

# Шпаргалка

## Поиск явных дубликатов: способ 1

```
import pandas as pd
data = pd.read_csv('data.csv')
print(data['column'].duplicated().sum())
```

## Поиск явных дубликатов: способ 2

```
import pandas as pd
data = pd.read_csv('data.csv')
print(data['column'].value_counts())
```

## Поиск неявных дубликатов с учётом регистра

```
import pandas as pd
data = pd.read_csv('data.csv')
data['column'] = data['column'].str.lower()
```

## Удаление дубликатов

```
data['column'] = data['column'].drop_duplicates()
```