

知 识 传 承 与 创 新 文 库

外语教学与研究 量化分析 自我管理

任正来◎编著

Management on the
quantity analysis in teaching
and research of foreign language

光明日报出版社

知 识 传 承 与 创 新 文 库

外语教学与研究 量化分析 自我管理

任正来◎编著

Management on the

quantity analysis in teaching

光明日报出版社

图书在版编目 (CIP) 数据

外语教学与研究量化分析自我管理 / 任正来编著.

—北京: 光明日报出版社, 2015. 9

ISBN 978-7-5112-9301-5

I. ①外… II. ①任… III. ①外语教学—教学研究

IV. ①H09

中国版本图书馆 CIP 数据核字 (2015) 第 232975 号

外语教学与研究量化分析自我管理

编 著: 任正来

责任编辑: 曹美娜 朱 然

责任校对: 张明明

封面设计: 中联学林

责任印制: 曹 净

出版发行: 光明日报出版社

地 址: 北京市东城区珠市口东大街 5 号, 100062

电 话: 010-67078251 (咨询), 67078870 (发行), 67019571 (邮购)

传 真: 010-67078227, 67078255

网 址: <http://book.gmw.cn>

E - mail: gmcbs@gmw.cn caomeina@gmw.cn

法律顾问: 北京德恒律师事务所龚柳方律师

印 刷: 北京天正元印务有限公司

装 订: 北京天正元印务有限公司

本书如有破损、缺页、装订错误, 请与本社联系调换

开 本: 710 × 1000 1/16

字 数: 322 千字

印 张: 18.5

版 次: 2016 年 1 月第 1 版

印 次: 2016 年 1 月第 1 次印刷

书 号: ISBN 978-7-5112-9301-5

定 价: 55.00 元

版权所有 翻印必究

前 言^{*}

外语研究教学统计是应用,是把数学用于实际,即将统计用于学生的成绩、研究者的调查得出的数据,实用性很强,需要对数据有着较为准确的理解,同时要有熟练的操作能力。这两者哪一条都不能少。

需要说明的是,外语教学研究统计也要有外语教学和研究的实践或经验。没有教学经验,没有研究实践,你就无法对统计结果进行专业解释,统计结果就会毫无意义。

学习了统计,就会统计操作,这是两个概念。要理解统计中的重要概念,不识那个字,不懂那家书,但是要会操作,还要学习一番。中国通史的作者有言“板凳要做十年冷,文章不著一字虚”,这种学术精神值得晚辈永远学习。向前辈学习,是要下功夫的,同行学习,向国外同行学习,更是此,所以笔者十几年来除了正常的教学之外,大部分时间都是研究统计如何外语教学研究结合在一起,以至于患有失眠高血压。身体虚了,但是并不后悔,以十年的所得与同行交流,是令人欣慰的,不为什么,就是对学术的一份追求。

经过多年的研究,在外语教学与研究中,使用的统计过程并不多,现总结如下。一是整理数据。要认识数据的性质,才能使用统计过程。如称名数据,只能做频次的统计,顺序数据,连续数据才能使用平均数及标准差。数据类型识别错误,进行的统计毫无意义。二描述数据。数据的描述,要看是什么样的数据,如果是连续性的数据,就要看连续数据的集中量数和离散量数,还有百分位。就这些基本的统计参数,也需要很好的理解,才能使用正确。统计过程很可能被误用,因为误用后,不懂统计的人,是没法识别统计的错误。三是比较数据。学生的成绩

* 本课题已获得山西省 2004 年省级先进鉴定。

可以用来进行比较教学法的效果,如一组变量与总体变量的比较,观察自己教的班级与其他班级比,有没有统计学上的显著性差异,不仅仅比一个成绩大小,而是这种差异是不是在95%上为显著差异。如果不是,尽管有差异,但差异不显著。四是试验检验。要比较一种教学法的效果,要通过试验的方法,当然也可以通过历史的方法,进行比较。试验检验,要用到多组变量,要通过人工计算,相当麻烦,使用计算机上统计检验,简便多了,效果也有效多了。五是相关分析。相关分析有些用处,外语教学的教学效果,永远是教学研究的目标,永远是统计分析中的因变量。自变量就多了,我们做外语教学研究,就是探查寻找这种潜在自变量。在相关分析中研究最具相关的变量,是直线相关,还是曲线相关,可以做进一步的深入或因果研究的初步。六是回归分析。也有些用处,其中的预测没有用处。七是因子分析和可靠性分析,在我们分析试卷或试题的信度和效度特别有用,但掌握好不太容易。

不是全部统计对于外语教学都有用,要有经验,有取舍。我们当然要理解统计,但不是要你精通,会使用即可。

在桂诗春教授的倡导之下,及宁春岩教授的初步实践中,韩宝成教授将数学统计引入外语教学,华中理工大学的秦晓晴教授将计算机统计加入定量的分析,这就是清华大学统计学教授贾俊平和外语教学研究统计学的开拓者韩宝成教授所言之“会用”,计算机比人做得更好。

使用说明

本书与其他书不同,他有以下几个特点:

直观性 在过去,阅读有图的书籍,要一边看图,一边对照该图,现在与过去阅读此类书籍不同的是,阅读本书,图与书合一,图就在手下,书也在手下,这就是本书的直观性。

易理解 本书对统计的处理,易难为易,一边看图,一边理解,理解高效,而不是非常吃力。

易使用 本书将术语,浅化处理,一般使用者,不需要过深的统计知识,小学、初中和中学老师,直观易解,上手快捷,对照使用,能做出漂亮的表格和图形。

下面是对本书的阅读说明

笔者从 spss12 一直关注和学习使用,现在到了 spss21 了,spss 已变成 IBM-spss,但是基本统计技术没有多大变化。所以,我们学习的是统计基本技术,特别是教师写论文,使用 spss,又懂得教育,就能撰写出优秀的论文,有理论,有经验,有数据。

1. 本书是为学习外语的学生和老师来编写的。只要懂得基本的统计技术,及一些统计基本术语,就可以使用。
2. 图形的标识在相应的章节下标有 A1,A2.....
3. 表格的标识在相应的章节下标有 B,B2.....
4. 统计量的标识符号#1,#2.....
5. 统计图 C1,C2.....

大数据对于统计学有重大的影响,但是缺乏对这方面的研究。我们提两个问题:如果技术发达,那么还需不需要抽样,这样样本技术就彻底颠覆。大数据技术能不能替代样本研究?需要统计学家深入的研究,找出答案。

目 录

CONTENTS

第一章 外语教学研究与统计	1
第一节 外语教学研究与 SPSS	1
第二节 外语教学研究的数据类型与统计方法	4
第二章 数据的输入与整理	6
第一节 数据的输入与输出	6
第二节 数据的整理	10
第三节 数据重构	11
第四节 分类汇总	23
第五节 加权记录	26
第六节 分类整理文件	31
第三章 数据的转换	33
第一节 运算	33
第二节 重新编码	42
第三节 记录排秩(Rank Cases)	46
第四节 替代缺失值	48
第五节 数据转换	50
第四章 变量描述	59
第一节 汇总	59
第二节 描述性统计	84
第三节 频率性统计	92
第四节 表格式描述统计	97
第五节 报告统计过程的区别	130
第六节 简述统计图形	134

第五章 数据结构的探索与统计力	139
第一节 数据结构分析	139
第二节 统计力和效应大小	147
第六章 小变量	149
第一节 小变量分析综述	149
第二节 卡方检验	150
第三节 二项分布检验	154
第四节 单样本 K-S 检验	156
第五节 两个独立样本检验	160
第六节 K 个独立样本检验	165
第七节 两个相关样本检验	167
第七章 双变量比较	170
第一节 比较综述	170
第二节 单样本与总体比较	174
第三节 测前测后成对样本比较	175
第四节 因变量的分组比较	178
第五节 一元方差分析	184
第八章 多因素变量比较	194
第一节 多样本比较	194
第九章 多因素协变量分析	201
第十章 相关分析	224
第十一章 线性相关	246
第十二章 外语教学调查分析	256
第十三章 试卷分析	261
第十四章 试题分析	273
参考文献	283
后 记	285

第一章

外语教学研究与统计

第一节 外语教学研究与 SPSS

一、外语教学研究特点

外语教学研究的类型

外语教学研究类型从成果的形式上看可分为三种。经验性研究、理论性研究和统计性研究。

经验性研究,指研究者只具有初步的研究能力,没有能够掌握该研究的主要基本理论,只具有一定的实践经验,而进行的研究。这是一种不成熟的研究,是初步的研究者必经的阶段。

理论研究是一种概念的逻辑推演,概念的定义,概念间的关系,以一定的前提进行逻辑的推理,从而得出研究结果。理论性研究也是一种间接性研究。

根据研究对象和资料信息的来源的方式的划分,可分为间接性研究和原始性研究。

间接性研究的资料来源于书刊杂志。学生的课程论文属于此类。这种类型的研究的优点在于能了解相关领域的研究状况,并能获得有价值的研究成果。

原始性研究的资料主要来源于被研究的本身。可分为个案研究或定性研究和统计研究或定量研究。个案研究以某个学生或某几个学生为研究对象,在相对稳定的一段时间研究这个或这些学生的语言发展状况。这是纵向研究。统计研究是指在某一时期收集成组数据,进行横向研究。这类研究通常采用调查方法或实验的方法。调查主要是用问卷的方式,调查一组受试者的态度、意见或特征。

统计研究是在受控的条件下,调查成组的语言行为,估计它的可能性——即概率。统计研究要求研究要有系统性、逻辑性、可视性、可复制性和归纳性,具有这五个特征的研究结果才有较高的可信度和可靠性。

外语教学研究与各种教学法的研究

外语教学研究与语言的结合

外语教学研究与学习者心理个性的结合

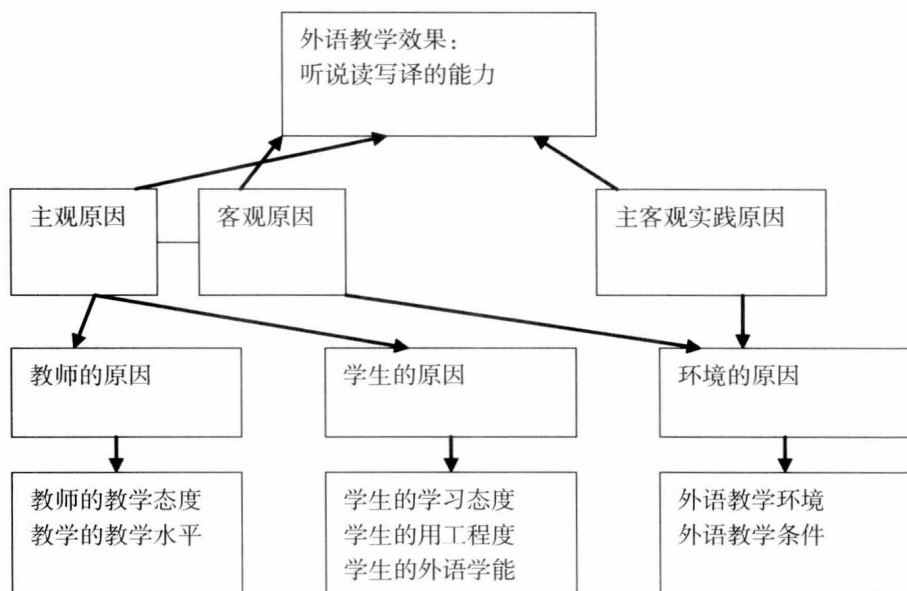
最早的外语教学研究法仍是语法翻译法,是西方人研究拉丁语文学的外语教学方法。这种外语教学研究法的特点是进行两种语言对比,找出两种语言的同异之处,所以有词汇和语法的对比,从而能够将外语准确地翻译为母语。尽管如此,语法翻译法仍旧被认为是没有理论的外语教学法,因为它还没有与心理学、语言学、教学理论学科的相关研究。这种教学法在欧洲存在约一百年(1840s—1940s)由于早期的外语教学的目的是学习文学,所以语法翻译是文学上的研究,而不是外语教学和应用语言学的研究。随着社会的发展、交往的需要,对外语的要求也不一样,不仅是阅读书面语言,而且要求能够听说,因此其它的教学法也随之提出。十九世纪中叶,翻译法受到质疑,同时由于欧洲人之间交往的要求,对于口语提出了要求。当时的研究还是私人性质的,但是已经与教学相结合,观察儿童的外语学习,提出了与上下文和情景结合在一起的教学法。这种教学法后来发展为情景教学法。这种教学法的提出是基于对儿童学习外语的观察。所以情景教学法的研究可以说是基于经验的教育理论。与此同时,外语教学也恰恰与语言的研究偶然结合在一起。十九世纪八十年代,产生外语教学改革运动,外语教学研究与语言学相结合,使得外语研究进入科学方法时代。如对于语音的科学分析和描写。1886年成立国际语音协会。国际语音字母的设计,能够精确为任何语言注音。与此同时外语教学研究争论最佳外语教学法。也就此时,外语教学将语言的科学分析和心理学的研究结合在一起了,外语教学的研究受到应用语言学的指导。随后提出的直接法,是用心理学的方法与比照的思维,将外语学习与母语学习等同起来。因此它缺乏应用语言学理论的支承。二十世纪二十年代,直接法式微,做了重大修正,与语法翻译结合在一起。与此同时外语教学的研究得出了重大的理论结论——没有任何单独教学法能保证外语教学的成功。

教学法的提出并不能解决外语学习的效果问题。哪一种教学法能够产生最大的教学效果,研究的结果是并没有一种百验百灵的外语教学法,于是在外语学习者本身寻找效果的方法。这便是以学生为中心的教学方法。

以学生为中心的外语教学法,是一种新兴起的外语教学法。它是以学生如何

学习外语来研究外语教学,与心理学结合较为紧密。

外语研究的最终归宿,是寻求提高外语教学效果的方法和途径。我们知道外语教学的效果与很多外语教学有关。外语教学效果是外语研究最大的因变量,而其它因素则是影响外语教学效果的自变量。下面我们画出一张外语教学效果与影响外语教学效果的各种因素的关系图,要说明的是它只是一种简略的关系图,要确定较为精确的关系,需要做深入和严谨的研究。



二、中国外语教学研究方法的回顾

外语教学研究与语法的结合

外语教学研究与文学的结合

外语教学研究与语言的结合

中国外语教学研究与语法有着不解之缘,同时与文学不可分离,是注重翻译的研究。与语言学的结合,是20世纪九十年代的事情,即运用语言学的方法来研究外语教学。

三、中国外语教学研究如何与概率结合在一起

中国外语的研究,在学习西方的语言学的同时,也学习使用语言学来研究外语教学,同时也用统计方法来做研究,特别是有了计算机及其统计软件,尤其要提

出的是社会科学统计软件 SPSS。

四、外语教学研究在 SPSS 的实现研究

要做外语教学研究,可通过翻译的研究、语言学的研究、以及统计研究,结合外语教学实践,来进行中国的外语教学研究。在使用统计研究外语教学研究时,我们具有外语教学学生的成绩和外语教学研究的调查数据,我们要使用的是计算机程序来处理,而不必再做具体的计算。但是把外语教学与调查的数据用统计软件来实现,必须要经过一个学习的过程。SPSS 软件在我国有着普及性的影响,所以我们选择用它进行外语教学研究的统计研究。SPSS 软件可分为两个部分,一是它统计过程,二是它的数据挖掘技术。我们使用它的统计过程以外,还要对外语的海量数据做挖掘,发现外语教学海量数据后的有用价值。在我国医用 SPSS 的统计研究做得非常好,但是外语教学研究用的 SPSS 统计研究却不那么有大进步(所谓的外语教学研究用 SPSS,主要是针对外语教学研究工作者的背景,使用外语教学研究的数据来学习统计方法。)。原因主要是,一是外语教学研究工作者在知识结构上的缺陷,即对于数学有心理障碍—数学恐惧症。二是外语教学研究工作者的语言思维与逻辑思维是不同的,外语教学研究工作者要掌握逻辑研究手段,学会理工科思维来研究外语教学。所以用统计来研究外语教学是外语教学研究工作者急需需要补的一课。

第二节 外语教学研究的数据类型与统计方法

一、数据类型与外语数据的类型

在统计中有四种类型的数据。一是称名数据、二是定序数据、三是间隔数据、四是比率数据。外语教学中也有分析数据如男女性别的二分类,有小学、中学、大学、研究性学院四分类数据,有一、二、三、四年级的国分类数据。分类正如其名,是表示名称的,不论是什么样的名字,有多少类别,当数据量化时,数据是不能相加或相乘的,如 1 表示女性,2 表示男,1+2 或女+男不表示任何意义。定序数据如外语学生的学习成绩排名,从 1—35 名,就是一种定序数据。定序数据可用来相加,做统计处理。间隔数据在外语教学中就是学生的外语成绩。如 35,56,77,87 等。间隔数据可用于相加、相乘,进行统计处理。比率数据,是指一个数字是另

一个的多少倍。因为比率数据有零值,能表示倍数,所以比率数据不用于行为科学研究。因此在外语教学研究中不使用比率数据。

在外语教学研究中有三类数据,即称名数据,定序数据,间隔数据。前两类称为离散数据,后一类为连续性数据。往往是三类数据在外语研究中一起使用,这时要用到分层或不同水平的概念。但是对于不同类型的数据要用不同的统计处理方法。

二、外语数据类型与统计方法

对于称名数据,我们可以统计它的频率,如有多少男性,或有多少女性,男女在整体中占多少比例。

对于定序数据,我们不仅可以统计它的频率,还可以通过秩和检验进行差异比较等。

对于间隔数据,可以使用不同多种统计方法,如相关检验、因果关系、多元回归等。

特别要提的是量表数据,可以看作是定序数据,也可以视作是间隔数据,可以相加也可以相乘,但是它要求在一定的条件,也要求在量表调查时就要考虑量表的设计。

另外三个类型的数据可以互相转换,但是这种转换要求在一定的条件之下。分类数据可变为连续性数据,连续数据也可变 8 为分类数据。

第二章

数据的输入与整理

第一节 数据的输入与输出

一、数据的输入

数据的输入是如何将外语教学的数据输入 SPSS 的窗口中。它的横格是什么,表示什么;纵格是什么,表示什么。

数据的整理是数据输入之后,根据统计要求对数据做预先处理,以便为后续的统计做数据上准备。

外语教学研究有三大类型的数据。

1. 一类数据 称名数据
2. 二类数据 定序数据
3. 三类数据 连续数据

这是从学理上讲,从实际用途上看,也有 3 类数据。

1. 学生的成绩,其中可以包含称名数据和定序数据,三类性质的数据。
2. 多响应调查表的数据,是一种定序数据。
3. 排名数据,可视作是定序数据。

从量上看,有两类数据。

1. 大量数据,如学生的成绩;2 小量数据,如小班的成绩,或临时性的成绩。

输入数据要有两大窗:数据窗口和变量窗口

(一)数据窗

在 SPSS 的数据输入窗,横向表示记录,纵向表示变量。具体讲,横向表示

学生的名字,某学生有听说读写译各个技能的成绩。纵向上看,他的各个技能成绩是多少。如果有一个班成绩 35 名,这 35 个学生是 35 个横向记录,纵向是各个技能的成绩。纵横组成一个表。有了这个表就可以进行做统计处理。

分类数据同样的方法,即横向是学生的名字和序号,纵向是性别,如男或女。

在数据窗中,数据横向是受试者的记录,纵向是各项的成绩或调查值,也是变量。

我们在输入学生成绩时,就不会出错。如果反向输入,可以用置换的方法来转换。但有时置换会丢失某些数据,所以最好是一开始数据就输入正确。

横为记录

纵为数据
也是变量

	name	sex	class1	class2	class3	Val1	Val2
1	李玉清	1.00	88.00	89.00	56.00		
2	刘明	2.00	67.00	66.00	76.00		
3	任文群	2.00	86.00	86.00	74.00		
4	柴泽民	1.00	88.00	78.00	68.00		
5	赖可欣	2.00	89.00	89.00	75.00		
6	王永民	1.00	87.00	87.00	65.00		
7	潘江清	1.00	76.00	88.00	69.00		
8	林玛杨	2.00	67.00	77.00	63.00		
9	闫和明	2.00	56.00	87.00	77.00		
10	石磊	1.00	68.00	67.00	86.00		

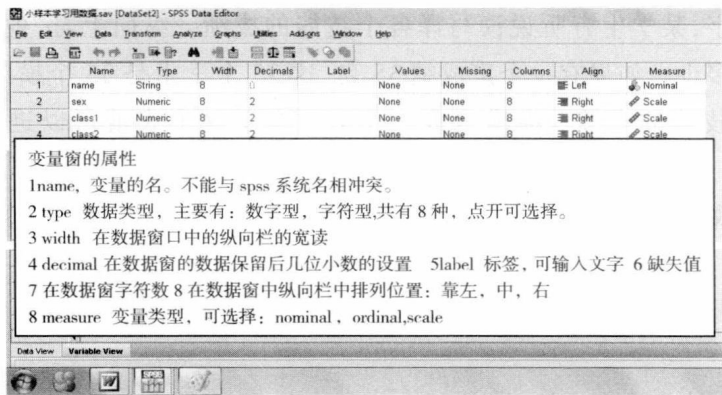
说明:

在数据窗中草,横格排列 5 种数据:NAME、SEX、CLASS1、CLASS2、CLASS3,这 5 个数据组成一个记录。

如果是纵向看,这个 5 种数据分别组成 5 组变量。

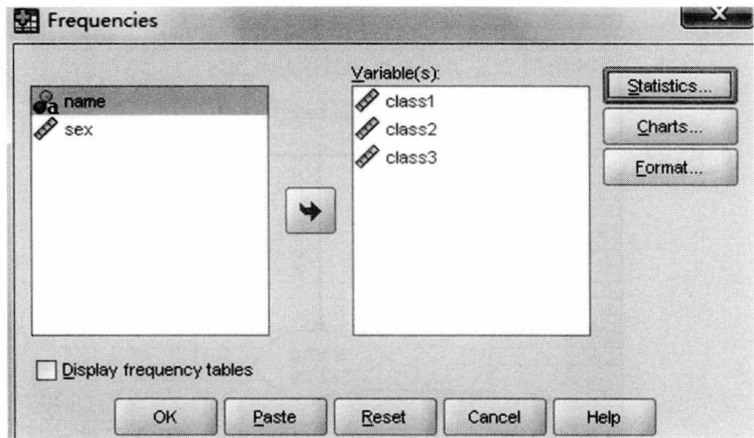
(二) 变量窗

点击“variable view”变量窗,显示如下:



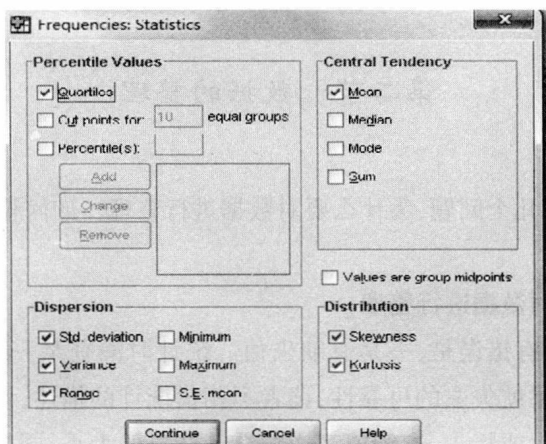
(三)简单的统计过程

1. 输入变量于变量窗



2. 设置各种统计

设置种种统计,就是要根据需要进行设置。这种设置就是象计算前的准备,OK 后就象等于号“=”,一点结果就出来了。SPSS 的好处就是我们不用计算。但是必须设置好。要设置好,就必须懂得的的各种设置。这是困难的所在,要花大量的时间学习设置。第二大难点就是把统计结果与设置结合在一起,即把统计与自己的专业经验相结合,对统计结果做出合理的解释。



点击“CONTINUE”，回到 31111 变量输入窗，再点击“OK”，看如下的输出窗。

(4) 结果输出窗(统计结果显示)

前面是数据的输入,这里严格说不是数据的输出,听起来好象对称,实际上是统计结果的显示。为了方便学习,权且表示为输出窗。

我们看懂了“输入窗”的“数据窗”和“变量窗”。即把数据输入后,要进行计算,算出的结果显示为输出窗。

3. 输出窗

```

DATASET ACTIVATE DataSet1.
FREQUENCIES VARIABLES=V5 V6 V7 V8 V9
  /FORMAT=NOTABLE
  /STATISTICS=STDDEV VARIANCE RANGE MEAN MEDIAN MODE SUM SKEWNESS SE
  /ORDER=ANALYSIS.
    
```

Frequencies

[DataSet1] E:\外语教学研究用SPSS\外语教学与研究量化分析管理\数据库\080030141+14

		V5	V6	V7	V8	V9
N	Valid	47	47	47	47	47
	Missing	0	0	0	0	0
	Mean	16.66	3.96	21.11	4.55	10.30
	Median	16.00	4.00	20.00	4.50	10.00
	Mode	16	4	18*	6	10
	Std. Deviation	3.772	2.519	4.061	1.396	1.488

结果就显示在输出窗。

4. 解释

我这里只解释主要统计内容。有效记录均为 47,缺失值没有,V5V6V7V8V9 的均值、中值、众数、标准差。

一个简单的统计结果就完成了。