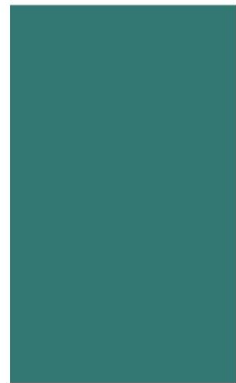


От чего зависит заработная плата



Участники команды

Шагиева Алия

Быковский Николай

Петрова Маргарита

Агранович Вика

Вот такие пироги



Информация о данных



3871 – число наблюдений

Результаты опроса
населения РМЭЗ НИУ ВШЭ
в **2017** году

Данные содержат:

- Выборку из индивидов, находящихся в трудоспособном возрасте
- Информация о заработной плате, образовании, семейном положении и другие характеристики

Исследовательский вопрос

От чего зависит заработная плата?



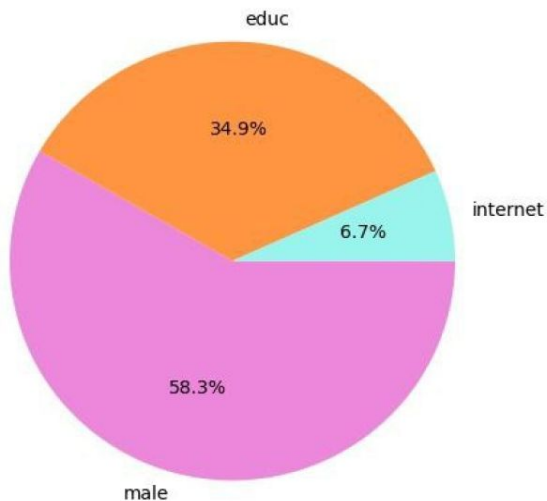
Гипотеза

Человек, пользующийся интернетом получает более высокую заработную плату.

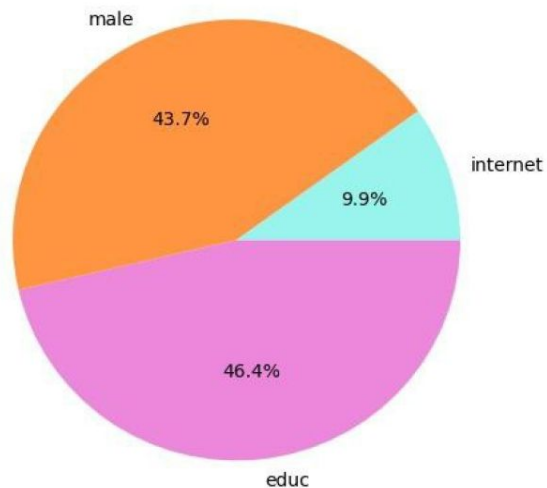
Альтернативная гипотеза:

Интернет не обладает наибольшим влиянием на заработную плату.

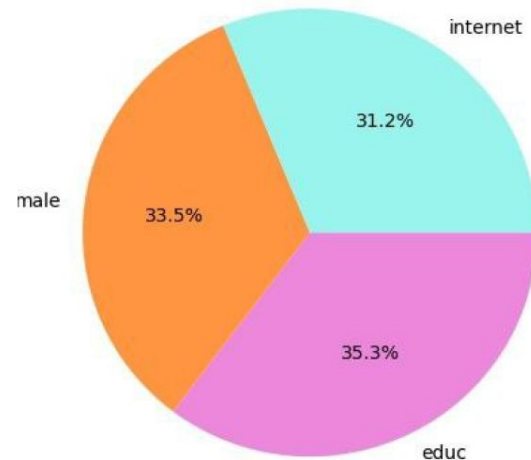
Круговые диаграммы



Влияние факторов
на зарплату людей
до 35 лет



Влияние факторов
на зарплату людей
от 35 до 45 лет

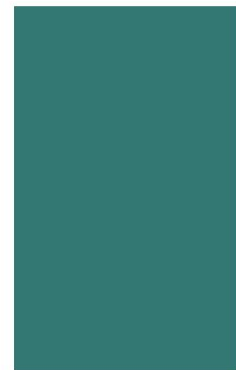


Влияние факторов
на зарплату людей
от 45 лет

Механизм

- 1 Люди в возрасте от 45 до 59 лет владеют навыками пользования интернетом
- 2 Благодаря возможности использования интернета они повышают свою квалификацию
- 3 Становясь более компетентными и востребованными сотрудниками
- 4 Тем самым являются важными сотрудниками для работодателей
- 5 Работодатель замечает потенциал сотрудника и повышает заработную плату

Приложение



Используемые переменные



age – возраст в годах;

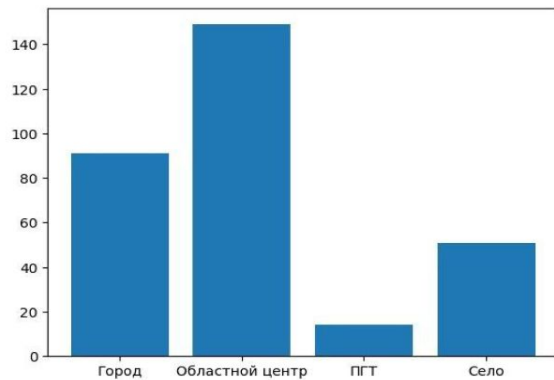
male – пол, 1 для мужчин,
0 для женщин

industry – отрасль занятости

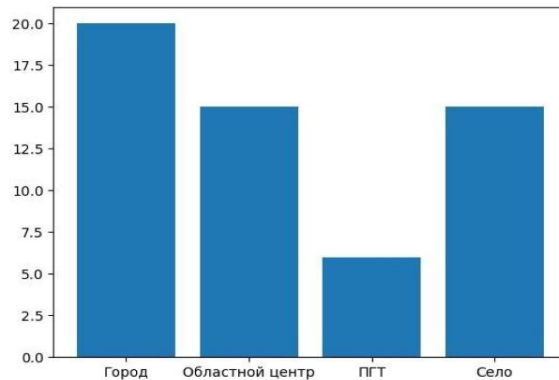
internet – приходилось ли
респонденту в течение последних
12 месяцев пользоваться
Интернетом, 1 если да, 0
в противном случае

urban – 1 респондент проживает
в городе, 0 в сельской местности

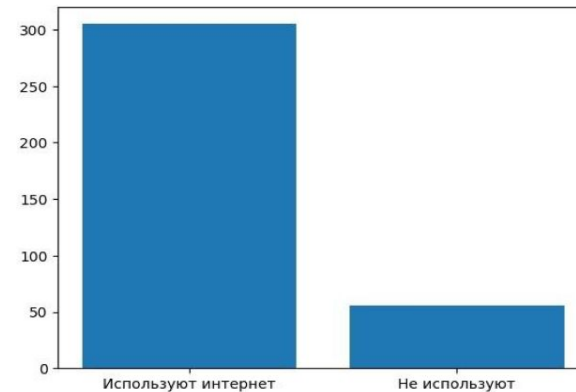
Диаграммы



Зависимость людей
использующих
интернет от места
жительства



Зависимость людей
не использующих
интернет от места
жительства



Зависимость людей
использующих и
неиспользующих
интернет в сфере
транспорта

Мат модель

$$b = \frac{\sum_{i=1}^n x_i y_i - n\bar{x}\bar{y}}{\sum_{i=1}^n x_i^2 - n\bar{x}^2} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{n \sum_{i=1}^n (x_i - \bar{x})^2} \quad d = \bar{y} - b\bar{x}$$

где $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$ – среднее значение случайной величины Y ;

$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ – среднее значение величин x_i .

Цель:

оценить взаимосвязь различных факторов на зависимость при получении зарплаты

Метод:

линейная регрессия с двунаправленными фиксированными эффектами

<https://github.com/GentleMann08/DANO>

-Hackathon-2024

15:37

Основные выводы:

Наша гипотеза не подтвердилась, в результате работы мы выявили, что использование интернета оказалось не самым влиятельным фактором на заработную плату.

А наиболее важным фактором оказалась гендерная принадлежность и образование.

Используя полученные результаты мы поняли, что при достаточных данных можно применить математические алгоритмы для выявления зависимостей между самыми разными аспектами жизни и заработной платой.