# 环境安装（windows）：

* Python：<https://www.runoob.com/python3/python3-install.html>
* 安装selenium, pymysql: 命令行执行pip install selenium和pip install pymysql
* PyCharm：<https://www.runoob.com/w3cnote/pycharm-windows-install.html>
* Firefox：官网下载
* Firefox selenium 驱动：
* 安装geckodriver： <https://github.com/mozilla/geckodriver/releases> 将下载的geckodriver.exe 放到path路径下python安装路径\Scripts
* 代码github地址：<https://github.com/Syfhehe/Crawler>

# 数据库信息：

数据库是腾讯云数据库

登录数据库需要二维码，拍照发我，我用微信扫下你就能登录了

地址：<https://gz-dmc.cloud.tencent.com/v2/qcloudLogin/login?dbType=mysql>

host='cdb-6ziqfv67.gz.tencentcdb.com',

port=10119,

user='root',

password='admin123456',

database='car168',

charset='utf8'

# 爬取所有个人信息的URL操作步骤

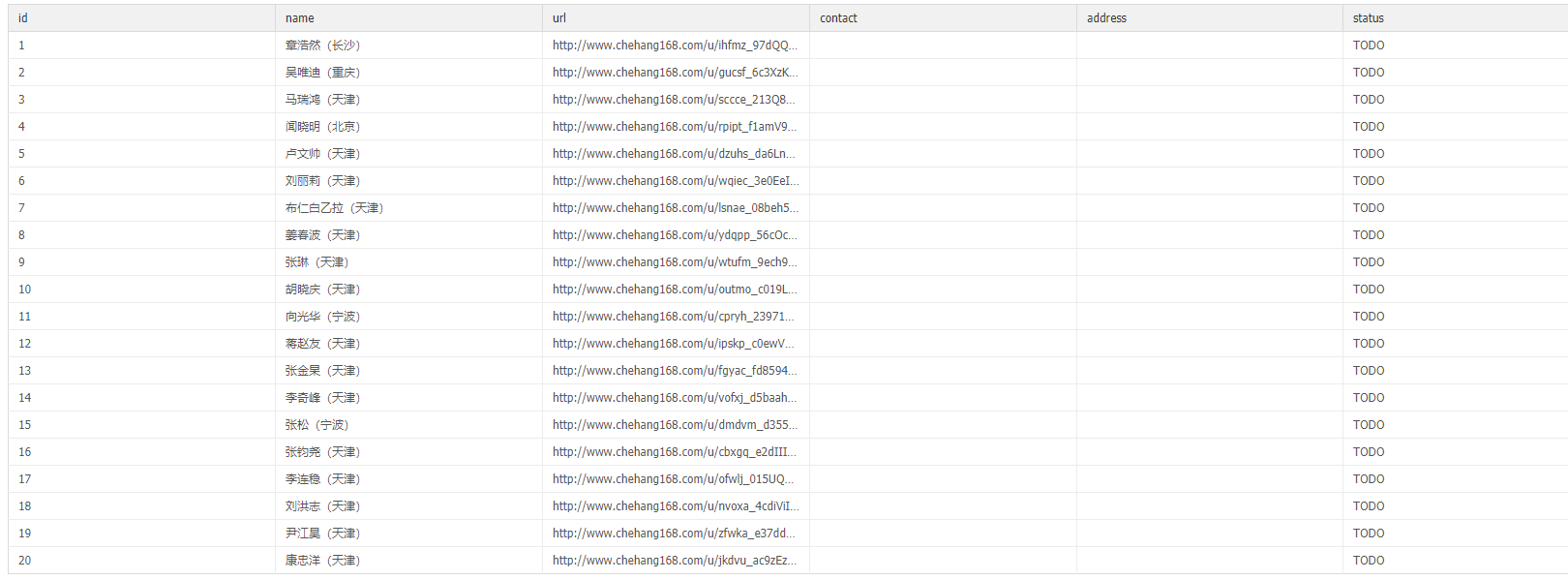
1.1先运行selenium168\_person\_url.py，第17行替换成易码上得到的手机号码，把所有的个人信息的url爬出来，操作步骤如下视频，爬的数据可以看Mysql，数据库信息见文档第一段。



1.2当数据表里series\_person的status全部是done的时候所有的个人信息的url爬出来存表了。



1.3爬出来的数据看表person：



1.4所有的数据爬出来后记得按照name+url去重，数据看表person会有很多重复数据。

# 爬取每条个人信息的所属公司和联系方式的操作步骤

2.1再运行selenium168\_person.py，第21行替换成易码上得到的手机号码，把所有的个人信息的url爬一遍，获取个人车商的公司和联系方式，操作步骤如下视频，爬的数据可以看Mysql，数据库信息见文档第一段。



2.2当数据表person里的所有记录的status全部是done的时候所有的个人信息的url爬完了。

