

5. Домашнее задание по теме

Pandas – продвинутый

Формулировка задания:

Необходимо загрузить и обработать предложенный датасет (Титаник), провести аналитику данных с помощью фреймворка Python Pandas.

Планируемый результат:

1. Датасет загружен в Colab
2. В датасете отсутствуют пустые ячейки
3. Представлена простая аналитика по датасету

Описания плана работы:

1. Создать новый ноутбук в Colab
2. Сохранить датасет в Google диске и загрузить в ноутбук Colab
<https://www.kaggle.com/datasets/yasserh/titanic-dataset/code>
*Загрузить датасет по ссылке из интернета (репозитория Kaggle)
3. Перечень и описание столбцов:

Survived	выжил (1-да, 0- нет)
----------	----------------------

Pclass	Класс круиза
--------	--------------

Name	ФИО
------	-----

Sex	Пол
-----	-----

Age	Возраст
-----	---------

SibSp	число братьев, сестер или супругов на борту у человека
-------	--

Parch	количество родителей или детей, с которыми путешествовал каждый пассажир
-------	--

Ticket	Номер билета
--------	--------------

Fare	Цена билета
------	-------------

Cabin	Номер каюты
-------	-------------

Embarked	Порт посадки
----------	--------------

4. Определить количество пустых ячеек
5. *Заполнить пустые ячейки используя любую логику замещения данных
6. Если пункт 5 не выполнен, то удалить строки имеющие пустые ячейки.
7. По каждому признаку произвести аналитику:
 1. Количество уникальных значений
 2. Минимальное и максимальное значение

8.*С помощью аналитики определить влияние всех признаков на признак Survived(выживание).

Пример: Parch и SibSP отрицательно влияли на выживание при крушении , так как чем больше эти параметры тем ниже процент выживания (одиноким выжить проще).

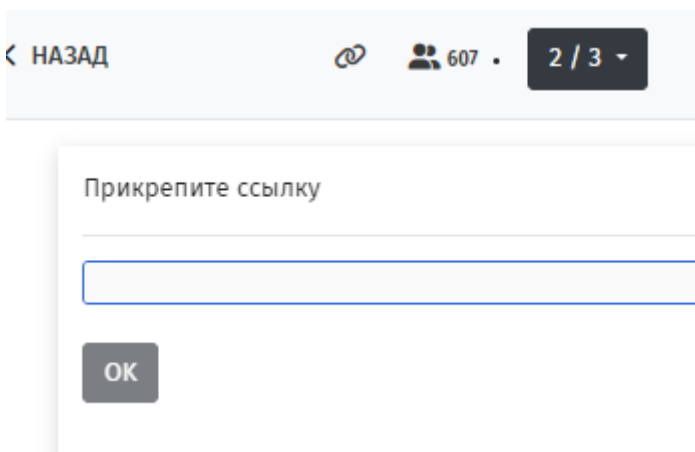
Результатом домашнего задания будет Таблица в Colab с комментариями в качестве ответов на пункты ДЗ.

Перечень инструментов, необходимых для реализации деятельности:

- 1) Google Colab <https://colab.research.google.com/>
- 2) *PyCharm

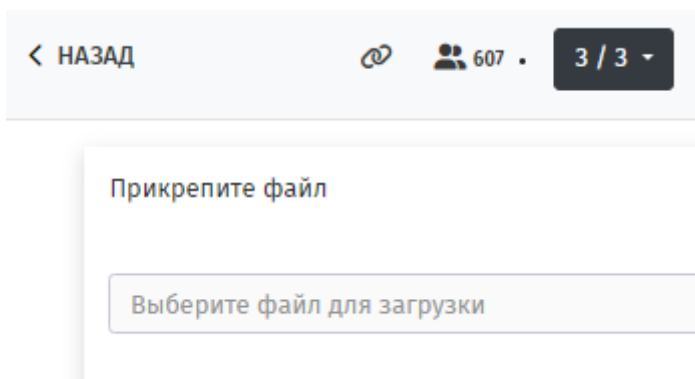
Форма загрузки:

- **В поле ссылки (2 страница задания)** загрузить ссылку на ноутбук google colab или github репозиторий.



The screenshot shows the Google Colab interface at the top with a navigation bar containing a back arrow, the text 'НАЗАД', a link icon, a user icon with '607', and a page indicator '2 / 3'. Below the navigation bar, a dialog box titled 'Прикрепите ссылку' (Attach link) is displayed. It features a text input field and an 'ОК' button.

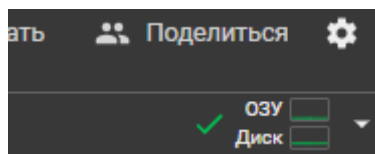
- **В поле файла (3 страница задания)** загрузить ноутбук с решением (файл с расширением .ipynb).



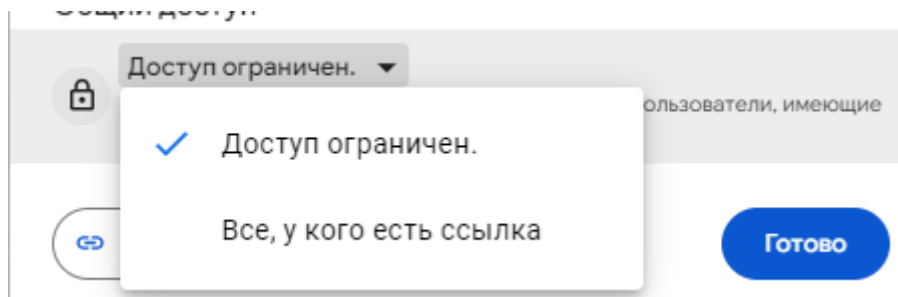
The screenshot shows the Google Colab interface at the top with a navigation bar containing a back arrow, the text 'НАЗАД', a link icon, a user icon with '607', and a page indicator '3 / 3'. Below the navigation bar, a dialog box titled 'Прикрепите файл' (Attach file) is displayed. It features a text input field with the placeholder text 'Выберите файл для загрузки' (Select file for upload).

Инструкция по получению ссылки на ноутбук google colab:

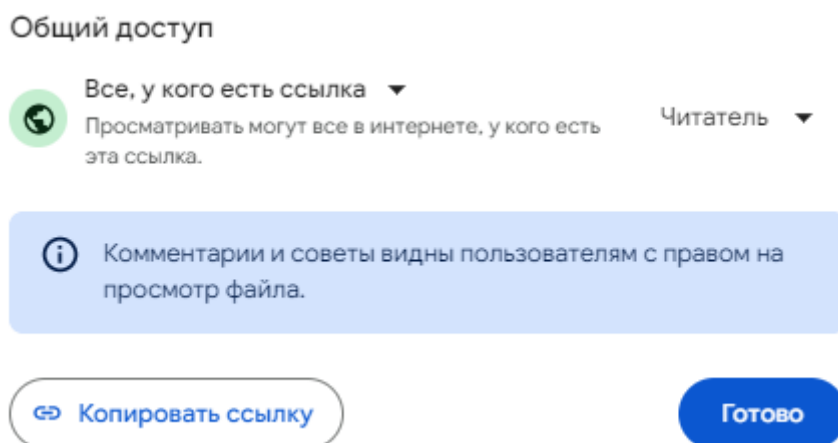
1. Нажмите **“Поделиться”** в правом верхнем углу экрана, рядом с лого вашего google аккаунта



2. В поле **“Общий доступ”** вместо **“Доступ ограничен”** выберите **“Все у кого есть ссылка”**

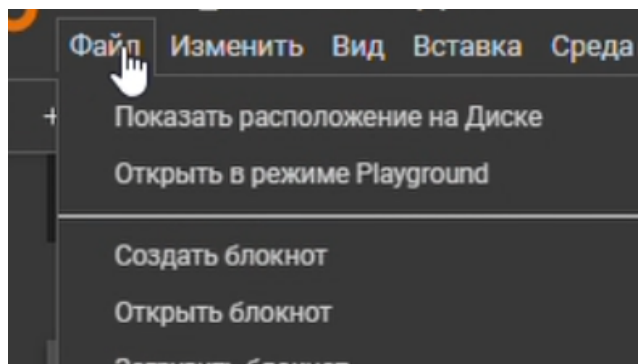


3. Нажмите **“Копировать ссылку”** и вставьте ее в поле ссылки

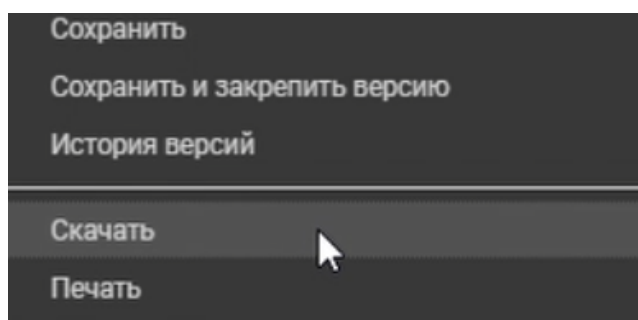


Инструкция по скачиванию файла с google colab:

В меню “Файл”



Выбрать пункт “Скачать”



Выбрать пункт “IPYNB”

