

## Лабораторная работа №2

(Подготовка данных в Pandas. Часть 2)

*Цель:* Закрепление основных команд форматирования данных в Pandas.

В файле BikeData.csv приведены данные за каждый час по аренде велосипедов в двух сервисных центрах Partner 1 и Partner 2.

- 1) Создайте новую переменную Total\_count, в которую записать общее количество арендованных велосипедов за каждый час.
- 2) Удалите ряды Partner 1 и Partner 2.
- 3) Преобразуйте переменную Date к стандартной форме (**to\_datetime**).
- 4) Определите сколько велосипедов было выдано в день, месяц и число которого совпадает с днем выполнения этого задания. Определите, какой это был день недели.
- 5) Перекодируйте категориальную переменную Functioning Day(рабочий или нерабочий день) в переменную типа boolean, т.е. "Yes"=True и "No" = False. Выведите количество нерабочих дней (не часов). День считается нерабочим, если пункт выдачи велосипедов работал менее 12 часов.
- 6) Перекодируйте переменную Holiday (являлся ли день праздничным) в 0, если No Holiday, и 1, если Holiday. (используйте анонимную функцию **lambda**).
- 7) Введите новую категориальную переменную Temperature category, которая будет равна:
  - "Freezing", если температура < 0;
  - "Chilly", если 0 <= температура < 15;
  - "Nice", если 15 <= температура < 26;
  - "Hot", если 26 <= температура.
- 8) Создайте новую переменную Good weather, которая будет равна 1 (т.е. хорошая погода), если:
  - Temperature category = "Nice",
  - Humidity в диапазоне от 40 до 60 ,
  - скорость ветра Wind speed меньше 5.4,
  - нет дождя (Rainfall)
  - нет снега (Snowfall).Определите, сколько процентов наблюдений соответствует хорошей погоде.
- 9) Сгруппируйте данные по сезонам (Seasons), внутри сезона по категории погоды (Temperature category) и вывести для полученной группировки количество арендованных велосипедов. Вывести результаты в виде сгруппированной таблицы (через **groupby**) и в виде сводной таблицы (**pivot\_table**).

10) Определите сезон и температуру, при которых выдается наибольшее и наименьшее количество велосипедов.

Отчет устный по результатам выполненной работы.

При выполнении работы рекомендуется использовать следующие методы библиотеки **pandas**: `drop`, `to_datetime`, `apply`, `pivot_table`, `groupby`.