

Домашнее задание №7

Оформите решение задания в виде маркдауна, сохранив документ в pdf или html.

Решение каждого задания на ggplot должно содержать:

- номер задания
- формулировка задания
- код, с помощью которого выполнено задание
- получившаяся картинка
- номер рисунка внизу справа, для этого используйте labs (caption = 'fig.N')) (см. слайд 26 в презентации к лекции)

Итоговый файл прикрепите в форму: <https://forms.gle/YNrVphAAvBitKbY29>

Задание 0. Посмотрите и проработайте выложенную к лекции презентацию. В ней примерно то же, что было на лекции, но на другом наборе данных. Оформлять это задание в маркдаун НЕ нужно.

Задание 1 (АЖ). По данным mtcars нарисуйте точечную диаграмму: mpg от hp (geom_point) с настройками по умолчанию - fig.1

Задание 2 (АЖ). Дайте fig.1 название, а также измените подписи осей - fig.2

Задание 3 (АЖ). По данным iris нарисуйте точечную диаграмму: Petal.Width от Petal.Length с настройками по умолчанию - fig.3

Задание 4 (АЖ). В fig.3 измените цвет точек на синий, а форму точек в зависимости от значений в столбце Species, назовите легенду - Species Type - fig.4

Задание 5 (АЖ). В fig.4 уберите легенду (используйте поиск в интернете) - fig.5

Задание 6 (АЖ). В fig.4 перенесите легенду вниз рисунка, используйте theme(legend.position="bottom") - fig.6

Задание 7 (АЖ). На fig.4 добавьте непрерывную горизонтальную линию (y=среднее по столбцу Petal.Width) красного цвета (используйте geom_hline) и пунктирную вертикальную линию (x=среднее по столбцу Petal.Length) синего цвета, увеличив толщину линии (используйте geom_vline). Координаты по осям должны идти от нуля - fig.7

Задание 8 (АЖ). По данным iris изобразите значение столбца Sepal.Length в виде гистограммы (geom_histogram), число столбцов - 20, измените цвет заливки столбцов и цвет контура столбцов - fig.8

Задание 9 (АЖ). По данным iris отберите значения Sepal.Length строго больше 5. Из получившихся данных нарисуйте barplot (geom_bar) по столбцу Species. Столбец для вида setosa должен быть посередине - fig.9

Задание 10 (АЖ). Для fig.9 поменяйте местами оси, используйте + coord_flip()

Задание 11 (1 допбалл).

Создайте 1000 выборок размера 45 из нормально распределения с средним 0 и стандартным отклонением 2.

Объедините объекты из всех выборок. Удалите все значения, которые больше 99% квантили или меньше 1% квантили полученного набора.

Для каждой выборки посчитайте среднее

На одном графике постройте:

- 1) постройте гистограмму значений, число бинов 100. Цвет границ бинов сделайте black, заполнение - "lightgreen", прозрачность 0.5. В aes поставьте `y=..density..`
- 2) Для каждой выборки (изначальных до удаления объектов) посчитайте среднее и постройте гистограмму средних, число бинов 100. Цвет границ бинов сделайте black, заполнение - "forestgreen", прозрачность 1. В aes поставьте `y=..density..`
- 3) Постройте вертикальную линию красного цвета, проходящую через точку $x = 0$ при помощи `geom_vline`

Задание 12 (1.5 допбалла)

- 1) Создайте набор1 из 1000 выборок размера 45 из нормально распределения с средним 0 и стандартным отклонением 2.
- 2) Создайте набор2 из 1000 выборок размера 45 из нормально распределения с средним 3 и стандартным отклонением 4.

Для каждого из набора для каждой выборки посчитайте среднее.

Для каждого из наборов постройте на одном графике гистограмму распределения полученных средних.

Для первого набора заполните бины цветом "forestgreen", для второго "darkred".

Для обоих наборов границы бинов сделайте серого цвета, прозрачность бинов - 0.5, число бинов - 100.

Постройте вертикальную прямую, проходящую через 95% квантиль средних первого набора.

Посчитайте число средних второго набора, которые находятся слева от прямой.

Выведите долю таких средних

Задание 13 (0.5 допбалла)

С помощью средств R посчитайте теоретическую долю **средних** выборок из второго распределения (со средним 3 и стандартным отклонением 4), которые меньше теоретической квантили **среднего** выборок из первого распределения (с средним 0 и стандартным отклонением 2)

Сравните с значением, полученным в предыдущем пункте.