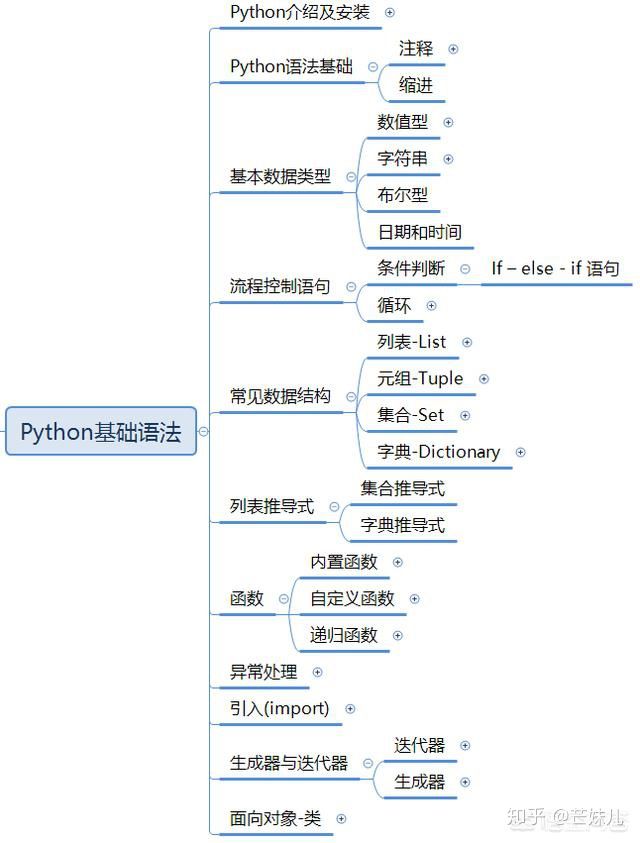
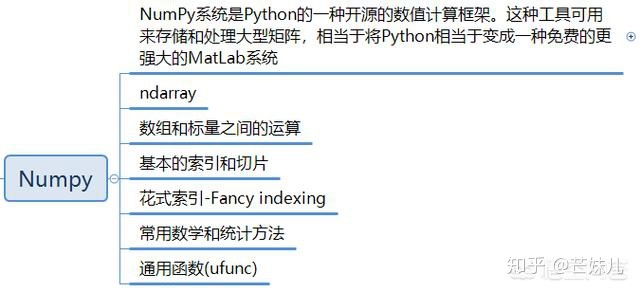
Python数据分析，主要需要学习以下内容

作者：芒妹儿  
链接：https://www.zhihu.com/question/412426667/answer/1388906233  
来源：知乎  
著作权归作者所有。商业转载请联系作者获得授权，非商业转载请注明出处。

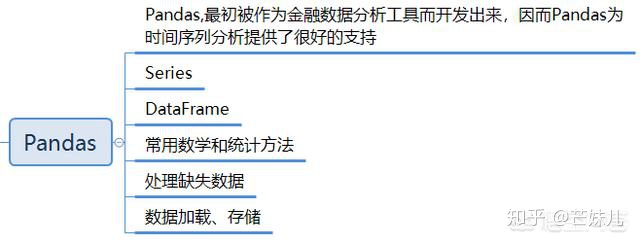
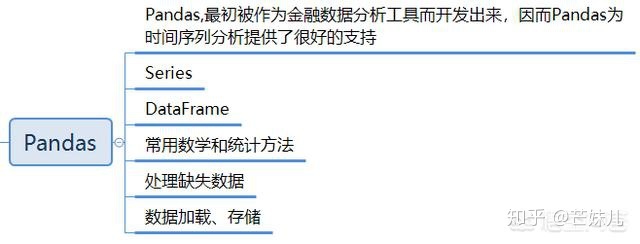
Python数据分析，主要需要学习以下内容：  
1、Python语法基础  
2、Python数据分析扩展包：Numpy、Pandas、Matplotlib等  
3、Python爬虫基础（非必须，但可以提升兴趣）  
4、Python数据探索及预处理  
5、Python机器学习  
Python[语法基础](https://www.zhihu.com/search?q=%E8%AF%AD%E6%B3%95%E5%9F%BA%E7%A1%80&search_source=Entity&hybrid_search_source=Entity&hybrid_search_extra={"sourceType":"answer","sourceId":1388906233}" \t "_blank)  
Python作为一门编程语言，当然需要先学习其语法基础，如果学习过其他编程语言，上手Python会比较快。Python语法基础需要掌握以下内容：



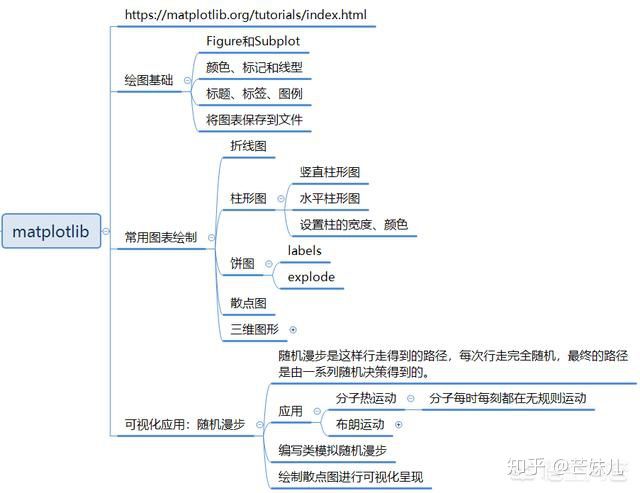
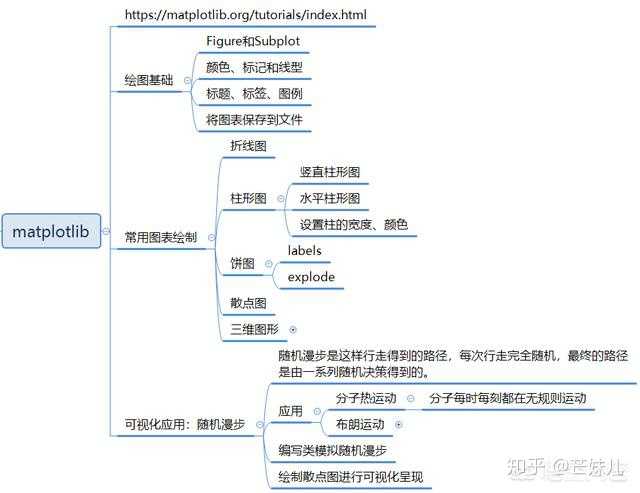
网上有很多学习资料，随便买本书就可以，或者私信小编了解更多。  
Python数据分析扩展包  
有了Python基础后，接下来就需要学习Python数据分析扩展包了，常用的有3个：Numpy、Pandas和Matplotlib。  
**1、Numpy**  
NumPy系统是Python的一种开源的数值计算框架。这种工具可用来存储和处理大型矩阵，相当于将Python相当于变成一种免费的更强大的MatLab系统。



**2、Pandas**  
Pandas，最初被作为金融数据分析工具而开发出来，因而Pandas为[时间序列分析](https://www.zhihu.com/search?q=%E6%97%B6%E9%97%B4%E5%BA%8F%E5%88%97%E5%88%86%E6%9E%90&search_source=Entity&hybrid_search_source=Entity&hybrid_search_extra={"sourceType":"answer","sourceId":1388906233}" \t "_blank)提供了很好的支持。另外Python中常用的DataFrame，及用读取外部数据文件的方法都属于Pandas。



**3、Matplotlib**  
Matplotlib是一个很强大的Python可视化库，可以很轻松地绘制各种数据图表，包括[三维图表](https://www.zhihu.com/search?q=%E4%B8%89%E7%BB%B4%E5%9B%BE%E8%A1%A8&search_source=Entity&hybrid_search_source=Entity&hybrid_search_extra={"sourceType":"answer","sourceId":1388906233}" \t "_blank)。



Python爬虫基础（非必须）  
严格来说，[Python爬虫](https://www.zhihu.com/search?q=Python%E7%88%AC%E8%99%AB&search_source=Entity&hybrid_search_source=Entity&hybrid_search_extra={"sourceType":"answer","sourceId":1388906233}" \t "_blank)不属于Python数据分析的范畴，但是可以作为一个可以提升自己兴趣，以及提升自己Python功底的工具，当然有些时候，数据分析师也需要自己爬取一些数据。



Python数据探索及预处理方法  
在学习了以上内容后，还需要学习一些常用的数据探索及预处理方法，才能够用Python进行一些基础统计分析，因为很多时候在分析数据前，还需要对数据进行探索及预处理。

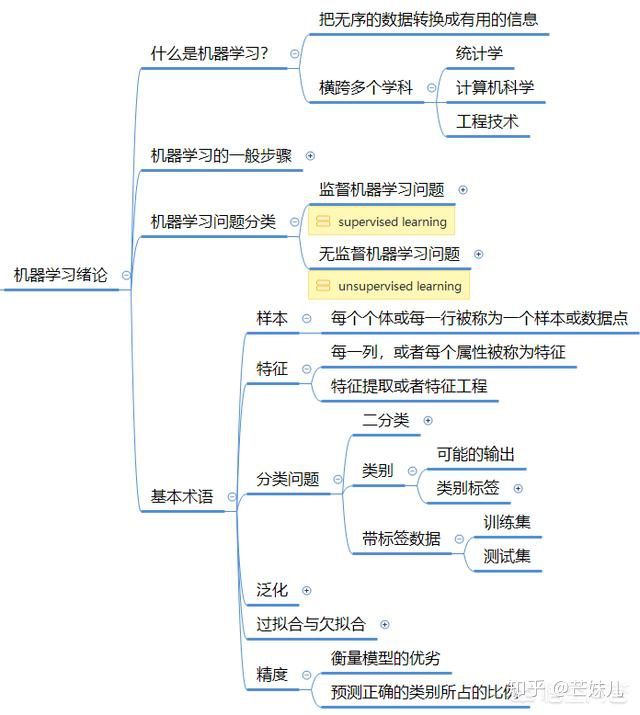
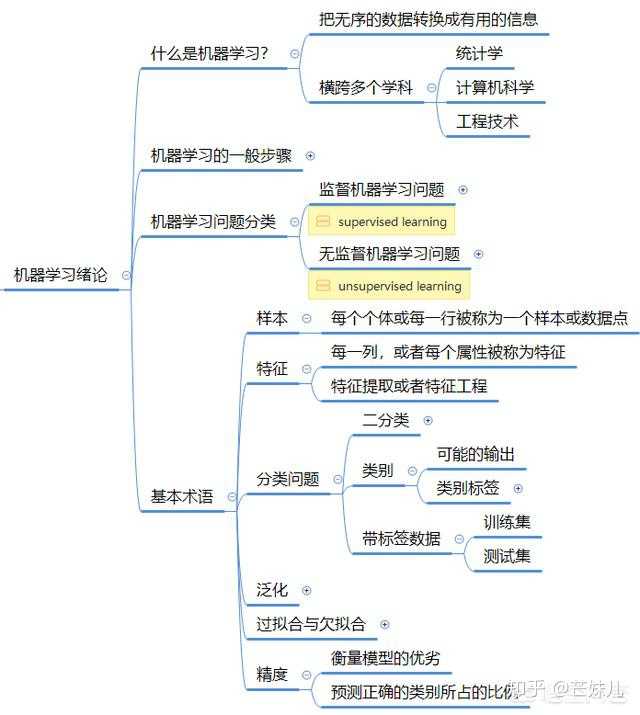


Python机器学习

在学习了以上内容后，就可以学习更强大也更复杂的分析方法了，也就说所谓的数据挖掘，主要工具就是机器学习。

**1、机器学习绪论**

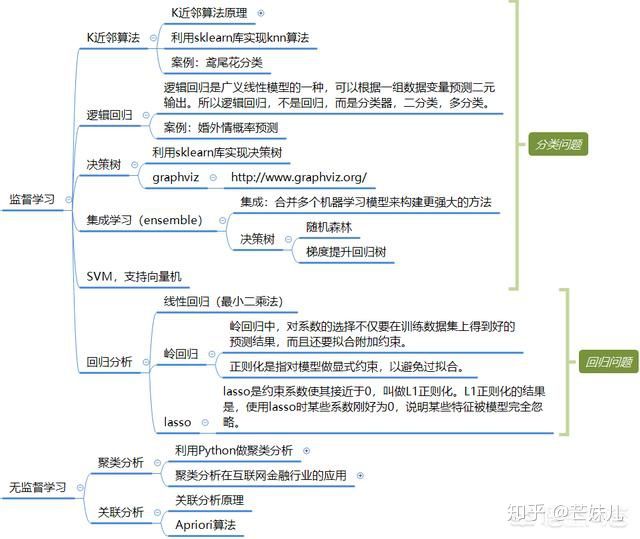
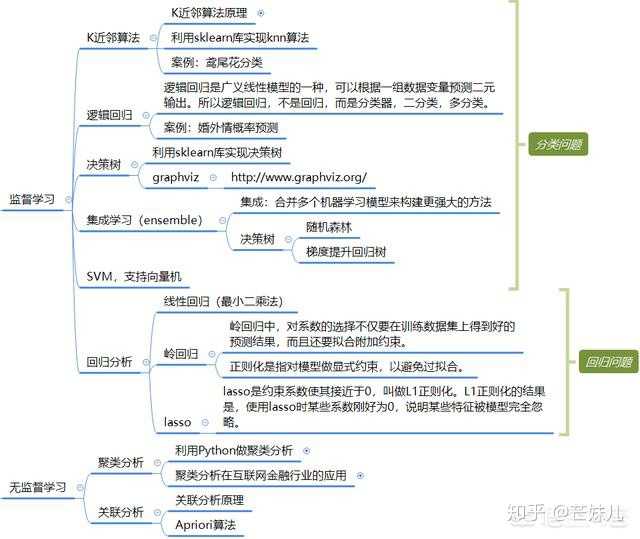
首先需要了解机器学习，及其常见术语。



**2、机器学习常用算法**

机器学习常用算法分为两类，[监督学习](https://www.zhihu.com/search?q=%E7%9B%91%E7%9D%A3%E5%AD%A6%E4%B9%A0&search_source=Entity&hybrid_search_source=Entity&hybrid_search_extra={"sourceType":"answer","sourceId":1388906233}" \t "_blank)和无监督学习。

大部分算法可以通过调用Scikit-learn中的现成算法来实现，当然可以自己编写算法，前提是数学功底要好，而且要对算法的原理掌握得很透彻。



回答完毕！

[发布于 2020-08-06 15:53](//www.zhihu.com/question/412426667/answer/1388906233" \t "_blank)