

# Python编程 & 数据科学入门

## 第 1 课 Python Hello World

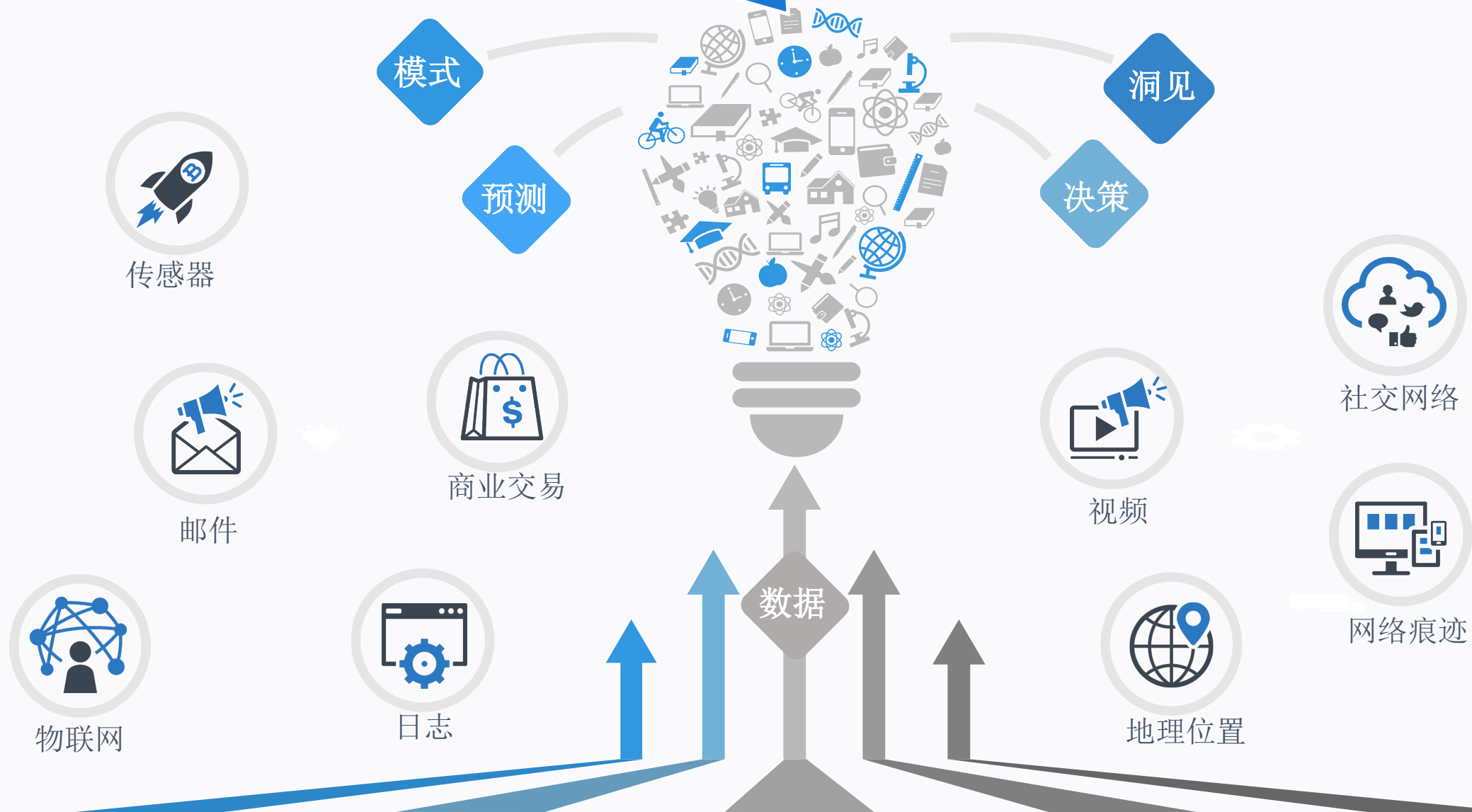


解密大数据团队出品



# 什么是数据科学？

发掘数据的价值



# 数据科学的应用领域



## 模式发现

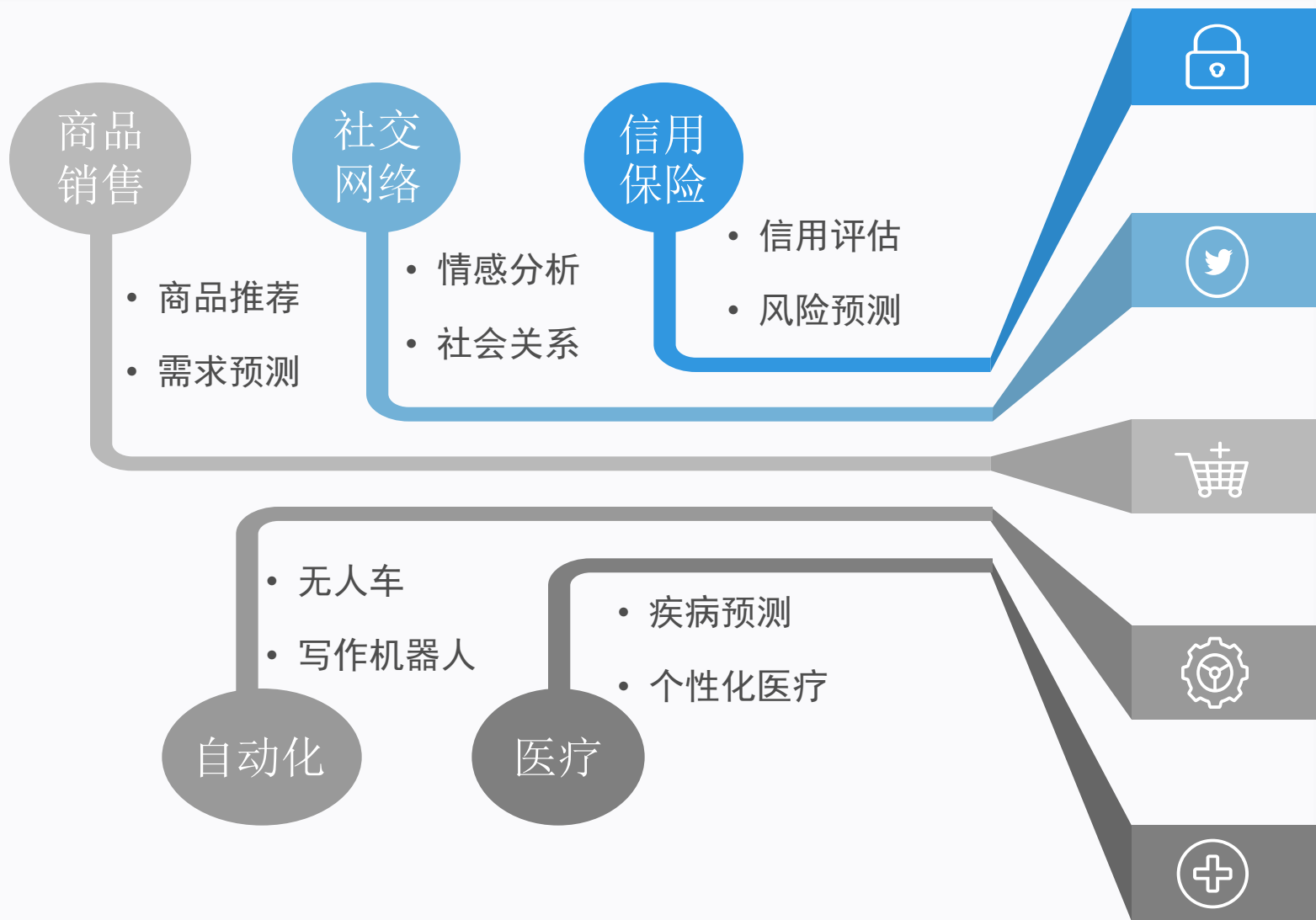
购物网站，根据用户浏览、购物记录，结合用户个人信息，绘制用户画像，并在合适的时间推荐合适的商品。

## 预测分析

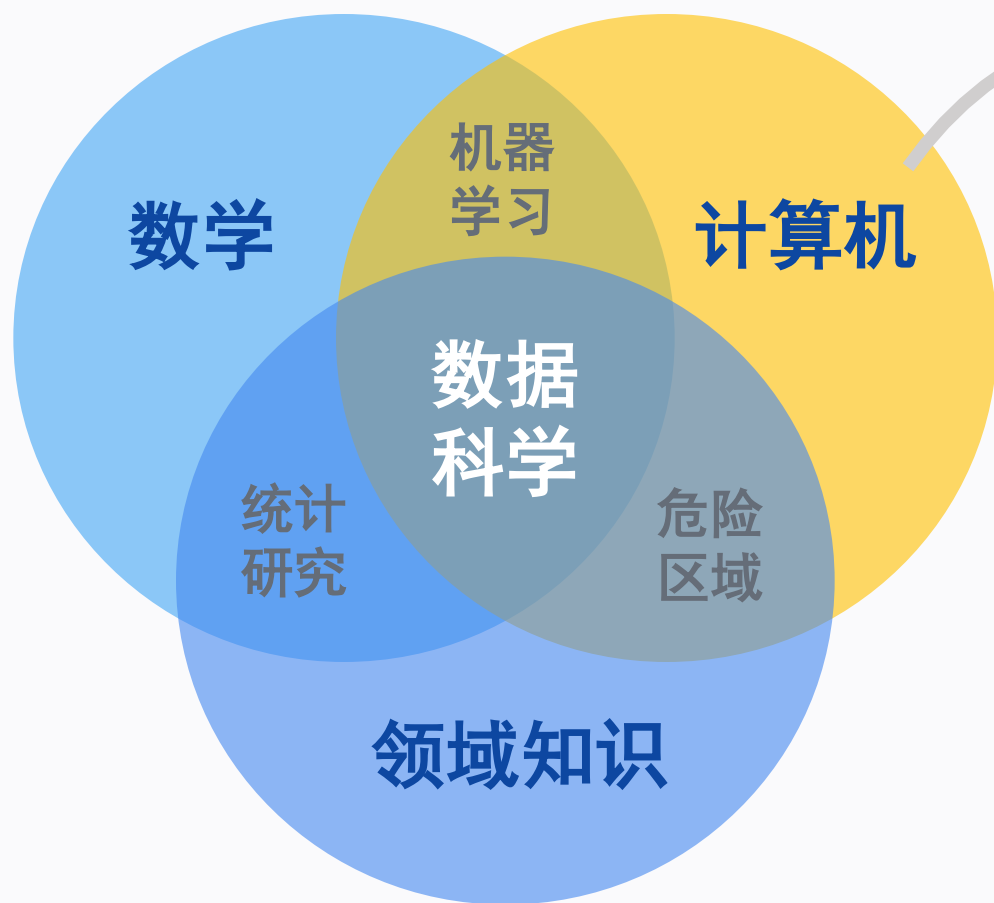
天气预报，收集来自船只、飞机、雷达、卫星上的数据，通过建模分析，来预测天气甚至是自然灾害。

## 智能决策

无人驾驶车，通过收集实时的传感器数据，构建周围环境的地图，从而进行加速、减速、转弯等操作。



# 数据科学的技能和Python



## 简单易学

美国顶尖大学最受欢迎的入门编程语言。



## 功能强大

不仅能进行数据分析，还能用于数据获取、清洗、数据库、高性能计算等领域。



## 丰富的数据分析工具

包括 Pandas、Numpy、Matplotlib、scipy、scikit-learn等。



# 课程大纲

START

介绍 Python 环境的搭建以及 Jupyter Notebook 的使用，并通过美国婴儿姓名分析的案例，讲解 Python 中的变量、表达式、语句等基本知识。

01

Part1 : Python 基础

Python Hello World

Part1 : Python 基础

数据结构，函数、条件和循环 (1)

02

通过文本单词计数的例子，讲解如何调用内置函数，如何实现条件判断和 for 循环，并介绍字符串和列表这两类数据结构。

通过创建学生成绩打分系统，学习如何自定义函数；通过编写九九乘法表，掌握嵌套循环的使用；并介绍元组和字典这两类数据结构。

03

Part1 : Python 基础

数据结构，函数、条件和循环 (2)



通过**红酒品质数据分析**的案例，学习Numpy科学计算包的基本运算和函数，掌握Numpy数组的操作。

## Part2 : Python 数据分析 Pandas 入门

05

通过**美国农业部食品数据分析**的案例，进一步学习pandas中DataFrame操作的高级功能，包括数据的合并、重塑、转换以及缺失值的处理。

## Part2 : Python 数据分析 Matplotlib 入门

07

04

## Part2 : Python 数据分析 Numpy 入门

通过**美国婴儿姓名演化分析**的案例，讲解数据分析包pandas的基本用法，包括Series和DataFrame数据结构及其操作。

06

## Part2 : Python 数据分析 Pandas 进阶

通过**波士顿房价的可视化分析**，介绍Matplotlib绘图库，学习各类数据图表的绘制。



通过泰坦尼克号事故数据的分析，介绍数据分析的总体流程，包括数据导入、清洗、探索等。

08

## Part2 : Python 数据分析 数据分析流程的总结和实操

## Part3 : 数据挖掘入门 线性回归-数据挖掘的经典模型

09

介绍机器学习的基本概念，并通过波士顿房价分析的案例，学习线性回归模型的应用。

通过贷款违约分析的案例，学习如何使用逻辑回归解决分类问题，以及分类模型的评估。

10

## Part3 : 数据挖掘入门 分类-数据挖掘的重要用途

END

# ◎ 本课内容



01

- Anaconda 安装
- Jupyter Notebook 介绍



PYTHON  
环境搭建

02

- 计算
- 数据类型
- 变量
- 函数和方法
- Packages 导入



PYTHON  
基础

03

- 通过美国婴儿姓名分析的案例，展示Python数据分析的强大功能。



案例演示



# Python 环境搭建

- Anaconda 安装
- Jupyter Notebook 介绍

01



# ② Anaconda 介绍



# 🌀 Jupyter Notebook 介绍



文学编程

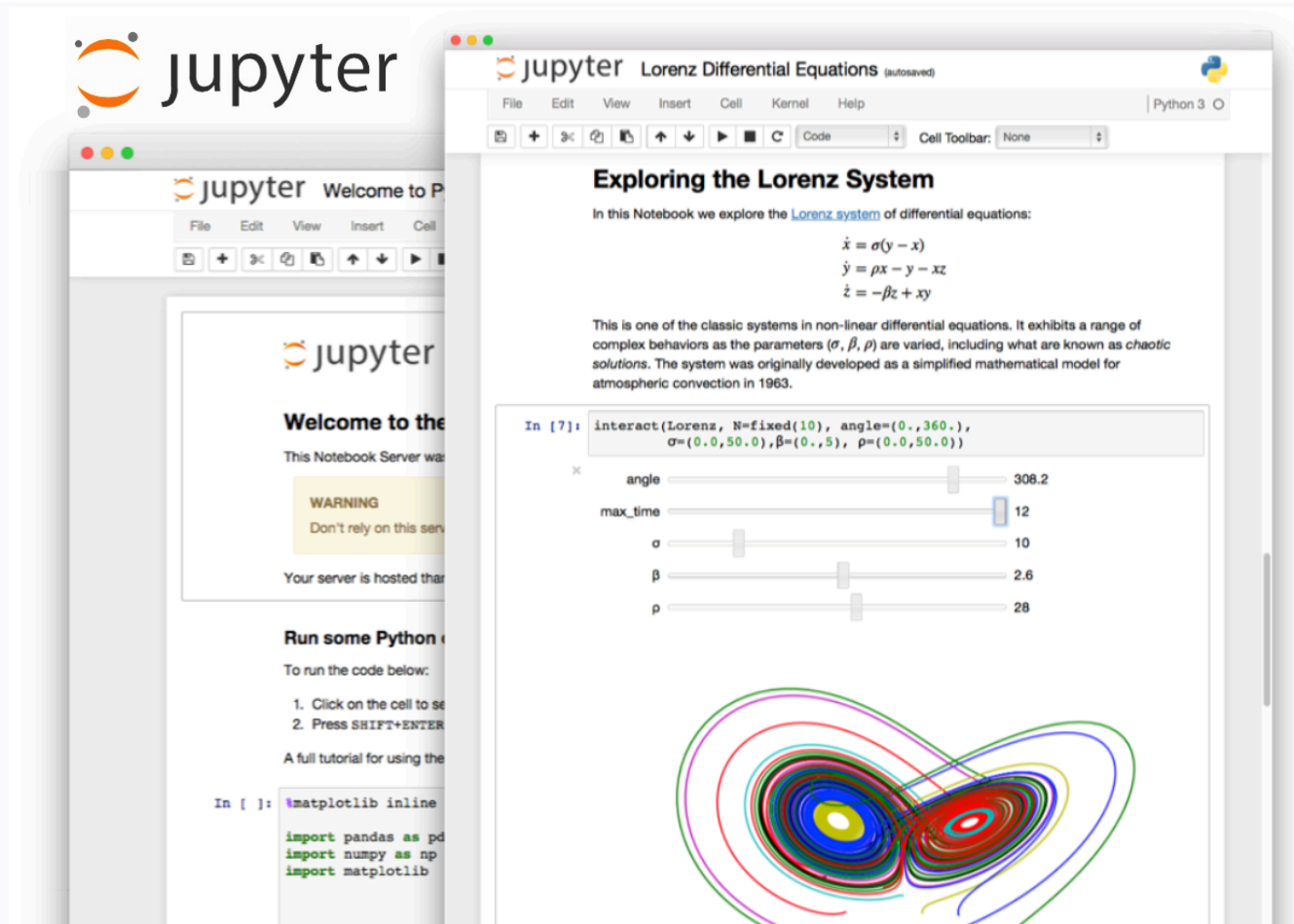
Literate programming



Coding 技能



Markdown 写作



# Python 基础

02

- 计算
- 变量
- 数据类型
- 函数和方法
- Packages 导入



# 案例演示

- 通过美国婴儿姓名分析的案例，展示Python数据分析的强大功能。

03



# 答疑时间

