

Изучение срезов данных

Фильтрация строк датафрейма с помощью логических операторов

Чтобы отфильтровать данные по нескольким условиям, объединяем их логическими операторами:

```
df[(условие_1) & (условие_2) & ...]
```

&	логическое И
	логическое ИЛИ
~	логическое отрицание

```
# ищем билеты из Москвы ценой менее 21 000 рублей – логическое И  
df[(df['From'] == 'Moscow') & (df['Price'] < 21000)]
```

```
# ищем либо прямой, либо дешёвый рейс – логическое ИЛИ  
df[(df['Is_Direct'] == True) | (df['Price'] < 20000)]
```

```
# ищем билеты не в Барселону – логическое отрицание  
df[~(df['To'] == 'Barcelona')]
```

Поиск элементов списка lst в столбце

```
# получаем билеты с вылетом из Москвы или Петербурга  
df[df['From'].isin(['Moscow', 'St. Petersburg'])]
```

Извлечение составляющих даты

```
data['datetime'].dt.date # дата  
data['datetime'].dt.year # год  
data['datetime'].dt.weekday # день недели
```

```
# номер недели в году находим немного по-другому:  
data['datetime'].dt.isocalendar().week
```

Изучение срезов данных

Быстрое получение срезов данных через query()

```
df.query('From == "Moscow"')  
df.query('Price < 21000')
```

```
max_price = 20000  
data.query('"Price" < @max_price')
```

Сдвиг даты и времени

```
data['shifted_dt'] = data['datetime'] + pd.Timedelta(hours=10) # добавить 10 часов
```

Округление времени

```
data['datetime'] = data['datetime'].dt.round('1H') # округлить до 1 часа  
data['datetime'] = data['datetime'].dt.round('1D') # до 1 дня  
data['datetime'] = data['datetime'].dt.round('5T') # до 5 минут  
data['datetime'] = data['datetime'].dt.round('10S') # до 10 секунд  
data['datetime'] = data['datetime'].dt.floor('1H') # округление в меньшую сторону  
data['datetime'] = data['datetime'].dt.ceil('1H') # округление в большую сторону
```

Построение графиков по датафрейму

```
data.plot(x='column1', # столбец значений для горизонтальной оси  
         y='column2', # столбец значений для вертикальной оси  
         style='o-', # стиль заполнения: 'o'(точечный) или 'o-'(точечно-линейный)  
         xlim=(0, 30), # границы по оси X  
         ylim=(30, 0), # границы по оси Y  
         figsize=(4, 5), # размеры картинки: (x_size, y_size)  
         grid=True) # отображать сетку или нет
```

Глоссарий

Срез данных — часть данных из набора, отобранная по определённым условиям