

Учебный проект №12

Построение дашборда на основе выгруженных данных (SQL)

Загрузка данных из БД для проекта

```
In [2]: # импортируем библиотеки
import pandas as pd
!pip install psycopg2
try:
    from sqlalchemy import create_engine
except:
    !pip install sqlalchemy
    !pip install psycopg2
    from sqlalchemy import create_engine

db_config = {'user': 'praktikum_student', # имя пользователя
            'pwd': '*****', # пароль
            'host': 'rc1b-wcoijxj3yxfsf3fs.mdb.yandexcloud.net',
            'port': 6432, # порт подключения
            'db': 'data-analyst-zen-project-db'} # название базы данных

connection_string = 'postgresql://{user}:{pwd}@{host}:{port}/{db}'.format(db_config['user'],
                                                                           db_config['pwd'],
                                                                           db_config['host'],
                                                                           db_config['port'],
                                                                           db_config['db'])

engine = create_engine(connection_string)
```

Defaulting to user installation because normal site-packages is not writeable
Requirement already satisfied: psycopg2 in c:\users\xrust\appdata\roaming\python\python311\site-packages (2.9.9)

```
In [3]: query = '''
        SELECT * FROM dash_visits
        '''
```

```
In [4]: df = pd.io.sql.read_sql(query, con = engine)
```

```
In [5]: df.head(5)
```

```
Out[5]:
```

	record_id	item_topic	source_topic	age_segment	dt	visits
0	1040597	Деньги	Авто	18-25	2019-09-24 18:32:00	3
1	1040598	Деньги	Авто	18-25	2019-09-24 18:35:00	1
2	1040599	Деньги	Авто	18-25	2019-09-24 18:54:00	4
3	1040600	Деньги	Авто	18-25	2019-09-24 18:55:00	17
4	1040601	Деньги	Авто	18-25	2019-09-24 18:56:00	27

```
In [6]: df.info(5)
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 30745 entries, 0 to 30744
Data columns (total 6 columns):
#   Column          Non-Null Count  Dtype
---  -
0   record_id       30745 non-null  int64
1   item_topic      30745 non-null  object
2   source_topic    30745 non-null  object
3   age_segment     30745 non-null  object
4   dt              30745 non-null  datetime64[ns]
5   visits          30745 non-null  int64
dtypes: datetime64[ns](1), int64(2), object(3)
memory usage: 1.4+ MB
```

```
In [7]: df.to_csv('dash_visits.csv', index=False)
```

Исходные данные загружены

Дашборд Tableau доступен по [ссылке](#)

Презентация доступна по [ссылке](#)

```
In [ ]:
```