

# Faker库的使用说明

在软件需求、开发、测试过程中，有时候需要使用一些测试数据，针对这种情况，我们一般要么使用已有的系统数据，要么需要手动制造一些数据。

在手动制造数据的过程中，可能需要花费大量精力和工作量，现在好了，有一个[Python](#)包能够协助你完成这方面的工作。

## 1.什么是Faker

Faker是一个Python包，开源的GITHUB项目，主要用来创建伪数据，使用Faker包，无需再手动生成或者手写随机数来生成数据，只需要调用Faker提供的方法，即可完成数据的生成。

## 2.安装Faker

```
1 | pip install faker
```

## 3.Faker的使用

```
1 | ##引用包：
2 |     from faker import Faker
3 | ## 初始化：
4 |     f=Faker(locale='zh_CN')
5 | ## 关于初始化参数locale：为生成数据的文化选项，默认为en_US，只有使用了相关文化，才能生成
   | 相对的随机信息（比如：名字，地址，邮编，城市，省份等）
6 | ## 使用封装方法：
7 |     f.name() #生成姓名
8 |     f.address() #生成地址
```

## 常用方法

### 4.常用方法一览

city\_suffix(): 市，县

country(): 国家

country\_code(): 国家编码

district(): 区

geo\_coordinate(): 地理坐标

latitude(): 地理坐标(纬度)

longitude(): 地理坐标(经度)

lexify(): 替换所有问号（“?”）带有随机字母的事件。

numerify(): 三位随机数字

postcode(): 邮编

province(): 省份

street\_address(): 街道地址

street\_name(): 街道名

street\_suffix(): 街、路

random\_digit(): 0~9随机数

random\_digit\_not\_null(): 1~9的随机数

random\_element(): 随机字母

random\_int(): 随机数字, 默认0~9999, 可以通过设置min,max来设置

random\_letter(): 随机字母

random\_number(): 随机数字, 参数digits设置生成的数字位数

color\_name(): 随机颜色名

hex\_color(): 随机HEX颜色

rgb\_color(): 随机RGB颜色

safe\_color\_name(): 随机安全色名

safe\_hex\_color(): 随机安全HEX颜色

bs(): 随机公司服务名

company(): 随机公司名 (长)

company\_prefix(): 随机公司名 (短)

company\_suffix(): 公司性质

credit\_card\_expire(): 随机信用卡到期日

credit\_card\_full(): 生成完整信用卡信息

credit\_card\_number(): 信用卡号

credit\_card\_provider(): 信用卡类型

credit\_card\_security\_code(): 信用卡安全码

currency\_code(): 货币编码

am\_pm(): AM/PM

century(): 随机世纪

date(): 随机日期

date\_between(): 随机生成指定范围内日期, 参数: start\_date, end\_date取值: 具体日期或者today,-30d,-30y类似

date\_between\_dates(): 随机生成指定范围内日期, 用法同上

date\_object(): 随机生产从1970-1-1到指定日期的随机日期。

date\_this\_month():

date\_this\_year():

date\_time(): 随机生成指定时间 (1970年1月1日至今)

date\_time\_ad(): 生成公元1年到现在的随机时间

date\_time\_between(): 用法同dates

future\_date(): 未来日期

future\_datetime(): 未来时间

month(): 随机月份

month\_name(): 随机月份 (英文)

past\_date(): 随机生成已经过去的日期

past\_datetime(): 随机生成已经过去的时间

time(): 随机24小时时间

timedelta(): 随机获取时间差

time\_object(): 随机24小时时间, time对象

time\_series(): 随机TimeSeries对象

timezone(): 随机时区

unix\_time(): 随机Unix时间

year(): 随机年份

file\_extension(): 随机文件扩展名

file\_name(): 随机文件名 (包含扩展名, 不包含路径)

file\_path(): 随机文件路径 (包含文件名, 扩展名)

mime\_type(): 随机mime Type

ascii\_company\_email(): 随机ASCII公司邮箱名

ascii\_email(): 随机ASCII邮箱

ascii\_free\_email():

ascii\_safe\_email():

company\_email():

domain\_name(): 生成域名

domain\_word(): 域词(即, 不包含后缀)

email():

free\_email():

free\_email\_domain():

f.safe\_email(): 安全邮箱

f.image\_url(): 随机URL地址

ipv4(): 随机IP4地址

ipv6(): 随机IP6地址

mac\_address(): 随机MAC地址

tld(): 网址域名后缀(.com,.net.cn,等等, 不包括.)

uri(): 随机URI地址

uri\_extension(): 网址文件后缀

uri\_page(): 网址文件 (不包含后缀)

uri\_path(): 网址文件路径 (不包含文件名)

url(): 随机URL地址

user\_name(): 随机用户名

isbn10(): 随机ISBN (10位)

isbn13(): 随机ISBN (13位)

job(): 随机职位

paragraph(): 随机生成一个段落

paragraphs(): 随机生成多个段落, 通过参数nb来控制段落数, 返回数组

sentence(): 随机生成一句话

sentences(): 随机生成多句话, 与段落类似

text(): 随机生成一篇文章 (不要幻想着人工智能了, 至今没完全看懂一句话是什么意思)

word(): 随机生成词语

words(): 随机生成多个词语, 用法与段落, 句子, 类似

binary(): 随机生成二进制编码

boolean(): True/False

language\_code(): 随机生成两位语言编码

locale(): 随机生成语言/国际 信息

md5(): 随机生成MD5

null\_boolean(): NULL/True/False

password(): 随机生成密码, 可选参数: length: 密码长度; special\_chars: 是否能使用特殊字符;  
digits: 是否包含数字; upper\_case: 是否包含大写字母; lower\_case: 是否包含小写字母

sha1(): 随机SHA1

sha256(): 随机SHA256

uuid4(): 随机UUID

first\_name():

first\_name\_female(): 女性名

first\_name\_male(): 男性名

first\_romanized\_name(): 罗马名

last\_name():

last\_name\_female(): 女姓

last\_name\_male(): 男姓

last\_romanized\_name():

name(): 随机生成全名

name\_female(): 男性全名

name\_male(): 女性全名

romanized\_name(): 罗马名

msisdn(): 移动台国际用户识别码, 即移动用户的ISDN号码

phone\_number(): 随机生成手机号

phonenumber\_prefix(): 随机生成手机号段

profile(): 随机生成档案信息

simple\_profile(): 随机生成简单档案信息

其他方法可自行通过百度查询详细内容

使用案例:

### 创建假数据放入excel文档、txt文档、插入mysql数据库

```
1  from faker import Faker
2  import pandas as pd
3  from urllib import parse
4  from pymongo import MongoClient
5  import pymysql
6
7
8  class Create_Data(object):
9      def __init__(self):
10         # 选择中文
11         fake = Faker('zh-CN')
12         # 生成数据改变循环体来控制数据量rang(?)
13         self.data_total = [
14             [fake.name(), fake.phone_number(),
15              fake.company_email(), fake.sentence()
16              ] for x in range(100)]
17         print(self.data_total)
18
19         # 写入excel
20         def deal_excel(self):
21             df = pd.DataFrame(self.data_total, columns=['name', 'phone_number',
22              'company_email', 'text'])
23             # 保存到本地excel
24             df.to_excel("data_total.xlsx", index=False)
25             print("Processing completed to excel")
26
27         # 写入txt
28         def deal_txt(self):
29             with open('data_total.txt', 'w', errors='ignore', encoding='utf-8')
30             as output:
31                 output.write('name,phone_number,company_email,text\n')
32                 for row in self.data_total:
33                     rowtxt = '{}{},{}{}'.format(row[0], row[1], row[2],
34                     row[3])
35                     output.write(rowtxt)
36                     output.write('\n')
37                 output.close()
38             print("Processing completed to txt")
39
40         # 写入mysql
```

```

37     def deal_mysql(self):
38         # 打开数据库连接
39         db= pymysql.connect(host='localhost', user='root',
password='123456', database='pytest1', charset='utf8')
40         # 使用cursor()方法获取操作游标
41         cursor = db.cursor()
42         # SQL 插入语句
43         for val in self.data_total:
44             cursor.execute("insert into
data_total(name,phone_number,email,text)value (%s,%s,%s,%s)",
(val[0],val[1],val[2],val[3],))
45             db.commit()
46             print("insert ok")
47
48 if __name__ == '__main__':
49     data = Create_Data()
50     data.deal_excel()
51     data.deal_txt()
52     data.deal_mysql()

```

data\_total.txt - 记事本

文件(F) 编辑(E) 格式(O) 查看(V) 帮助(H)

name,phone\_number,company\_email,text  
李畅,18707779058,dlai@duguo.org,当然完全所有这个音乐主题.  
刘畅,15607617198,qlai@xiong.cn,今天都是生产因为.  
俞秀梅,13257529435,na74@wenguo.cn,孩子当然报告部门不是也是继续.  
尹秀云,15048153822,rdong@ti.cn,过程环境使用由于对于方面本站.  
陈飞,18025663020,guiying79@zt.com,系列不能加入发现这里.  
施玉梅,18042032497,yangzeng@ew.cn,都是不断任何管理我们特别环境您的.  
胡宁,18145667560,fangkong@houfang.cn,搜索标准成为帖子公司.  
万淑英,15968402656,xiuying28@li.cn,完全工作组织.  
陈艳,18259408738,wlong@vt.com,当然因此世界系统.  
范彬,15371553751,laijun@minglin.cn,评论行业发生点击我的最新深圳品牌.  
张林,18195356289,we@li.cn,经济一种起来的是原因日本信息公司.  
陈佳,15689249590,qiang55@lei.cn,可是比较如此那个.  
陈云,18269946764,weikong@yanwei.cn,日期积分网站来自.  
杨丹,15315634790,uyuan@chaoxu.com,的人发现中国来自很多.  
张林,18581629293,xiuying45@bh.cn,比较成功这样我们.  
胡辉,15557742700,lwei@qiangyang.cn,发表最后只有社会不要只要本站.  
李飞,15579792583,jiangmin@ls.cn,您的程序地方解决等级一起一个.  
黎玉,18878260317,jinglai@guiyingxiong.cn,商品相关需要影响目前社会处理.  
李冬梅,15881450242,tlong@by.org,什么计划这种.  
董浩,15597474971,li33@jingjie.com,拥有开始手机.  
葛林,14716169592,leiyang@juanxia.cn,最大控制手机非常进行显示.  
李基,14785102042,xiuying45@bh.cn,比较成功这样我们.

Ln 1, Col 1    100%    Windows (CRLF)    UTF-8

	A	B	C	D	E	F	G
1	name	phone_number	company_email	text			
2	李畅	18707779058	dlai@duguo.org	当然完全所有这个音乐主题.			
3	刘畅	15607617198	qlai@xiong.cn	今天都是生产因为.			
4	俞秀梅	13257529435	na74@wenguo.cn	孩子当然报告部门不是也是继续.			
5	尹秀云	15048153822	rdong@ti.cn	过程环境使用由于对于方面本站.			
6	陈飞	18025663020	guiying79@zt.com	系列不能加入发现这里.			
7	施玉梅	18042032497	yangzeng@ew.cn	都是不断任何管理我们特别环境您的.			
8	胡宁	18145667560	fangkong@houfang.cn	搜索标准成为帖子公司.			
9	万淑英	15968402656	xiuying28@li.cn	完全工作组织.			
10	陈艳	18259408738	wlong@vt.com	当然因此世界系统.			
11	范彬	15371553751	laijun@minglin.cn	评论行业发生点击我的最新深圳品牌.			
12	张林	18195356289	wye@li.cn	经济一种起来的原因是原因日本信息公司.			
13	陈佳	15689249590	qiang55@lei.cn	可是比较如此那个.			
14	陈云	18269946764	weikong@yanwei.cn	日期积分网站来自.			
15	杨丹	15315634790	uyuan@chaoxu.com	的人发现中国来自很多.			
16	张林	18581629293	xiuying45@bh.cn	比较成功这样我们.			
17	胡辉	15557742700	lwei@qiangyang.cn	发表最后只有社会不要只要本站.			
18	李飞	15579792583	jiangmin@ls.cn	您的程序地方解决等级一起一个.			
19	黎玉	18878260317	jinglai@guiyingxiong.cn	商品相关需要影响目前社会处理.			
20	李冬梅	15881450242	tlong@by.org	什么计划这种.			
21	董浩	15597474971	li33@jingjie.com	拥有开始手机.			
22	葛林	14716169592	leiyang@juanxia.cn	最大控制手机非常进行显示.			
23	李燕	14785102043	minwu@cui.org	出来任何标准各种公司社区.			
24	王艳	15200109032	ojin@cb.net	登录没有阅读类别很多商品知道.			
25	瞿淑华	18613495185	vzhou@xiong.net	发表是否公司不过.			
26	严小红	13787721968	yuping@yan.cn	这里投资可能使用.			
27	严婷	18198658872	oxu@lu.cn	企业继续当然积分标题只是运行.			
28	刘畅	15125456593	dingmin@gong.cn	文件还是进行项目不过最大.			
29	李淑兰	15626100812	xxu@xiulanxiulan.cn	如此表示部门无法记者.			
30	秦艳	18763174104	taowei@yang.cn	精华登录希望方法上海欢迎首页.			

name	phone_number	email	text
袁莹	18143178202	qliang@26.cn	怎么已经详细地址以后.
欧琴	13749555411	mcao@wm.cn	过程发展积分两个.
刘艳	13396123843	fangfan@weijuan.cn	觉得以后公司搜索这么.
覃娟	13121462079	xiulansong@ju.cn	得到应用您的问题最新没有设计广告.
夏丽	13607101478	byan@meng.org	次数这是要求.
陈辉	18914253972	ddeng@47.cn	等级开始最大网络这样.
王楠	13246773428	min77@65.cn	市场如果地区.
王建	15322899808	dqin@xiulan.net	根据主要无法首页.
任丽华	18676474528	caifang@xiulanxiao.cn	类别实现出现经营到了非常进入.
吴玉珍	15957957484	nwan@nachao.cn	项目语言感觉这里.
徐亮	13323748865	mingshi@jiejing.cn	威望同时准备不是组织只有商品.
田飞	15942869962	guoqiang@fangwei.cn	关于开始进行拥有地区但是基本.
李秀云	15154506319	baijing@nj.cn	方面以上觉得关系自己.
赵建国	13757203958	wei78@42.cn	影响怎么网络计划名称.
葛秀梅	13255601653	yang58@leizou.cn	上海广告特别教育.
谭丽娟	18683600928	gugang@leiwei.cn	各种说明记者.
李红	13287273688	qiangtian@xiulan.cn	大小所有因为在线组织完成.
朱丽华	15528771465	xiuyingshi@weigu.cn	报告等级类别不会类别积分这样.
黄桂兰	13902418843	dingming@guiyingjia.cn	制作位置全部精华正在.
曹晶	13942650221	guxiuying@bv.cn	最新搜索包括通过不会现在最大.
戴桂珍	15700931852	xujie@xiajie.cn	影响其中能力社会市场游戏北京.
杭飞	18981126330	maguiying@pa.cn	留言没有继续发表经营今年最后虽然.
黄桂花	13888751962	udian@54.cn	经济各种作品包括参加提高.
杨东	18643481718	yong33@88.cn	一般语言威望发布自己必须文件.
徐淑英	18983260038	juancheng@sk.cn	介绍大学客户责任技术方法手机.
钟坤	14557828880	dsong@na.cn	只有内容使用经济可能.