13. Urban\_or\_Rural\_Area(Городская или сельская местность)
14. Longitude(Долгота)
15. Latitude(Широта)

Объем выборки: 2000

**Фрагмент таблицы данных:**

****