

# Валидация результатов

## Построение столбчатой диаграммы

```
# если не указывать x, по умолчанию ось абсцисс займет индекс датафрейма
data.plot(x='column_1', y='column_2', kind='bar')
```

## Построение круговой диаграммы

```
# названия категорий (секторов) берутся из индекса датафрейма
data.plot(y='column', kind='pie')
```

## Выборочное изменение значения

```
# если не выполняется условие в первом параметре, то значение заменяется на new_value
data['column'].where(s > control_value, new_value)
```

```
# в новом столбце с оценками студентов все двойки заменили на тройки
data['new_mark_system'] = data['mark'].where(data['mark'] >= 3, 3)
```

## Работа с категориями из одного столбца

```
for group_name, group_data in data.groupby('column'):
    # group_name – строка с названием группы
    # group_data – датафрейм со всеми данными из неё
```

```
for name, group_data in good_data.groupby('group_name'):
    group_data.hist('time_spent', bins=50) # строим гистограмму для очередной АЗС
    plt.title(name)                        # подписываем график названием этой АЗС
```