

▼ Подготовка

```
1 import pandas as pd
2 import matplotlib
3 import plotly.express as px
4 from sqlalchemy import create_engine
5 import ipywidgets as widgets
6 from ipywidgets import interact, interact_manual
7 import datetime
8 from datetime import timedelta
9 from plotly.subplots import make_subplots
```

```
1 HOST = '37.139.42.145'
2 DBNAME = 'game-analytics'
3 USER = 'analytics'
4 PASSWORD = 'BrtQyYiJyr29WXN'
5 TABLE_SCHEMA = 'data_viz_1068.project_dataset'
```

```
1 engine = create_engine(f'postgresql://{USER}:{PASSWORD}@{HOST}/{DBNAME}')
```

```

/usr/local/lib/python3.7/dist-packages/psycpg2/__init__.py:144: UserWarning: The psycpg2 wheel package will be renamed from release 2.8; in order to keep installing from binary please use "pip insta
    """

```

```
1 project_dataset = pd.read_sql(f"""
2     SELECT *
3     FROM {TABLE_SCHEMA}
4 """, con=engine)
5 project_dataset
```

[illegible]

```

1 project_dataset.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1176998 entries, 0 to 1176997
Data columns (total 14 columns):
#   Column      Non-Null Count  Dtype
---  -
0   event_time  1176998 non-null  datetime64[ns]
1   event_date  1176998 non-null  object
2   event_name  1176998 non-null  object
3   revenue_usd 1176998 non-null  float64
4   region      1176998 non-null  object
5   country     1176977 non-null  object
6   device_type 1176998 non-null  object
7   platform    1176998 non-null  object
8   cohort_date 1176998 non-null  object
9   user_id     1176998 non-null  int64
10  user_type   1176998 non-null  object
11  content_id  13333 non-null   object
12  currency    13285 non-null   object
13  revenue     1176998 non-null  float64
dtypes: datetime64[ns](1), float64(2), int64(1), object(10)
memory usage: 125.7+ MB

```

```

1 project_dataset['event_date'] = pd.to_datetime(project_dataset['event_date'])
2 project_dataset['cohort_date'] = pd.to_datetime(project_dataset['cohort_date'])
3
4 project_dataset.info()

```

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1176998 entries, 0 to 1176997
Data columns (total 14 columns):
#   Column      Non-Null Count  Dtype
---  -
0   event_time  1176998 non-null  datetime64[ns]
1   event_date  1176998 non-null  datetime64[ns]
2   event_name  1176998 non-null  object
3   revenue_usd 1176998 non-null  float64
4   region      1176998 non-null  object
5   country     1176977 non-null  object
6   device_type 1176998 non-null  object
7   platform    1176998 non-null  object
8   cohort_date 1176998 non-null  datetime64[ns]
9   user_id     1176998 non-null  int64
10  user_type   1176998 non-null  object
11  content_id  13333 non-null   object
12  currency    13285 non-null   object
13  revenue     1176998 non-null  float64
dtypes: datetime64[ns](3), float64(2), int64(1), object(8)
memory usage: 125.7+ MB

```

```

1 project_dataset['event_name'].unique()

```

```

array(['LaunchApp', 'achieve_level_8', 'achieve_level_4',
       'achieve_level_12', 'achieve_level_10', 'af_purchase',
       'achieve_level_14', 'achieve_level_5', 'FirstLaunchApp',
       'achieve_level_9', 'achieve_level_6', 'achieve_level_15',
       'achieve_level_7', 'achieve_level_13', 'achieve_level_16',
       'achieve_level_17', 'achieve_level_19', 'achieve_level_11',
       'achieve_level_18'], dtype=object)

```

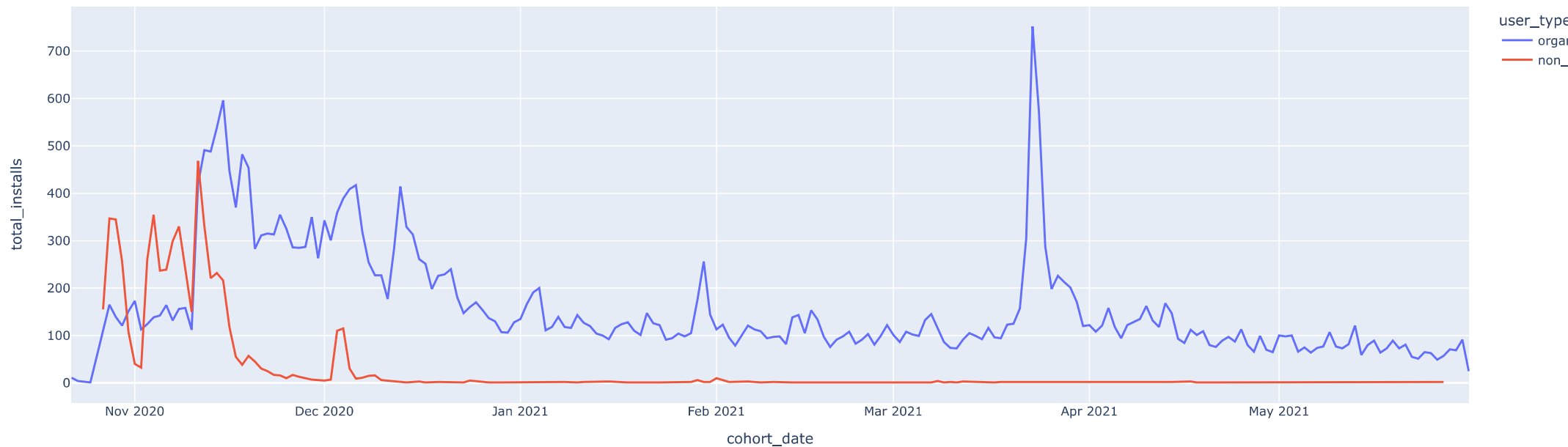
1. Анализ динамики количества ежедневных инсталлов.

Подготовка

```
1 #preinstalls = project_dataset.query('event_date==cohort_date')
2 preinstalls = project_dataset[project_dataset['event_date'] == project_dataset['cohort_date']]
```

По типу пользователя.

```
1 installs_by_user_type = (preinstalls
2                             .groupby(by=['cohort_date', 'user_type'], as_index=False)
3                             .agg(total_installs=('user_id', 'nunique')))
4 fig = px.line(installs_by_user_type, x='cohort_date', y='total_installs', color='user_type')
5 fig.show()
```



Неорганические.

С конца Октября и по вторую треть Декабря видны инсталлы неорганических пользователей с пиком в 11 Ноября.

Это похоже на начало рекламной компании (27 Октября).

К 25 Декабря деньги у рекламной компании практически закончились. 3 Декабря докинули денег на рекламу.

7 Декабря решили забить полностью на рекламу (/платное продвижение).

После 7 декабря неорганические пользователи практически не приходят.

И 27 Мая (ровно через 7 месяцев) платные инсталлы прекращаются полностью.

Органика.

28 Октября - 10 Ноября органические пользователи регулярно приходят - примерно 140 в день.

11 Ноября происходит резкий скачок органических инсталов с пиком 15 Ноября.

В это время так же был пик неорганических пользователей, возможно они рассказали своим знакомым об игре и это привело к увеличению органики.

Может быть в этот день была еще либа какая акция направленная на привлечение новых пользователей.

Далее число органических инсталов медленно но верно падает. (этому так же способствует прекращение рекламной компании). 23 Марта происходит какое то событие которое приводит к максимальному всплеску инсталов. Который впрочем так же быстро спадает. (может это был заказ рекламы у какого нибудь известного стримера (трафик при этом получился органическим так как игру просто скачали из магазина). А может какой нибудь 1 дневный featuring :).

Далее инсталлы медленно уменьшаются.

► По региону, стране.

[] ↳ Скрыто 3 ячейки.

► По платформе.

[] ↳ Скрыто 2 ячейки.

► 2. Анализ удержания по региону.

[] ↳ Скрыто 19 ячеек.

► 3. Анализ поведения в первые 24 часа

[] ↳ Скрыто 30 ячеек.

