**Задания по файлу «Descriptive\_Statistics.R»**

**Задание №1**

Используя данные mtcars, рассчитайте средний расход топлива (mpg) для автомобилей c числом лошадиных сил (hp), большим 120 и у которых вес менее 4000 фунтов.

Получившийся результат (среднее значение) сохраните в переменную result.

**Задание №2**

При помощи функции aggregate рассчитайте стандартное отклонение переменной mpg (расход топлива), переменной disp (вместимости двигателя)  у машин с автоматической и ручной коробкой передач.

Полученные результаты (результаты выполнения функции aggregate) сохраните в переменную descriptions\_stat.

**Задание №3**

Воспользуемся встроенными данными airquality. В новую переменную сохраните subset исходных данных, оставив наблюдения только для месяцев 5, 6 и 7.

При помощи функции aggregate рассчитайте количество непропущенных наблюдений по переменной Solar.R (солнечная радиация) в 5, 6 и 7 месяце. Для определения количества наблюдений используйте функцию length().

Результат выполнения функции aggregate сохраните в переменную result1.

Подсказки:

1. Не забудьте сделать subset, чтобы отобрать наблюдения только по нужным месяцам, вам может пригодиться следующая конструкция:

> x <- 5

> x %in% c(3, 4, 5)

2. Для подсчета числа непропущенных наблюдений воспользуйтесь записью с помощью формулы, при которой пропущенные значения не учитываются:

aggregate(y ~ x + z , data, FUN)

**Задание №4**

Примените функцию describeBy к количественным переменным данных airquality, группируя наблюдения по переменной Month.  Чему равен коэффициент асимметрии (skew) переменной Wind в седьмом месяце?

**Задание №5**

В переменной my\_vector сохранен вектор с пропущенными значениями. Вам нужно создать новый вектор fixed\_vector, в котором все пропущенные значения вектора my\_vector будут заменены на среднее значение по имеющимся наблюдениям.

При этом исходный вектор оставьте без изменений!

Ниже небольшой код, который может создать случайный вектор my\_vector (выборка из нормального распределения) с пропущенными значениями.

my\_vector <- rnorm(30)

my\_vector[sample(1:30, 10)] <- NA # на десять случайных позиций поместим NA

Задача для самостоятельной работы:

Изучите справку по функции replace. Вызвать справку можно исполнив команду:

?replace

Попробуйте решить это задание также при помощи этой функции.

**Задания по файлу «Plots.R»**

**Задание №6**

При помощи функции ggplot() или boxplot() постройте график boxplot, используя встроенные в R данные airquality. По оси x отложите номер месяца, по оси y — значения переменной Temp.

На графике boxplot отдельными точками отображаются наблюдения, отклоняющиеся от 1 или 3 квартиля больше чем на полтора межквартильных размаха. Сколько таких наблюдений присутствует в июле (месяц №7)?

Обратите внимание, что для корректного отображения графика ggplot ожидает факторную переменную по оси x.

**Задание №7**

Используя данные mtcars, нужно построить scatterplot с помощью ggplot из ggplot2, по оси x которого будет hp, по оси y – qsec, а цветом отобразить переменную (mpg). Полученный график нужно сохранить в переменную plot1.

**Задание №8**

Основываясь на данных iris постройте график Scatterplot (диаграмма рассеивания), где по оси X будет отложена переменная Sepal.Length,  по оси Y переменная  Sepal.Width. За цвет точек будет отвечать переменная  Species, а за размер точек переменная Petal.Length.