

Задание №1

Pandas, seaborn.

1. Определите количество мужчин и женщин среди испытуемых. Однозначной кодировки пола у вас нет, поэтому подумайте, как можно это посчитать с помощью других переменных. Для уточнения – данные собраны в европейской популяции. (2 балла)
2. Создайте новый признак --- BMI (body mass index, индекс массы тела). Для этого разделите вес в килограммах на квадрат роста в метрах. (2 балла)
3. Визуализируйте распределение переменной BMI. Есть ли выбросы? (2 балла)
4. Создайте новый датафрейм, где вы удалите пациентов, у которых диастолическое давление выше систолического. Какой процент от общего количества пациентов они составляли? (2 балла)
5. Визуализируйте распределение уровня холестерина для различных возрастных категорий. (2 балла)
- 6*(optional). Как соотносятся ИМТ и наличие ССЗ? Придумайте подходящую визуализацию. (4 балла)