



江西财经大学
JIANGXI UNIVERSITY OF FINANCE AND ECONOMICS

课程名称: Python语言与数据分析

课 程 报 告

项目名称 笔记本电脑数据集分析

班 级 金融203

学 号 0204890

姓 名 黄诗雨

任课教师 肖 泉

开课学期: 2020 至 2021 学年 第 2 学期

完成时间: 2021 年 7 月 4 日

《项目名称》数据分析报告

目 录

目录

1 概述.....	1
2 数据描述.....	1
2.1 数据来源.....	1
2.2 数据特点.....	1
2.3 数据清洗.....	1
2.4 数据整理.....	2
3 数据分析内容.....	2
3.1 各公司产品多样性对比	2
3.2 低、中低、中高、高端市场分层分析	2
3.2.1 各公司在低、中低、中高、高端市场的产品研发布局分析.....	2
3.2.2 低、中低、中高、高端市场各公司产品占比份额分析.....	3
3.3 市场畅销度分析.....	3
3.3.1 各产品类型走俏度分析	3
3.3.2 各操作系统受欢迎度分析	3
3.4 产品性能与市场价格关系	3
4 数据分析图表.....	4
4.1 各公司产品多样性对比	4
4.2 低、中低、中高、高端市场分层分析	5
4.2.1 各公司在低、中低、中高、高端市场的产品研发布局分析.....	5
4.2.2 低、中低、中高、高端市场各公司产品占比份额分析.....	5
4.3 市场畅销度分析.....	6
4.3.1 各产品类型走俏度分析	6
4.3.2 各操作系统受欢迎度分析	7
4.4 产品性能与市场价格关系	7
4.4.1 笔记本电脑的尺寸与价格的关系	8
4.4.2 笔记本电脑的重量与价格的关系	8
5 数据分析结果.....	8
5.1 各公司产品多样性对比	8
5.2 各公司产品多样性对比	9
5.2.1 各公司在低、中低、中高、高端市场的产品研发布局分析.....	9
5.2.2 低、中低、中高、高端市场各公司产品占比份额分析.....	9
5.3 市场畅销度分析.....	10
5.3.1 各产品类型走俏度分析	10
5.3.2 各操作系统受欢迎度分析	10
5.4 产品性能与市场价格关系	10
5.4.1 笔记本电脑的尺寸与价格的关系	10
5.4.2 笔记本电脑的重量与价格的关系	10
6 总结.....	11

附录-数据分析代码	11
-----------------	----

1 概述

本数据分析报告聚焦于一个小型笔记本电脑数据集，通过在细分为都下对笔记本电脑市场静态市分布数据集的分析，反向推导笔记本电脑市场所具有的一些特性，为该笔记本电脑市场中的主体企业未来产品研发战略布局提供建设性大方向意见，同时预测市场未来的走势以及主流，为明晰该笔记本电脑市场的现行特性以及未来属性提供帮助。

2 数据描述

2.1 数据来源

数据来源于阿里云天池

网址为：<https://tianchi.aliyun.com/dataset/dataDetail?dataId=92728>

2.2 数据特点

总共 1303 条数据，13 个字段。

```
Data columns (total 13 columns):
Unnamed: 0      1303 non-null int64
Company         1303 non-null object
Product         1303 non-null object
TypeName        1303 non-null object
Inches          1303 non-null float64
ScreenResolution 1303 non-null object
Cpu             1303 non-null object
Ram             1303 non-null object
Memory          1303 non-null object
Gpu             1303 non-null object
OpSys           1303 non-null object
Weight          1303 non-null object
Price_euros     1303 non-null float64
dtypes: float64(2), int64(1), object(10)
```

以上为笔记本电脑市场中 1300 个笔记本电脑型号的数据集，其中包含笔记本电脑的公司名称、产品名称、笔记本电脑类型、屏幕英寸、屏幕分辨率、CPU 型号、RAM 特性、记忆、GPU 特性、操作系统、笔记本电脑重量以及笔记本电脑的价格。接下来将通过以上数据集，对笔记本电脑的市场整体的产品多样性以及笔记本电脑的价格影响因素进行分析。

2.3 数据清洗

通过 info 函数了解到数据集数据较为良好，无缺失值。

使用 describe() 函数查看，输出结果指标包括 count、mean、std、min、max 及下四分位数，中位数和上四分位数，发现数据并无明显异常值。

2.4 数据整理

根据 describe 函数所给出的数据，通过对产品市场价格的四层分级，定义低端、中低端、中高端、高端四个产品等级，并增添一列“Class”属性，值为'低端','中低端','中高端','高端'，分别表示高端、中高端、中低端和低端产品。以便对不同层级的市场进行分析。

Class
中高端
中低端
低端
高端
高端
低端
高端

3 数据分析内容

3.1 各公司产品多样性对比

- 产品多样性，反映了企业的多角经营，体现市场的完善度以及对与消费者的服务度，对于企业、市场实力以及发展潜力的评估具有重要的参考意义。
- 通过每个公司品牌所设计研发的笔记本电脑的产品种类数量，以其占笔记本电脑市场产品种类总数的比例为依据，判断公司对于市场多样性的贡献，并通过饼图进行数据可视化，来分析该笔记本市场多样性现状，从而看出该笔记本市场的活力以及各公司的创新潜力，判断投资者或消费者未来进入该市场发展的可能性以及可收益预期。

3.2 低、中低、中高、高端市场分层分析

- 通过对产品价格进行分段，对整个市场进行低端、中低端、中高端和高端的层次的细分，对于微观层次的市场分析，以及企业个体方向指定、战略调整、建立优势竞争力具有重要的指导意义，也有利于分析消费者需求以及购买力的现状，预测未来市场的发展潜力所在。（在此因数据显示与可读性的问题，选择市场产品多样性贡献前四名的公司作为典例分析）

3.2.1 各公司在低、中低、中高、高端市场的产品研发布局分析

- 通过产品数据集，寻找企业在低端、中低端、中高端、高端四个层级的研发产品的种类数量，并通过柱状图将数据进行可视化，了解各企业创新创收能力的集中所在，以及企业目前发展战略布局现状，判断企业的主要目标市场以及目标消费人群，简单评判企业的类型为全面发展型亦或者是目标导向型，并且对未来可能的扩张方向进行预测导向，于企业有重要的指导意义。

3.2.2 低、中低、中高、高端市场各公司产品占比份额分析

- 与各公司不同层级的战略布局分析不同，以低端、中低端、中高端、高端为横轴纵向分析各市场公司扩张性及产品研发市场占领性，并以饼图形式对数据进行可视化，有助于剖析每一个市场，为不同层级的消费者提供更加丰富且完整的参考信息，利于供销对接，促进市场良性发展。

3.3 市场畅销度分析

- 以数量上的多少为依据，通过对市场上目前现存的产品类型以及某些特性进行分析，评估产品的畅销程度，并联系产品类型以及操作系统等产品本身特性的关联因素，分析其走俏的原因，通过多子图柱状图进行数据可视化，为企业的扩大再生产以及销售策略的改进提供材料以及数据指导。

3.3.1 各产品类型走俏度分析

- 通过分组分析每种产品类型在现存市场产品容量中的数量占比，从而看到某些产品类型在笔记本电脑市场上的覆盖程度，预测产品类型未来发展的趋势是向上向下，为企业未来的产品研发方向提供整体指导，达到高效率，高质量的发展。

3.3.2 各操作系统受欢迎度分析

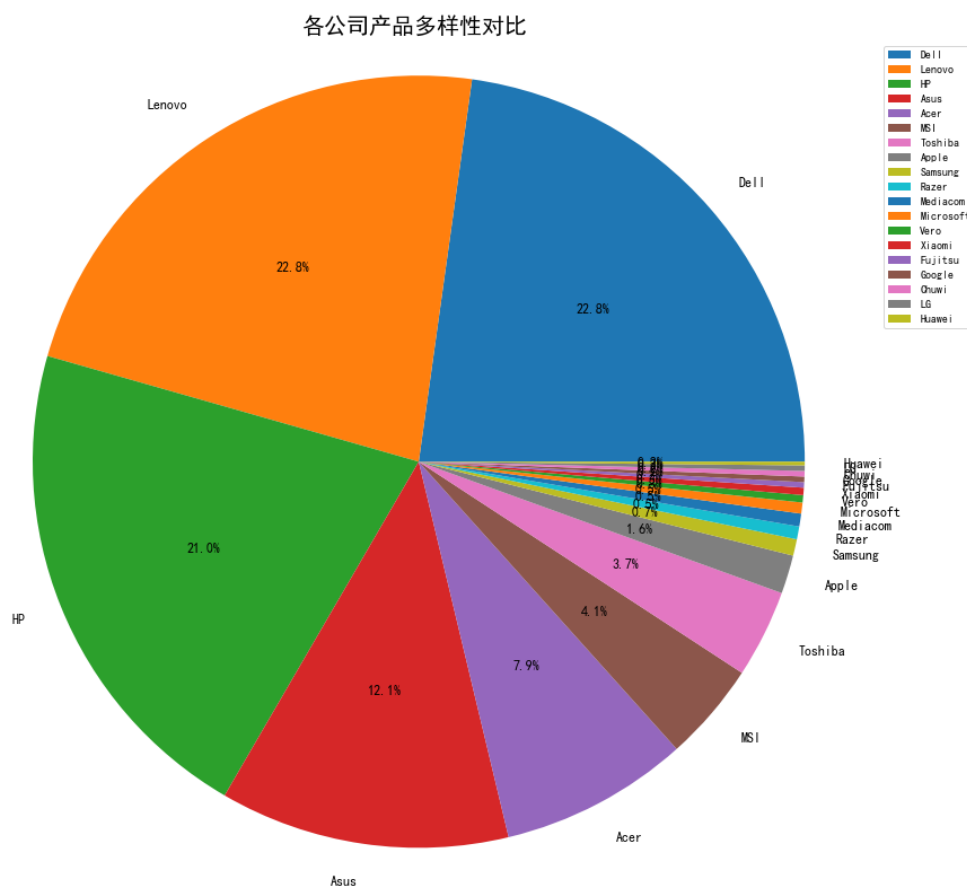
- 通过对目前市场上所拥有的所有产品使用的操作系统总体量进行分析，并将数据可视化，得出其中哪些操作系统的占比较大，从而反向推出市场上主流的操作系统，从消费者信任度的角度出发，结合未来科技发展实际和软件系统开发公司的发展前景，反向为公司未来产品研发的方向提供指导。

3.4 产品性能与市场价格关系

笔记本市场产品的定价还有多方面的影响因素，其中最主要的就是笔记本本身的特性，即英寸及重量，从这个角度分析，更加有利于抓住消费者对于产品外观的喜好，并且分析外部特性与价格的关系，通过关系系数的大小，判断内部功能研发与外部特性加强二者孰轻孰重。通过散点图将数据进行可视化。

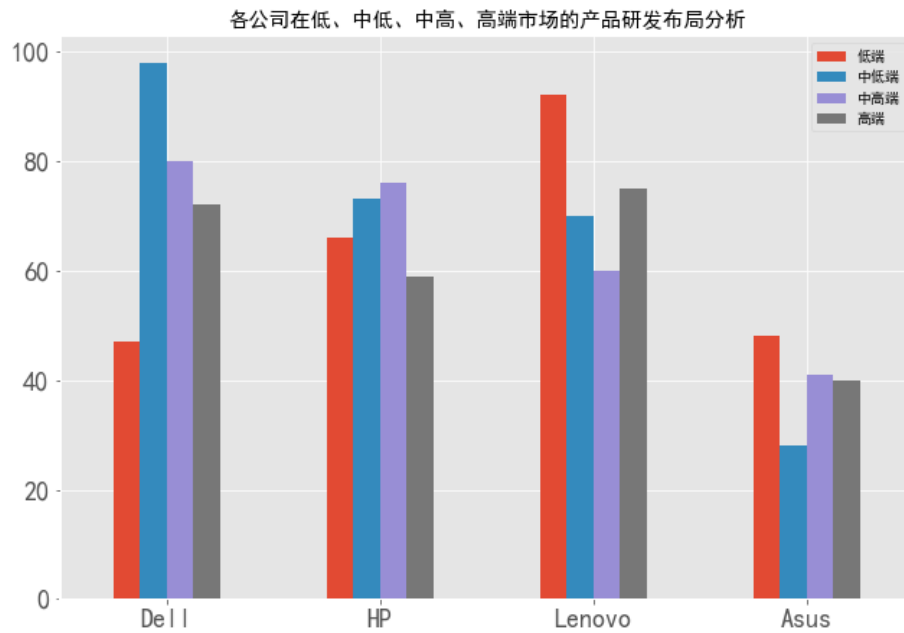
4 数据分析图表

4.1 各公司产品多样性对比

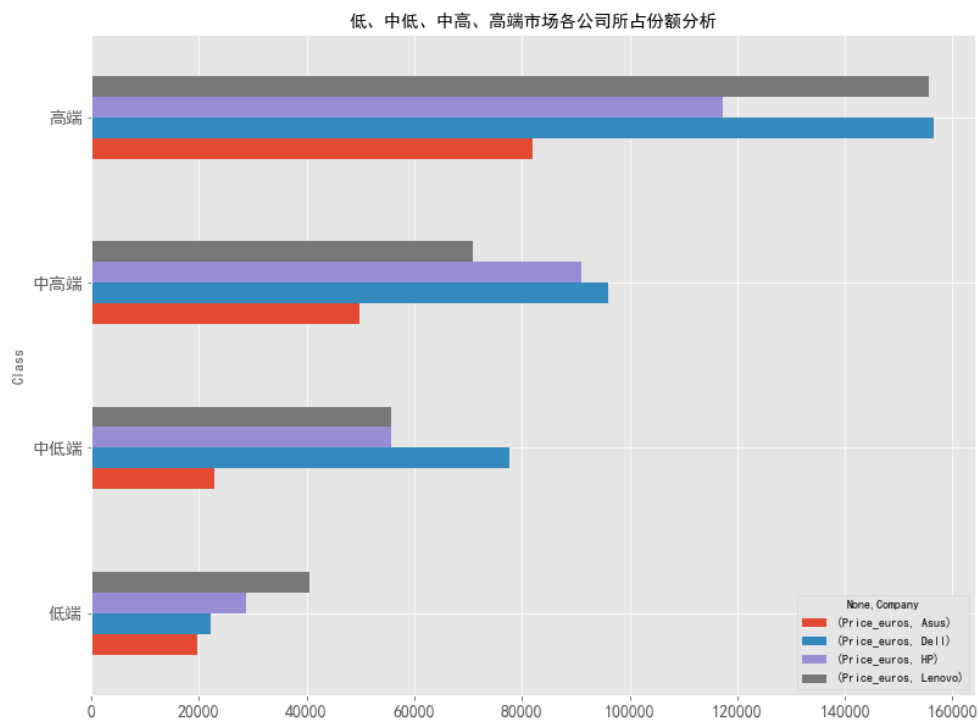


4.2 低、中低、中高、高端市场分层分析

4.2.1 各公司在低、中低、中高、高端市场的产品研发布局分析



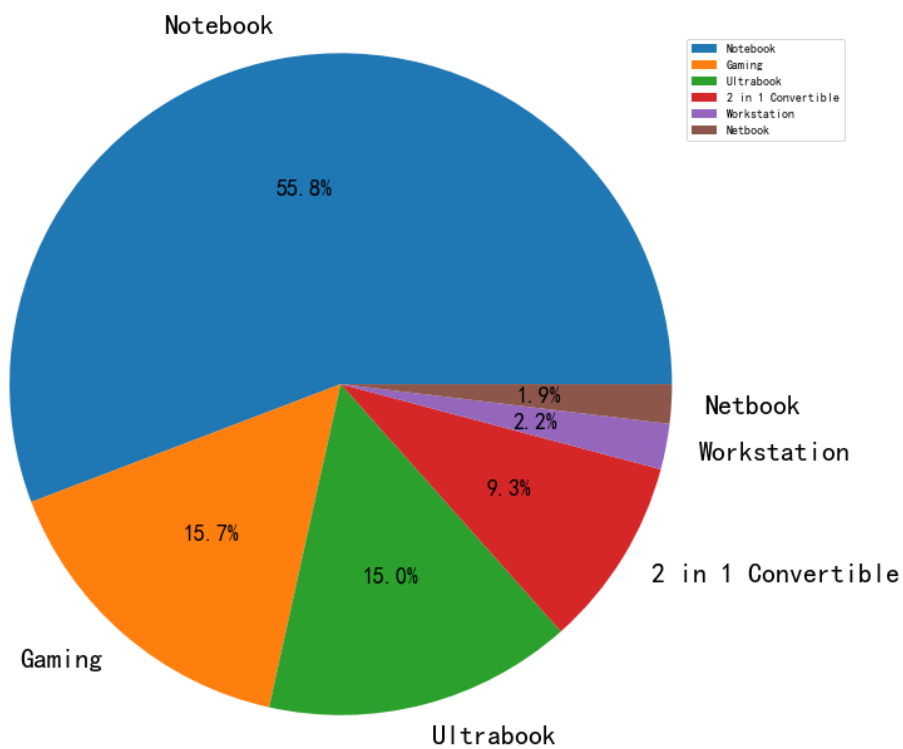
4.2.2 低、中低、中高、高端市场各公司产品占比份额分析



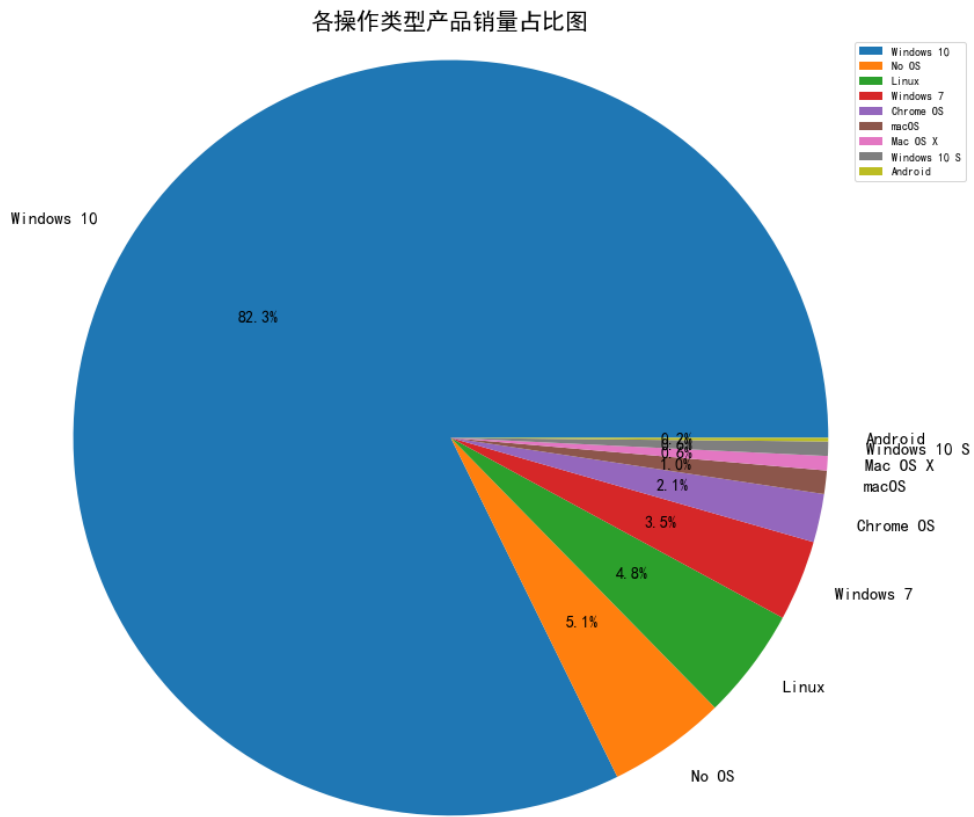
4.3 市场畅销度分析

4.3.1 各产品类型走俏度分析

各产品类型销量占比图

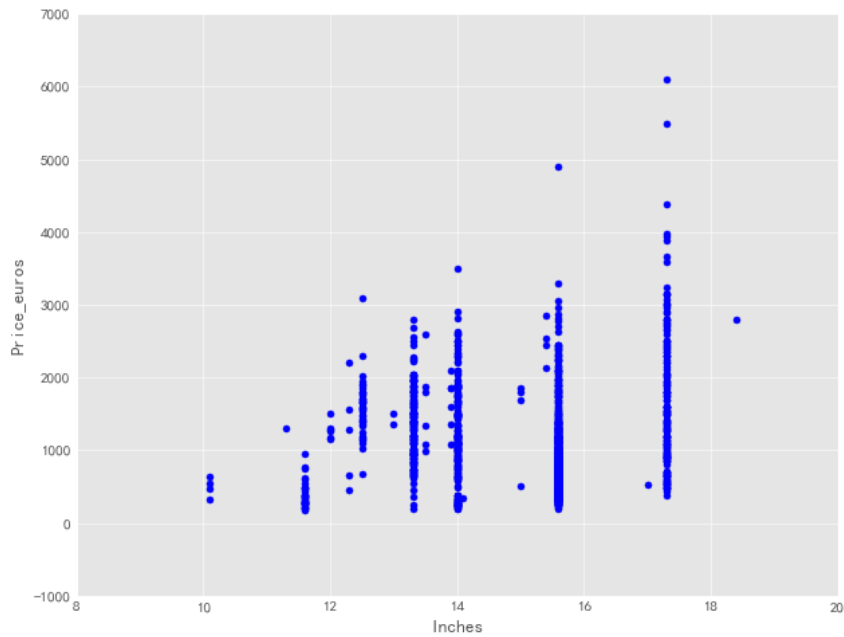


4.3.2 各操作系统受欢迎度分析

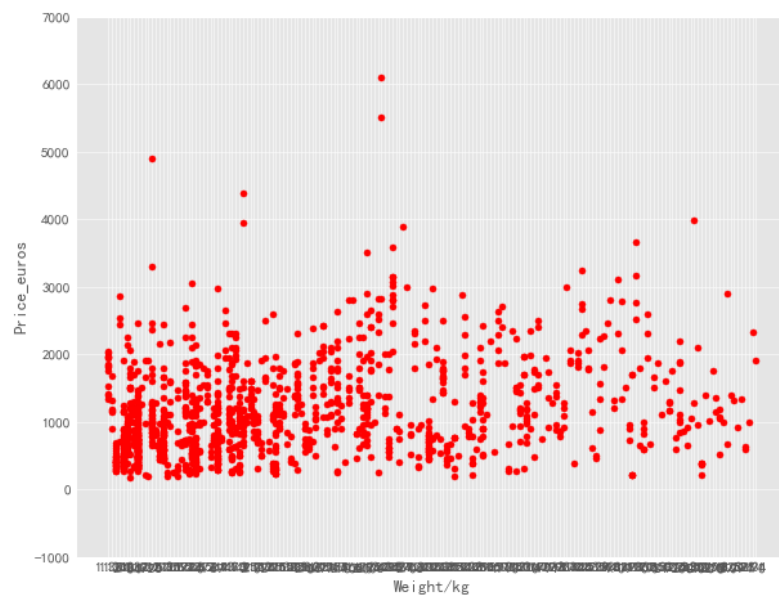


4.4 产品性能与市场价格关系

4.4.1 笔记本电脑的尺寸与价格的关系



4.4.2 笔记本电脑的重量与价格的关系



5 数据分析结果

5.1 各公司产品多样性对比

- 从饼图中可以看出，该笔记本电脑市场中，从高到低依次是Dell,Lenove,HP,Asus 这四个公司品牌为市场多样性作了主要贡献，产品多样

性，反映了企业的多角经营，体现市场的完善度以及对与消费者的服务度，对于企业、市场实力以及发展潜力的评估具有重要的参考意义。

- Dell,Lenove,HP,Asus 这四家公司在笔记本电脑的市场发展历史悠久，类型丰富，研究全面，产品多样，在市场上拥有强大的竞争实力，以及创新潜力，掌握多项专利技术与相关研究成果，有难以撼动的地位。但同时，这四家企业占比差别相距无几，市场上尚未出现强烈的垄断征兆，使得该笔记本市场在产品多样性的创新领域有良好的竞争氛围，有利于整个市场的创新发展、良性盘活。未来以投资者的身份进入该市场创收预期收益是较大的，而若以创业者的身份进入市场，由于行业市场内已有实力强大的商业大拿坐镇，技术的发展也需要一定时间和金钱的累计研发，难以在短时间内与之匹敌，是不明智的。

5.2 各公司产品多样性对比

5.2.1 各公司在低、中低、中高、高端市场的产品研发布局分析

- 摘取市场多样性贡献前四家公司 Dell,Lenove,HP,Asus 进行低端、中低端、中高端、高端的分级分析。从中可以看出，Dell 公司品牌主要在中低端产品进行多样性研发，但并不注重低端市场，中高端市场以及高端市场的扩张较为中等，目标人群可能主要是在中等收入阶层且日常家庭使用为主要需求。之后可以将一些精力放于低端市场，可以以中高端市场打下的名誉为基础，增加在低端市场的享誉度，获得更好的发展机遇。
- HP 公司品牌在低端、中低端、中高端、高端的布局都较为均衡，可以说是全面发展型企业，综合能力高，市场扩张潜力大，整体竞争实力较强，目前可以增加在高端市场的专注度，提升短板，以综合实力的信誉度，开展高端市场业务，同时更加注重高端产品的研发，打出自己的标志性产品。
- Lenovo 公司品牌，在低端产品的研发力度较大，高端市场也有兼顾，但中端市场的整体实力相比于其他公司品牌较为薄弱，可以适当加大终端产品研发，开拓中端市场，斩获中产阶层日渐扩大的发展机遇。
- Asus 公司品牌相较于其他公司品牌是较弱，应更加注重产品类类型的研发，率先追求突出优势，以此带动综合实力，在其他大公司品牌主导下，打出自己的专属品牌，在某一细分市场，以功能或价格优势占据优势地位。

5.2.2 低、中低、中高、高端市场各公司产品占比份额分析

- 从图中可以看出，Dell 和 Lenove 作为实力品牌，在产品的丰富和研发上具有强大的综合竞争优势，占据了低端、中低端、中高端、高端四个市场的大部优势，在各个层级消费者的选择中，强势挤占较高地位，与消费者关系较为良好
- 二者的优势尤其体现在对高端产品的功能研发改进以及高端消费者客户的累积上。而在低端、中低端、中高端三个层级的市场，四者相差较小。
- 同时从低端、中低端、中高端、高端四个层级的市场综合纵向来看，笔记本电脑的市场主要集中在中高端以及高端，因为大部分的电脑需求起源于工作性质需要抑或是，电子竞技类需要，这一需求出自收入较高的阶层，因此相应的业余更加多样性、个性化的产品服务，高端市场于笔记本电脑行业显然也是利润更加丰厚、创收更强的行业。未来，先进通讯及软件技术的发展

必定是此类高科技产品的生产导向，因此，专注新型产品的研发显然是符合未来发展趋势的。

5.3 市场畅销度分析

5.3.1 各产品类型走俏度分析

- 从饼图中可以看出，目前市场上所拥有的产品容量内，Notebook, Gaming, Ultrabook 占据绝大多数份额，由此可知，目前市场上笔记本电脑产品对于类型的需求具有固定性，大部分人出于某种需要的购买不会改变笔记本电脑的产品类型，因此，在产品类型的布局方面，可以依据市场现存的产品类型数量占比，来指定产品的研究以及推销方向，要避免库存过多带来的营业外支出损失，指定保险的风险调控应急方案。

5.3.2 各操作系统受欢迎度分析

- 但以产品类型为依据，指定产品研发方向的布局，较为单薄，因此，通向从笔记本电脑的显著特性——操作系统，来观察该笔记本市场目前产品种类存量中，各系统所占比例，推出系统的研发，将使得产品研发布局更为稳妥
- 从图中可以看出，Windows 10 系统目前占据了市场内笔记本电脑产品的据大多数操作系统类型，可以说是几乎垄断市场，反向推之，其消费者和市场信誉度必然较高，因此选择 Windows 10 系统是较为稳妥的产品研究方向基础，但同时大范围的使用也使得风险进一步扩大，若是 Windows 10 系统出现漏洞，将使得整个笔记本市场陷入风险，企业信誉受损，消费者信息安全遭到威胁，使整个市场的恢复运营信誉与实现良性发展受到强大的阻碍，因此研发拥有自我基础备用系统，不仅可能在未来的科技创新飞跃潮流抓住机会，也可以规避未来可能的巨大风险，一举两得，应该在有实力进行系统技术研发品牌的战略考虑范围内。

5.4 产品性能与市场价格关系

5.4.1 笔记本电脑的尺寸与价格的关系

相较来说，尺寸较大的笔记本电脑定价较高，但区别并不明显，主要还是因为内部配置与功能的原因。

5.4.2 笔记本电脑的重量与价格的关系

从散点图可以看出，笔记本电脑的重量和价格并无明显的线性相关关系，且也是由于电脑本身配置的精良性，导致重量较大的价格相对较高，以游戏高配置之机型为例，面向此类需求，便需要相应的重量。

- 因此，笔记本定价的主要因素仍然是内部的配置，而非外部形体条件，因此公司主要的战略导向依旧是产品研发，应避免过于注重外表的研发方向。（诚然，此处无销量的数据，因此也不能忽视外观对于消费者需求的导向作用，

但此处单指产品的研究方向。)

6 总结

该数据分析报告，仅针对一个小型欧洲市场的笔记本电脑产品数据集进行分析，维度非常有限，结论也不具有广泛实用性，但可以为笔记本电脑市场的分析提供参考，以及用于不同地域市场的对比（例如，欧洲市场与亚洲市场），仍具有一定的参考意义和价值。此分析报告主要从市场和企业品牌两个维度对数据进行解析，同时将市场分为低端、中低端、中高端、高端四个细分维度，提高分析的细致度以及严密度，使得分析报告的数据以及结论更加丰富。预期对于各公司产品研发战略布局，和对于产品类型以及操作系统研发的建议，具有一定的现实参考意义。但请注意这个数据集是针对笔记本电脑市场本身的产品特性以及定价的数据集，而不包括销售数量数据，因此本分析报告中相关的分析均是从无销量维度的单纯静态产品市场存在特性出发，进行反向分析推到，从而为现阶段动态的企业产品研发战略布局以及市场动态发展趋势提供指导以及预测意义。感谢观看数据报告，祝您生活愉快

附录-数据分析代码

```
#读取 csv 文件，并处理缺失值和异常值
import pandas as pd
data = pd.read_csv(r"d:\Spyder\laptops.csv", encoding='ISO-8859-1')
print(data.info())
print('各字段是否含有空值情况: \n', data.isna().any())
print(data.describe())

#数据整理
bins=[0,599,977,1487,6099] #定义区间段
labels=['低端','中低端','中高端','高端'] #设定标识文字
scut=pd.cut(data['Price_euros'],bins,labels=labels) #分段
data.insert(12,'Class',scut) #插入新列

#解决中文显示问题
import matplotlib.pyplot as plt
plt.rcParams['font.sans-serif'] = 'SimHei'
plt.rcParams['axes.unicode_minus'] = False
#整理数据，绘制图形

#各公司产品多样性对比
Company_marketheat=data['Company'].value_counts()
rate=Company_marketheat.values
labels=Company_marketheat.index
plt.figure(figsize=(18,18))
```

```

patches,ltext,ptext=plt.pie(rate,labels=labels,autopct='%1.1f%%',shadow=False)
for x in ltext:
    x.set_size(12)
for x in ptext:
    x.set_size(12)
plt.axis('equal')
plt.legend()
plt.title('各公司产品多样性对比',fontsize=20)

#各公司在低、中低、中高、高端市场的产品研发布局分析
#由于公司过多，这里分析只取占市场多样性前四名的公司
#设置绘图风格
plt.style.use('ggplot')
from pandas import DataFrame, Series
Dell=data[data.Company=='Dell']['Class']
Dell_Class=pd.value_counts(Dell)
HP=data[data.Company=='HP']['Class']
HP_Class=pd.value_counts(HP)
Lenovo=data[data.Company=='Lenovo']['Class']
Lenovo_Class=pd.value_counts(Lenovo)
Asus=data[data.Company=='Asus']['Class']
Asus_Class=pd.value_counts(Asus)
data3={'低端':[47,66,92,48],'中低端':[98,73,70,28],
        '中高端':[80,76,60,41],'高端':[72,59,75,40]}
Class=DataFrame(data3,index=['Dell','HP','Lenovo','Asus'])
Class.plot(kind='bar',rot=0,
            title='各公司在低、中低、中高、高端市场的产品研发布局分析',fontsize=18)

#低、中低、中高、高端市场各公司产品占比份额分析
data2=data[(data.Company=='Dell')|(data.Company=='HP')|(data.Company=='Lenovo')|
            (data.Company=='Asus')]
Class_distribution=data2.groupby(['Class','Company'])['Price_euros'].sum()
#将分组后的多层索引设置成列数据
Class_distribution=Class_distribution.reset_index(level=[0,1])
#使用数据透视表重新整理数据
Class_distribution=pd.pivot_table(Class_distribution,
                                   index='Class',
                                   columns='Company')

#绘制柱形图
Class_distribution.plot(kind='barh',title='低、中低、中高、高端市场各公司所占份额分析',
                        fontsize=14,figsize=(13,10))

#各产品类型走俏度分析
Typename_marketheat=data['Type Name'].value_counts()

```

```
rate=Typename_marketheat.values
labels=Typename_marketheat.index
plt.figure(figsize=(10,10))
patches,ltext,ptext=plt.pie(rate,labels=labels,autopct='%1.1f%%',shadow=False)
for x in ltext:
    x.set_size(22)
for x in ptext:
    x.set_size(18)
plt.axis('equal')
plt.legend()
plt.title('各产品类型销量占比图',fontsize=20,y=1.1)
```

#各操作系统受欢迎度分析

```
OpSys_marketheat=data['OpSys'].value_counts()
rate=OpSys_marketheat.values
labels=OpSys_marketheat.index
plt.figure(figsize=(20,20))
patches,ltext,ptext=plt.pie(rate,labels=labels,autopct='%1.1f%%',shadow=False)
for x in ltext:
    x.set_size(15)
for x in ptext:
    x.set_size(14)
plt.axis('equal')
plt.legend()
plt.title('各操作类型产品销量占比图',fontsize=20)
```

#笔记本英寸大小与定价分析

#设置风格

```
plt.style.use('ggplot')
x1=data['Inches']
y=data['Price_euros']
#新建 figure 对象
fig=plt.figure()
#新建子图 1
ax1=fig.add_subplot(1,1,1)
ax1.scatter(x1,y,color='blue')
ax1.set_xlabel('Inches')
ax1.set_ylabel('Price_euros')
ax1.set_title('笔记本电脑的尺寸与价格的关系')
```

#笔记本重量与定价关系分析

```
x2=data['Weight']
```

```
lst=[]
```

```
for i in x2:
```



```
i=str(i)
i=i[0:len(i)-2]
lst.append(i)
x2=lst
y=data['Price_euros']
#新建 figure 对象
fig=plt.figure()
#新建子图 1
ax2=fig.add_subplot(1,1,1)
ax2.scatter(x2,y,color='r')
ax2.set_ylabel('Price_euros')
ax2.set_xlabel('Weight/kg')
ax2.set_title('笔记本电脑的重量与价格的关系')
```