

# 课程名称: Python语言与数据分析

# 课程报告

项目名	名称 _	沪深股票数据分析	
班	级 _	金融203班	
学	号_	0204863	
姓	名 _	谢鑫平	
任课	教师 _	肖 泉	
开课等	学期:	20 至 21 学年 第 2 学期	
完成田	付间.	2021 年 7 月 3 日	

# 《沪深股票数据》分析报告

# 目 录

1	概述	Error! Bookmark not defined.
2	股票数据描述	Error! Bookmark not defined.
	2.1 数据来源:	Error! Bookmark not defined.
3	股票数据分析内容	Error! Bookmark not defined.
	3.1 沪深股票整体市场情况	2
	3.2 沪深特色股票举例分析	3
	3.3 沪深股票的基本收益	3
4	股票数据分析图表	4
	4.1 沪深股票行情数据	4
	4.2 沪深上市公司基本面数据	
	4.4 沪深股票波动 k 线图	
5	股票数据分析结果	Error! Bookmark not defined.
	5.2 沪深公司销售股票结果分析	
	5.3 股票规模相关性分析	
6 ,	总结	7
	<del></del>	

《沪深股票》数据分析报告

## 1 概述

随着我国经济的不断发展,大量资金在市场上流通;然而股票的发展可以集中社会闲散资金,扩大生产建设规模,推动经济发展;股票可以扩大我国利用外资的渠道和方式,增强对外的吸纳能力,有利于更好的利用外资。上海和深圳地区作为我国经济发展的强力引擎,沪深股票的发行更是关系着两个地区企业的发展。对沪深地区股票的数据进行现代化分析更是当代企业发展不可忽视的必备技能。

## 2 数据描述

#### 2.1 数据来源

数据的来源:发现数据平台

https://www.fxbaogao.com/dt?keywords=%E6%B2%AA%E6%B7%B1%E8%82%

A1%E7%A5%A8&order=2

特点: 沪深股票主要数据以图表的形式呈现, 清晰明了

典型股票数量: 20 支

股票代码:

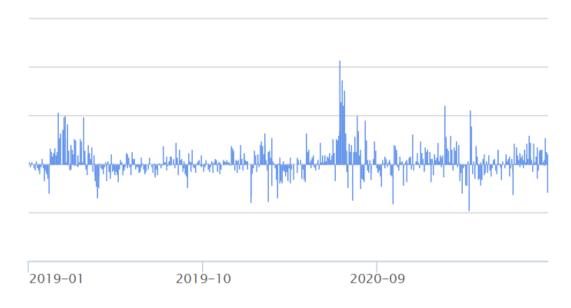
沪市: http://www.sse.com.cn/assortment/stock/list/share/

深市: <a href="http://www.szse.cn/market/companys/company/index.html">http://www.szse.cn/market/companys/company/index.html</a>

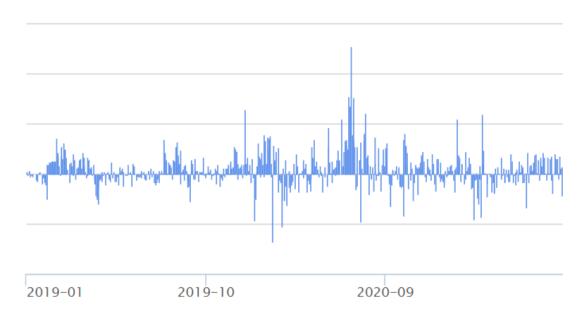
# 3 数据分析内容

### 3.1 沪深股票整体市场情况

沪市:



#### 深市:



总体而言,两个市场的股票发展情况有起有落,保持持续发展的趋势。整体 来看,有跌有落,大宗交易多,交易金额大。

#### 3.2 沪深特色股票举例分析

股票	最新价	涨跌幅	流通市值	换手率
中恒集团	3.98	9.94%	138 亿	12.03%
中远海特	6.03	9.95%	135 亿	6.57%
野马电池	44.11	3.13%	15亿	31.8%
神州数码	22.41	10.01%	112亿	10.93%
通合科技	16.43	20.1%	21 亿	14.52%
嘉益股份	23.54	<u>-9.56%</u>	6亿	30.75%

# 注: 沪市股票为 蓝色 深市股票为 黑色

分析:沪市股票大部分每支价格低,涨跌幅小,波动变化小,风险较低;流通市 值较大,换手率低,市面上的选择度高,比较受欢迎;而深市股票相比起沪市股 票来,单价较高,涨跌幅较大,稳定性能差,具有一定程度的风险;市面选择度 低于沪市,换手率较高。

## 3.3 沪深股票的基本收益

沪深 300, 又名中证 300, 由国内 A 股中规模排在前 300 的公司的股票组成 具有竞争力强、经营稳健、业绩优良的特点

中证 500 由国内 A 股中规模排在 300-800 之间的 500 家公司的股票组成具有成长性好、发展潜力大的特点

	中证500	沪深300	20%中证500+80%沪深300	80%中证500+20%沪深300
总回报	245.23%	216.65%	222.37%	239.51%
年化收益率	31.70%	29. 19%	29.69%	31.20%
波动率	38.73%	34.09%	35.02%	37.80%

数据来源:中证指数有限公司,统计区间:2004/12/31至2009/6/30

# 4 数据分析图表

## 4.1 沪深股票行情数据

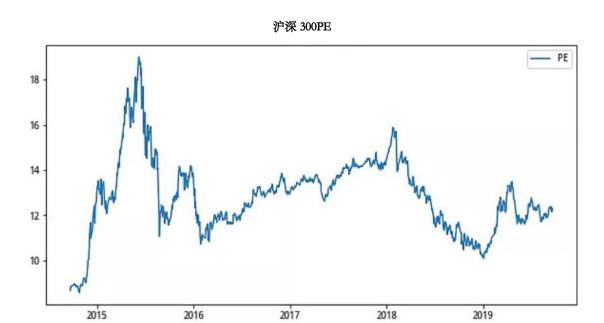
id	name code now rise		changehands	amplitude	priceearnings	marketrate			
	1	浦发银行	600000	12.94	3.94	0.44	5.62	5.92	0.82
	2	白云机场	600004	22.58	-3.91	1.10	4.89	54.57	2.98
	3	东风汽车	600006	4.62	1.09	0.41	1.31	18.78	1.28
	4	中国国贸	600007	17.75	-0.50	0.20	2.91	18.32	2.52
	5	首创股份	600008	3.46	0.00	0.30	1.45	30.66	1.80
	6	上海机场	600009	79.23	-1.28	1.40	3.03	28.28	5.38
	7	包钢股份	600010	1.49	0.68	0.47	1.35	32.61	1.27
	8	华能国际	600011	5.85	-0.17	0.32	1.02	12.02	1.21
	9	皖通高速	600012	5.60	0.90	0.29	1.80	8.48	0.95
	10	华夏银行	600015	7.77	2.10	0.46	3.29	5.67	0.62
	11	民生银行	600016	6.27	1.62	0.46	3.08	4.34	0.67
	12	日照港	600017	3.02	0.67	0.63	1.33	12.60	0.81
	13	上港集团	600018	5.96	1.02	0.16	1.86	15.79	1.78
	14	宝钢股份	600019	6.00	0.67	0.14	1.01	10.80	0.78
	15	中原高速	600020	4.28	1.90	0.66	1.90	5.60	0.97
	16	上海电力	600021	8.14	0.74	0.19	1.11	17.06	1.36
	17	山东钢铁	600022	1.51	1.34	0.49	2.01	14.31	0.80

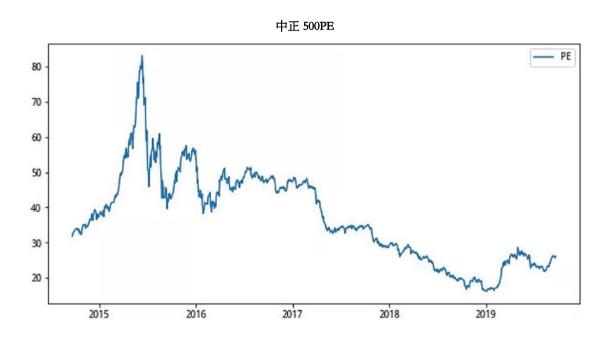
## 4.2 沪深上市公司基本面数据

Α	В	C	D	E	F	G	Н		J	K	L	M	N	0	P	Q	R	S	1
股票代码	股票名称	最高	最低	今开	胜收	涨停	跌停	换手率	振幅	成交量	成交额	内盘	外盘	委比	涨跌幅	市盈率(动	市净率	流通市值	总市值
000002	万 科A	28.19	27.82	27.87	28.02	24.95	20.41	0.32%	1.32%	30.73万	86154.637	15.46万	15.26万	48.71%	0.39%	70.93	70.93	2732.88(2	3179.29(2
000001	平安银行	12.99	12.78	12.93	12.92	9.60	7.86	0.37%	1.63%	63.49万	81674.097	34.8万	28.69万	-10.46%	-0.54%	7.41	7.41	2206.38(2	2206.4
019583	18国债01	0.00	0.00	0.00	103.00	0.00	0.00	0.00%	0.00%	0万	0万	0万	0万	0.00%	0.00%	0.00	0.00	0(Z	0(Z
200152	山航B	9.58	9.47	9.47	9.58	13.73	11.23	0.05%	1.15%	0.07万	62.23万	0.04万	0.02万	12.97%	-0.63%	23.98	23.98	13.33(Z	38.08(Z
000001	平安银行	12.99	12.78	12.93	12.92	9.60	7.86	0.37%	1.63%	63.49万	81674.097	34.8万	28.69万	-10.46%	-0.54%	7.41	7.41	2206.38(2	2206.4Z
000002	万 科A	28.19	27.82	27.87	28.02	24.95	20.41	0.32%	1.32%	30.73万	86154.637	15.46万	15.26万	48.71%	0.39%	70.93	70.93	2732.88(2	3179.29(2
000004	国农科技	21.88	21.35	21.88	21.40	19.51	15.97	1.00%	2.48%	0.83万	1787.47万	0.51万	0.32万	91.61%	0.09%	84.05	84.05	17.76(Z	17.99(Z
000005	世纪星源	3.33	3.16	3.16	3.17	3.38	2.76	1.17%	5.36%	12.37万	3991.57万	5.62万	6.75万	34.23%	2.52%	-48.45	-48.45	34.38(Z	34.4(2)
000006	深振业A	5.76	5.67	5.70	5.73	5.96	4.88	0.74%	1.57%	10.01万	5727.3万	3.99万	6.02万	-48.81%	0.52%	17.53	17.53	77.63(Z	77.76(Z
000007	全新好	6.80	6.60	6.70	6.78	15.31	12.53	1.85%	2.95%	5.7万	3799.1万	3.68万	2.03万	51.79%	-1.92%	64.80	64.80	20.55(Z	23.04亿
800000	神州高铁	4.05	3.99	4.02	4.02	5.46	4.46	0.74%	1.49%	19.28万	7754.2475	10.59万	8.69万	0.00%	0.00%	238.04	238.04	104.48{Z	111.79(Z
000009	中国宝安	5.95	5.75	5.78	5.81	5.47	4.47	1.20%	3.44%	25.41万	14935.987	11.43万	13.98万	-28.91%	1.72%	48.05	48.05	125.28(Z	127.03{Z
000010	美丽生态	2.93	2.83	2.83	2.92	4.75	3.89	0.76%	3.42%	3.97万	1142.58万	1.82万	2.15万	-28.03%	-0.34%	206.99	206.99	15.2亿	23.86(Z
000011	深物业A	10.42	10.27	10.38	10.34	10.96	8.96	1.39%	1.45%	2.45万	2525.46万	1.41万	1.04万	43.54%	-0.39%	19.57	19.57	18.11{Z	61.39(Z
000012	南玻A	4.80	4.67	4.69	4.67	5.26	4.30	0.70%	2.78%	12.35万	5862.93万	5.4万	6.96万	-54.90%	2.78%	25.99	25.99	85.27(2	137.44Z
000014	沙河股份	10.13	9.93	9.97	9.97	13.78	11.28	1.68%	2.01%	3.38万	3393.85万	1.62万	1.76万	-51.62%	0.70%	-74.88	-74.88	20.25(Z	20.25(Z
000016	深康佳A	4.51	4.46	4.46	4.47	5.29	4.33	0.78%	1.12%	12.38万	5550.05万	6.53万	5.84万	-7.88%	0.45%	31.59	31.59	71.69(Z	108.12{Z
000017	深中华A	4.66	4.51	4.55	4.55	3.84	3.14	1.17%	3.30%	3.55万	1631.01万	1.96万	1.59万	-16.43%	1.32%	-1732.50	-1732.50	13.97亿	25.42亿
000018	神州长城	1.61	1.55	1.56	1.58	3.52	2.88	2.83%	3.80%	28.14万	4471.65万	16.18万	11.96万	-27.06%	0.63%	-5.23	-5.23	15.83(Z	27{Z
000019	深深宝A	7.75	7.52	7.57	7.55	10.63	8.69	1.99%	3.05%	8.2775	6319.47万	3.82万	4.45万	25.15%	1.19%	18.14	18.14	31.78(Z	88.05亿
000020	深华发A	12.09	11.94	12.09	12.08	13.84	11.32	0.48%	1.24%	0.88万	1050.34万	0.53万	0.34万	12.92%	-0.17%	3306.09	3306.09	21.85(Z	34.15{Z
000021	深科技	8.11	7.98	8.11	8.12	7.49	6.13	1.11%	1.60%	16.26万	13068.527	9.97万	6.29万	-12.95%	-0.99%	28.48	28.48	118.14亿	118.29亿
000022	深赤湾A	0.00	0.00	0.00	16.46	19.13	15.65	0.00%	0.00%	0万	0万	0万	0万	-7.03%	0.00%	48.85	48.85	76.52(Z	295.2(Z
000023	深天地A	13.93	13.60	13.93	13.88	14.44	11.82	3.01%	2.38%	4.18万	5722.0375	2.53万	1.64万	-21.56%	-1.37%	31.88	31.88	19{Z	19{Z
000025	特力A	22.73	20.52	20.84	20.66	34.56	28.28	1.52%	10.70%	5.98万	13260.537	4.08万	1.9万	53.73%	10.02%	135.95	135.95	89.28(Z	97.98(Z
000026			7.84	7.85	7.88	9.04	7.40	0.71%	2.41%	2.53万	2006.96万	1.2万	1.32万	-13.18%	0.76%	13.66	13.66	28.32亿	35.17{Z
000027			6.19	6.20	6.22	5.30	4.34	0.14%	0.96%	5.7万	3546.9万	2.62万	3.08万	-44.57%	0.16%	16.78	16.78	246.99{Z	246.99(Z
000028	国药一致	45.91	43.63	44.00	43.90	50.41	41.25	0.93%	5.19%	2.87万	12870.377	1.56万	1.31万	15.59%	4.53%	16.36	16.36	141.22(Z	196.47{Z
000029	深深房A	0.00	0.00	0.00	11.17	12.29	10.05	0.00%	0.00%	0万	0万	0万	0万	0.00%	0.00%	33.62	33.62	99.6亿	113(Z
000030	富卑股份	5.13	5.05	5.11	5.11	8.25	6.75	0.18%	1.57%	3.15万	1600.83万	1.89万	1.26万	-45.83%	0.00%	11.36	11.36	89.76(Z	92.52(Z
000031	中粮地产	7.07	6.86	6.95	6.97	6.27	5.13	1.94%	3.01%	18.56万	12934.627	9.31万	9.25万	12.45%	0.29%	5.73	5.73	66.85(7)	274.4247
000032	深桑达A	8.84	8.65	8.70	8.74	8.43	6.89	0.53%	2.17%	2.19万	1919.47万	0.96万	1.23万	4.62%	1.03%	33.20	33.20	36.32{Z	36.49(Z
000034		13.75	13.52	13.72	13.64	17.26	14.12	0.87%	1.69%	4.2775	5838.37万		2.04万	34.62%	0.29%	16.43	16.43	66.96(Z	89.48(Z
000035			4.57	4.61	4.61	5.09	4.17	0.42%	1.52%	5.32万	2454.34万		2.7万	-12.84%	0.43%	35.31	35.31	59.31(Z	112.91{Z
000036	华联控股		7.25	7.29	7.29	6.68	5.46	0.81%	2.88%	9.1777	6774.8375		4.8477	-5.68%	1.78%	7.79	7.79	84.25(7)	84.717

#### 4.3 沪深股票指数数据

## 对其中几只股票分别进行股票指数数据分析,获取其折线图。如下所示:





#### 4.4 沪深股票波动 k 线图



## 5 数据分析结果

#### 5.1 股票行情数据分析

沪深股票均有一定程度的波动,时高时低,但总体上收益保持向好趋势发展。两 者均有不稳定性,有一定程度的风险,在挑选购买股票时应当谨慎,不能盲目偏 执,应当全面的总结市场行情,选择适合自己的股票。

### 5.2 沪深公司销售股票结果分析

从以上部分对沪深股票的发行的信息数据图来看,股民们更加偏向于购买稳定性 较强,波动性不大,风险较小的股票。

由图可知,沪市股票在一定程度上要比深市股票销量较高,有很大一部分原因是 沪市股票稳定性强,满足股民的低损失心理偏好。

### 5.3 股票规模相关性分析

一支股票的购买和很多因素具有相关性。

首先,上市公司的权衡。应当关注公司长远的发展情况,股票盈利与否与公司的 经营状况由密切联系,根据公司的发展情况判断自己的股票收益是否会受到波及。 其次,公司技术面和基本面的考量。一支好的股票一定要由好的技术面和基本面 做支撑。技术面就是说要买入趋势向上的股票,因为趋势线是一切的一切,趋势 上就是股价在上升通道,可能受消息面等影响或主力震仓洗盘,使股票有回落, 但总趋势向上。

最后,投资目的即股民心理的存在。股民购买股票是出于什么目的也在一定程度上影响了这支股票的最终受益。

#### 6 总结

在整个沪深股票数据分析中,我学到了运用代码绘制股票 k 线图,运用 Python 数据语言获取股票基本信息,比起一般的搜索方法来,Python 操作无疑是最省时省力又准确的高效方法。

首先,随着经济高速增长,股票发行不可忽视。我对沪深股票做了概述和深入研究。在此过程中,我又通过数据图的绘制,来帮助自己理解并且把数据分析做得更加透彻。

其次,这份数据报告从开始到结束花费了很多时间、精力和心血。当然也遇到了 很多困难。比如在编写代码的过程中,会有某种函数的运行不熟练,导致忽视了 简洁的方法,使整体数据更加繁杂甚至出错。但是最终通过查询教材资料和同学 的帮助等,困难都一个个迎刃而解。股票的大数据分析是当今经济发展以及企业 运行不可忽视的重要技术。

### 附录-数据分析代码

#### 1. 股票指数的获取

```
import tushare as ts
df=ts. get stock basics ()
df.info()
<class 'pandas.core.frame. DataFrame'>
Index: 3676 entries, 603662 to 603115
Data columns (total 22 columns):
Name
           3676 non-null object
Industry
           3676non-null object
Area
           3676 non-null object
Reserved
           3676 non-null float64
           3676 non-null float64
Share
           3676 non-null float64
Esp
Market
           3676 non-null float64
Undp
            3676 non-null float64
Profit
            3676 non-null float64
             3676 non-null float64
Gprnpr
dtypes:float64 (18),int64 (1), object (3)
memory usage: 660. 5+ KB
```

#### 2. 股票 k 线图的绘制

```
import tushare as ts
import pandas as pd
import numpy as np
import matplotlib.pyplot as pltfrom matplotlib
import tickerfrom mpl finance
import candlestick ochlplt.rcParams ['font.sans-serif'] = ['SimHei']pro=ts.pro api ()code='600004.
SH'
df = pro.daily (ts code=code, start date='20181201')
df = df.sort values (by='trade date', ascending=True)df ['trade date2'] = df['trade date'].copy()
df ['dates'] =np.arange (0, len (df))
def format date (x,pos):
if (x<0) or (x>len(date tickers)-1):
return'
return date_tickers [int (x)]
fig.subplots adjust(bottom=0.2)
candlestick ochl (ax, quotes=df2 [ ['dates', 'open', 'close', 'high', 'low']J.values,width=0. 55,
colorup='r', colordown='g', alpha=0. 95)
ax.xaxis.set major formatter (ticker. FuncFormatter (format date))
ax.set yiabel('交易价格')
```

```
plt.title (code)
plt.grid(True)
plt.xticks(rotation=30)
plt.show ()
```

#### 3. 将交易所网上的股票代码下载下来汇总到一个文件中

```
def share_code():
    with open('sh_info.txt', 'rU') as file:
        for code in file.readlines():
            code_list.append(code.strip())
    print(code_list)
```

#### 4.针对某些的空白问题的解决

```
def format_date(x,pos):
    if x<0 or x>len(date_tickers)-1:
        return "
    return date_tickers[int(x)]

date_tickers = df.trade_date2.values
fig, ax = plt.subplots(figsize=(10,5))
ax.xaxis.set_major_formatter(ticker.FuncFormatter(format_date))
mpl_finance.candlestick_ochl(
    ax=ax,
    quotes=df[['dates', 'open', 'close', 'high', 'low']].values,
    width=0.7,
    colorup='r',
    colordown='g',
    alpha=0.7)
ax.set_title('上证综指 K 线图(2018.9-)', fontsize=20);
```

#### 4. 爬取沪深所有股票数据并生成 Excel 文件

```
import requests
from bs4 import BeautifulSoup
import re
import xlwt
import time
def getHTMLText(url, code="utf-8"):
    try:
        r = requests.get(url)
        r.raise_for_status()
```

```
r.encoding = code
         return r.text
    except:
         return ""
def getStockList(lst, stockURL):
    html = getHTMLText(stockURL, "GB2312")
    soup = BeautifulSoup(html, 'html.parser')
    a = soup.find all('a')
         for i in a:
         try:
              href = i.attrs['href']
              lst.append(re.findall(r"[S][HZ]\d{6}", href)[0])
def get txt():
    stock list url = 'https://hq.gucheng.com/gpdmylb.html'
    stock info url = 'https://hq.gucheng.com/'
    output file = '\\文件\\
    slist=[]
    getStockList(slist, stock list url)
    getStockInfo(slist, stock info url, output file)
def T excel(file name,path):
    fo = open(file_name,"rt",encoding='utf-8')
    file = xlwt.Workbook(encoding='utf-8', style compression=0)
    sheet = file.add sheet('stockinfo')
    line num = 0
    title = ['股票代码', '股票名称', '最高', '最低', '今开', '昨收',
               '涨停', '跌停', '换手率', '振幅', '成交量', '成交额',
               '内盘', '外盘', '委比', '涨跌幅', '市盈率(动)', '市净率',
               '流通市值', '总市值']
    for i in range(len(title)):
         sheet.write(0, i, title[i])
    for line in fo:
         stock txt = eval(line)
         line num += 1
         keys = []
         values = []
         for key, value in stock txt.items():
              keys.append(key)
              values.append(value)
         for i in range(len(values)):
              sheet.write(line num,i,values[i])
              i = i+1
    file.save(path)
```

main()