

Работа с Excel

Лабораторная работа 7

1. Загрузите данные из файлов `reviews_sample.csv` (ЛР2) и `recipes_sample_with_tags_ingredients.csv` (ЛР5) в виде `pd.DataFrame`. Обратите внимание на корректное считывание столбца(ов) с индексами. Оставьте в таблице с рецептами следующие столбцы: `id`, `name`, `minutes`, `submitted`, `description`, `n_ingredients`

In [15]: `import pandas as pd`

In [16]: `reviews_csv = pd.read_csv(r'C:\Users\sunya\Desktop\6семестр\интернетвещей\excel\reviews_sample.csv')`
`recipes_csv = pd.read_csv(r'C:\Users\sunya\Desktop\6семестр\интернетвещей\excel\recipes_sample_with_tags_ingredients.csv')`

In [4]: `reviews_csv`

Out[4]:

	id	user_id	recipe_id	date	rating	review
0	370476	21752	57993	2003-05-01	5	Last week whole sides of frozen salmon fillet ...
1	624300	431813	142201	2007-09-16	5	So simple and so tasty! I used a yellow capsu...
2	187037	400708	252013	2008-01-10	4	Very nice breakfast HH, easy to make and yummy...
3	706134	2001852463	404716	2017-12-11	5	These are a favorite for the holidays and so e...
4	312179	95810	129396	2008-03-14	5	Excellent soup! The tomato flavor is just gre...
...
126691	1013457	1270706	335534	2009-05-17	4	This recipe was great! I made it last night. I...
126692	158736	2282344	8701	2012-06-02	0	This recipe is outstanding. I followed the rec...

In [6]: recipes_csv

Out[6]:

	name	id	minutes	contributor_id	submitted	n_steps	description	n_in
0	george s at the cove black bean soup	44123	90	35193	2002-10-25	11	an original recipe created by chef scott meska...	
1	healthy for them yogurt popsicles	67664	10	91970	2003-07-26	3	my children and their friends ask for my homem...	
2	i can t believe it s spinach	38798	30	1533	2002-08-29	5	these were so go, it surprised even me.	
3	italian gut busters	35173	45	22724	2002-07-27	7	my sister-in-law made these for us at a family...	
4	love is in the air beef fondue sauces	84797	25	4470	2004-02-23	4	i think a fondue is a very romantic casual din...	
...
29995	zurie s holey rustic olive and cheddar bread	267661	80	200862	2007-11-25	16	this is based on a french recipe but i changed...	
29996	zwetschkuchen bavarian plum cake	386977	240	177443	2009-08-24	22	this is a traditional fresh plum cake, thought...	
29997	zwiebelkuchen southwest german onion cake	103312	75	161745	2004-11-03	10	this is a traditional late summer early fall s...	
29998	zydeco soup	486161	60	227978	2012-08-29	7	this is a delicious soup that i originally fou...	
29999	cookies by design cookies on a stick	298512	29	506822	2008-04-15	9	i've heard of the 'cookies by design' company,...	

30000 rows × 11 columns



```
In [4]: recipes = recipes_csv.drop(columns=['contributor_id', 'n_steps', 'n_tags', 't
```

```
In [8]: recipes
```

```
Out[8]:
```

	name	id	minutes	submitted	description	n_ingredients
0	george s at the cove black bean soup	44123	90	2002-10-25	an original recipe created by chef scott meska...	18
1	healthy for them yogurt popsicles	67664	10	2003-07-26	my children and their friends ask for my homem...	3
2	i can t believe it s spinach	38798	30	2002-08-29	these were so go, it surprised even me.	8
3	italian gut busters	35173	45	2002-07-27	my sister-in-law made these for us at a family...	9
4	love is in the air beef fondue sauces	84797	25	2004-02-23	i think a fondue is a very romantic casual din...	12
...
29995	zurie s holey rustic olive and cheddar bread	267661	80	2007-11-25	this is based on a french recipe but i changed...	10
29996	zwetschgengkuchen bavarian plum cake	386977	240	2009-08-24	this is a traditional fresh plum cake, thought...	11
29997	zwiebelkuchen southwest german onion cake	103312	75	2004-11-03	this is a traditional late summer early fall s...	13
29998	zydeco soup	486161	60	2012-08-29	this is a delicious soup that i originally fou...	22
29999	cookies by design cookies on a stick	298512	29	2008-04-15	i've heard of the 'cookies by design' company,...	10

30000 rows × 6 columns

- Случайным образом выберите 5% строк из каждой таблицы и сохраните две таблицы на разные листы в один файл `recipes.xlsx`. Дайте листам названия "Рецепты" и "Отзывы", соответствующие содержанию таблиц.

```
In [5]: recipes_save = recipes.sample(n=int(recipes.shape[0]*0.05))
reviews_save = reviews_csv.sample(n=int(reviews_csv.shape[0]*0.05))
```

```
In [6]: with pd.ExcelWriter(r'C:\Users\sunya\Desktop\6семестр\интернетвещей\excel\sam
        recipes_save.to_excel(writer, sheet_name="Рецепты")
        reviews_save.to_excel(writer, sheet_name="Отзывы")
```

3. Используя `xlwings`, добавьте на лист `Рецепты` столбец `seconds_assign`, показывающий время выполнения рецепта в секундах. Выполните задание при помощи присваивания массива значений диапазону ячеек.

```
In [19]: import xlwings as xw
```

```
In [20]: recipes_excel = xw.Book(r'C:\Users\sunya\Desktop\6семестр\интернетвещей\excel\
```

```
In [21]: recipes = recipes_excel.sheets['Рецепты']
```

```
In [4]: minutes = recipes.range('D2').expand('down').value
```

```
In [5]: import numpy as np
        minutes = list(np.array(minutes)*60)
```

```
In [6]: recipes.range('H1').value = 'seconds_assign'
```

```
In [8]: recipes.range('H2').options(transpose=True).value = minutes
```

4. Используя `xlwings`, добавьте на лист `Рецепты` столбец `seconds_formula`, показывающий время выполнения рецепта в секундах. Выполните задание при помощи формул Excel.

```
In [9]: recipes_excel.sheets['Рецепты'].activate()
```

```
In [10]: xw.Range('I2').formula = f'=D2*60'
        from xlwings.constants import AutoFillType
        xw.Range('I2').api.AutoFill(xw.Range('I2:I1501').api,
                                     AutoFillType.xlFillDefault)
```

```
Out[10]: True
```

```
In [11]: recipes.range('I1').value = 'seconds_formula'
```

5. Добавьте на лист `Рецепты` столбец `n_reviews`, содержащий кол-во отзывов для этого рецепта. Выполните задание при помощи формул Excel.

```
In [12]: recipes.range('J1').value = 'n_reviews'
```

```
In [13]: xw.Range('J2').formula = f'=COUNTIF(Отзывы!$D$2:$D$6335, C2)'  
from xlwings.constants import AutoFillType  
xw.Range('J2').api.AutoFill(xw.Range('J2:J1501').api,  
                             AutoFillType.xlFillDefault)
```

```
Out[13]: True
```

6. Сделайте названия всех добавленных столбцов полужирными и выровняйте по центру ячейки.

```
In [15]: recipes.range('H1').api.Font.Bold = True  
recipes.range('I1').api.Font.Bold = True  
recipes.range('J1').api.Font.Bold = True  
  
recipes.range('J1').api.HorizontalAlignment = xw.constants.HAlign.xlHAlignCenter  
recipes.range('H1').api.HorizontalAlignment = xw.constants.HAlign.xlHAlignCenter  
recipes.range('I1').api.HorizontalAlignment = xw.constants.HAlign.xlHAlignCenter
```

7. Раскрасьте ячейки столбца minutes в соответствии со следующим правилом: если рецепт выполняется быстрее 5 минут, то цвет - зеленый; от 5 до 10 минут - жёлтый; и больше 10 - красный.

```
In [6]: for a_cell in recipes.range('D2').expand('down'):  
    if a_cell.value < 5:  
        a_cell.color = (0,255,0)  
    elif a_cell.value > 10:  
        a_cell.color = (255,0,0)  
    else:  
        a_cell.color = (255,255,0)
```

8. Напишите функцию validate() , которая проверяет соответствие всех строк из листа Отзывы следующим правилам:

- Рейтинг - это число от 0 до 5 включительно
- Соответствующий рецепт имеется на листе Рецепты

В случае несоответствия этим правилам, выделите строку красным цветом

```
In [33]: %%writefile validate.py
import xlwings as xw

def validate():
    wb = xw.Book.caller()
    reviews = wb.sheets[1]

    for cell in reviews.range('F2').expand('down'):
        if cell.value < 0 or cell.value > 5:
            cell.color = (255,0,0)

    for cell in reviews.range('D2').expand('down'):
        if cell.value not in wb.sheets[0].range('C2').expand('down').value:
            cell.color = (255,0,0)
```

Overwriting validate.py