**建立个人数据分析系统**

## 工具系统

Python3.6+pycharm+tushare再加其他

## 数据来源

stocknotdealmarket

stockmarket

## 实现内容

## 系统基础表结构

Poolinfo

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 字段名称 | 字段类型 | 中文名称 | 字典项 | 来源 |
| product\_code\_min |  | 产品代码最小值 | 000000 |  |
| Product\_code\_max |  | 产品代码最大值 | zzzzzz |  |
| Trade\_day\_min |  | 交易日最小值 | 最小0 |  |
| Trade\_day\_max |  | 交易日最大值 | 最大99999999 |  |
| DB\_logicname |  | 数据库逻辑名 | 必填写 | 与配置必须相同 |
| db\_flag |  | 是否启用 |  |  |

datadownloadlog数据下载记录表

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 字段名称 | 字段类型 | 中文名称 | 字典项 | 来源 |
| Product\_code |  | 产品代码 |  |  |
| eventtype |  | 事件类型 | 1. 交易数据 2. Income 3. Balancesheet 4. Cashflow |  |
| dealstatus |  | 处理状态 | 0：准备开始  1：下载完成  2：导入正式表完成 |  |
| Sourcetype |  | 来源类型 | 0：各自下载  1：批量下载 |  |
| logdate |  | 日志插入日期 |  |  |
| logtime |  | 日志插入时间点 |  |  |

表结构设计

productbasicinfo

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 字段名称 | 字段类型 | 中文名称 | 字典项 | 来源 |
| product\_code |  | 产品代码 |  | stock\_basics.code |
| product\_name |  | 产品名称 |  | stock\_basics.name |
| product\_type |  | 产品类型 | 1:股票,2:基金3：指数 | Stock\_basics为股票 |
| money\_type |  | 币种类型 | 0:人民币1:美元 | Stock\_basics为人民币 |
| product\_area |  | 所在地域 |  | stock\_basics.area |
| product\_industry |  | 所在行业 |  | stock\_basics.industry |
| product\_fullname |  | 产品全称 |  | stock\_basics.name |
| market\_type |  | 市场类型 | 1.主板2.中小板3.创业板4.新三板5.科创板 | 股票命名规则参考[1] |
| exchange\_code |  | 交易所代码 |  | 交易所代码[2] |
| ipo\_status |  | 上市状态 | N:上市 D:退市 | 默认为N，退市手动维护 |
| listed\_date |  | 上市日期 |  | stock\_basics.timetomarket |
| delisted\_date |  | 退市日期 |  |  |

Openday 交易日

|  |  |  |  |
| --- | --- | --- | --- |
| 字段名称 | 字段类型 | 中文名称 | 字典项 |
| Exchange\_code |  | 交易所 SSE上交所 SZSE深交所 |  |
| Trade\_date |  | 日历日期 |  |
| trade\_flag |  | 是否交易日 | 0否,1是 |
| Pretrade\_day |  | 上个交易日 |  |

Exchangeinfo 交易所信息

|  |  |  |  |
| --- | --- | --- | --- |
| 字段名称 | 字段类型 | 中文名称 | 字典项 |
| Exchange\_code |  | 交易所 SSE上交所 SZSE深交所 |  |
| Exchange\_name |  | 交易所名称 |  |
| Country\_code |  | 国家 |  |

pddelistinfo产品停复牌信息

|  |  |  |  |
| --- | --- | --- | --- |
| 字段名称 | 字段类型 | 中文名称 | 字典项 |
| product\_code |  | 产品代码 |  |
| delistbegindate |  | 停牌起始日 |  |
| delistenddate |  | 停牌结束日 |  |
| delistreason |  | 停牌事由 |  |

Productdailybasicinfo 每日交易数据

|  |  |  |  |
| --- | --- | --- | --- |
| 字段名称 | 字段类型 | 中文名称 | 字典项 |
| Product\_code |  | 产品代码 |  |
| Trade\_date |  | 交易日期 |  |
| Close\_price |  | 收盘价 |  |
| turnover\_rate |  | 换手率 |  |
| turnover\_rate\_f |  | 换手率（自由流通股） |  |
| Product\_volume |  | 交易量 |  |
| volume\_ratio |  | 量比 |  |
| Pe\_static |  | 市盈率 | 总市值/净利润 |
| pe\_ttm |  | 市盈率TTM |  |
| Stock\_Pb |  | 市净率 | 总市值/净资产 |
| Ps\_static |  | 静态市销率 |  |
| ps\_ttm |  | 市销率TTM |  |
| total\_share |  | 总股本(万股) |  |
| float\_share |  | 流通股本(万股) |  |
| free\_share |  | 自由流通股本(万股) |  |
| total\_mv |  | 总市值(万元) |  |
| cicr\_mv |  | 流通市值(万元) |  |

Company\_balance\_sheet资产负债表

|  |  |  |  |
| --- | --- | --- | --- |
| 字段名称 | 字段类型 | 中文名称 | 字典项 |
| product\_code |  | 交易所 SSE上交所 SZSE深交所 |  |
| Announce\_date |  | 公告日期 |  |
| F\_announce\_date |  | 最终公告日期 |  |
| Report\_date |  | 报告日期 | 20180331,  20180630,  20180930,  20181231 |
| Report\_type |  | 报告类型 | 1 合并报表 上市公司最新报表（默认）  2 单季合并 单一季度的合并报表  3 调整单季合并表 调整后的单季合并报表（如果有）  4 调整合并报表 本年度公布上年同期的财务报表数据，报告期为上年度  5 调整前合并报表 数据发生变更，将原数据进行保留，即调整前的原数据  6 母公司报表 该公司母公司的财务报表数据  7 母公司单季表 母公司的单季度表  8 母公司调整单季表 母公司调整后的单季表  9 母公司调整表 该公司母公司的本年度公布上年同期的财务报表数据  10 母公司调整前报表 母公司调整之前的原始财务报表数据  11 调整前合并报表 调整之前合并报表原数据  12 母公司调整前报表 母公司报表发生变更前保留的原数据 |
| Company\_type |  | 公司类型 | 1一般工商业  2银行  3保险  4证券 |
| total\_share |  | 总股本(万股) |  |
| Cap\_rese |  | 资本公积金 |  |
| undistr\_profit |  | 未分配利润 |  |
| Surplus\_rese |  | 盈余公积金 |  |
| Special\_rese |  | 专项储备 |  |
| Money\_cap |  | 货币资金 |  |
| Trade\_asset |  | 交易性金融资产 |  |
| Notes\_willreceiv |  | 应收票据 |  |
| Accounts\_willreceiv |  | 应收账款 |  |
| Other\_willreceiv |  | 其他应收款 |  |
| Prepayment |  | 预付款项 |  |
| Dividend\_willreceiv |  | 应收股利 |  |
| Interest\_willreceiv |  | 应收利息 |  |
| Inventories |  | 存货 |  |
| Amor\_exp |  | 长期待摊费用 |  |
| Notcash\_within\_1y |  | 一年内到期的非流动资产 |  |
| Sett\_rsrv |  | 结算备付金 |  |
| Loanto\_oth\_bank\_fi |  | 拆出资金 |  |
| premium\_receiv |  | 应收保费 |  |
| reinsur\_receiv |  | 应收分保账款 |  |
| reinsur\_res\_receiv |  | 应收分保合同准备金 |  |
| pur\_resale\_fa |  | 买入返售金融资产 |  |
| oth\_cur\_assets |  | 其他流动资产 |  |
| total\_cur\_assets |  | 流动资产合计 |  |
| fa\_avail\_for\_sale |  | 可供出售金融资产 |  |
| htm\_invest |  | 持有至到期投资 |  |
| long\_stock\_invest |  | 长期股权投资 |  |
| invest\_real\_estate |  | 投资性房地产 |  |
| time\_deposits |  | 定期存款 |  |
| oth\_assets |  | 其他资产 |  |
| Long\_term\_rec |  | 长期应收款 |  |
| fix\_assets |  | 固定资产 |  |
| Constr\_in\_process |  | 在建工程 |  |
| const\_materials |  | 工程物资 |  |
| fixed\_assets\_disp |  | 固定资产清理 |  |
| produc\_bio\_assets |  | 生产性生物资产 |  |
| oil\_and\_gas\_assets |  | 油气资产 |  |
| intan\_assets |  | 无形资产 |  |
| Study\_spending |  | 研发支出 |  |
| Goodwill |  | 商誉 |  |
| Long\_term\_amor\_exp |  | 长期待摊费用 |  |
| defer\_tax\_assets |  | 递延所得税资产 |  |
| decr\_in\_disbur |  | 发放贷款及垫款 |  |
| other\_not\_cash |  | 其他非流动资产 |  |
| total\_not\_cash |  | 非流动资产合计 |  |
| cash\_reser\_cb |  | 现金及存放中央银行款项 |  |
| depos\_in\_oth\_bfi |  | 存放同业和其它金融机构款项 |  |
| prec\_metals |  | 贵金属 |  |
| deriv\_assets |  | 衍生金融资产 |  |
| rr\_reins\_une\_prem |  | 应收分保未到期责任准备金 |  |
| rr\_reins\_outstd\_cla |  | 应收分保未决赔款准备金 |  |
| rr\_reins\_lins\_liab |  | 应收分保寿险责任准备金 |  |
| rr\_reins\_lthins\_liab |  | 应收分保长期健康险责任准备金 |  |
| refund\_depos |  | 存出保证金 |  |
| ph\_pledge\_loans |  | 保户质押贷款 |  |
| refund\_cap\_depos |  | 存出资本保证金 |  |
| indep\_acct\_assets |  | 独立账户资产 |  |
| client\_depos |  | 其中：客户资金存款 |  |
| client\_prov |  | 其中：客户备付金 |  |
| transac\_seat\_fee |  | 其中:交易席位费 |  |
| invest\_as\_receiv |  | 应收款项类投资 |  |
| total\_assets |  | 资产总计 |  |
| long\_borrow |  | 长期借款 |  |
| Short\_borrow |  | 短期借款 |  |
| Central\_bank\_borrow |  | 向中央银行借款 |  |
| depos\_ib\_deposits |  | 吸收存款及同业存放 |  |
| loan\_oth\_bank |  | 拆入资金 |  |
| trading\_finance\_load |  | 交易性金融负债 |  |
| notes\_payable |  | 应付票据 |  |
| acct\_payable |  | 应付账款 |  |
| adv\_receipts |  | 预收款项 |  |
| sold\_for\_repur\_fa |  | 卖出回购金融资产款 |  |
| comm\_payable |  | 应付手续费及佣金 |  |
| payroll\_payable |  | 应付职工薪酬 |  |
| taxes\_payable |  | 应交税费 |  |
| interst\_payable |  | 应付利息 |  |
| dividend\_payable |  | 应付股利 |  |
| oth\_payable |  | 其他应付款 |  |
| acc\_exp |  | 预提费用 |  |
| deferred\_inc |  | 递延收益 |  |
| st\_bonds\_payable |  | 应付短期债券 |  |
| payable\_to\_reinsurer |  | 应付分保账款 |  |
| rsrv\_insur\_cont |  | 保险合同准备金 |  |
| acting\_trading\_sec |  | 代理买卖证券款 |  |
| acting\_uw\_sec |  | 代理承销证券款 |  |
| non\_cur\_liab\_due\_1y |  | 一年内到期的非流动负债 |  |
| oth\_cur\_liab |  | 其他流动负债 |  |
| total\_cur\_liab |  | 流动负债合计 |  |
| bond\_payable |  | 应付债券 |  |
| Long\_money\_payable |  | 长期应付款 |  |
| specific\_payables |  | 专项应付款 |  |
| estimated\_liab |  | 预计负债 |  |
| defer\_tax\_liab |  | 递延所得税负债 |  |
| defer\_inc\_non\_cur\_liab |  | 递延收益-非流动负债 |  |
| other\_not\_cur\_load |  | 其他非流动负债 |  |
| total\_not\_cur\_load |  | 非流动负债合计 |  |
| depos\_oth\_bfi |  | 同业和其它金融机构存放款项 |  |
| deriv\_liab |  | 衍生金融负债 |  |
| Depos |  | 吸收存款 |  |
| agency\_bus\_liab |  | 代理业务负债 |  |
| oth\_liab |  | 其他负债 |  |
| prem\_receiv\_adva |  | 预收保费 |  |
| depos\_received |  | 存入保证金 |  |
| ph\_invest |  | 保户储金及投资款 |  |
| reser\_une\_prem |  | 未到期责任准备金 |  |
| reser\_outstd\_claims |  | 未决赔款准备金 |  |
| reser\_lins\_liab |  | 寿险责任准备金 |  |
| reser\_lthins\_liab |  | 长期健康险责任准备金 |  |
| indept\_acc\_liab |  | 独立账户负债 |  |
| pledge\_borr |  | 其中:质押借款 |  |
| indem\_payable |  | 应付赔付款 |  |
| policy\_div\_payable |  | 应付保单红利 |  |
| total\_liab |  | 负债合计 |  |
| treasury\_share |  | 减:库存股 |  |
| ordin\_risk\_reser |  | 一般风险准备 |  |
| forex\_differ |  | 外币报表折算差额 |  |
| invest\_loss\_unconf |  | 未确认的投资损失 |  |
| minority\_int |  | 少数股东权益 |  |
| total\_hldr\_eqy\_exc\_min\_int |  | 股东权益合计(不含少数股东权益) |  |
| total\_hldr\_eqy\_inc\_min\_int |  | 股东权益合计(含少数股东权益) |  |
| total\_liab\_hldr\_eqy |  | 负债及股东权益总计 |  |
| lt\_payroll\_payable |  | 长期应付职工薪酬 |  |
| oth\_comp\_income |  | 其他综合收益 |  |
| oth\_eqt\_tools |  | 其他权益工具 |  |
| oth\_eqt\_tools\_p\_shr |  | 其他权益工具(优先股) |  |
| lending\_funds |  | 融出资金 |  |
| acc\_receivable |  | 应收款项 |  |
| st\_fin\_payable |  | 应付短期融资款 |  |
| Payables |  | 应付款项 |  |
| hfs\_assets |  | 持有待售的资产 |  |
| hfs\_sales |  | 持有待售的负债 |  |
| update\_flag |  | 更新标识 | 0：首次获取 1:数据更正 |

数据来源：参考：<https://zhuanlan.zhihu.com/p/32800066> http://quotes.money.163.com/service/zcfzb\_600675.html?type=year

company\_cashflow 现金流量表

|  |  |  |  |
| --- | --- | --- | --- |
| 字段名称 | 字段类型 | 中文名称 | 字典项 |
| product\_code |  | 产品代码 |  |
| Announce\_date |  | 公告日期 |  |
| F\_announce\_date |  | 最终公告日期 |  |
| Report\_date |  | 报告日期 | 201803,  201806,  201809,  201812 |
| Company\_type |  | 公司类型 | 1一般工商业  2银行  3保险  4证券 |
| Net\_profit |  | 净利润 |  |
| Finan\_exp |  | 财务费用 |  |
| C\_fr\_sale\_sg |  | 销售商品、提供劳务收到的现金 |  |
| Recp\_tax\_rends |  | 收到的税费返还 |  |
| N\_depos\_incr\_fi |  | 客户存款和同业存放款项净增加额 |  |
| n\_incr\_loans\_cb |  | 向中央银行借款净增加额 |  |
| n\_inc\_borr\_oth\_fi |  | 向其他金融机构拆入资金净增加额 |  |
| prem\_fr\_orig\_contr |  | 收到原保险合同保费取得的现金 |  |
| n\_incr\_insured\_dep |  | 保户储金净增加额 |  |
| n\_reinsur\_prem |  | 收到再保业务现金净额 |  |
| n\_incr\_disp\_tfa |  | 处置交易性金融资产净增加额 |  |
| ifc\_cash\_incr |  | 收取利息和手续费净增加额 |  |
| n\_incr\_disp\_faas |  | 处置可供出售金融资产净增加额 |  |
| n\_incr\_loans\_oth\_bank |  | 拆入资金净增加额 |  |
| n\_cap\_incr\_repur |  | 回购业务资金净增加额 |  |
| c\_fr\_oth\_operate\_a |  | 收到其他与经营活动有关的现金 |  |
| c\_inf\_fr\_operate\_a |  | 经营活动现金流入小计 |  |
| c\_paid\_goods\_s |  | 购买商品、接受劳务支付的现金 |  |
| c\_paid\_to\_for\_empl |  | 支付给职工以及为职工支付的现金 |  |
| c\_paid\_for\_taxes |  | 支付的各项税费 |  |
| n\_incr\_clt\_loan\_adv |  | 客户贷款及垫款净增加额 |  |
| n\_incr\_dep\_cbob |  | 存放央行和同业款项净增加额 |  |
| c\_pay\_claims\_orig\_inco |  | 支付原保险合同赔付款项的现金 |  |
| pay\_handling\_chrg |  | 支付手续费的现金 |  |
| pay\_comm\_insur\_plcy |  | 支付保单红利的现金 |  |
| oth\_cash\_pay\_oper\_act |  | 支付其他与经营活动有关的现金 |  |
| st\_cash\_out\_act |  | 经营活动现金流出小计 |  |
| n\_cashflow\_act |  | 经营活动产生的现金流量净额 |  |
| oth\_recp\_ral\_inv\_act |  | 收到其他与投资活动有关的现金 |  |
| c\_disp\_withdrwl\_invest |  | 收回投资收到的现金 |  |
| c\_recp\_return\_invest |  | 取得投资收益收到的现金 |  |
| n\_recp\_disp\_fiolta |  | 处置固定资产、无形资产和其他长期资产收回的现金净额 |  |
| n\_recp\_disp\_sobu |  | 处置子公司及其他营业单位收到的现金净额 |  |
| stot\_inflows\_inv\_act |  | 投资活动现金流入小计 |  |
| c\_pay\_acq\_const\_fiolta |  | 购建固定资产、无形资产和其他长期资产支付的现金 |  |
| c\_paid\_invest |  | 投资支付的现金 |  |
| n\_disp\_subs\_oth\_biz |  | 取得子公司及其他营业单位支付的现金净额 |  |
| oth\_pay\_ral\_inv\_act |  | 支付其他与投资活动有关的现金 |  |
| n\_incr\_pledge\_loan |  | 质押贷款净增加额 |  |
| stot\_out\_inv\_act |  | 投资活动现金流出小计 |  |
| n\_cashflow\_inv\_act |  | 投资活动产生的现金流量净额 |  |
| c\_recp\_borrow |  | 取得借款收到的现金 |  |
| proc\_issue\_bonds |  | 发行债券收到的现金 |  |
| oth\_cash\_recp\_ral\_fnc\_act |  | 收到其他与筹资活动有关的现金 |  |
| stot\_cash\_in\_fnc\_act |  | 筹资活动现金流入小计 |  |
| free\_cashflow |  | 企业自由现金流量 |  |
| c\_prepay\_amt\_borr |  | 偿还债务支付的现金 |  |
| c\_pay\_dist\_dpcp\_int\_exp |  | 分配股利、利润或偿付利息支付的现金 |  |
| incl\_dvd\_profit\_paid\_sc\_ms |  | 其中:子公司支付给少数股东的股利、利润 |  |
| oth\_cashpay\_ral\_fnc\_act |  | 支付其他与筹资活动有关的现金 |  |
| stot\_cashout\_fnc\_act |  | 筹资活动现金流出小计 |  |
| n\_cash\_flows\_fnc\_act |  | 筹资活动产生的现金流量净额 |  |
| eff\_fx\_flu\_cash |  | 汇率变动对现金的影响 |  |
| n\_incr\_cash\_cash\_equ |  | 现金及现金等价物净增加额 |  |
| c\_cash\_equ\_beg\_period |  | 期初现金及现金等价物余额 |  |
| c\_cash\_equ\_end\_period |  | 期末现金及现金等价物余额 |  |
| c\_recp\_cap\_contrib |  | 吸收投资收到的现金 |  |
| incl\_cash\_rec\_saims |  | 其中:子公司吸收少数股东投资收到的现金 |  |
| uncon\_invest\_loss |  | 未确认投资损失 |  |
| prov\_depr\_assets |  | 加:资产减值准备 |  |
| depr\_fa\_coga\_dpba |  | 固定资产折旧、油气资产折耗、生产性生物资产折旧 |  |
| amort\_intang\_assets |  | 无形资产摊销 |  |
| lt\_amort\_deferred\_exp |  | 长期待摊费用摊销 |  |
| decr\_deferred\_exp |  | 待摊费用减少 |  |
| incr\_acc\_exp |  | 预提费用增加 |  |
| loss\_disp\_fiolta |  | 处置固定、无形资产和其他长期资产的损失 |  |
| loss\_scr\_fa |  | 固定资产报废损失 |  |
| loss\_fv\_chg |  | 公允价值变动损失 |  |
| invest\_loss |  | 投资损失 |  |
| decr\_def\_inc\_tax\_assets |  | 递延所得税资产减少 |  |
| incr\_def\_inc\_tax\_liab |  | 递延所得税负债增加 |  |
| decr\_inventories |  | 存货的减少 |  |
| decr\_oper\_payable |  | 经营性应收项目的减少 |  |
| incr\_oper\_payable |  | 经营性应付项目的增加 |  |
| Others\_payable |  | 其他 |  |
| conv\_debt\_into\_cap |  | 债务转为资本 |  |
| conv\_copbonds\_due\_within\_1y |  | 一年内到期的可转换公司债券 |  |
| fa\_fnc\_leases |  | 融资租入固定资产 |  |
| end\_bal\_cash |  | 现金的期末余额 |  |
| beg\_bal\_cash |  | 减:现金的期初余额 |  |
| end\_bal\_cash\_equ |  | 加:现金等价物的期末余额 |  |
| beg\_bal\_cash\_equ |  | 减:现金等价物的期初余额 |  |
| im\_n\_incr\_cash\_equ |  | 现金及现金等价物净增加额(间接法) |  |
| update\_flag |  | 更新标识 |  |

Company\_income 利润表

|  |  |  |  |
| --- | --- | --- | --- |
| 字段名称 | 字段类型 | 中文名称 | 字典项 |
| product\_code |  | 产品代码 |  |
|  |  |  |  |
| Announce\_date |  | 公告日期 |  |
| F\_announce\_date |  | 最终公告日期 |  |
| Report\_date |  | 报告日期 | 20180331,  20180630,  20180930,  20181231 |
| Company\_type |  | 公司类型 | 1一般工商业  2银行  3保险  4证券 |
| basic\_eps |  | 基本每股收益 |  |
| diluted\_eps |  | 稀释每股收益 |  |
| total\_revenue |  | 营业总收入 |  |
| Revenue |  | 营业收入 |  |
| int\_income |  | 利息收入 |  |
| prem\_earned |  | 已赚保费 |  |
| comm\_income |  | 手续费及佣金收入 |  |
| n\_commis\_income |  | 手续费及佣金净收入 |  |
| n\_oth\_income |  | 其他经营净收益 |  |
| n\_oth\_b\_income |  | 加:其他业务净收益 |  |
| prem\_income |  | 保险业务收入 |  |
| out\_prem |  | 减:分出保费 |  |
| une\_prem\_reser |  | 提取未到期责任准备金 |  |
| reins\_income |  | 其中:分保费收入 |  |
| n\_sec\_tb\_income |  | 代理买卖证券业务净收入 |  |
| n\_sec\_uw\_income |  | 证券承销业务净收入 |  |
| n\_asset\_mg\_income |  | 受托客户资产管理业务净收入 |  |
| oth\_b\_income |  | 其他业务收入 |  |
| fv\_value\_chg\_gain |  | 加:公允价值变动净收益 |  |
| invest\_income |  | 加:投资净收益 |  |
| ass\_invest\_income |  | 其中:对联营企业和合营企业的投资收益 |  |
| forex\_gain |  | 加:汇兑净收益 |  |
| total\_cogs |  | 营业总成本 |  |
| oper\_cost |  | 减:营业成本 |  |
| int\_exp |  | 减:利息支出 |  |
| comm\_exp |  | 减:手续费及佣金支出 |  |
| biz\_tax\_surchg |  | 减:营业税金及附加 |  |
| sell\_exp |  | 减:销售费用 |  |
| admin\_exp |  | 减:管理费用 |  |
| fin\_exp |  | 减:财务费用 |  |
| assets\_impair\_loss |  | 减:资产减值损失 |  |
| prem\_refund |  | 退保金 |  |
| compens\_payout |  | 赔付总支出 |  |
| reser\_insur\_liab |  | 提取保险责任准备金 |  |
| div\_payt |  | 保户红利支出 |  |
| reins\_exp |  | 分保费用 |  |
| oper\_exp |  | 营业支出 |  |
| compens\_payout\_refu |  | 减:摊回赔付支出 |  |
| insur\_reser\_refu |  | 减:摊回保险责任准备金 |  |
| reins\_cost\_refund |  | 减:摊回分保费用 |  |
| other\_bus\_cost |  | 其他业务成本 |  |
| operate\_profit |  | 营业利润 |  |
| non\_oper\_income |  | 加:营业外收入 |  |
| non\_oper\_exp |  | 减:营业外支出 |  |
| nca\_disploss |  | 其中:减:非流动资产处置净损失 |  |
| total\_profit |  | 利润总额 |  |
| income\_tax |  | 所得税费用 |  |
| n\_income |  | 净利润(含少数股东损益) |  |
| n\_income\_attr\_p |  | 净利润(不含少数股东损益) |  |
| minority\_gain |  | 少数股东损益 |  |
| oth\_compr\_income |  | 其他综合收益 |  |
| t\_compr\_income |  | 综合收益总额 |  |
| compr\_inc\_attr\_p |  | 归属于母公司(或股东)的综合收益总额 |  |
| compr\_inc\_attr\_m\_s |  | 归属于少数股东的综合收益总额 |  |
| Ebit |  | 息税前利润 |  |
| Ebitda |  | 息税折旧摊销前利润 |  |
| insurance\_exp |  | 保险业务支出 |  |
| undist\_profit |  | 年初未分配利润 |  |
| distable\_profit |  | 可分配利润 |  |
| update\_flag |  | 更新标识 |  |

profitschema 分红送股数据

|  |  |  |  |
| --- | --- | --- | --- |
| Product\_code |  | 产品代码 |  |
| announce\_date |  | 公告年度 | 从上市开始记录每个半年度的日期。20181231,20180630 |
| div\_proc |  | 分红处理进度 | 0.不分配1.预案 2.股东大会通过 3.实施 |
| stk\_div |  | 每股送转 |  |
| stk\_bo\_rate |  | 每股送股比例 |  |
| stk\_co\_rate |  | 每股转增比例 |  |
| cash\_div |  | 每股分红(税后) |  |
| cash\_div\_inc\_tax |  | 每股分红(税前) |  |
| Register\_date |  | 股权登记日 |  |
| Dividend\_date |  | 除权除息日 |  |
| Pay\_interest\_date |  | 派息日 |  |
| div\_share\_listdate |  | 红股上市日 |  |
| imp\_ann\_date |  | 实施公告日 |  |
| base\_date |  | 基准日 |  |
| base\_share |  | 基准股本 |  |

暂时先实现以上的内容。下面的暂时不实现。

Namechange股票曾用名

new\_share IPO新股列表

60min行情

日线行情,同时记录PE

Producttradedata每日不复权行情

|  |  |  |  |
| --- | --- | --- | --- |
| 字段名称 | 字段类型 | 中文名称 | 字典项 |
| product\_code |  | 产品代码 |  |
| trade\_date |  | 交易日期 |  |
| open\_price |  | 开盘价格 |  |
| high\_price |  | 最高价格 |  |
| close\_price |  | 收盘价格 |  |
| low\_price |  | 最低价格 |  |
| product\_volume |  | 交易量 |  |
| product\_amount |  | 交易金额 |  |

30min交易数据prod30tradedata

|  |  |  |  |
| --- | --- | --- | --- |
| 字段名称 | 字段类型 | 中文名称 | 字典项 |
| product\_code |  | 产品代码 |  |
| trade\_date |  | 交易日期 |  |
| Trade\_time |  | 交易时间 |  |
| open\_price |  | 开盘价格 |  |
| high\_price |  | 最高价格 |  |
| close\_price |  | 收盘价格 |  |
| low\_price |  | 最低价格 |  |
| product\_volume |  | 交易量 |  |
| product\_amount |  | 交易金额 |  |
| Symbol\_code |  | 自己的代码 | 支持指数等，现在的product\_code不能支持指数 |

通过以下代码可以获取最近的30min交易数据，获取不到历史

import os

import time

df = ts.get\_stock\_basics()

codelist = list(df.index)

filename = 'E:\\pydevproj\\stockproj\\stockmarket\\finance\\test'

for oneindex in range(len(codelist)):

print(codelist[oneindex]+ " begin")

df = ts.get\_hist\_data(codelist[oneindex], ktype='30')

if df is None:

continue

realfilename = filename + "\\"+codelist[oneindex]+".csv"

if os.path.exists(realfilename):

df.to\_csv(realfilename, mode='a', header=None)

else:

df.to\_csv(realfilename)

print(codelist[oneindex]+ " end")

time.sleep(1)

每日指标daily\_basic 市盈率信息等。

外部字段和内部字段对应

Tfilefielddict

|  |  |  |  |
| --- | --- | --- | --- |
| 字段名称 | 字段类型 | 中文名称 | 字典项 |
| filechiname |  | 中文名 |  |
| l\_fileid |  | 文件字段序号 |  |

Tstfielddict

|  |  |  |  |
| --- | --- | --- | --- |
| 字段名称 | 字段类型 | 中文名称 | 字典项 |
| engname |  | 英文字段 |  |
| l\_stid |  | st字段序号 |  |

Tfilestfield

|  |  |  |  |
| --- | --- | --- | --- |
| 字段名称 | 字段类型 | 中文名称 | 字典项 |
| l\_fileid |  | 文件字段序号 |  |
| l\_stid |  | st字段序号 |  |
| filetype |  | 文件类型 |  |

中文名和数据库内的字段进行对应

get\_hist\_data(<http://tushare.org/trading.html>)

获取历史行情数据

PE获取 收盘价\*

## 基础数据

见Database下的内容。

### 缠论相关表

缠论中的表需要支持分库分表，为了后续更小级别的数据:

1. 分表的原则是每个库对于这类交易数据需要分32个子表，比如entdaytradedata，每个分库都拆分成entdaytradedata01,entdaytradedata02,entdaytradedata03,…,entdaytradedata32.

Entdaytradedata，通过产品代码取32余数

|  |  |  |  |
| --- | --- | --- | --- |
| 字段名称 | 字段类型 | 中文名称 | 字典项 |
| product\_code |  | 产品代码 |  |
| trade\_date |  | 交易日期 | K线合并后只记录最后一天。 |
| trade\_time |  | 交易时间 | K线合并后只记录最后一段 |
| Open\_price |  | 开盘价格 |  |
| High\_price |  | 最高价格 |  |
| Close\_price |  | 收盘价格 |  |
| Low\_price |  | 最低价格 |  |
| Merge\_flag |  | 合并标志 | 如果存在包含关系，由于做了合并，则该标志为1，否则为0 |
| updown\_flag |  | 顶底分型标志 | 笔的顶A（Apical typing），笔的底B（base typing）伪顶分型U，伪底分型D，四者都不是N |

### 买卖点代码实现

#### 2.2.1 一买代码实现思路

#### 2.2.2 二买代码实现思路

#### 2.2.3 三买代码实现思路

#### 2.2.4 在中枢中代码实现思路

#### 2.2.5 一卖代码实现思路

#### 2.2.6 二卖代码实现思路

## 微信与数据互通

Itchat模块使用，对于个人账号存在风险，就是该模块支持的是网页版，现在已经被腾讯封锁，不能使用。 现在新的方案是在手动完成登录后，需要通过模拟按键的功能来完成文字或者文件的传输(<https://blog.csdn.net/idecl/article/details/80505050>)。

1. 发送文件
2. 识别新消息

这块存在问题了。微信必须走公众号的流程。

### 存在的问题

1. swig.exe不存在(<http://www.swig.org/index.php>)，沟通C与python的接口，需要存在。
2. 将swig.exe复制到python安装文件夹(对于ex C：/ python27)
3. 打开swigwin-3.0.12 / Lib文件夹并将所有\* .swg文件复制到C：/ python27 / lib或等效的python路径
4. 打开swigwin-3.0.12 / lib / python并将所有文件复制到(C：/ python27 / lib)
5. 还要确保安装了Microsoft Visual C Compiler for Python

## 使用技术

### pdf读取

通过<https://www.cnblogs.com/gl1573/p/10064438.html>的分析，最后选取<https://github.com/jsvine/pdfplumber>来作为pdf的读取支持。同时pdfminer需要一定的版本，[pdfminer.six==20181108](https://github.com/jsvine/pdfplumber/commit/027d999ae66f5557a37346fd676f4c3d76139993" \o "Upgrade to pdfminer.six==20181108  https://github.com/pdfminer/pdfminer.six/releases/tag/20181108)。通过以下命令的顺序进行安装

Pip install pdfplumber

Pip install pdfminer.six==20181108

Pip install pdfminer

## 数据获取

### 上市产品数据获取

### ETF股票基金获取

fund\_basic

### 交易数据获取

交易数据调用tushare的pro接口按以下步骤进行获取。

1. 从表productbasicinfo获取所有要获取交易数据的产品和对应的需要获取交易数据的起始日期(startdate)和交易截止日期(enddate)。
2. 循环产品数据，插入表数据下载日志表(datadownloadlog),事件类型为交易数据下载(1),处理状态(0)的数据
3. 相同的startdate的股票数小于100只，则该startdate对应的产品获取自己的交易数据。
4. 更新表数据下载日志表(datadownloadlog),事件类型为交易数据下载(1),处理状态(1)的数据
5. 插入到表tradeData股票代码的数据插入到正式表中。
6. 更新表数据下载日志表(datadownloadlog),事件类型为交易数据下载(1),处理状态(2)的数据
7. 相同的startdate的股票数大于等于100只，则一起获取从startdate起始的数据。
8. 产品循环准备从临时表插入到正式表前，插入表数据下载日志表(datadownloadlog),事件类型为交易数据下载(1),处理状态(1)的数据
9. 数据插入完成后更新表数据下载日志表(datadownloadlog),事件类型为交易数据下载(1),处理状态(2)的数据

### 公司数据获取

## 数据整理

数据都是用更新后的数据，纠错之后的。像现在的通策医疗2015年的数据做了调整，但是库里面的还是2015年年报里面的数据。

每次发送公司代码或者公司名称，微信显示公司的以下资料。生产的html文件或者其他图片格式可以通过pyecharts-snapshot来转化成pdf。

### 数据整理

#### 6.1.1 重复数据

对应的语句：

select "company\_income" tablename,a.\*

from

(

select count(1) cntnum,a.`product\_code`,a.`report\_date`

from stockmarket.company\_income a

GROUP BY a.`product\_code`,a.`report\_date`

union all

select count(1) cntnum,a.`product\_code`,a.`report\_date`

from stocktrade1.company\_income a

GROUP BY a.`product\_code`,a.`report\_date`

union all

select count(1) cntnum,a.`product\_code`,a.`report\_date`

from stocktrade2.company\_income a

GROUP BY a.`product\_code`,a.`report\_date`

) a

where a.cntnum > 1;

问题：重复数据做了处理了，但是是数据库级别的，代码级别还未处理。

代码级别进行了处理。

#### 6.1.2 数据纠错

只能跟pdf来做比较了。关键字“差错更正”

只对年报的数据进行核准，其他季度的数据不做核对。

配置表tfilefielddict，tstfielddict，tfilestfield三个表，使得文件中的中文能匹配数据库中的字段。

### 消息推送

1. 还是先推送自己想要的数据先。
2. 可以定制自己想要推送的消息

### 基本数据(图的形式展示)

1. 营业收入, 营业利润，值和同比增长百分比。近5年(最多)。
2. 营业收入和现金流增长率
3. 营业收入，应收账款，应付账款，如果应付在营业收入中占比小，则提示。
4. 营业收入与存货，如果当年存货占比突增，则需要显示解释（爬虫的手段）。
5. 股东账户数变化和大股东股票数变化

### 友商数据对比

1. 哪些是友商，国内外，数据如何。这种友商的判断依据。共同竞标，他人研报中多次一起提及的。
2. 完成以上基础数据对比。

### 每日数据

每日PE计算

1. 计算每日动态PE

### 股票选取

可采用的筛选条件

1. 高负债，高现金剔除。
2. 股东信息

## 后台维护界面

管理人可以通过django完成web框架搭建。Python 3.6.0和django3.0