欢迎继续我们的学习。

在上一部分的内容中，我们学习了NumPy这个非常重要的package。

而从这一课开始，我们将进入Pandas的世界。当然，这里的Pandas并不是人人都爱的滚滚~

Pandas的全称是Python Data Analysis Library,是基于Numpy的一种开源数据分析包，其中纳入了大量与数据分析有关的库和标准数据模型，从而可以高效操作大型数据库。

Pandas的官方地址在这里：<http://pandas.pydata.org/>

Pandas有以下几个重要的特性：



1.Pandas是一个基于Numpy的开源库

2.使用Pandas可以快速进行数据分析

3.Pandans的运行效率非常高

4.Pandas自带可视化的数据分析

5.Pandas支持海量数据的大型数据库。

Pandas的安装

在进一步学习Pandas之前，首先需要安装Pandas。

如果之前已经安装过anaconda，那么只需要在终端输入命令

conda install pandas

如果没有，那么需要输入命令：

pip install pandas

在Pandas的这部分课程中，我们将了解以下内容：

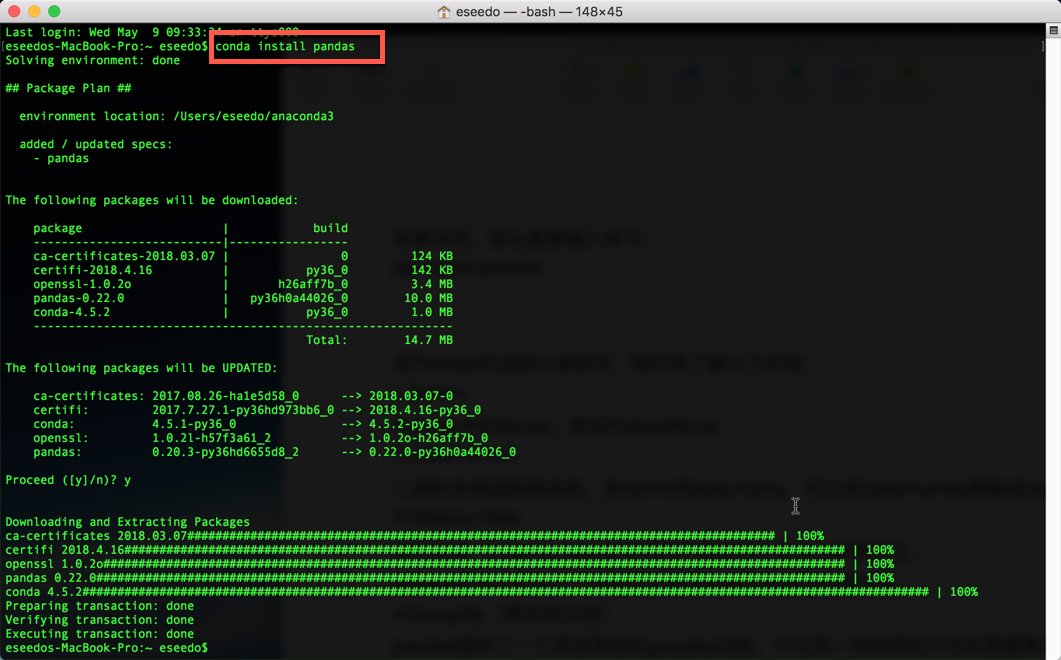
1.Series

Pandas中的Series，类似Python的List

2.DataFrames

二维的表格型数据结构，类似R中的data.frame，可以把DataFrames理解成Series的容器。

3.Missing Data  
缺失数据（missing data)在大部分的数据分析应用中都很常见。



pandas的设计目标之一就是尽可能简化缺失数据的处理任务。

4.GroupBy（聚合和分组）

pandas提供了一个灵活高效的groupby功能，可以用一种自然的方式对数据集进行切片、切块、摘要等操作。

5.Merging,Joining和Concatenating

Pandas提供了基于 series, DataFrame 和panel对象集合的连接/合并操作

6.Operations

这部分内容是关于Pandas的数据运算

7.数据输入和输出

这部分内容将带领大家来一起学习Pandas中的数据输入和输出

好了，本课的内容就到这里了。

我们下一课再见~

讨论群-笨猫学编程QQ群：

375143733

知乎专栏：

<https://zhuanlan.zhihu.com/kidscoding>

新浪博客:

<http://blog.sina.com.cn/eseedo>

Github:

<https://github.com/eseedo>

个人网站：

<http://icode.ai/>