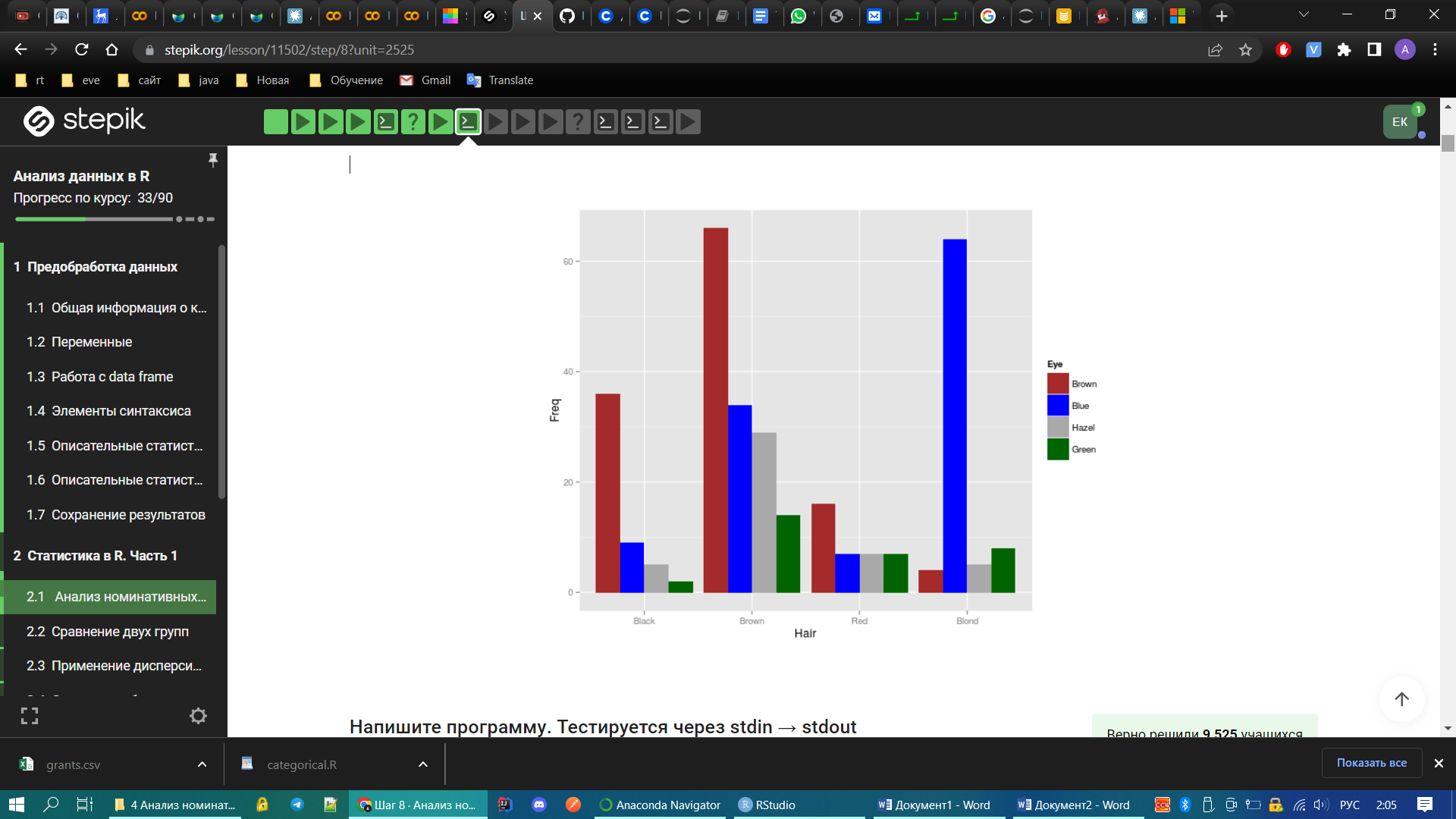
**Лабораторная работа 2**

**Задание 1**

Воспользуйтесь таблицей HairEyeColor. Ваша задача а) в переменную **blond\_men** сохранить долю блондинов (Blond) от общего числа кареглазых мужчин; б) посчитайте число голубоглазых женщин в наборе данных HairEyeColor; в) постройте столбчатую диаграмму распределения цвета глаз по цвету волос только у женщин из таблицы HairEyeColor. По оси X должен идти цвет волос, цвет столбиков должен отражать цвет глаз. По оси Y - количество наблюдений.

В случае, если все сделано правильно, он будет выглядеть так (обратите внимание на название осей и легенды):



**Задание 2**

На основе таблицы HairEyeColor создайте ещё одну таблицу, в которой хранится информация о распределении цвета глаз у женщин-брюнеток (Hair = 'Black'). Проведите тест равномерности распределения цвета глаз у брюнеток и выведите значение хи-квадрата для этого теста.

**Задание 3**

Воспользуемся данными из таблицы diamonds. При помощи критерия Хи - квадрат проверьте гипотезу о взаимосвязи цены (price) и каратов (carat) бриллиантов. Для этого сначала нужно перевести эти количественные переменные в формат пригодный для Хи - квадрат. Создайте две новые переменные в данных diamonds:

factor\_price - где будет 1, если значение цены больше либо равно чем среднее, и 0, если значение цены ниже среднего цены по выборке.

factor\_carat - где будет 1, если число карат больше либо равно чем среднее, и 0, если ниже среднего числа карат по выборке.

Важный момент - на больших данных цикл for() работает довольно медленно, постарайтесь решить эту задачу без его использования!

Используя эти шкалы при помощи Хи - квадрат проверьте исходную гипотезу. Сохраните в переменную main\_stat значение критерия Хи - квадрат.

**Задание 4**

Воспользуйтесь таблицей mtcars. При помощи точного критерия Фишера проверьте гипотезу о взаимосвязи типа коробки передач (am) и типа двигателя (vs) в данных mtcars. Результат выполнения критерия сохраните в переменную. Получившийся p - уровень значимости сохраните в переменную fisher\_test.

**t-test**

**Задание 5**

Воспользуйтесь таблицей ToothGrowth. Данные позволяют исследовать рост зубов у морских свинок в зависимости от дозировки витамина C и типа потребляемых продуктов. Сравните среднее значение длины зубов свинок, которые потребляли апельсиновый сок (OJ) с дозировкой 0.5 миллиграмм, со средним значением длины зубов свинок, которые потребляли аскорбиновую кислоту (VC) с дозировкой 2 миллиграмма. Значение t - критерия сохраните в переменную t\_stat.

**Задание 6**

В этом задании нужно проверить гипотезу о равенстве средних двух выборок, загрузив набор данных из файла «dataset\_11504\_15.txt» и выполнив все необходимые операции на вашем компьютере. В скачанных данных вы найдете две переменные: количественную переменную, и номинативную переменную с двумя градациями (которая разделяет наблюдения на две группы).

Сначала с помощью теста Бартлетта проверьте гомогенность дисперсий двух выборок. В случае, если дисперсии значимо не отличаются (с уровнем 0.05), примените тест Стьюдента, иначе - непараметрический тест (Манна-Уитни). Выведите получившийся p-value, с точностью четыре знака после запятой.