

# Лабораторная работа №1

В этой лабораторной работе, нам необходимо было научиться работать с данными. Для этого были приведены 10 заданий, которые мы успешно выполнили:

Задание 1. Сколько человек спаслись и сколько погибли?

```
1 print("Выжило: ")
2 df[df['Survived'] == 1].PassengerId.count()
```

Out:

```
1 Выжило:
2 342
```

```
1 print('погибло')
2 df[df['Survived'] == 0].PassengerId.count()
```

Out:

```
1 Погибло
2 549
```

Задание 2. Выведите распределение переменной Survived отдельно для мужчин и отдельно для женщин

```
1 print('Муж выжило: ')
2 df[(df['Sex'] == 'male') & (df['Survived'] == 1)]['Survived'].count()
```

Out:

```
1 Муж выжило:
2 109
```

```
1 print('Муж погибло: ')
2 df[(df['Sex'] == 'male') & (df['Survived'] == 0)]['Survived'].count()
```

Out:

```
1 Муж погибло:
2 468
```

```
1 print('Жен выжило: ')
2 df[(df['Sex'] == 'female') & (df['Survived'] == 1)]['Survived'].count()
```

Out:

```
1 Жен выжило:
2 233
```

```
1 print('Жен погбило: ')
2 df[(df['Sex'] == 'female') & (df['Survived'] == 0)][['Survived']].count()
```

Out:

```
1 Жен погбило:
2 81
```

Задание 3. Каковы медиана и стандартное отклонение цен за билет?

```
1 print('median: ')
2 df['Fare'].median()
```

Out:

```
1 median:
2 14.4542
```

```
1 print('std: ')
2 df['Fare'].describe()['std']
```

Out:

```
1 std:
2 49.6934285971809
```

Задание 4. Подсчитайте долю выживших среди людей моложе 25 лет и старше 55.

```
1 df[((df['Age'] < 25) | (df['Age'] > 55)) & (df['Survived'] == 1)]
   ['PassengerId'].count()/df[(df['Age'] < 25) | (df['Age'] > 55)]
   ['PassengerId'].count()
```

Out:

```
1 0.4088050314465409
```

Задание 5. Подсчитайте долю выживших среди мужчин и женщин.

```
1 df[(df['Sex'] == 'male') & (df['Survived'] ==
   1)].PassengerId.count()/df[df['Sex'] == 'male'].PassengerId.count()
```

Out:

```
1 0.18890814558058924
```

```
1 | df[(df['Sex'] == 'female') & (df['Survived'] == 1)].PassengerId.count()/df[df['Sex'] == 'female'].PassengerId.count()
```

Out:

```
1 | 0.7420382165605095
```

Задание 6. В каком городе больше всего село пассажиров?

```
1 | df['Embarked'].mode()[0]
```

Out:

```
1 | 'S'
```

Задание 7. Подсчитайте средний возраст выживших

```
1 | df[df.Survived == 1]['Age'].mean()
```

Out:

```
1 | 28.343689655172415
```

Задание 8. Подсчитайте средний возраст пассажиров каждого из трех классов

```
1 | df[df['Pclass'] == 1]['Age'].mean()
```

Out:

```
1 | 38.233440860215055
```

```
1 | df[df['Pclass'] == 2]['Age'].mean()
```

Out:

```
1 | 29.87763005780347
```

```
1 | df[df['Pclass'] == 3]['Age'].mean()
```

Out:

```
1 | 25.14061971830986
```

Задание 9. Подсчитайте средний возраст выживших мужчин, купивших билет во 2 класс

```
1 df[(df['Pclass'] == 2) & (df['Survived'] == 1) & (df['Sex'] == 'male')][  
  'Age'].mean()
```

Out:

```
1 16.022000000000002
```

Задание 10\*. Оцените, как влияет на выживаемость пассажиров наличие их родственников на борту.

In :

```
1 cold = ["Survived", "SibSp", "Parch"]  
2 df[cold].corr()
```

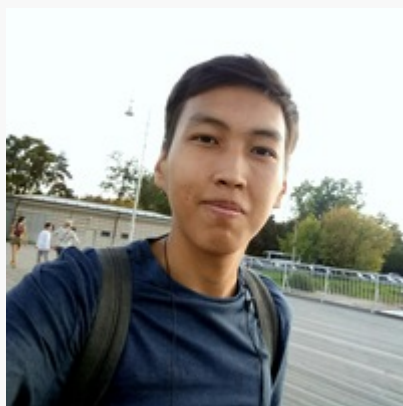
Out:

	Survived	SibSp	Parch
Survived	1.000000	-0.035322	0.081629
SibSp	-0.035322	1.000000	0.414838
Parch	0.081629	0.414838	1.000000

## Вывод

В этой лабораторной работе мы научились работать с массивами данных с помощью библиотеки pandas. Это был первый наш опыт использования языка Python как программы для работы с данными. Действительно, это мощная библиотека, которая может изменить взгляд на Python, так как с ней он становится другим.

## Authors



[Arthur Kupriyanov](#)



[Artyom Kolokolov](#)

Группа: P3212