

Лабораторная работа №5. Корреляционный анализ в Python

Контрольные вопросы

1. Обратите внимание на диагональ слева направо в таблице корреляции. Почему диагональ заполнена единицами, это совпадение?
 - Тот факт, что диагональ заполнена единицами (1.000000), указывает на то, что существует положительная общая корреляция между переменными, которые находятся в абсциссе, и переменными, которые являются в порядке :
 - Согласно коэффициенту Пирсона, который указывает, насколько линейно коррелированы две переменные, это не совпадение
2. Можно видеть, также что значения в таблице являются зеркальными, значения ниже единичной диагонали имеют зеркальный аналог выше этой диагонали. Почему?
 - Коэффициент Пирсона равен частному лицу ковариации двух переменных и произведению двух стандартных отклонений. Это частное в 1 означает, что оба значения эквивалентны, что дает этот зеркальный эффект части массива.
3. Многие переменные пары представляют корреляцию, близкую к нулю. Что это обозначает?
 - Согласно коэффициенту Пирсона, это означает, что между этими переменными существует низкая корреляция.
4. Зачем было разделять данные по полу?
 - Разделение данных по полу гарантирует, что результаты не искажаются из-за различий между мужским и женским телами;
5. Какие переменные имеют более сильную корреляцию с размером мозга (MRI_Count)? Это ожидаемо?
 - Для глобального фрейма данных (brainFrame) сортировка по убыванию корреляций переменных :
 - ✓ **Height – 0.601712;**
 - ✓ **Weight – 0.513378;**
 - ✓ **PIQ – 0.386817;**
 - ✓ **FSIQ – 0.357641;**
 - ✓ **VIQ – 0.337478;**
 - Для фрейма данных женщин (womenDf), нисходящий рейтинг корреляций переменных :
 - ✓ **Weight – 0.446271;**
 - ✓ **PIQ – 0.396157;**
 - ✓ **FSIQ – 0.325697;**
 - ✓ **VIQ – 0.254933;**
 - ✓ **Height – 0.174541;**
 - Для кадра данных мужчин (menDf), нисходящая классификация корреляций переменных :
 - ✓ **PIQ – 0.568237;**
 - ✓ **FSIQ – 0.498369;**
 - ✓ **VIQ – 0.413105;**
 - ✓ **Height – 0.301543;**

✓ **Weight – -0.076875;**

<https://github.com/malko1278/labo-analyse-big-data.git>