

Автоматизация - подготовка к проекту по средству tableau

```
Ввод [1]: #!/pip install psycopg2-binary
# импортируем библиотеки
import pandas as pd
from sqlalchemy import create_engine

db_config = {'user': 'praktikum_student', # имя пользователя
             'pwd': 'Sdf4$2;d-d30pp', # 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)

query= ''' SELECT *
          FROM dash_visits
        '''

dash_visits = pd.io.sql.read_sql(query, con = engine)
```

Requirement already satisfied: psycopg2-binary in /opt/conda/lib/python3.9/site-packages (2.9.2)

```
Ввод [2]: dash_visits.to_csv('dash_visits.csv')
```

```
Ввод [3]: dash_visits
```

Out[3]:

	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
...
30740	1071337	Юмор	Финансы	36-40	2019-09-24 18:57:00	2
30741	1071338	Юмор	Финансы	36-40	2019-09-24 19:00:00	1
30742	1071339	Юмор	Финансы	41-45	2019-09-24 18:54:00	1
30743	1071340	Юмор	Финансы	41-45	2019-09-24 18:56:00	1
30744	1071341	Юмор	Финансы	41-45	2019-09-24 19:00:00	1

30745 rows × 6 columns

```
Ввод [4]: dash_visits['item_topic'].unique()
```

```
Out[4]: array(['Деньги', 'Дети', 'Женская психология', 'Женщины', 'Здоровье',  
              'Знаменитости', 'Интересные факты', 'Искусство', 'История',  
              'Красота', 'Культура', 'Наука', 'Общество', 'Отношения',  
              'Подборки', 'Полезные советы', 'Психология', 'Путешествия',  
              'Рассказы', 'Россия', 'Семья', 'Скандалы', 'Туризм', 'Шоу', 'Юмор'],  
            dtype=object)
```

```
Ввод [6]: dash_visits['source_topic'].unique()
```

```
Out[6]: array(['Авто', 'Деньги', 'Дети', 'Еда', 'Здоровье', 'Знаменитости',  
              'Интерьеры', 'Искусство', 'История', 'Кино', 'Музыка', 'Одежда',  
              'Полезные советы', 'Политика', 'Психология', 'Путешествия',  
              'Ремонт', 'Россия', 'Сад и дача', 'Сделай сам',  
              'Семейные отношения', 'Семья', 'Спорт', 'Строительство',  
              'Технологии', 'Финансы'], dtype=object)
```

```
Ввод [7]: dash_visits['dt'].max()
```

```
Out[7]: Timestamp('2019-09-24 19:00:00')
```

```
Ввод [8]: dash_visits['dt'].min()
```

```
Out[8]: Timestamp('2019-09-24 18:28:00')
```

```
Ввод [9]: persent_item_1 = (dash_visits
    .groupby('dt')
    .agg({'visits':['sum']}))
    .sort_values(by=('visits', 'sum'), ascending=False)
    .copy()
persent_item_1
```

Out[9]:

	visits
	sum
dt	
2019-09-24 18:58:00	61247
2019-09-24 18:57:00	60759
2019-09-24 18:56:00	54168
2019-09-24 18:59:00	53366
2019-09-24 18:55:00	32696
2019-09-24 19:00:00	17704
2019-09-24 18:54:00	14140
2019-09-24 18:31:00	2592
2019-09-24 18:32:00	2587
2019-09-24 18:30:00	2442
2019-09-24 18:53:00	2057
2019-09-24 18:29:00	1817
2019-09-24 18:33:00	1726
2019-09-24 18:52:00	1112
2019-09-24 18:28:00	835
2019-09-24 18:34:00	763
2019-09-24 18:35:00	196

```
Ввод [11]: persent_item_topic = (dash_visits
    .groupby('item_topic')
    .agg({'visits':['sum']}))
    .sort_values(by=('visits', 'sum'), ascending=False)
    .copy()
)/dash_visits['visits'].sum()*100
```

Ввод [12]:

persent_item_topic

Out[12]:

visits	
sum	
item_topic	
Наука	7.006934
Отношения	6.662003
Интересные факты	6.428611
Общество	6.331256
Подборки	5.729078
Россия	5.469251
Полезные советы	4.975710
История	4.960881
Семья	3.835181
Женщины	3.706880
Дети	3.542473
Рассказы	3.516684
Здоровье	3.352278
Деньги	3.317462
Культура	3.289739
Красота	3.163694
Туризм	3.066340
Юмор	3.029590
Скандалы	2.996064
Путешествия	2.985103
Искусство	2.745264
Психология	2.590528
Женская психология	2.494141
Шоу	2.421286
Знаменитости	2.383570

Ввод []: