

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

ДИСЦИПЛИНА:  
«Инструменты для хранения и обработки больших данных»  
**пр\_01-2**  
**Тема:**  
**«Product analytics»**

Выполнила: Николаева С.Г., АДЭУ-201  
Преподаватель: Босенко Тимур Муртазович

Москва  
2023

Задача:

## Самостоятельная работа

- Реализовать графическое оформление полученных результатов NPS методами Matplotlib.
- Реализовать выгрузку полученных результатов в csv или xlsx форматы.
- Визуализировать входные и выходные результаты в Yandex DataLens и Google dashboard.

Необходимо показать соотношение количества пользователей, которые довольны сервисом, относятся нейтрально и не довольны.

```
Самостоятельная работа
Реализовать графическое оформление полученных результатов NPS методами Matplotlib.

q4_nps.head()

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

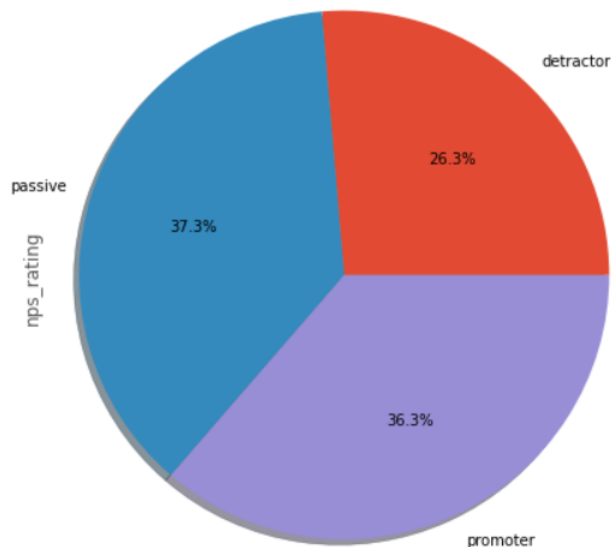
[27] %matplotlib inline
import matplotlib.pyplot as plt
import matplotlib as mpl

[28] mpl.style.use(['ggplot'])

[35] count = q4_nps.groupby('nps_group', axis=0).count()
```

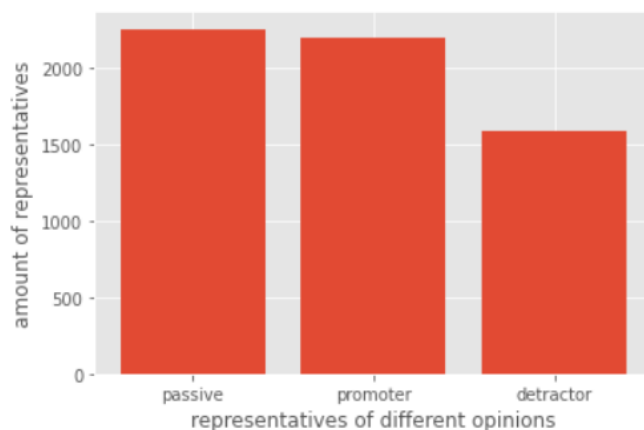
```
count['nps_rating'].plot(kind='pie', figsize=(7,7),
                        autopct='%1.1f%%',
                        shadow=True)
plt.title('The ratio of representatives of different opinions')
plt.axis('equal')
plt.show()
```

The ratio of representatives of different opinions



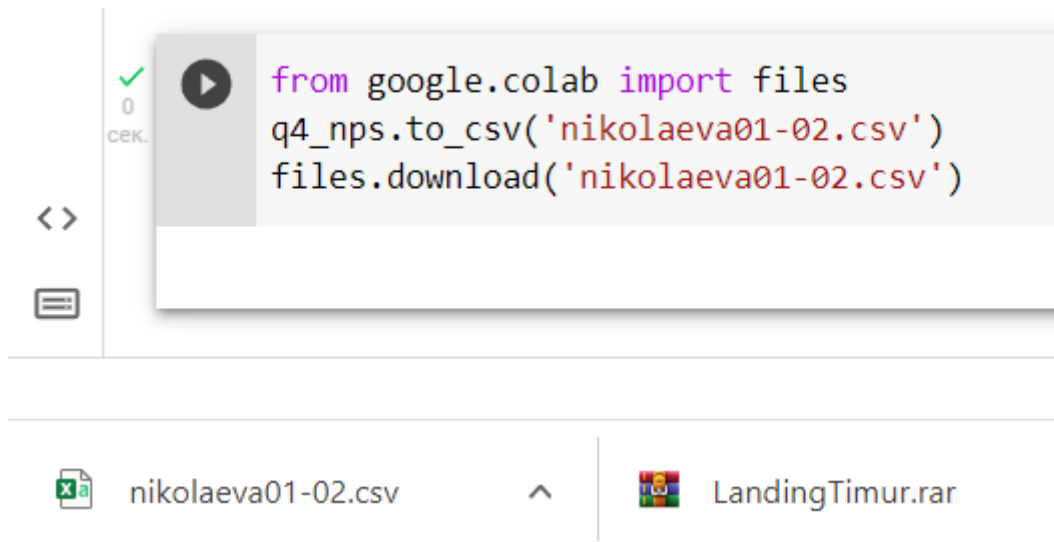
```
opinions = q4_nps['nps_group'].value_counts().head(7)
plt.xlabel('representatives of different opinions')
plt.ylabel('amount of representatives')
plt.bar(x=opinions.index, height=opinions.values)
```

<BarContainer object of 3 artists>



**ВЫВОД:** по данным графикам отчетливо видно, что большинство нейтрально относятся к работе сервиса, при этом почти такое же количество пользователей довольны работой сервиса, а меньшее количество не довольны. Плохим звоночком является то, что представителей таких мнений примерно равное количество, а компании стремятся, чтобы довольных пользователей было гораздо больше, чем остальных.

**Выгрузка полученных результатов в csv или xlsx форматы.**



Ссылка на файл python: [https://github.com/SofiaNikolaeva-adey-201/Big-Data-Storage-and-Processing-Tools/blob/main/Nikolaeva\\_pr\\_01\\_02.ipynb](https://github.com/SofiaNikolaeva-adey-201/Big-Data-Storage-and-Processing-Tools/blob/main/Nikolaeva_pr_01_02.ipynb)

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

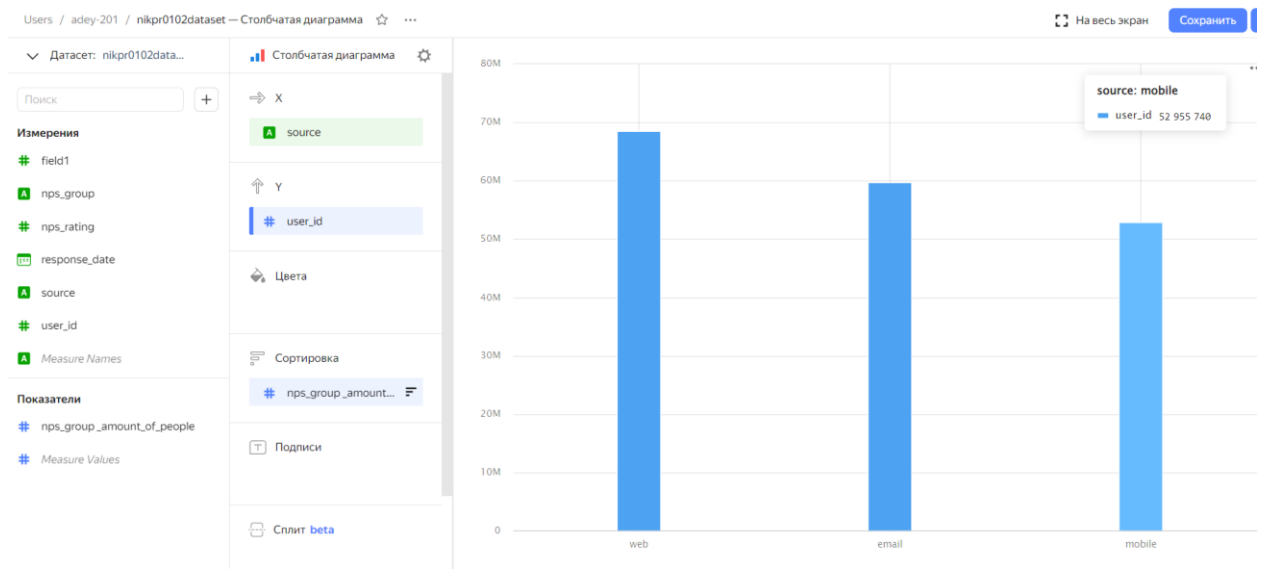
Users / adey-201 / nikpr0102dataset ☆ ... Созда

Источники Поля Параметры Фильтрация Обновить поля Предпросмотр + Добавить поле Имя поля

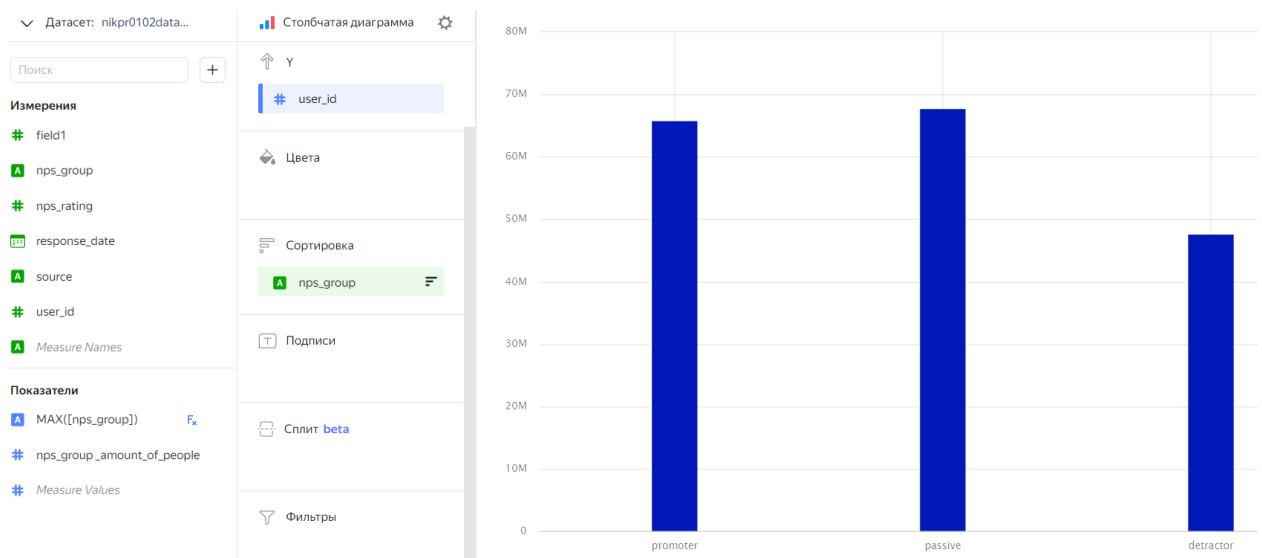
#	Имя ↓	Источник поля ↓	Тип ↓	Агрегация ↓	Описание ↓
1	field1	csv.field1	# Целое число	Нет	
2	response_date	csv.response_date	📅 Дата	Нет	
3	user_id	csv.user_id	# Целое число	Нет	
4	nps_rating	csv.nps_rating	# Целое число	Нет	
5	source	csv.source	📄 Строка	Нет	
6	nps_group	csv.nps_group	📄 Строка	Нет	
7	nps_group_amount_of_people	csv.nps_group	📄 Строка	Количество уника...	

Предпросмотр Количество строк: 10 не больше 1 000

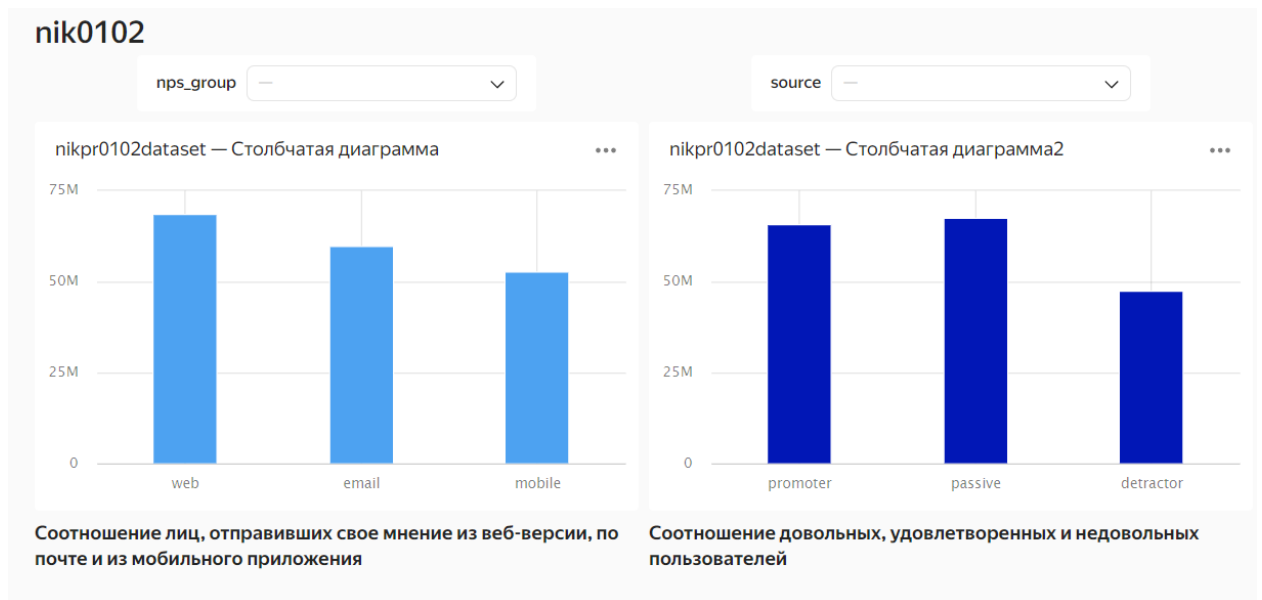
#	field1	response_date	user_id	nps_rating	source	nps_group	nps_group (1
1	0	2020-11-06	11037	7	email	passive	passive
2	1	2020-12-24	34434	9	email	promoter	promoter
3	2	2020-12-03	49547	8	email	passive	passive
4	3	2020-10-04	13821	7	email	passive	passive



Такое представление отражает соотношение количества проголосовавших из различных источников, больше всего – из веб-версии.

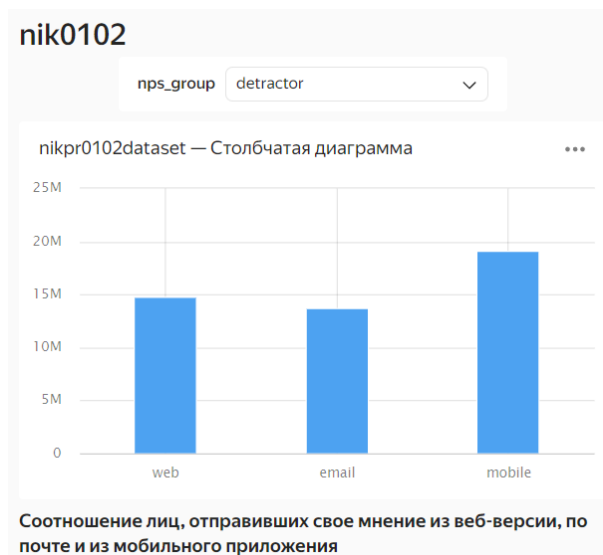


На данном чарте видно соотношение количества лиц, которые довольны, не довольны работой сервиса, а также относятся нейтрально. Больше всего лиц, относящихся нейтрально, немного меньше довольных и чуть меньше недовольных.

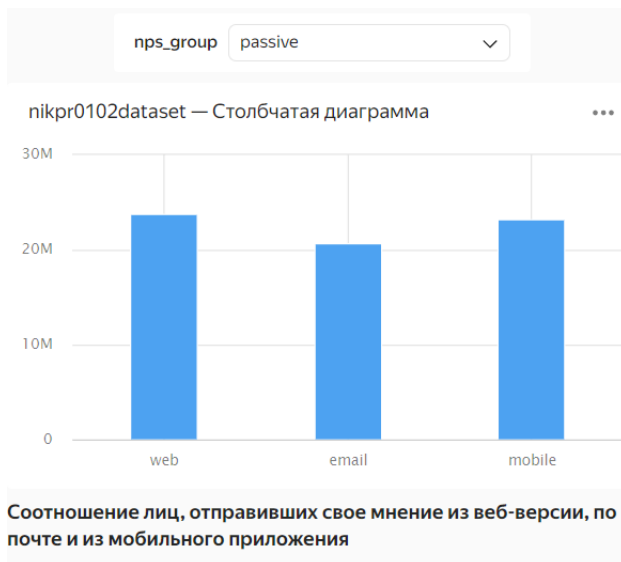


Селекторы, добавленные на итоговый, дашборд позволяют отследить соотношение довольных, удовлетворенных и не довольных пользователей в разрезе источника мнения.

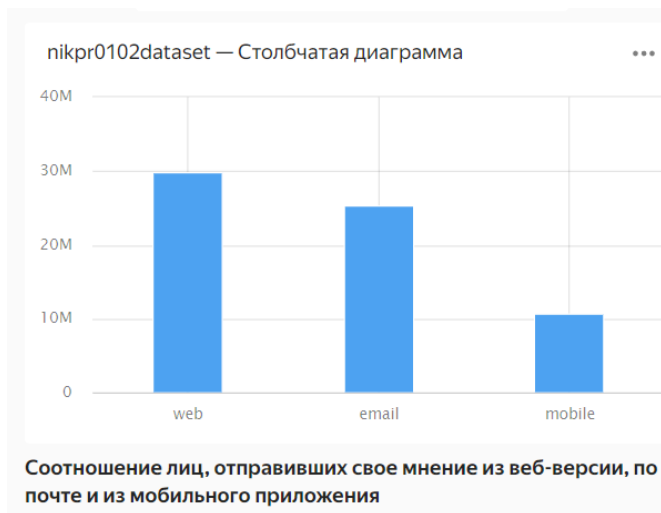
Так, недовольных больше в мобильном приложении:



Удовлетворенных больше в веб-версии и в мобильном приложении:



Большинство довольны веб-версией сервиса:



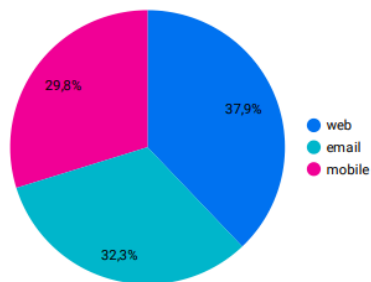
Результат по ссылке:

<https://datalens.yandex/rr6juh9fg7qh>

---

**Google dashboard**

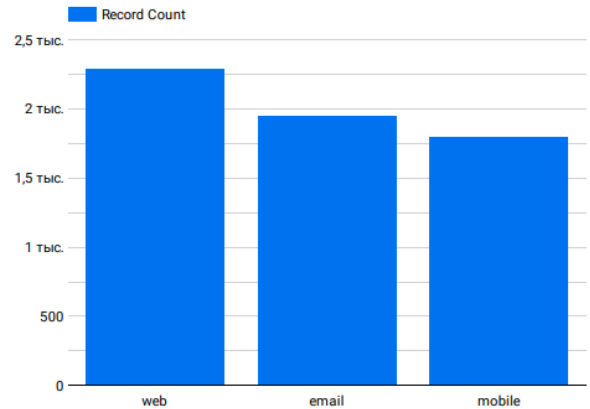
соотношение количества  
проголосовавших из различных  
источников



	source	Record Count ▾
1.	web	2 290
2.	email	1 952
3.	mobile	1 801

1 - 3 / 3 < >

соотношение количества лиц, которые  
довольны, не довольны работой  
сервиса, а также относятся нейтрально



source ▾

Данный дашборд аналогичен дашборду, созданному в Yandex DataLens .

Результат: <https://lookerstudio.google.com/reporting/76efa6e2-5346-4303-8dec-cb896bbafe42>

Ссылка на файл python: [https://github.com/SofiaNikolaeva-adey-201/Big-Data-Storage-and-Processing-Tools/blob/main/Nikolaeva\\_pr\\_01\\_02.ipynb](https://github.com/SofiaNikolaeva-adey-201/Big-Data-Storage-and-Processing-Tools/blob/main/Nikolaeva_pr_01_02.ipynb)

Ссылка на даталенс: <https://datalens.yandex/rr6juhx9fg7qh>