

Python 101: Lesson 2

震惊！不打开文件就能读写数据！

Today's Agenda

#1 文件的读写

- 如何创建一个txt文件?
- 如何读取文件内容?

#2 文字的自动生成

- for循环
- list数据类型的运用
- format()函数

1 – 文件的读写

How to read/write a txt file?

A photograph showing a person's hands typing on a laptop keyboard. A smartphone lies next to the laptop, and an open notebook is visible in the foreground.

Let's create your first txt file!

文件操作三部曲：

Step 1: 打开

```
file = open('file_name', w)
```

Step 2: 使用（读、写、whatever）

```
file.write('your content here')
```

Step 3: 关闭

```
file.close()
```

將以下信息
ak, 32, m, it
elsa, 22, f, hr
tiger, 35, m, it
lisa, 24, f, hr
寫入hr.txt

2- 文字的生成

How to generate texts ?

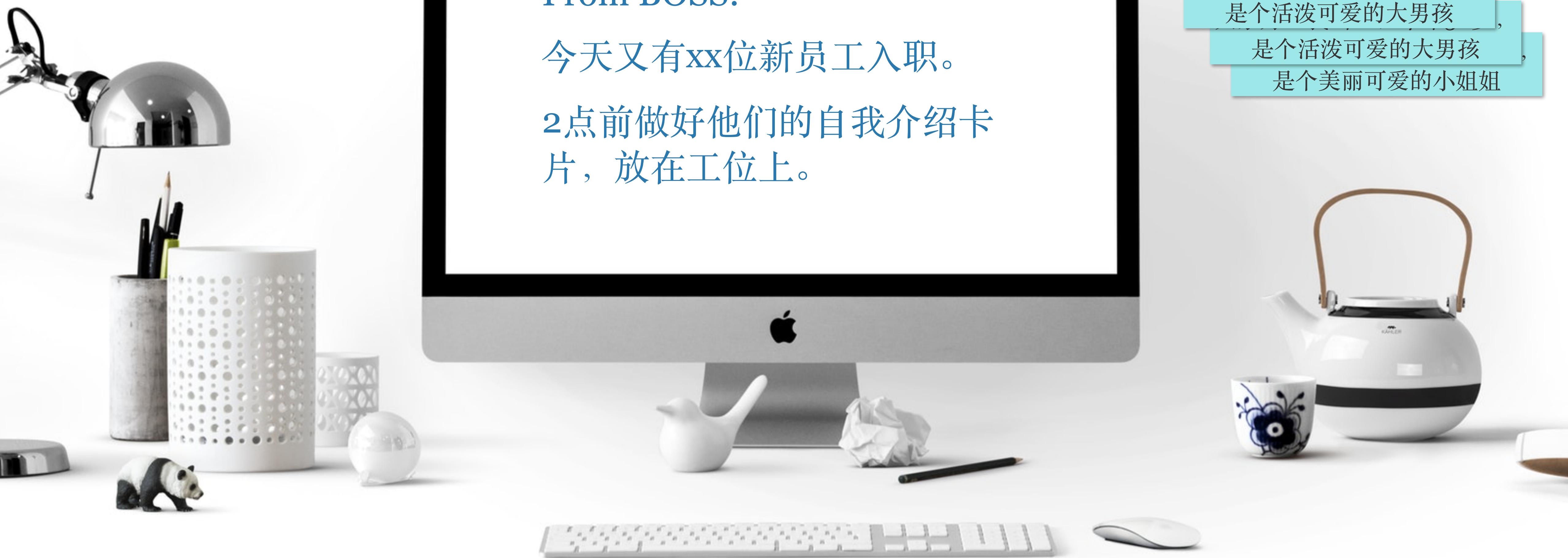
虚拟工作场景：Junior HR的日常

From BOSS:

今天又有xx位新员工入职。

2点前做好他们的自我介绍卡片，放在工位上。

大家好！我叫ak，今年30岁，
是个活泼可爱的大男孩，
是个活泼可爱的大男孩，
是个活泼可爱的大男孩，
是个美丽可爱的小姐姐



循环 Loops

while 循环

只要达到某条件，就循环执行某段程序

for 循环

对列表中的每个元素循环重复执行某段程序

```
for line in data:
```

data 是一个列表



常用的文字处理函数

我拆~

split()

我砍~

strip(), lstrip(), rstrip()

我填填填~

format()

读取hr.txt
生成欢迎语句

大家好，我叫XX，今年x岁，是个活泼
可爱的大男孩/美丽冻人的小姐姐
并保存到report.txt

读取hr.txt

生成模版

姓名 先生/小姐，就职于xx部门
打印到屏幕

读取hr.txt
生成模版

姓名 先生/小姐， 就职于xx部门
保存到tutorial.txt

Recap - 敲黑板，记重点！

open()函数

- ✓ ‘r’, ‘w’, ‘a’ 三种模式
- ✓ 默认为‘r’ 模式
- ✓ ‘w’会清空文件内容，谨慎使用

list列表

- ✓ 用[]表示，元素间用逗号隔开
- ✓ 列表的元素可以是各种数据类型，也不必为相同类型

file.close()

- ✓ 文件操作三部曲，打开、使用、关闭
- ✓ with open() as file:

for循环

- ✓ 通常与列表结合使用
- ✓ for i in range(x)

Tips: 试一下快捷键 Ctrl + D



From BOSS Again:

今天之内要做完
3,000个名牌

Who怕who啊!

Powered by Saint Shine Education



Lesson 3

Python 101

Powered by Saint Shine Education

建立一个列表，里面包含1-10
打印3
打印4到8
打印7-9
打印1,3,5

读取hr.txt
利用for循环，
打印出每个人的名字，性别和年龄

字典

{key : value , key:value}

读取hr.txt
将每条记录生成字典：
姓名： XXX
信息： 【年龄， 性别， 部门】
所有字典存储到一个list里
打印出所有男性员工

读取hr.txt

生成字典为：

姓名： XX, 年龄： xx, 性别： xx 部门： xx

将所有字典存储到一个list里

打印出所有30岁以下的员工



Lesson 4

Python 101

确保你的机器上有**lesson3.json**文件
并且里面有大量内容
否则请运行左边代码获取

读取lesson3.json文件
中国 直辖市/省份的境外输入疫情数据
在屏幕上打印：
国家/地区： 直辖市/地区： 境外输入
xxx例

读取lesson3.json文件
获得海外国家/地区疫情数据

保存文件如下：

国家/地区： 确诊 确诊新增 累计确诊，
治愈 治愈新增， 死亡 死亡新增

读取lesson3.json文件
获得中国直辖市/省份疫情数据

保存文件如下：

直辖市/省份：无病例最长 x 天 现存确诊
诊 累计确诊，治愈 治愈新增，死亡 死亡
新增

PANDAS的DataFrame 类似excel的结构

读取lesson3.json文件
获得中国直辖市/省份疫情数据
获得如下数据并转换为DataFrame：
[{'直辖市/省份': xxx, '城市/地区':
xxx, '现存确诊': xxx, '累计确诊': xxx,
'治愈': xxx, '死亡': xxx, '无病例天数':
xxx}, ...]

设置数据类型 写入excel与筛选

设置 确诊，治愈，死亡为整数类型
写入l4_report.xlsx



Lesson 5

获取所有省份下城市/地区的数据
将确诊，治愈，死亡为整数类型
将关键列的表头改为中文
写入l5_data.xlsx的疫情sheet下

读取l5_data.xlsx
只选择必要的列
按照省区分数据
将省份数据保存到l5_report.xlsx，以
省份命名的sheet下

读取l5_data.xlsx
获取全部省份汇总信息
将汇总信息写入到汇总sheet

读取l5_data.xlsx
获取分省汇总信息
将汇总信息加入到各省份sheet最后一行

读取l5_data.xlsx
获取累计确诊>1000 且 现存确诊<100
的缓解地区数据
将缓解地区信息写入到缓解区sheet



Lesson 6

Python 101

Powered by Saint Shine Education

篩選1列或多列
用df[]

前X条 / 后X条
head / tail

符合某个条件的筛选
确诊人数大于? >
省份为香港或澳门? isin
病例清零天数不为空? notna / != “

符合某个条件的筛选

清零天数大于10的地区

`df.loc[df[rows]>xxx, columns]`

第10-100行， 第2到5列

`df.iloc[10:100, 1:5]`

清零参数为空的改为0
df.loc[row, columns] = ‘xxx’

插入新列
死亡率与治愈率
死亡人数/累计确诊
治愈人数/累计确诊

插入一个新列

累计确诊-累计治愈 ≥ 50000

累计确诊-累计治愈 < 5000

累计确诊-累计治愈 < 1000

累计确诊-累计治愈 < 100

危险， 中等， 良好， 优秀

排序

df.sort_values(by=col,
ascending=T/F)

groupby() 根据状态统计求和与平均值

根据状态列统计记录条数

多表合一
pd.concat()

读取世界地区数据，
创建top10 sheet。将累计确诊最多的十
个国家的数据写入
创建缓解区sheet，将：
累计确诊 >20000 ，且现存确诊 $<$
5000的国家数据写入