МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное

образовательное учреждение высшего образования

«ЧЕРЕПОВЕЦКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_Институт информационных технологий\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

наименование института (факультета)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_МПО ЭВМ\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

наименование кафедры \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_Основы Data Science \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

наименование дисциплины в соответствии с учебным планом

ЛАБОРАТОРНАЯ РАБОТА № 5

Визуализация данных

Исполнитель

студент \_\_\_1ПИб-02-1оп-22\_\_\_\_

группа

\_\_\_Микуцких Г. А.\_\_\_

Фамилия, имя, отчество

Руководитель \_\_ Юдина О. В. \_

Ф.И.О. преподавателя

\_\_\_\_\_\_\_доцент\_\_\_\_\_\_

Оценка \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Подпись \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

2024 год

Цель: Знакомство с этапом понимание данных стандарта CRISP-DM.

Задача этапа – представить набор данных, использованный в лабораторных работах. Описать основные закономерности, которые в них содержатся, связи между этими данными, фактами, знаниями, получить что-то новое, понять, что они могут дать и визуализировать свои представления

Задание

Вам нужно представить результаты изучения этих данных руководству предприятия, чтобы решить в каком направлении строить дальнейшую кадровую политику или проводить исследование. Представьте в диаграммах имеющиеся данные, на основании которых возможно построить исследование.

Вам нужно показать:

1. Закономерности;
2. Структуры;
3. Возможные связи;
4. Влияние факторов друг на друга;
5. Если вам удалось построить модели – покажите их.

Все диаграммы должны иметь комментарии: что вы хотели ею показать, какие выводы сделали.

Требования

Диаграммы должны:

1. Иметь подписи по осям и заголовки, отражающие их суть;
2. Размерности должны быть нанесены на оси;
3. Быть разными;
4. Их тип должен быть пригодным для представления типа данных, которые вы показываете;
5. Они должны отражать ваши данные, их связи, закономерности, которые вам удалось обнаружить.

Ход работы:

Диаграмма на рис. 1 показывает количество мужчин и женщин, участвовавших в опросе.

Рис. 1. Участники опроса

Была исследована длительность сна среди мужчин и женщин (рис. 2).

Рис. 2. Длительность сна среди мужчин и женщин

Явно виден разброс среди групп, а линии тренда не совпадают. Для того, чтобы понять от чего существует разброс, была исследована зависимость сна от наличия депрессии. Выдвинуты гипотезы:

* H0: Люди, у кого есть депрессия, спят также, как и люди без депрессии.
* H1: Люди, у кого есть депрессия, спят не так, как люди без депрессии.

Исследование показано на рис. 3.



Рис. 3. Исследование зависимости сна от наличия депрессии

Наглядное представление на рис. 4.

Рис. 4. Зависимость сна от наличия депрессии

Выяснили, что время сна зависит от наличия депрессии. Чтобы выяснить есть ли связь между полом и возникновением депрессии, выдвинуты гипотезы:

* H0: Возникновение депрессии среди женщин такое же, как и у мужчин;
* H1: Есть связь между полом и депрессией.

Результаты исследования на рис. 5.



Рис. 5. Исследование связи между полом и депрессией

Выяснили, что есть связь между гендером и наличием депрессии. Наглядное представление на рис. 6.

Рис. 6. Зависимость пола и депрессии

Однако, если исследовать зависимость наличия трудностей с выполнением заданий от пола, то получим следующие результаты (рис. 7).



Рис. 7. Исследование связи между наличием трудностей от пола

Связи между полом и наличие трудностей при выполнении заданий нет (сумма меньше критического значения). Была исследована зависимость успеваемости от пола (рис. 8-9).



Рис. 8. Исследование зависимости успеваемости от пола

Рис. 9. Зависимость успеваемости от пола

Успеваемость не зависит от пола. То же самое можно увидеть исходя из исследования связи между наличием депрессии и успеваемостью (рис. 10-11).



Рис. 10. Зависимость успеваемости от статуса депрессии

Вывод: вполне возможно, что внешние негативные факторы сильнее влияют на женщин, снижая их общее благосостояние. Однако на общую эффективность это не оказывает сильного влияния.