

Департамент образования и науки города Москвы

**Государственное автономное образовательное учреждение высшего  
образования города Москвы**

**«Московский городской педагогический университет»**

Институт цифрового образования

Департамент информатики, управления и технологий

Башкатова Анна Денисовна

## **ОТЧЕТ**

по дисциплине «Инструменты для хранения и обработки больших данных»

**Тема: «01-02 Product analytics»**

Направление подготовки (специальность) 38.03.05 – бизнес-информатика

Направленность (профиль) образовательной программы «Аналитика данных  
и эффективное управление»

Курс обучения: 3

Форма обучения: очная

Руководитель: Босенко Т. М.

Москва  
2023

**Ссылка на Colab:**

<https://colab.research.google.com/drive/1nEo7MQEmGAKwFy8OMzDHRLKN-AgAzO8B?usp=sharing>

**Ссылка на ДАШБОРД1:** <https://datalens.yandex/jt8mvgqxkm6u9>

**Ссылка на ДАШБОРД2:**

<https://lookerstudio.google.com/reporting/bf9b4d89-bc90-43eb-9491-d2ae3ff8bd12>

1. Воспользовавшись кодом из задания сделали датафрейм q4\_nps.

```
[16] q4_nps.head(10)
```

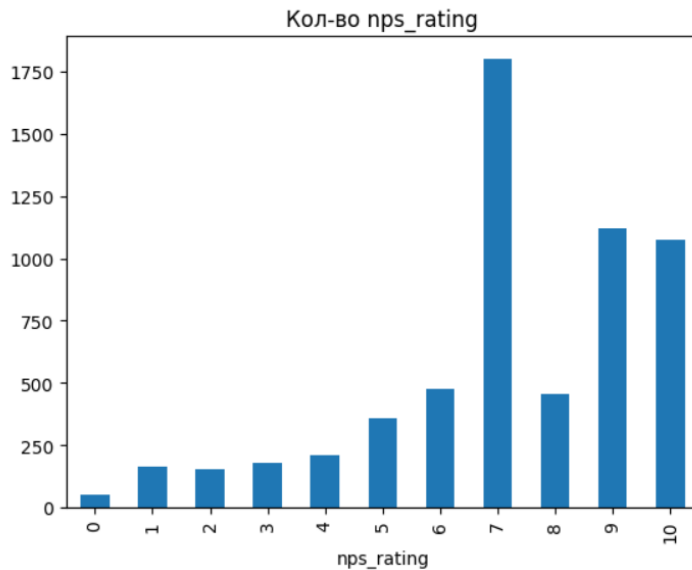
	response_date	user_id	nps_rating	source	nps_group
0	2020-11-06	11037	7	email	passive
1	2020-12-24	34434	9	email	promoter
2	2020-12-03	49547	8	email	passive
3	2020-10-04	13821	7	email	passive
4	2020-10-23	29407	9	email	promoter
5	2020-12-03	13041	10	email	promoter
6	2020-10-12	37166	9	email	promoter
7	2020-10-15	14489	8	email	passive
8	2020-12-31	44317	7	email	passive
9	2020-12-05	24127	6	email	detractor

2. Задание 1. Реализовать графическое оформление полученных результатов NPS методами Matplotlib.

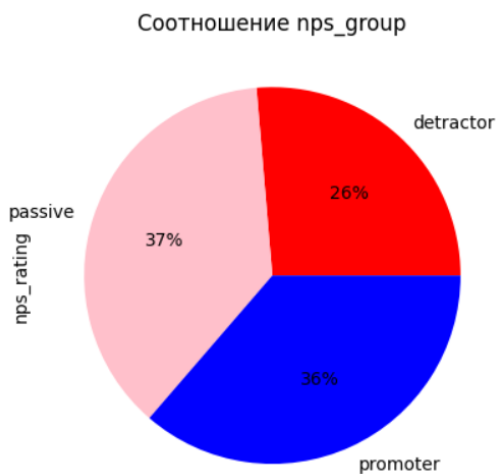
Для реализации подключаем библиотеку и создаем графическое оформление.

```
[17] import matplotlib.pyplot as plt
```

```
[19] q4_nps.groupby('nps_rating')['nps_rating'].count().plot(kind='bar')  
plt.title("Кол-во nps_rating")  
plt.show()
```



```
[30] q4_nps.groupby('nps_group')['nps_rating'].count().plot(kind='pie', autopct='%1.0f%%', colors=['red', 'pink', 'blue'])  
plt.title('Соотношение nps_group')  
plt.show()
```



1750 и 37% это нейтралы, а следовательно большая часть nps рейтингов является "passive", а меньшая "detractor".

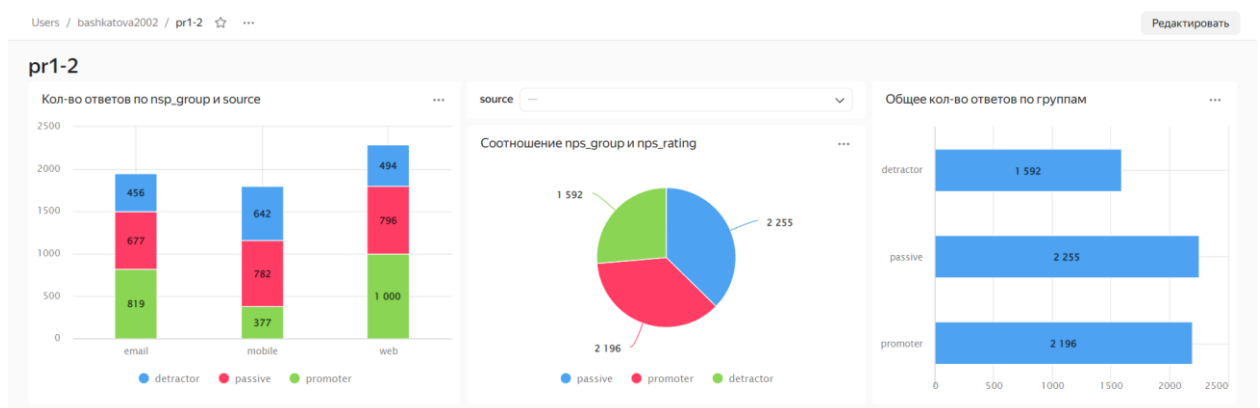
### 3. Реализовать выгрузку полученных результатов в csv или xlsx форматы.

```
[31] from google.colab import files  
q4_nps.to_csv('bashkatova_pr_01_2.csv')  
files.download('bashkatova_pr_01_2.csv')
```

```
[32] from google.colab import files  
q4_nps.to_excel('bashkatova_pr_01_2.xlsx')  
files.download('bashkatova_pr_01_2.xlsx')
```

#### 4. Визуализировать входные и выходные результаты в Yandex DataLens и Google dashboard.

##### Визуализация в Yandex DataLens.



##### Визуализация в Google dashboard.

