



南京大學
NANJING UNIVERSITY

管理统计学——SPSS统计分析

	名称	特点	适用场景
1	Excel	一般非大量数据分析的人员，可以满足大部分需求	财务、金融、产品经理等一般数据量处理需求
2	R语言	兼容性强，语言程序化也强，在编程语言方面需要投入的精力比Python大，但适用面较广	最常用的数据分析工具之一，兼容性强
3	Python	以语言简单，注重数据分析的高效著称，尤其是在文本处理等数据结构化方面有很大优势	编程类数据分析，如文本字符等非结构化数据的处理
4	SQL	数据库处理和分析的必备技能，属于数据库方面的基本工具	侧重数据库方面，如数据仓库等，作为Oracle等数据库方面的基础知识不可或缺
5	SPSS	统计分析功能强大，侧重于统计分析类模型	建模能力已经不局限于统计了，在预测、机器学习方面也有很多包
6	SAS	金融大数据分析	金融风控建模较多
7	Matlab	矩阵计算等数学专用建模工具	强大的各种工具包，仿真能力



南京大學
NANJING UNIVERSITY

第1章 SPSS入门

本章学习目标

- 了解SPSS的发展历程
- 了解SPSS的特点
- 熟悉数据窗口和输出窗口
- 了解数据分析的步骤
- 了解统计分析软件的作用

主要内容

- 1. SPSS概述
- 2. SPSS操作入门
- 3. SPSS的窗口和结果输出
- 4. 数据分析方法论概述
- 5. 统计分析软件的作用

1. SPSS概述

- 软件名称
 - Statistical Package for Social Science（始于20世纪70年代），社会科学统计软件包；
 - Statistical Product and Service Solutions（始于2002年），统计产品与服务解决方案；
 - 注意：SPSS公司于2009年4月曾更名为PASW（Predictive Analytics Software，即预测统计分析软件），但用户还是习惯用SPSS。2009年9月被IBM收购后，重新将软件命名为IBM SPSS。

1. SPSS概述

- 软件发展

- 20世纪60年代：美国斯坦福大学三位研究生开发出SPSS；
- 20世纪70年代：SPSS总部成立于芝加哥，推出SPSS中小型机版—SPSSX；
- 20世纪80年代：微机版（V1~4）SPSS/PC+；
- 20世纪90年代：Windows版（v6~10）；
- 本世纪：11.0版~

1. SPSS概述

- 软件特点

- **SPSS的定位：**全球领先的统计分析、数据挖掘、预测建模产品及解决方案。
- **SPSS Statistics定位：**针对常规统计分析应用的统计软件。
- **功能强大：**系统地集成了多种成熟的统计方法与模型；强大的数据整理技术；提供自由灵活的图表功能；无处不在的帮助系统。
- **兼容性好：**能方便地和其他软件进行数据交换，Excel、文本、数据库表等；图表可直接复制粘贴到WORD、PPT等文件中。
- **易用性强：**人际界面友好、操作简单。
- **扩展性高：**直接和强大的R语言进行对接，通过直接调用R语言的各种统计模块，直接实现对最新统计方法的调用。

2. SPSS操作入门

- SPSS的操作方式

- 菜单对话框方式

- 最常用，最好记，重用性差，效率较低。

- 命令行/程序方式

- 适合于批量数据分析。
 - 可以利用对话框中的“粘贴”按钮自动生成程序。
 - 通过宏、Include命令等方法使得已有代码得到充分利用。
 - 程序全自动执行的“生产工作（Production Facility）”方式。

3. SPSS的窗口和结果输出

- SPSS的4种窗口
 - a. 数据窗口（最常用）
 - 数据表格
 - 标准的电子表格
 - 数据视图：数据界面，是主要操作界面
 - 变量视图：用于定义变量
 - 单元格编辑器
 - 显示精确数值
 - 不能直接输入公式
 - 菜单，快捷工具栏，状态栏

3. SPSS的窗口和结果输出

- SPSS的4种窗口
 - b.输出窗口（最常用）
 - 用于存放分析结果
 - 结果输出美观
 - 提供方便的类似资源管理器的界面
 - c.语法窗口
 - 编写SPSS程序时使用
 - 功能集中体现在菜单“**RUN**”中
 - 适用于高级数据分析人员
 - d.脚本窗口
 - 用Sax Basic语言编写的程序，可以构建和运行SPSS命令。

3. SPSS的窗口和结果输出

- SPSS的4种结果输出
 - 枢轴表/轻量表
 - 增强的多维立体表格，拥有动态展示功能
 - 文本格式
 - 标准的RTF文本，结果中较少出现
 - 标准图与交互图
 - 具有动态展示功能
 - 模型
 - 非参数检验、树模型、最邻近元素分析等的结果用模型展示。

3. SPSS的窗口和结果输出

- 分析结果的保存和导出
 - 直接保存
 - 直接保存为 “. spv” 格式的文件。
 - 导出
 - 将分析结果导出到Word格式、Excel格式、Text格式、HTML格式的文件中。
 - 直接复制粘贴
 - 选择输出窗口中的若干个对象直接复制粘贴到目标文件中。

4. 数据分析方法论概述

- 指导数据分析工作的3种方法论体系

- 严格设计支持下的统计方法论

- 经典统计方法论，研究者在整个研究体系中可以掌控一切。
 - 7大步骤：试验设计、数据收集、数据获取、数据准备、数据分析、结果报告和模型发布。

- 半试验研究支持下的统计方法论

- 经典统计方法论的现实情况。
 - 7大步骤可能不被严格遵循，试验设计无法理想化、数据收集无法理想化。

- 偏智能化、自动化分析的数据挖掘应用方法论

- 计算机技术飞速发展。
 - 针对海量数据发现有意义的行为和规则、找到关联与趋势。
 - 不是线性的7大步骤，而是周而复始的循环。

5. 统计分析软件的作用

- 统计分析软件是数据分析的主要工具。
- 在统计软件中实现：
 - 数据获取
 - 数据准备
 - 数据分析
- 统计学为数据分析过程提供一套完整的科学的方法论。
统计分析软件为数据分析提供了实现手段。

练习

- 浏览网站 (IBM SPSS):
<https://www.ibm.com/analytics/cn/zh/technology/spss/>
- 浏览“经管之家”的“SPSS论坛”
– <http://bbs.pinggu.org>
- 在自己的电脑上安装SPSS
- 准备教材《SPSS统计分析基础教程》(第3版), 现在提供的是第2版教材的电子档。
- 对照教材熟悉软件。

THE END