附件2：期末大作业格式



2021— 2022学年（2）学期

大作业名称： **上市公司毛利率分析**

姓 名：\_\_郝建民 许晓明 李佳宜\_蒋曼\_

年 级: \_\_\_\_2021级\_\_\_\_

专业班级：\_\_\_\_1班\_\_\_\_

学 号：\_\_202199001\_\_\_\_202199002\_\_\_

\_202199003\_\_\_\_202199004\_

|  |  |  |  |
| --- | --- | --- | --- |
| **评分标准** | **各项标准满分** | **学生得分** | **总分** |
| 设计目的明确，研究内容有实际价值 | 10 |  |  |
| 格式准确，语言通顺流畅，逻辑性强 | 15 |  |
| 小组分工明确，对数据分析有一定认识 | 15 |  |
| 程序代码运用得当，结果分析正确 | 30 |  |
| 分析过程全面、详细 | 30 |  |
| **备注** | 如有大段抄袭或雷同文章均为不合格 | |

注:评分标准和各项标准满分仅供教师参考

目录

[一、总体介绍 3](#_Toc88901137)

[二、数据采集及清洗 3](#_Toc88901138)

[三、数据分析 3](#_Toc88901139)

[四、数据可视化 3](#_Toc88901140)

[五、运行结果及分析 4](#_Toc88901141)

[六、感想及体会 4](#_Toc88901142)

# 一、总体介绍

（一）本次项目背景及目的

每股收益即每股盈利（EPS），每股收益通常被用来反映企业的经营成果，衡量普通股的获利水平及投资风险，是投资者等信息使用者据以评价企业盈利能力、预测企业成长潜力、进而做出相关经济决策的重要的财务指标之一。

本学期《Python程序设计》课程是为进行大数据分析奠定了基础，通过Pandas库和Matplotlip库对数据进行分析和可视化展示。本文将运用Pandas对沪深上市公司的股票进行统计，找出当季每股盈利最好的十只股票，并通过Matplotlip进行展示。。

（二）数据介绍

本文数据来源Tushare财经数据库，通过Tushare数据库下载2019年第一季度所有股票成长能力数据。数据属性做如下说明：

code,代码

name,名称

roe,净资产收益率(%)

net\_profit\_ratio,净利率(%)

gross\_profit\_rate,毛利率(%)

net\_profits,净利润(万元)

eps,每股收益

business\_income,营业收入(百万元)

bips,每股主营业务收入(元)（三）技术说明

本文实验环境为Windows2010操作系统，编程语言为Python3.10，编程工具是Jupyter Notebook。

# 二、数据采集及整理

（一）数据采集

1.导入tushare库。

代码：import tushare as ts

2.下载2019年1-3月数据。

代码：ts\_data=ts.get\_profit\_data(2019,1)

# 三、数据分析

（一）数据观察

1.导入pandas库

代码：import pandas as pd

2.将数据转换为DataFrame数据

代码：df1=pd.DataFrame(ts\_data)

3.检查数据

3.1重复数据

代码：df1.duplicated().sum()

3.2统计数据空值

代码：df1.isna().sum()

（二）数据清洗

1.重复数据处理

代码：df1=df1.drop\_duplicates()

2.空值处理

代码：df1=df1.dropna(axis=1, how='all')

（三）数据分析

1.数据筛选

代码：df1=df1.dropna(axis=0, subset=['gross\_profit\_rate'])

2.对数据进行排序，并选择最好的十只股票

代码：df1=df1.sort\_values('gross\_profit\_rate')

df1=df1[-10:]

# 四、数据可视化

1.导入matplotlib库

代码：import matplotlib.pyplot as plt

2. 绘图设置

代码：

#设置X轴,Y轴数据

y=data\_sort['gross\_profit\_rate']

x=data\_sort['name']

#解决汉字显示乱码

plt.rcParams['font.sans-serif'] = ['SimHei']

plt.rcParams['axes.unicode\_minus'] = False

#设置画布大小

plt.figure(figsize=(10,6),dpi=100)

# 添加标签

plt.xlabel('股票名称',fontsize='16')

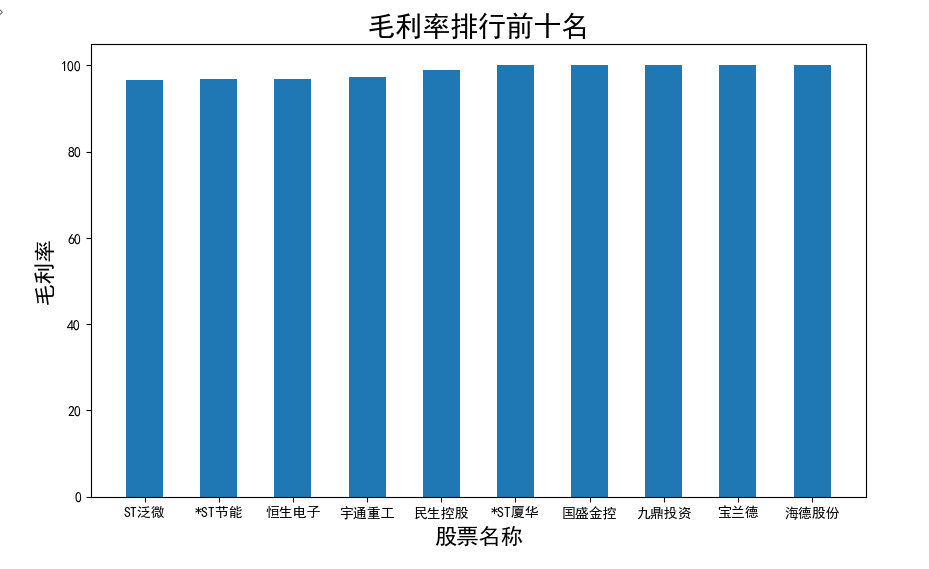
plt.ylabel('毛利率',fontsize='16')

plt.title('毛利率排行前十名',fontsize='20')

plt.bar(x,y,width=0.5)

# 五、运行结果及分析

通过分析，长航凤凰。。。等十只股票是2014年第三季度每股收益最好的十只股票。



# 六、感想及体会

**郝建民（202199001）**：本人负责数据收集和相关资料收集工作，通过本次项目锻炼，。。。。。

**蒋曼（202199002）**：

**李佳宜（202199003）**：

**许晓明（202199004）**：