Институт информационных технологий

Кафедра: Математическое и программное обеспечение ЭВМ

Дисциплина: Основы Data Science

ЛАБОРАТОРНАЯ РАБОТА № 2

Тема: Структура и описание данных. Описательные характеристики

Выполнил:

студент гр. 1ПИб-02-2оп-23

Кринкин Олег Алексеевич

Проверил:

Юдина Ольга Вадимовна

# 1. Задание

**Цель**. На основе результатов предыдущей работы составить представление о структуре и распределении данных.

Задача этапа – понять, что именно содержится в этих данных, какие факты они описывают, понять, что они могут дать.

**Задания**

Вспомните формулировки задач (исследований), которые вы сделали на прошлом занятии. Запишите предположения, которые нужно проверить. Подумайте, что может их подтвердить или опровергнуть.

Например, мы предполагаем, что в игры «с головой» чаще уходят дети, у которых нет интереса к учебе, а родители тоже в учебе не были заинтересованы.

Что может подтверждать эти предположения?

*Например, определим чрезмерную увлеченность. Как это можно сделать? Посмотрим* *на распределение – можно ли выделить группу «игроков» по времени? А другие группы?* *Можем в этих группах посчитать среднее время в игре, среднюю частоту…Посчитать* *доверительные интервалы – сопоставить*. Нужна проверка гипотез о среднем?

Используя результаты прошлой работы по структурированию данных. а также с учетомпостроенных описаний:

Для выделенных для сравнения групп, вычислите (количественных) подходящие средние значения, доверительные интервалы меры разброса

Для интересующих категорий – доли и их доверительные интервалы.

*Например, совпадают или нет доли самых активных игроков среди лучших и не* *лучших? С начала посчитаем доли и доверительные интервалы, а затем примем решение,* *нужна ли проверка статистических гипотез*

**Помните, что в исследовании должны быть задействованы максимум переменных**

Заключение: какие выводы можно сделать по вашей работе

# 2. Ход работы