

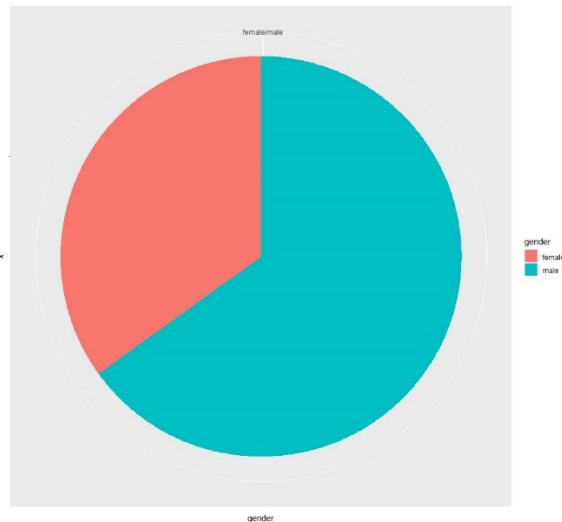
Отчет по итоговому заданию

1. Я сформулировала следующие гипотезы:

- 1) оценки по математике (math's score) будут лучше у парней чем у девушек.
- 2) письменные оценки (writing score) связаны с оценками по чтению (reading score).

2. Визуализирование переменных:

1) гендер: 482 девушек; 518 юношей



2) рассчитала описательные статистики для переменных: письменные оценки, оценки по чтению и оценки по математике

Таблица 1. Описательные статистики для оценок по математике

Среднее	Медиана	Станд. Откл.	Мода
66,089	66	15,163	65

Рисунок 1. Математические оценки у девушек и юношей

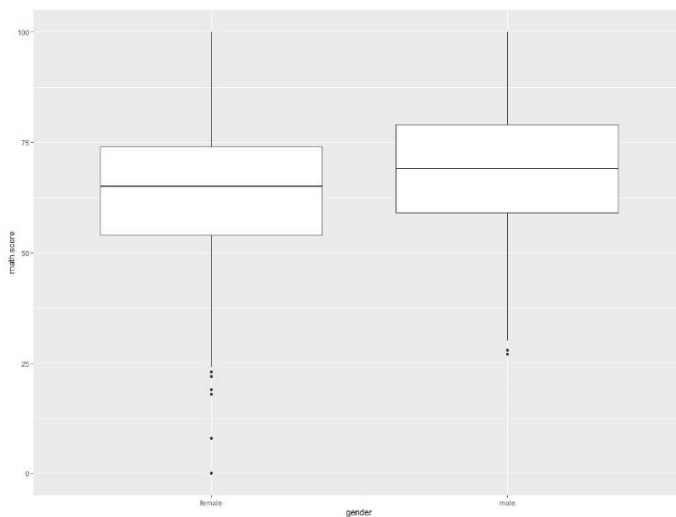


Таблица 2. Описательные статистики для оценок по чтению

Среднее	Медиана	Станд. Откл.	Мода
69,169	70	14,6	72

Рисунок 2. Оценки по чтению у девушек и юношей

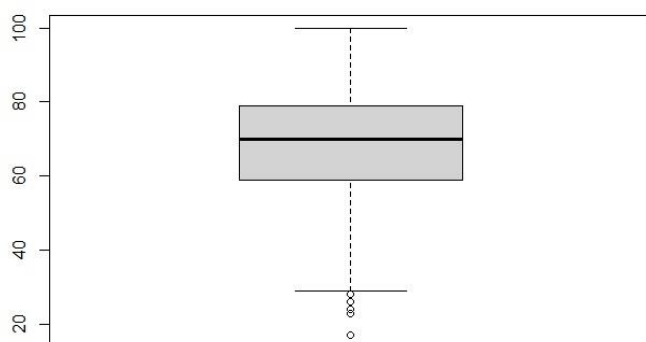
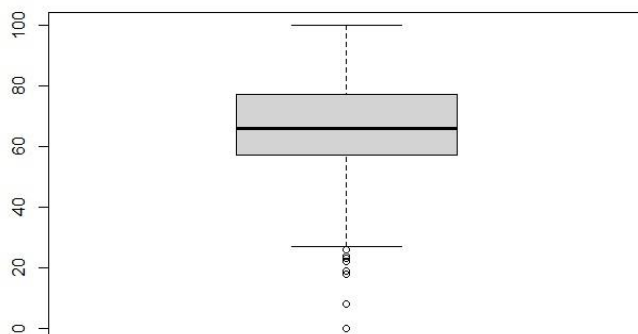


Таблица 3. Описательные статистики для оценок по письму

Среднее	Медиана	Станд. Откл.	Мода
68,054	69	15, 196	74

Рисунок 3. Оценки по письму у девушек и юношей



3. Все три переменных оценок были проверены на нормальность распределения через Шапиро-Уилко тест.

Переменные: письменные оценки, оценки по чтению и оценки по математике – распределены не нормально ($p < 0,05$)

4. Проверка первой гипотезы.

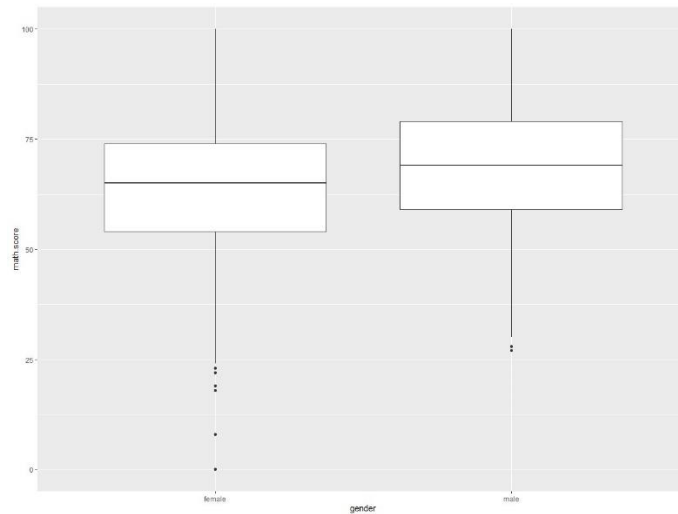
Первая гипотеза подтвердилась. Действительно, оценки по математике (math's score) выше у парней, чем у девушек.

Использовался Т-критерий Манна-Уитни, поскольку нужен был непараметрический статистический тест (выборка оценок по математике распределена не нормально).

Оценки по математике у юношей значительно отличаются от оценок по математике у девушек ($W = 101769$, $p = 4.279e-07$, $p < 0,05$).

По графику видно, что у юношей оценки по математике выше, чем у девушек.

Рисунок 4. Оценки по математике у девушек и юношей



Я бы сказала, что это отличие небольшое, однако при проверке, оно оказалось значимым. Возможно, это связано с тем, что в исследовании принимало большое количество участников – выборка большая.

5. Проверка второй гипотезы о связи между двумя переменными.

Поскольку обе выборки распределены не нормально, то использовался критерий Спирмена.

Гипотеза подтвердилась - письменные оценки (writing score) связаны с оценками по чтению (reading score).

Причем связь достаточно сильная ($S = 8507905$, $p < 2.2e-16$, $p < 0,05$), что также отражено на рисунке.

Рисунок 5. Корреляция между письменными оценками с оценками по чтению.

