

正确决策

警务专业人员基于证据的决策指南



加拿大警察局
长联盟警务研
究基金会

保罗·马克西姆，任·加里斯，戴雷·普利卡斯，梦娜·戴维斯

译：徐秋荻，刘仁菲，陈韞韬

校对：徐秋荻



正确决策

警务专业人员的基于证据决策指南

保罗·马克西姆， 任·加里斯， 戴雷·普利卡斯， 梦娜·戴维斯

译：徐秋荻，刘仁菲，陈韞韬

校对：徐秋荻

© 2015 保罗·马克西姆
(2015年5月修订版)

作者介绍

保罗·马克西姆 (Paul Maxim) ， 劳里埃大学教授

保罗·马克西姆于渥太华大学获犯罪学硕士学位，之后于宾夕法尼亚大学获社会学博士学位，专攻犯罪学和研究方法。目前，他在位于加拿大安大略省滑铁卢的劳里埃大学 (Wilfrid Laurier University) 巴尔西利国际关系学院经济系担任教授，主要研究领域为人口和劳动经济学。

任·加里斯 (Len Garis) ， 菲沙河谷大学客座教授

任·加里斯是卑诗省素里市消防局局长，同时也被菲沙河谷大学 (University of the Fraser Valley) 犯罪学及刑事司法学院聘任为客座教授，他还是纽约江杰刑事司法学院 (John Jay College of Criminal Justice) 附属研究院成员，以及西蒙菲沙大学 (Simon Fraser University) 加拿大城市研究所成员。他的关注点在于如何通过进行基于证据的决策和创新来应对公共安全领域面临的挑战。

戴雷·普利卡斯 (Darryl Plecas) ， 菲沙河谷大学荣誉教授

戴雷·普利卡斯是菲沙河谷大学犯罪学及刑事司法学院的荣誉教授，在他退休前，曾担任加拿大皇家骑警大学资深研究员。他曾获大学杰出教授荣誉，并于2003年荣获第十四届大学教学国际会议颁发的教学及科研杰出创新人才奖。他发表了大量文章论述加拿大刑事司法领域的相关问题，同时在应用政策、项目评估以及有效决策的研究上也具备丰富的经验。

梦娜·戴维斯 (Mona Davies) ， 法律分析师

梦娜·戴维斯获伦敦大学法学学士学位，随后于多伦多大学获政治科学硕士学位，她还获得特许秘书及行政人员研究所 (Institute of Chartered Secretaries and Administrators) 颁发的专业证书。她在国际法领域具有超过10年的工作经验，工作单位包括私营和公立机构。她在投资者-国家仲裁方面颇具专长。

目录

- 2 前言
- 3 简介
- 15 确定问题
- 27 批判性思考
- 37 搜集证据
- 47 统计学
- 57 试验设计
- 67 成本分析
- 83 进行决策

如果需要本书的其他信息，请与云南警官学院办公室外事科徐秋荻女士联系：

电话：+86 871 6519 7410

电子邮件：rogella1031@aliyun.com

本书英文版及菲沙河谷大学其他研究成果请参看大学网站：

www.cjr.ufv.ca



加拿大警察局长联盟警务研究基金会

本项目是受加拿大警察局长联盟警务研究基金会（Canadian Association of Chiefs of Police Research Foundation）委托，该基金会一直致力于通过研究，创新和发展最具成效的标准化执法。同时，该项目的合作伙伴还包括卑诗省素里市政府以及卑诗省菲沙河谷大学。

本手册是改写自《正确决策：消防专业人员的基于证据决策指南》，该书获加拿大安全及保卫项目支持，由加拿大安全科学中心（Canada's Centre for Security Science）防卫研究及发展部牵头，与加拿大公共安全部（Public Safety Canada）协作完成。

前言

加拿大的警务工作是捍卫公共秩序和安全的关键，但是在此项工作上投入的资金和人力成本却可能使公共财政面临巨大挑战。警务专家们进行关键决策的依据主要是警方能够提供给社区的服务，以及需要对社区提出的要求。加拿大警方一直致力于使警务工作与时俱进并且增进成效，他们已经开始采用一些创新型的策略，目的是：（1）重新思考那些工作是应当放在首要位置的；（2）带来更好的成效；以及（3）更好地控制工作成本¹。在大多数城市，警务工作都是一项主要的预算开支项目。

虽然公众和政府官员都了解警方所担负的重要角色，但他们还是要求重大决策必须是建立于确凿证据之上。因此，类似于“警方的重点策略价值何在”以及“相关成本和收益分别是怎样”这样的问题不绝于耳。

本手册就是为了帮助决策者们回答这些疑问而编写的。基于证据的决策（Evidence-based decision making）是最为有效的工具之一，你可以利用它合理地解释为什么要选择执行某个特定策略或项目。不过，基于证据的决策并不是什么新鲜事物，它是一个能够将策略性计划和社会及经济成本分析相结合的、公开透明的决策模型框架。



本手册对基于证据的决策中的一些重要组成部分进行了概述。其中一些材料一开始可能令人觉得望而怯步，不过，我们建议你将其分成几部分慢慢研读，然后回过头再温故几次。这种类型的材料通常都需要和某些你所在机构的特殊问题或议题相联系才能彰显出其意义所在。总而言之，我们希望本手册能够对你的决策过程以及验证说明有所帮助。

谨此，我们也对以下人员所提供的帮助致以谢意：萨尼治市警察局局长罗伯特·唐尼（Robert Downie），渥太华警察局伊安·马特雅斯（Ian Matyas）警官，素里市消防局策略计划分析师亚历克斯·雅科夫（Alex Tyakoff），以及美国纽约江杰刑事司法学院雷根哈德应急研究中心主任查尔斯·杰宁斯（Charles Jennings）博士。

1. 《警方绩效标准研究简报》第31期 <http://www.publicsafety.gc.ca/cnt/rsrscs/pblctns/plc-prfrmnc-mtrcs/plc-prfrmnc-mtrcs-eng.pdf>

简介

在不断变化的世界中进行有效决策

虽然警务部门的首要职责一直以来都没有改变——即保护人民的生命和财产安全——但是，警方该如何履行自己的职责这一问题变得日益复杂。在过去的20年间，犯罪率在整体上有所下降，但是公众要求警方进行响应的方面却越来越多。要处理这些报警电话通常需要更为先进的设备和接受过相应专业训练的人才。而且，很多辖区还要求他们的警察部门将自己的职能与其他紧急情况响应部门相结合，包括消防和医疗急救服务等。

因此，领导们和管理人员们不断遇到下面这个问题：在社会对警方的需求变得越来越复杂，而资金支持又颇受限制的情况下，我们到底该如何提供高质量的服务呢？这时候就需要做出选择和权衡，并且考虑可能带来的后果。由于所做决策会受到政治家、政府人员以及最终由社会公众所进行的严格检验，决策者们面临着越来越大的压力。那种可以随心所欲地向政府和纳税人索要更多设备和人员的日子——如果确实有过这样的好日子的话——那么，这些日子现在也已经过去了。政府官员、市级领导以及高层执行官们也不得不面临在紧张的资金预算下进行决策的情形。

更进一步说，警务机构的领导们需要使决策过程更加公开透明、有理有据。在我们看来，一个好的决策需要依靠尽可能多的证据、研究以及方方面面的信息。

那种政府和纳税人对我们增加设备和人员的申请有求必应的日子——如果确实有过这样的好日子的话——那么，现在也已经是过去式了。

我们将这种方式称为“基于证据的决策”。我们通过研究可独立进行自证的多个事实，进行以证据为基础的决策并证明它的合理性，可以帮助我们确定所做出的决策是合理且经得起推敲的。运用有效的、以证据为基础的方法可以帮助寻找到您所觅求的答案。

那么究竟为何基于证据的研究会如此重要？为什么这种方法有助于解决警方所面临的问题？关键原因有以下几点：

- 缺乏合理证据支持的政策和项目通常会造成对成本和资源的浪费，或者效果不佳，或者造成未曾预料的后果；
- 如果所提交的资金请示缺乏有说服力的证据，那么对机构预算有审批权的外部决策者可能会因此认定该需求不合理；
- 基于证据的政策和策略通常会带来更好的效果，因为它能全面提高您所在机构的公信度并进而获得支持。

本手册会帮助您了解怎样寻找必要的信息和研究，并利用它们进行基于证据的决策。同时，它还将帮助你在一个强有力的框架下进行决策，使它们更具有说服力。

当然，并非所有的决策都是可以基于事实来制定的。无论是从专业角度出发，还是从我们的日常生活出发，都需要考虑道德标准、价值标准、选择偏好以及政治选择等。如果不考虑这些，就等于拒绝承认社会生活的复杂性。因此，即使在这些情况下，基于证据的决策也能够帮助您将那些对部门具有指导意义的价值标准、社会原则以及意识形态与独立证据和辅助研究相结合。

我们可以从多种来源获得这些信息，并将在之后的章节中学习如何对其加以利用。有些信息是可以通过政府部门或其他正规组织的常规统计来获取的行政性数据，有些是在执行政策法规和项目评估的过程中获取的，而有些则是通过政府和学术科研工作者的工作来获取的；其他信息来源还包括您所在的机构，而且特别是您所任职的科室或部门。

本手册能帮助你：

1. 找准并运用相关信息和研究来进行基于证据的决策；
2. 在一个强有力的框架下进行决策，使它们具有说服力。

学会在证据的世界中游刃有余

本手册将帮助你在证据的世界中游刃有余，不受桎梏。正如我们将探讨的，并非所有的证据或数据都具有同等价值。即使是优质数据，我们也需要对其准确性和意义进行综合评估。换句话说，您不用成为这方面的科学家或学者，本手册将帮助您了解关于数据资料自己需要知道的究竟是什么。

除了学习怎样评估证据，我们还将讨论如何使用证据来形成具备说服力的材料。数据本身并不足以成为你的决策依据。在就某项政策或项目进行决策的过程中，我们需要具备公开论证。那些缺乏合理依据的策略性或商务计划将很难获得成功。我们将看到，很多摆在桌面上的理由或论证其实根本没有任何意义；此外，我们还会对一些无论如何都该避免的重大逻辑错误进行论述。

本手册还会解释如何进行环境扫描（environmental scan）以及SWOT分析（即对一所机构的优势、劣势、外在机会以及威胁进行评估的方法）。您将了解到为什么应当在制定新政策或新项目之前，甚至是确定策略重点之前就进行信息搜集；您还将了解到成本效益分析以及成本核算的相关知识，这些对于在资金有限的情况下制定战略性计划非常重要。

本书以世界各地的警务机构为案例，将为您展示如何对一个问题进行定性。它将帮助您对问题进行批判性和创造性思考，并且找到必要证据来支持您的决策。除此之外，它还对各种类型的研究进行了简单易懂的说明，方便您了解何时该采用以及该如何采用这些研究来进行决策支持。

不过，在我们开始前，如果您能更加深入地考虑一下几个问题的答案，那么将事半功倍：基于证据的决策究竟有多重要？为什么这么重要？您自己，或者其他任何人，为什么应该重视这样的决策过程呢？

医疗和康复专家所开创的先河

关于基于证据进行决策的起源，可以追溯到上世纪八十年代。当时，在严重的财政危机面前，英国政府开始重视那些得到有力证据和实证研究支持的政策和最佳实践。决策者们认为他们在那些没有证据支撑而进行的选择上浪费了太多的资源，他们常常根据个人喜好、传统做法，或是那些流行一时但其实缺乏支撑的理念来进行决策。这一认识产生了深远的影响，包括日后警务情报（police intelligence）的流行，促使传统警务向情报主导警务（Intelligence-Led Policing，简称ILP）模式转变，并成为世界各国竞相采纳的理念。

任何人只要在其专业领域工作达到一定时长后都知道，这世界上哪里都有江湖骗子，无论是管理还是实践，他们都可以兜售“灵丹妙药”。英国政府很清楚他们需要做出投资，但是他们也需要确保这些投资能够兼具效率和效益¹。



这一措施影响到其他多个领域，特别是医疗健康科学，因为这一领域的研究人员们可以把实践经验的缺乏直接联系到患者受到伤害程度的增加上来。为了弥补存在于学术研究和临床实践之间的差距，基于证据的医疗应运而生。从理论上而言，它可以保障尽可能获得最佳的结果，并为患者们提供最恰到好处的医护服务。研究人员和医疗专业人员们对政策和程序进行严格审查，找寻医疗机构更具效率和效益的运营模式²。

他们需要改变的是，是当时医疗界存在的一些行事规则弊端日益明显。比如说，一项主要研究显示，大约需要15年的时间才能将研究成果和倡导政策相结合；这个例子颇具戏剧性，我们可以假设，这项研究是关于某个已有的、针对某一特别类型癌症的治疗药剂，然而，从发明该药剂到部分投入临床实践居然间隔了约十五年；甚至在投入临床实践后，也仅仅40%的临床医生在应用这一资讯³。

与此同时，就因为这一资讯没能及时传递给医护专业人员，导致那些原本可以从研究成果中获益的病人们，不得不继续受病痛折磨甚至死亡。更糟的是，如果一些组织在“掌控疾病”上看到更大的收益和经济利益，他们甚至有可能刻意推延将研究成果投入临床实践的时间，而不是真正的去治愈疾病。

基于证据的方法尝试通过利用从研究、试验、观察和其他可靠来源所获取的最佳信息，帮助人们尽可能作出最优决策和政策。不过，正如前面的假设案例所展示的，有时候，这种方法会和其他力量、价值观和利益集团发生直接冲突。

案例分析

基于证据的警务是指警务活动中决定“怎样做最有用”的方法，也就是采用什么样的做法或策略能够在最具成本效益的前提下完成任务。与基于理论、假设、传统或社会习俗的决策相比，基于证据的决策方法会持续利用实证研究的结果对问题进行检验。关于警务活动各方各面的研究在上世纪末发展迅速，因此本世纪初这些研究在警务实践中的应用也得到加强。这一点在警务活动“3T”策略中表现尤为明显，“3T”指的是警务活动中的三项任务，即设定目标，验证方法和追踪结果。基于证据的目标设定需要就危害的程度进行系统排序和比较，这种危害程度是与警方能够依法处理的不同地点、时间、人物和场景联系在一起。基于证据的验证有助于确保警方的活动既不会增加犯罪率也不会造成金钱的浪费。在未来几年，对警务人员是否按照长官的指令行事所进行的追踪将会得到快速发展。我们已经看到GPS在追踪警务人员行踪方面的应用以及运用可携带式视频装置来记录与普通民众接触过程中所发生的一切。

二战后的发展

警务活动在二战后发生了巨大的变化。这种变化源于技术的革新以及管理方法的变革。其中一个主要的因素就是运用“硬数据”或证据来评估警方的行为带来的影响。在业界，大家普遍认为基于证据的警务的发展要归功于劳伦斯·谢尔曼⁴。在其一篇有重大影响的评论文章中，谢尔曼对很多变革进行了总结。他写道，最初“带警用车辆的随机巡逻得到推广使用是基于警察无处不在能够震慑犯罪分子的说法。上世纪60年代，

三位数应急号码的使用，比如911，把随机巡逻转变为应急反应等待模式，像机场的‘航线等待’一样，这种做法也是基于威慑理论”⁵。由此就带来了我们现在所说的应对式（reactive）警务和先发式（proactive）警务的划分。

应对式警务指的是警方等待报警电话，随时提供警务支援的行为。正如谢尔曼所说，也就是“警方在接到公民的报警电话后，行动几乎不用受上级警察机构的命令或分析的制约，他们需要做的主要就是尽可能迅速地到达现场，进行处置，然后离开”⁶。

到上世纪70年代中期，应急式警务已经成为组织响应的主要形式。这种形式的典型做法可以用“3R”来概括，也就是随机巡逻（random patrol）、快速反应（rapid response）和应对调查（reactive investigation）⁷。大多数国家的警务处置活动依然采用应急模式，也就是接到报警电话后进行处置。大多数情况下，报案的自然是受害者。不管是偶然发现犯罪行为还是警方在巡逻过程中通过蛛丝马迹判断有罪案发生的情况都不普遍。正如谢尔曼所说，在传统的“3R”模式下，“几乎没有对犯罪类型或社会不稳定因素的预测，也没有人做任何测试来寻找预防犯罪或解决危机的最佳办法，更别提会有警察机构对警察的行为，也就是警察与特定对象在什么时间、什么地点做了什么进行追踪和管理”⁸。

由于当时大多数警务部门面临技术上的局限，因此应对式警务模式在很大程度上意义重大。收集和分析罪案数据是一项需要大量人手的费力工作。直到上世纪的最后十年，计算机的使用成本都很高，很难做到及时地把罪案发生地点和中央数据库直接联网。但这并不代表警方在罪案分析方面未发挥其主动性和积极性。在大多数较大的警察机构中，在插针板上着重标出罪案“高发地点”是很普遍的做法，警官们也经常互相交换情报信息。然而，这里要重点指出的是，大部分的罪案分析通常是由个别警官或是业务小组在处理备受关注的重大案件时手动完成的：定期的对罪案进行分析并没有成为常规工作。随着计算机的大量普及，非专业使用者也能更为容易地操作更复杂的软件程序，这种情况才开始得以改变。

谢尔曼指出“到2012年，‘3R’模式已逐渐转变为我所提到的‘3T’模式，即锁定目标，验证方法，追踪结果。但是标准模式还不会消失，其资源也会越来越依赖于统计数据的指导。随着‘3T’策略的涌现和发展，不论是巡逻部门还是侦查部门的长官们，都已更加趋向于对警务资源的先发式管理⁹。”

根据这一1998年提出的基本框架，正如谢尔曼所说，“3T”策略构成了现代基于证据警务的基础。这一点成为可能是由于大量的资料能够得到及时的、主动的搜集、储存和分析。警务部门不仅关注罪案资料本身，而且还可以把那些资料和罪案发生现场的地理和社会特征联系起来。同样，通过寻找受害者的共同特征¹⁰，也有助于高级受害者心理学获得进一步发展。

现代刑事侦查学使各部门能够运用硬数据来验证犯罪构想或是所怀疑的犯罪模式，而在某些可根据潜在受害者或目标而将警力资源集中或进行重新配置的情况下，就形成了预测性警务。这并不是说传统警务中就没有先发式警务的元素，我们所强调的重点是，现代科技使得预测型警务活动变得更加常规化，而且警官们还可以节约在寻找实际记录或案件档案之间联系上所耗费的精力，把更多时间花在验证猜想和尝试找出隐藏在犯罪模式背后的深层原因上。

锁定目标，验证方法，以及追踪结果

正如谢尔曼所总结的，基于证据的警务需要关注以下三个原则¹¹：

“1. 警方应当进行并运用优质的研究，将稀缺资源有的放矢地配置到那些可预测的犯罪和治安问题集中的区域。

2. 一旦警方选定他们的首要目标，就应当对警务方法进行审视（review）或验证（testing），以确保所选择的方法可以从最大程度上有助于减少危害。

3. 一旦警方利用研究结果来锁定其试验性实践，他们就应当建立并运用内部证据（internal evidence）来追踪（track）这些实践工作每天的运行情况和效果，包括公众对于该警务合法性的看法。”

基于证据的警务不仅有利于犯罪应对和预防工作，同样的数据和分析步骤还能够被应用于评估人才素质和整体管理方法的绩效上。在当下，公众和财政部门对警务工作的监督审查日益严苛，因此，证据和“硬数据（hard data）”可以帮助警方来证明所做出的执法选择和实践，并且获得公众的理解。

“要想理解一事物，首先得认识它的模式”

——以赛亚·伯林（Isaiah Berlin），理论家，哲学家

绘制图示资料，为决策提供依据

使用地理信息系统（GIS）绘制和分析犯罪地图是全世界警察机构的重大变革。它使得犯罪分析师、在职警官、侦查人员、行政管理人和警务部门的领导们能够通过地图对资料进行评估和分析，从而做出更为正确的决策。简单来说，表格数据可以被录入到地理信息系统中，然后以地图的形式立体而直观地展现出来。这些地图能够体现您在社区各类犯罪的分布情况，或是交通流量和建筑环境对警方响应时间的影响；同时，它们还可用于重新划分警方辖区范围和展示社区警署的位置。正如世上很多事那样，没它办不到的，只有您想不到的。

地理信息系统一开始被引入现代警务是为了帮助罪案分析师和犯罪学家们更好地了解犯罪现场的地理特征。



事实上，一个叫约翰·斯诺（John Snow）的内科医生在1854年首先使用最初的地图演示技术阻止了一场在英国伦敦爆发的霍乱疫情。他通过地图演示寻找到这种通过水来传播的疾病的传染源是位于索霍区百老汇街的一处公共水泵后，就移除了水泵的把手，禁止了这个水泵的使用。正是他的这一行为帮助阻止了这次霍乱的蔓延。最初用于疾病防控的地图演示技术随后被应用于警务管理；业内主流观点认为犯罪行为 and 传染性疾病有着惊人的相似性。

警方可以运用地图演示技术来确定某个城市中一小部分罪案频发的特定区域，这样警方就可以集中警力，重点解决这些区域的犯罪问题¹²。在微观层面，这些犯罪热点是显而易见的，比如一个或几个特定的地址，或者处于某条街道两个拐角之间的一块特定区域。警方可以依据犯罪类型把这些热点区域做成图表，对其进行排序、分类，也可以运用和警务活动密切相关的多种方法对其进行分析。实际上，在北美地区，地图演示技术在大多数运行情报引导警务模式的警察机构中处于核心地位。

相较于描述性的或标注犯罪热点的地图演示，分析式地图和预测式警务更具意义。过去，研究人员可以通过被盗车辆的弃置和发现地点绘制地图，从而在数据中把隐藏的嫌疑人找出来。通常，嫌疑人都会把所盗车辆丢弃在离家不远的地方，这在警方的调查中对寻找线索间的关联是很有用的。警方在处理谋杀、非法拘禁等重大刑事案件时，会同时采取地图演示方法和关联性分析来进行电话模式分析。通过把“案发时间”信号塔发出的关于电话打给谁（由谁接听）以及拨出地点的信息制成图表，警方就可以依据三点连线对犯罪情况进行推断。地理信息剖析技术在理论上和操作方法上都有所发展，而且由能手操作的话，可以非常精确的预测到连续作案的嫌疑人会在什么时间和什么地点再次作案。运用地理信息技术从空间和时间的角度来发现犯罪模式，警方就能更加高效地部署人手，节约成本，维护社区安全。

没多少年前，警官们还在墙面地图上插入图钉来标识罪案发生的地点；更具创造性的做法是，使用图钉与图钉之间的连线来揭示犯罪事件或事件时间顺序之间的联系。这与警官们在餐巾纸上随手记录的行为很相似，但我们并不赞同这种做法。幸运的是，随着使用成本低廉但功能强大的计算机和地理信息系统软件（GIS）的出现，大部分警察机构都能更容易地获取和使用地图演示技术来分析、了解和解决出现的问题，从而做出更具说服力的正确决策。

可喜的是，大部分的警察机构，包括一些规模较小，资金也不充裕的警察机构，都可以相对容易地获取地理信息技术的使用方法并能够承担其使用成本，而且可供他们选择的针对此项技术的优秀培训课程比比皆是。此外，要想了解此项技术，您只需要到本地某个大学的地理系或某个理工专科学校去调研一下就行。对于那些不太熟悉地理信息软件的运作方式的人来说，我们建议您可以先到您所在城市的建设规划部门去了解一下情况。这些市政部门都具备必要的专业技术，并且有基础性的地图可以帮助您迅速熟悉地理信息系统。最后，别忘了向犯罪分析师以及计算机专业技术人员咨询在您的部门配置地理信息系统的最佳方式。

有效决策：优秀领导者和管理人员的任务

决策是领导者和管理人员最重要的任务之一。他们的决策好坏不仅影响到整个部门的发展方向，也影响到下属员工的工作士气和切身利益。错误决策不仅会激化矛盾，还会降低员工士气。正确决策实际的积极意义在于提升员工的士气，使整个部门向良性方向发展。

然而，即便是领导者和管理人员意识到以证据为基础的方法的价值所在，在具体实施过程中依然会受到诸多因素的阻碍。比如，一些管理人员会感觉在信息不完整的情况下很难做到快速决策，而另外一些则会使用过时的信息。警务工作的快节奏氛围要求一切行动都要快速。但是，快速行动之前，必须要对最新数据进行冷静而深入地思考，从而指导决策和行动。此外，大多数人在不得不做出决策时总会依赖自身的经验、观察所得或是直觉。作为接受过训练的专业人士，我们个人的经验和对事物的判断通常是有效的，但那也只是决策过程的一部分而已。认知科学认为我们会倾向于看到我们所期望看到的事物。当我们在执法过程中想要有效地克服自身的内在不确定性和由具有多面性的复杂问题所引发的不确定性时，我们的大脑就会“短路”。基于证据的研究有助于我们把个人的经验和观点植入更广泛而更具说服力的信息环境中。除此之外，警务活动获得了长足的发展，新千年的警务工作与上世纪早期的警务工作将不可同日而语。

错误决策不仅会激化矛盾，还会降低士气。正确决策实际的积极意义在于提升员工的士气，使整个部门向良性方向发展。

在制定新的策略或政策前，最好先对您所了解的、您周围的人所了解的以及更广泛的研究领域所告诉您的一切进行一番评估。同样，在您开始制定新的政策或计划后对其在进行评估也是明智之举，这样就能形成您自己的能充分体现有效性的证据。这对本领域的整体发展大有帮助，您所在的机构就可以通过自身的研究告诉其他警察机构什么方法可行，什么方法不可行以及这样做的原因。通常情况下，我们不愿意对一个项目或是一项行动作出评估，因为我们认为那样做可能没用。但这并不是问题，毕竟不管是作为个人也好，社会也好，我们从失败中获取到的经验远比从成功中获取到的要多。

实证研究的性质

当我们使用“证据”这个词语的时候，我们真正要讨论的到底是什么呢？与警方通过犯罪现场侦查所得到的证据不同，这里所说的证据有其特定的含义，指的是对系统数据收集所进行的实证研究的结果，这些数据是来源于正式的评估、实验或其他研究模型。这是一个解决研究中出现的问题的系统方法，这些问题能够产生大量可复现、可观察、可检验、可支持并且很可靠的信息和事实。

您所进行的实证研究分为：

- 定量研究，即对生成的数字和统计数据进行分析。
- 定性研究，即通过受访者的答案形成对其个人偏好、价值取向和看待事物的观点的主观判断。

两种研究方法都能产生有效数据。关键是我们要知道在什么时间和什么地点使用什么样的证据，并且能够判断出这些证据是否足以帮助我们达到目的。

虽然我们有众多的良好来源获取可提供支持的证据，但是学术研究依然有其特殊的优势，那就是研究结果会接受同行的评议。意思是其他独立学者和研究人员对某项研究进行检验，看结果是否可靠，设计过程是否合理。这并不是说这样做就不存在任何瑕疵或任何错误。它能确实实地提升您通过检验来获取对结果的信任的能力。任何一项研究，其研究成果在学术期刊发表之前都必须接受同行的评议。

对于外行人来说，一些学术期刊上发表的内容其专业程度简直高得吓人。幸运的是，很多期刊还是会对学术界比较重要的研究成果进行总结或是使用普通人容易理解的浅显易懂的语言来进行表达。

普遍研究方法

随机对照试验是医学领域研究的黄金法则。研究人员会随机选取一些病人接受各种各样的预防式、治疗式或介入式干预，并随时跟进了解介入效果。还有一种可能的介入方式就是没有任何介入。这使得研究人员可以在两个试验对照组，即控制组和测试组之间进行比较，控制组指的是研究对象未接受任何的介入治疗，而测试组的研究对象则接受了众多的介入治疗。毒品测试大多采用这种方式来进行。在随后的章节中我们将讨论证据收集的不同框架模式的利弊，并且寻求研究人员对随机对照试验如此推崇的原因。

在社会科学领域，由控制组参与的随机测试同样可行。比如，我们可以随机地安排一些人，在其家里安装报警器，这些人就构成了测试组，把测试组和同样是随机选择的家里未安装报警器的人组成的控制组进行比较。这是解决诸如“是否未安装报警器的家庭发生入室盗窃案件的机率高于那些安装了报警器的家庭？”此类研究问题的方法。研究人员会运用这样的实验来“控制”可能使研究结果产生偏差的因素。

此举使研究的有效性大大增加，从而使你对研究的方法和结果抱有更大的信心。研究人员同样很关心他们的研究结果是否具有可靠性——意思就是也就是说，假如我们把同样的研究再重复一次，能否得到相同的结果？假如我们在不同的社区进行此项测试，结果又是否相同？或者，这种方法是否只适用于某个特定的社区？如果是的话，原因是什么？研究需要同时具备有效性和可靠性两个特征，这样才能保证研究的结果是合理的、可信的，而不是出于侥幸或纯属巧合。

做出更好的决策

迄今为止，您大概能够看到翔实可信的研究能够给决策过程带来怎样的好处了。其实这非常简单，在您完成了前期功课的前提下，就可能作出更加有理有据的决策。同时，由于这些步骤都更加透明，并且是基于某些事实而非凭借直觉、碰运气的猜测或个人看法，所以您的决策也就不容易被推翻。

但是，我们必须认识到基于证据的决策最适合用于解决客观问题。而其他类型的决策，正如我们在本章一开始提到的那样，会受到我们个人的偏好、价值观或信仰的影响，与研究结果的联系不大。

尽管如此，当我们想要寻求最有效的方法来解决那些最终会跟我们的价值观联系在一起的问题时，我们就会把两种决策方法融合起来。通过提供高质量的警务服务和尽力减少犯罪来提升我们生活质量的社会价值理念能够带给我们使用不同方法解决问题的动力。基于证据的研究能够帮助我们了解该做什么以及怎样做。

注释及参考文献

1. Tranfield, D., Denyer, D. and Smart, P. (2003) "Towards a methodology for developing evidence-informed management knowledge by means of systematic review," *British Journal of Management*, Vol. 14, No.3, pp. 207-222.
2. Vishwanath V. Baba, Farimah HakemZadeh, (2012) "Toward a theory of evidence based decision making," *Management Decision*, Vol. 50 Issue 5, pp. 832-867.
3. Antman, E.M., Lau, J., Kupelnick, B., Mosteller, F., Chalmers, T.C. (1992) "A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts. Treatments for myocardial infarction." *Journal of the American Medical Association*, Vol. 268, No. 2, pp. 240-248.
4. Sherman, Lawrence W. (1997). "Policing for crime prevention." In L. Sherman, D. Gottfredson, D. MacKenzie, J. Eck, P. Reuter, and S. Bushway (Eds.), *Preventing crime: What works, what doesn't, what's promising*. Washington, DC: National Institute of Justice, U.S. Department of Justice. Also see Sherman, L. (1998). *Evidence-based policing*. Ideas in American Policing Lecture Series. Washington, DC: Police Foundation.
5. Sherman, Lawrence W. (2013) "The Rise of Evidence-Based Policing: Targeting, Testing, and Tracking" *Crime and Justice* 42: 377-451.
6. Sherman (2013), p. 378.
7. Sherman (2013), p. 378.
8. Sherman (2013), p. 378.
9. Sherman (2013), p. 379.
10. Sherman, L. W. (1998), *Evidence-Based Policing*. Washington, DC: Police Foundation.
11. Sherman (2013), p. 382-383.
12. Sherman, Lawrence W., Patrick R. Gartin, and Michael E. Buerger. 1989. "Hot spots of predatory crime: Routine activities and the criminology of place." *Criminology* 27: 27-56
Weisburd, David, Elizabeth R. Groff and Sue-Ming Yang (2012) *The Criminology of Place: Street Segments and our Understanding of the Crime Problem*. New York: Oxford University Press.

定义问题

不是所有决策都相似

我们每天要做出数以百计的决定，从早餐吃什么到究竟在哪个房间召开会议，以及是否要买一辆新车之类。这些决定中，很多都是非正式的，相较而言不是那么重要而且带来的后果也比较单一；而一些决定则会导致一系列的风险或不确定性。我们每个人每天都会面对风险，对我们中的大多数人而言，合理的风险并不能阻止我们着手每天的任务和日常工作。而通过一系列正规程序，我们可以有效应对那些可能会导致我们生活遭受打击并停滞不前的问题。

另一方面，无论是在我们的私人生活还是职业生涯中，我们也会面临重要决策，它们带来的影响非同凡响，有时我们还需要请他人一起加入到决策过程中。我们值得投入时间和精力来仔细研究这些问题，并执行一些列的正规程序。一般说来，正规程序包括对问题进行清晰明确的定义，对各个解决方案进行概述，以及对选择任意方案将带来的成本和效益进行估算。

基于证据的决策能够帮助我们审时度势地做出具备经济、社会或政治意义的决策。基于证据的决策所具备的优势在于使我们能够利用已知结论来估算成效。好消息是，对于任何事物都可以进行测算，无论这个测算有多“含糊”，它终究是可以向您展示出一些此前不曾知晓的点¹。

基于证据的决策能够帮助我们审时度势地做出具备经济、社会或政治意义的决策。

虽然我们永远无从知晓某个决策可能引发的后续影响，不过，通过从经验中总结教训以及依靠可用的证据，我们其实能够对特定的结果进行合理且站得住脚的预测。

在现实生活中，我们都有可能做出一些带来始料未及后果的决策。实际上，做出错误选择和做出失败决策是两码子事。我们将进一步展开说明，错误决策是可以避免的。正确的决策有时也会带来不理想的结果，只不过多数是由于我们很难控制的意外事件所导致的。那么问题来了，如何才能将正确决策和错误决策区分开呢？简单说来，正确决策来自于对问题清晰明确的定义。

在进行某些决策的过程中，我们会尽可能多地参考那些恰当的、可供使用的证据。一个好的决策意味着当您回过头去看时，也可以坚信，在相同情况下，具备相同证据时，您还是会得出一模一样的结论。

如果一个好的决策带来了不尽如人意的结果，那的确令人遗憾，不过，做出正确决策的另一个好处就是我们可以从中汲取教训。如果决策过程是透明的，那么就可以考虑为什么决策会带来负面的影响：是因为我们进行了一些错误的设想？抑或是我们遗漏了某些重要信息？

在本章中，我们将会探讨以下问题：

- 所面临的情况是什么以及我们是否能将其定义为问题？
- 我们如何确定现有的选择和替代方案？
- 我们如何才能进行创造性思考以得出新见解？
- 我们如何才能制订出替代解决方案？

什么才是问题？

一般来说——或者理性的来说，有条理的决策者们会先列出一系列的备选方案。列表不仅是一种好的做法，而且毫无疑问应当在决策过程中具备一席之地。当然啦，正如约翰·洛克菲勒(John D. Rockefeller)在其他场合下曾经说过的，“列表可不是计划”。

在我们着手制订各选项时，我们需要问的是：决策的目的是什么？我们期待达到的目标是什么？我们需要带着这些问题来进行议题分析。这类分析的框架通常是战略规划或商业企划。没有进行先期规划而直接进行决策也是很常见的，正如老话说的，“凡事预则立，不预则废”。如果缺乏一个清晰明确的计划，那么通常我们就无从知晓原来不理想的结果是由于一个糟糕的方法造成的，或者说是由于我们遇到的是新的或不同的情况。一套公开且规范的程序能够帮助我们来积累经验，减少将来我们再犯同样的错误的机率。

通常，缺乏规划的决策并不会得到理想的结果。进行规划有利于我们做出逻辑清晰、条理分明的决策。合理的规划会使得决策过程更加简单，而且公开透明，也就是说，我们能够像挑剔的旁观者们展示出我们在当前情况下所做出的选择是合情合理的。即便最终的结果不符合预期值，合理的规划也有利于我们为所作出的决策辩护，而这一点在当前警方日益强化的责任制度下尤为关键。当我们提出“这个事情的实质是什么”这一问题时，我们实际上需要回答的是“我们所面临的决策是否符合并能推进本机构的职责履行？”

着手任何事情之前，先问句“为什么？”

在看似简单的“到底该选A还是选B”这类问题面前，我们经常会发现自己居然束手无策；当希望推进某件事情时，这通常意味着战略性的一步。比如，一个员工可能会要求当面就绩效和工资进行谈判，并开门见山地提出“您是否准备按照去年的增薪幅度给我加薪？还是说可以根据我为公司做出的贡献给我升职？”

显然，这位员工试图趋使上级做出错误的选择。我们把这样的情况称为“伪二分法” (false dichotomy)，意思是问题表面看来只有A和B两个选项，但实际上可能还存在着多种选择。在对多种可能性进行思考前，需要先对该员工对本机构的贡献做出评估，如果已经有绩效考核标准，那就是最理想的；但是如果没有，您就需要问五个“什么”，分别是“为什么” (Why) 根据你的表现应当获得奖励？你为提高所在部门的工作效率“做了什么” (What)？我们“根据什么” (Where) 可以看出你所做出的贡献？你本年度为所在部门的“什么人” (Who) 提供了帮助或支持？“什么时候” (When) 我们才能看到你工作所带来的回馈？

也许在某些情况下，这些未必是最为恰当的问题，但是您可以先有初步的概念。这些问题旨在将某个要求和本单位或组织根本的目标相关联，从而确保我们所考虑的各个方案是和这些目标保持一致的。通常，我们都试图确保做出的某个决定是有据可循的。如果说我们决定给某些员工嘉奖仅仅是因为他们待人友善、按时上班或者着装得体，那这些理由都是很难禁得住推敲的。

如果决策依据都站不住脚，您就得扪心自问：“我能够向单位里的其他同事、领导或者公众，对我所作出的决策做出合理解释吗？”正如笔者一位同事曾经说过的：“我在进行每一次重大决策的时候，都会假象它会成为明天报纸的头条新闻。如果我觉得这没什么不妥，那么我就有理由相信自己在一个合理的依据上做出了一个合理的选择。”

这一要求通常是我们战略计划的重要组成部分，有时它也会在我们的操作方案或标准操作模式（standard operating procedures，简称SOPs）中得以体现。

举个例子，让我们假设一位中等规模警察局的局长刚刚参加完一个关于人群管控以及大型公共事件或游行示威管控的会议。会议的基本共识是，最近几次类似事件的发生，比如温哥华的斯坦利杯暴乱以及多伦多G20峰会召开期间的骚乱，都意味着执法机构需要做好应对高危情况的准备；而这一话题，由于新闻媒体对这些事件的广泛报道，获得了很高的关注度。虽然从这位局长所在警局的角度出发，大多数这些公共事件都是有计划的和平聚会，给公共秩序带来的负面影响也控制在最小范围内；只有少部分的的活动，比如游行示威，会扰乱公共秩序并且导致警方逮捕部分人员。虽然本警察局的人员都接受了基础的人群管控培训以应对此类事件，但是局长现在认为他的部下们接受培训的频率和级别似乎还不够。

局长所面临的问题是他是否应当创建自己的危机应对小组（Emergency Response Team，简称ERT），还是说就依托临近一所更大规模警察局的此类服务。从表面上看，我们很难驳斥说他们没必要投入更多资源以准备应对类似事件；但是实际上，在这所警察局制定其战略计划时，需要综合平衡考虑本地社区不同的要求。

根据本警局的战略规划，首要目标是保障社区的安全及和谐，并致力于进一步优化自身机构建设和服务职能。因此，警局的真正需求应当是强化机动巡逻支队，或者升级派遣指挥系统。

通过将决策置于预先制定好的计划框架下——比如战略规划——那么做出的选择才站得住脚，因为它是以战略评估为基础的。

通过反思此前的规划方针，我们可以看到，组建一支危机应对小组（ERT），并非本部门的头等大事；而且很有可能的是，在该领域不断投入的资金就是白白打了水漂，或者仅仅是回馈少许成效。

但是换个角度看，组建一支ERT队伍也符合本警局的整体目标，即提升社区安全；因此，这个问题的关键在于，所建议的投资是否符合社区的切实和直接需求。这里的重点并不是能否提高社区的整体安全，而是如何最好地应对本社区最有可能遇到的安全风险。

毫无疑问，这位局长可以将本警局面临的选择一一罗列出来，包括组建ERT以及其他备选方案。不过，进行抉择可不是这里需要考虑的唯一问题，关键是要反思本部门的职能重心，或者最好是能从战略规划的高度进行思考。

再次强调，通过将决策置于预先制定好的计划框架下——比如战略规划——那么做出的选择才站得住脚，因为它是以战略评估为基础的。在本案例中，推迟实施某个为ERT服务的新项目是有理有据的。

获取灵感

通常，面前的可选项是显而易见的。我们是否在器材或人才上投入了过多资源？我们那些通讯器材是不是物尽其用了？不过，在其他一些情形下，选项却并不清晰易见，我们所面临的不是那种“选A或者不选A”的情况。在后面的几章中，我们会研究如何进行环境扫描（Environmental Scans）以及进行SWOT分析，这些都是相对正规的程序，可以对相似情况下其他人是如何做或者可能会做的事情进行系统化的审视。

不过，在求助于那些方法之前，我们可以先运用几个现成的简便方法来制定备选方案。我们建议您考虑采取以下方法：

和您常规圈子外的人士交流

多数时候，我们将自己局限于社交圈或职业圈，仅仅和那些我们已经认识或共事的人交往。通常，这会导致团体共性思维的形成，也就是我们不断强化自己对数量有限的几个选择的执念；更糟的是，同事和下级可能更关心您说的是什么或者只讲那些符合您心意的话，而不是提供独到的见解。而圈外人在面对相似情形的时候，可能会提出全然不同的应对方法。

参与头脑风暴团队讨论

撇开团体共性思维的可能，有时候您周边的人员正是好意的最佳来源，因为他们熟知本机构而且明白问题之所在，再者，由于他们是机构成员，也就无需支付像外聘顾问那样高的酬劳。

所以，您当请他们提供个人建议。有时候，组织一个小组讨论有助于我们督促员工提出“疯狂”的备选方案。

制造一点点竞争性的气氛，可以帮我们得到新的见解。请您记住，今天的“创新”就是昨天的“不可能”。头脑风暴（Brainstorming）即可以是轻松随性的也可以是精心组织的，旨在于收获尽可能多的构思来解决那些长期以来令人烦恼的问题。

多读书籍杂志；以及上网冲浪

您读的书越多，能得到的新思路就越多，尤其是阅读那些在您从事的警务工作领域之外的书籍。商科的书籍毫无疑问是个好选择，不过有时候阅读小说也可以得到很大启发。我们大都愿意进一步充实自己，而阅读专业期刊就是我们紧跟新思潮的好办法。一直以来，互联网就是一锅大杂烩，而且也正应了那句老话“一分钱一分货”，不过，里面也有珍宝亟待您发掘，而最有用的工具就是现代化的搜索引擎。正如史蒂芬·柯维（Stephen Covey）曾说过的，任何时候都别忘了“磨砺锯刃”²。

关注您所服务的人群 —— 注重内部沟通和外在联系

无论是通过内部沟通还是外在联系，请您从所服务人群的角度出发看待事物，要知道，他们对您所在机构的看法很可能与您及您身边同事们的看法大相径庭。除了您所服务的人群，您还可以从合作伙伴那里获取很多好主意，比如专业协会、社区伙伴、教育机构以及其他政府机构等。这些相关机构通常能够提供一些有价值的意见，而了解这些外在人士的看法能够给予您很大启发。

聘任一位有声望的咨询师

一般来说，您在当地就是所从事行业或工作领域的专家，这也是您担任了领导职务的原因；但是另一方面，并不是所有您需要做出的决策都和您工作的核心职责相关联，很多事务涉及外部的设计公司、营销机构、网络设计师或是管理咨询师等。关键在于，您需要确认所必需的专业服务有哪些；一旦定下来，您就应当询问相关人员他们是否能够推荐咨询公司或是咨询师。通常，小公司更具创造性而且收费也更低，话说回来，创造性本就是一桩生意。

当然，您得愿意洗耳恭听新观点，千万别为成见所左右，不能因为您对某些人看法不佳就笃定认为他们提不出什么好意见；同时，也别因为某人能够提出比您更胜一筹的想法就紧张惶恐，特别是如果您的下属提出了建设性意见，更是能提高您的威望，因为大家会觉得这就是强将手下无弱兵。

最后，您得做好准备去接受一个事实，即有时最好的选择恰好就是那些最浅显易见的，因此，如果咨询师给您的报告里全是您一早知晓的事情，也别就此认为他偷懒或是缺乏创意，很可能您觉得显而易见的办法就是最佳方案，您可以把这份报告当作一个佐证来消除疑虑。

制订计划

任何一所机构，无论其规模或复杂程度如何，都能够从制定计划中获益。而不管是将其命名为战略规划、组织规划还是商业企划，重点都是一样的：一所机构需要清楚它为何要做自己正在从事的事情，它的下一步是要走向何方，以及它如何才能到达自己的目的地。

如果没有计划，人们就容易武断地做出决策。往好里说，这些决策缺乏统一性；而往坏里讲，它们则很可能相互矛盾。有计划并不是就意味着机构的成功或效能；但是，缺乏计划，则通常是一所机构将会碌碌无为或是遭遇失败的凶兆。

我们可以通过书店和互联网找到很多材料，主题可能是项目管理实践和原则或是最新提出的主要案例管理，从中学习如何制定出一个组织计划。您在研读这些材料上所投入的时间，将会得到丰厚的回报。

简要说来，一个计划包括以下四个要素：

1. 关于组织价值体系的概述；
2. 关于远景目标和近期目标的陈述；
3. 关于本机构如何实施或达到远景目标的纲要；
4. 关于如何评估成效的标准。

计划在复杂程度上会有所区别，但是总的说来保持简明扼要是很有好处的。复杂的计划通常不容易记忆，而且可能限制太多。正如很多身经百战的将军们所知道的，一旦行动开始，计划就没有变化快，可以说，几乎没有战斗计划是实战开始后无需调整的。通常，将军们所期望的最佳状况，就是军队知道自己是为何而战，而且记得总体目标和近期目标，同时一线的军官们都训练有素，可以应对计划外的战术挑战和挫折。关于如何保持简明扼要，真不是三言二语就能说清的。

计划四要素：



保持简明扼要

下面这个关于福尔摩斯和华生医生的小笑话，可以告诉大家应当保持简明扼要，并且真诚看待那些一览无遗的事实：

夏洛克·福尔摩斯和华生医生外出露营。在享用过一顿美餐和一瓶美酒之后，他们返回帐篷睡觉了。

大概夜里3点，福尔摩斯推了推华生，并问到：“华生啊，抬头看看天空，然后告诉我你看到了什么？”

华生说：“我看到浩瀚星河。”

福尔摩斯接着问道：“那么，你从中领会到什么吗？”

华生答到：“从天文学的角度来看，它们告诉我宇宙中有成百上千个银河系，而且可能有亿万颗星球；而从占星学来说，这意味着土星落入狮子座；而从神学出发，我则感受到了上帝的伟大和我们的渺小；要说到计时学，那我能看出现在是夜里3点；而在气象学上，可以推断明天将是晴朗的一天。那么你呢，福尔摩斯，你从中领会到什么？”

福尔摩斯叱责道：“华生你这个大笨蛋——我们的帐篷被偷了！”

精心归纳陈述的使命、愿景或价值体系可以鼓舞士气；而胡乱编写的陈词只会让人觉得一所机构很肤浅。

申明机构价值体系

最近，在管理学大师间流行的观点是，应当投入大量的时间来确定本机构的基本价值体系。通常，我们会通过以下的某个或几个方面来对机构价值体系进行概述，即：使命声明（mission statement），愿景声明（vision statement），以及价值体系声明（values statement）。精心归纳陈述的使命、愿景或价值体系可以鼓舞士气；而胡乱编写的陈词只会让人觉得一所机构很肤浅。一如既往的，最为实用的建议还是保持简明扼要和直接了当。简单、明确的声明不仅易于记忆，而且还便于遵循。

从根本上说，一份价值声明应当对本机构存在的原因进行概述，这也被称为使命声明。对于很多机构而言，比如警务部门，承担的使命可能显而易见，存在的意义就是保护人民的生命和财产安全，或者更广义的说，是创建一个更安全的社会。因此，使命声明其实就是要您回答这个至关重要的问题：“你存在的目的是什么”？

价值声明还应当一定程度上体现出本机构未来3到5年的发展规划，换句话说，也就是说明本机构的中期愿景是什么。比如说，您可以预计本机构将成为区域性的行业标杆。

最后，一份价值声明还应体现出该机构的核心理念。核心理念需要即是立足之本，又能鼓舞人心。在谷歌（Google）公司，它是“与人为善”（Do No Harm）；而在您的机构，它可能是“服务社区”。虽然这似乎是老生常谈，但是在决策过程的初期，当大家太纠结于究竟什么才是本机构的首要利益时，回头想想核心价值，将会很有帮助。在这种情况下，您并不是要从本机构出发考虑，而是要从服务社区出发考虑。

声明远景目标和近期目标

一所机构对于远景目标和近期目标的声明涵盖了本机构为自己所设定的工作目标。机构的远景目标要更为广泛；而近期目标则是在达成这些远景目标之前需要完成的中期步骤。设定类似于减少社区财产和暴力犯罪数量，或是提高本单位在应对更多警务服务需求方面的能力这样的工作目标，就是远景目标。要达到减少犯罪这样的远景目标，通常需要通过制定一系列的近期目标来一步步实现，比如，近期目标之一可以是进行研究，购买并使用最佳科技装备和基础设施来帮助警官们发现并调查犯罪。

实施步骤

正如我们之前说过的，一份列表可不是一份计划。仅仅对本机构的远景和近期目标进行概述只是策划的一个必要组成部分，但绝不是一个完整的计划过程。

一个真正的计划还需要说明我们如何才能达到这些目标，为了达到预期效果，我们设置了什么机制或者哪些步骤？比如，一个近期目标是减少工伤的发生，那我们就需要将这一目标和提高员工安全系数的总体目标相结合。



拿破仑的远景目标和近期目标

有时候很容易混淆远景目标和近期目标，在很多情况下，两者被混为一谈。虽然它们彼此关联，但是却是截然不同的两个概念。我们可以通过拿破仑·波拿马（Napoleon Bonaparte）在1799年的意图来作为一个绝佳例子。

远景目标

统治全欧洲

近期目标

成为法国领袖
 征服意大利
 征服西班牙
 击败普鲁士军队
 击败奥匈帝国军队
 将波兰纳入法兰西帝国
 征服俄国

令人感叹的是，波拿马就差最后一步就能完成所有近期目标了。因此，虽然他战绩惊人，却终究棋差一招，没能达到远景目标——他没有料到，俄国战斗民族的勇猛和寒冬的凶残。

关于计划的执行，一个重要的议题是：我们怎样做才能达到目标？很明显，我们是根据具体情况来进行机制选择的。在应对违法的精神病患者时，也许我们应当投入更多资源来进行人才培养以做

好更充分的准备；另一方面，人们可能已经得到充分训练，但是却没有足够多的机会来实践所学，因此，还应再引入以增进警务机构和外部社区伙伴之间的信息共享和工作关系为目标的机制。

这一点适用于所有在计划中明确指出的远景目标和近期目标，无论它们是提升员工士气这样的“软”目标，还是减少工伤或警用车辆受损这样的“硬”目标。实施步骤(Implementation procedures)是我们计划中的可供执行的条款。很多时候，战略规划和商业企划体现了某一机构力求取得的成绩，但是却并没有指出它希望通过什么途径来取得这些成绩。

换句话说，如果远景目标和近期目标是一个句子中的名词，那么实施目标就是行为语素或动词。

评估成效

从根本上而言，进行绩效考核需要进行分数记录。不过，在您着手此项工作前，必须先您的方案中提出具体的、将投入使用的绩效评估指标(indicators)。您应当将这些指标和具体的工作目标相结合，并且从宏观上说明这些指标与整体方案勾勒出的蓝图有何关联。正如杰出的管理学大师彼得·德鲁克(Peter Drucker)指出的：“被评估即被管理”。

显然，明晰的定量考核机制要更便于使用，比如说警方接处警数量、响应时间、结案率等方面的变化；但是，我们也不能就忽略了定性评估，比如社区对警方工作的满意度或者对犯罪的担忧度这类指标，虽然不容易进行定量考核，但仍然是评估警方服务质量的关键因素。

通常，绩效评估包含一系列的考核项目，从内部绩效度量值，到提供服务的评级，再到财务职责的考核等等。

很多时候，战略规划和商业企划体现了某一机构力求取得的成绩，但是却并没有指出它希望通过什么途径来取得这些成绩。

针对战略计划进行的论述中，很多人建议创建一张表格，在第一栏里列出本机构的所有工作目标，然后在接下来的一栏里就列出它们所对应的成功标准。这种类型的关联其实带有很强的主观性，因为相比起那些简单直接的工作目标，复杂的目标一般需要更为多样化的绩效指标来进行考核。

而远景目标，顾名思义，实现的时间就要更为长远，所给予的思考也要更为深入，因此通常也就更难确定具体考核标准。进一步说来，和中短期目标相比，对远景目标的绩效进行考核时，更多的需要定性考核。不过，也请您注意，虽然说中短期目标和远景目标的绩效考核标准上应当具备一致性，但是这并不是说两者间必须要完全契合。对我们来说，要完成所有的中短期目标并非难事，但是实现远景目标就不好说了；同样的，如果某个单位没能完成一个或几个中短期目标，也并不意味着它就无法达成自己的远景目标。具备可行性的战略计划或商业企划有时也会涵盖其他情况，或者在某些考核维度上会给出更为详实的细则。

考核细则可能也会体现在本机构对社区的“增值服务”上，或者是说明本机构与同行业或者相似服务提供方有何不同。计划里提及的这些情况是否具有相关性，取决于该机构自身所处的具体环境和所面临的情况。不管怎样，我们在进行关键决策的过程中，都需要将这些情况纳入考虑范围。

不少时候，我们会模棱两可地提出问题或是决定；而一个优秀的决策者，则会清晰明确地说明问题并且指出它对本机构的规划实施有何影响。只有做到这一点，您才能接着提出以下问题：对这个问题需要采取措施吗？如果要，那么我们合适着手解决？这个问题紧急吗？重要吗？还是说无关紧要？



基于证据的决策

要想做出基于证据的正确决策，您需要紧密结合本机构的各类计划。不过，这并不是说我们就不能就那些超出我们计划范畴的事情进行重要决策，毕竟环境不断变化，新问题也层出不穷；但我们还是必需紧扣我们的核心工作职能。这个世界并非一成不变，而这一点在警务领域更是毫无争议的事实。

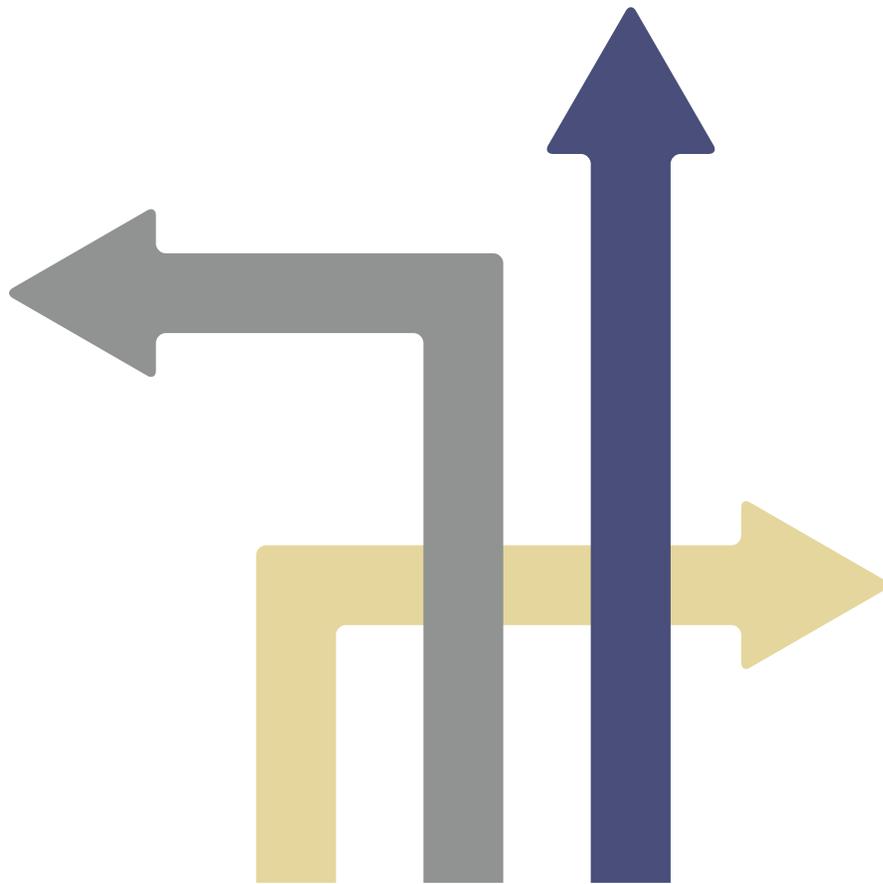
因此，卓有成效的管理者们需要具备足够的灵活性来应对这些情况。但无论怎样，通过制定计划，在面对那些“凭空出现”的挑战时，我们通常会具备更为全面的视角或足够的指导纲领，从而在我们整体规划的大框架内来应对这些挑战。

一个好的计划能够带来的首要益处是，它可以帮助决策者们知道自己该如何评定和选择各个备选方案，并轻而易举地对某个备选方案提出质疑。实际上，在很多决策过程中，备选方案可谓数不胜数；而这时候，我们可以指出某些建议颇有道理，但是却超出了我们战略计划的范畴，从而有效地为备选列表“瘦身”。

所以说，一个好的计划能够帮助我们知晓哪些问题是相关联的，哪些备选方案时值得认真思考的，以及哪些证据是我们在评估这些策略时需要采用的。

注释及参考文献

1. Douglas W. Hubbard (2014) *How to Measure Anything: Finding the Intangibles in Business*. New Jersey, NJ: John Wiley & Sons.
2. Stephen R. Covey (1989). *7 Habits of Highly Effective People*, New Jersey: Simon & Schuster.



一个好的计划能够帮助我们知晓哪些问题是相关联的，哪些备选方案时值得认真思考的，以及哪些证据是我们在评估这些策略时需要采用的。

批判性思考

清醒思考

仅凭证据和数据本身，还不足以作为做出正确有效决策的充分依据。如何进行驳论或解释，和我们如何得到高质量的有效信息一样重要。当我们进行基于证据的决策时，我们需要谨记两件事情。首先，和其他阐明观点的情况一样，你所要进行的论述需要建立在明晰的逻辑基础之上。如果您的论述得出的结论不止一个，那总体上而言，这个论述就不是很有意义。其次，虽然很多人认为证据或证明是一个论述的支柱，但其实这通常并不是阐明论点最有力的方法。

如果我们需要通过两件事来概括人性特点，那么就是，第一，人类喜欢辩驳；第二，即使有人指出他们的观点是错误的或者不符合逻辑的，一般情况下他们也坚决不改变自己的世界观。人类真是一种冥顽不化的生物——即使在一大堆真凭实据面前，他们也总是倾向于去相信自己梦寐以求的谎言。

就自己进行的论述不断列出貌似相关的证据，是人类最为喜爱的社交活动之一。您要是在周六晚上随便去一家运动风格的酒吧，就将见识到什么是乐此不疲的论证。

而且，现在我们有互联网。它的出现，为那些不成熟的观点、阴谋论以及他们的支持者提供了一个绝佳的平台，可谓是在墙壁涂鸦之后最妙的发明了。

人类真是一种冥顽不化的生物——即使在一大堆真凭实据面前，他们也总是倾向于去相信自己梦寐以求的谎言。

从根本上讲，证据和严谨的逻辑都很难撼动人们的已有观点。比如，您大可回忆一下上次听到这句话是什么时候，即有人听完了您的论述说，“感谢您指出我的逻辑错误。我意识到我在这个问题上想得不对，我会根据您所说的改变我对此事的看法。”通常，比较正面的结局是，人们转变了话题；而比较负面的结局则是对方骂骂咧咧转身离去，而且觉得那些跟您意见一致的人也都是傻瓜。

现实情况是，在一些辩论中，无论是从逻辑上还是从经验上讲，都得不出一个终极结论。诸如是否存在神；谁才是长得最好看的女演员或男演员；或者，海伦阿姨做的蛋糕好吃到举世无双——此类的论题永远没法争辩出结论。

总的来说，涉及价值观的事情是建立在个人情感喜好的基础之上的。

另一方面，也存在证据和理性可以改变我们（或者，至少是我们中的某些人）想法的情况。比如，经济问题通常就会促使我们进行理性思考；涉及健康及生死的事情——像是为您的小孩接种麻疹疫苗——也会让人们得出更为理性的结论。

不过，我们还是得承认，即使在这些领域，也存在着一票颇有市场的江湖骗子。

无论是个人还是团体，在一些事宜上都趋向于进行理性的、基于证据的思考，并进行决策过程，而本章节所关注的就是此类问题。由于这些情况仅占人类所面临各类事情中的一小部分，因此我们更应当把握住机会进行有效决策，而不是被错误的逻辑所误导了。

逻辑谬误

逻辑论述一般都遵循以下模式：如果A能得出B，B能得出C，那么出现A则意味着能得出C。逻辑谬误（Logical fallacies）是指，在论述中存在一些错漏百出、前后矛盾或是就是毫无关联的点，但却都被忽视了，或者没有受到质疑。一些逻辑学家和哲学家终身致力于罗列层出不穷、形形色色的谬误（还是请大家参看互联网）。不过，多数的逻辑谬误可以归结一个类别。鉴别下面这些问题，也许能帮助您以后不被人牵着鼻子走——无论是自觉的还是被动的。

迎合权威

我们中没有一个人具备仅靠自己涂涂写写就能掌握各项人类知识的能力。当我们年少时，父母、老师和其他的“专家”是我们获取知识的来源，这点基本上没有异议。能够积累知识并代代相传，是人类和其他动物的区别，这使得我们可以研发出抗生素，建造起摩天大

对用以下句子开头的话要特别留神：“但是，一直以来都是那么做的”，或“专家们认同……”

楼，还有在那种被我们称为手机的小盒子上收发垃圾邮件。如果不是得益于不断接收权威人士传递而来的知识，人类文明可能也不复存在。

不过，虽然我们心甘情愿从身边的“尤达大师！”那里汲取各种智慧，但是也不要盲目相信尤达大师就绝对不会犯错。追寻进一步的证据来论证某个权威的观点，绝对没有任何错误。

虽然我们没有那么多时间来质疑所有权威，不过某些迎合性的论述您需要予以怀疑。

通常，如果出现以下面这些句子开头的話，您就得特別留神了：

- “但是，一直以來都是這麼做的啊。”
- “你懂的。”
- “你（我們）知道什麼？那誰誰誰可是這方面的行家啊。”
- “科學告訴我們……”
- “專家們認同……”

在這些情況下，您絕對可以說，“如果事實如此，那應該会有一些過硬的證據來證明，也許我們應該得到進一步的細節。”或者，“哎呀，有意思的是另外一些人（比如科學家、專家等）說的剛好相反呢。我們怎麼解決這個問題？”

通常情況下，迎合權威的潛台詞是“我懶得去求證”或“我在拍馬屁呢”。

個人觀點或人身攻擊

人身攻擊的英文單詞“Ad Hominem”來源於拉丁文，指針對個人進行攻擊。簡單說來，人身攻擊指某人在對另一個人進行駁斥時，完全從個人角度出發表達意見，通常在論述中質疑對方的心智、道德或是出身。人身攻擊的目的是通過貶損對手的信譽來“踢走”對方。下面列舉的是一些還算禮貌的人身攻擊：

- “從這幾個法西斯分子（或自由主義者、學院派）嘴里，你還指望能聽到啥？”
- “這完全就是脫離現實的人才會有的典型論調。”
- “簡直就是大男子主義/女權主義的反應。”

- “天啊，你能想像就憑他/她現在搞的這一套，他/她居然還會是這方面的專家？”
- “好吧，那你在這行當里幹多久了？”

這裡的關鍵是將發言者本身和他們所給出的論述或觀點相剝離。不能因為對所駁斥對象的印象不好，就認定他們說出的話一定是錯誤的或者不相干的。雖然這有時確實很難做到，但還是請您盡量尊重這些相左的個人觀點。

“紅鯡魚”

“紅鯡魚”(Red herrings)這一習語指的是在討論中有人提出了毫不相關的議題。比如，在依次理事會上，大家正在討論的重点是行政人員們在差旅或裝備上花銷太大，這時候有人發言說要是我們早幾年就配置適當的科技設備，那就沒有現在這些麻煩了。

這裡大家在討論的問題是由於缺乏財政監管而導致出現不當支出，而有效的財務監管制度在計算機普及前就有了，甚至在羅馬建成前就有。投資適合的科技也許有助於監管過程，但是科技不是監管制度本身。類似這樣提出一大堆解決方案的例子有很多，但其實這些意見不過是使問題疊加，而不是解決問題本身。

應對這些“紅鯡魚”的技巧是，詢問提出論點的人員，他的論述和正在探討的問題有何關聯。以上述情況為例，使用科技怎麼就能增強財務監管了？有合適的可供使用的軟件嗎？審計人員們具備利用這類科技來提高績效的相關技能嗎？畢竟，計算機還是要靠人來操作的。

“粉红鲱鱼”

真正的“红鲱鱼”指的是那些毫无疑问和当前议题不相干的话题。但是有时候，某些人提出的论点可能也契合正在讨论的问题，但是却并非解决之道，我们把这样的论点称为“粉红鲱鱼（pink herring）”。最常见的粉红鲱鱼，恐怕莫过于行政人员们动辄就说本机构存在的问题是由于缺乏财政资源所导致的。

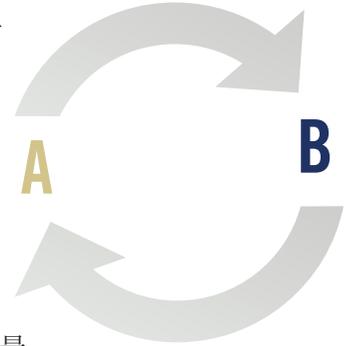
当然，钱是能够买到各类资源的。但是，很多时候，金钱的增加不过是导致相似问题也在增加而已。钱本身解决不了问题。恰当的监管，对现有的物质和人力资源更为有效的配置，或者针对问题的创造性解决方案，可能都比单纯地投钱解决问题要更见成效。应对某个问题的关键在于，直接利用现有的或即将到位的资源来发展或加强各相关机制。

正如上面“红鲱鱼”中提到的例子，在应对这些针对资金所提出的开放式问题上，我们需要询问对方这些钱的使用途径。大多数答案会类似于“购买更多设备”或者“聘请更多警官”，那么接下来的问题就是：这些新增设备或警官如何能够弥补当前存在的不足或是效率低下？

循环论证

循环论证（Circular arguments）指的是A与B互为因果，在政治辩论中，循环论证尤为多见。老师们最喜闻乐见的就是学生们在考试结束后宣称，自己绝对不会得C，因为他们是A等学生。（所以，请解释一下，如果你是一个A等生，你怎么会在考试中得了个C?）

另一个好的例子常见于薪酬谈判中。要求加薪的一方有时会坚称，比起自己的同事，他们需要得到更大幅度的加薪，因为自己所任职部门的薪水有史以来都是整个警队里最高的。如果你们不提高薪水，那他们怎么会成为薪水最高的部门呢？



通常，我们在劳资谈判中所面临的多数棘轮效应（ratcheting effects）²都是源于循环推论。A组认为合同中应明确指出他们的薪酬比其他辖区要高出10%，原因是他们所在区域的生活成本较高。而B组则认为，为了保持竞争性，所以无论出于业绩还是别的什么因素，A组的额外薪酬都不能高过10%。这两个组别中，任何一方的补贴改变，都会导致另外一方的酬劳自动出现棘性变动。

有时候我们需要通过提问来解决循环论证。这类论证的形式基本上都一样：“你知道这个行动是非法的，因为它违背法律了。”类似的，在市政府经常能听到特定的某一方说不支持增税提案，原因是他们要捍卫自己的根基，如果这是问他们为什么增税会影响其根基，他们的回答是因为增税不符合人民的利益。

要打破这种循环，我们需要知道立法主体在通过某部法律时候的初衷：即制定该法律的目的是什么？同样的，我们也需要思考，如果不提高赋税，是否真的能够给选民们带来利益？哪种经济机制才能够切实发挥作用？

其他谬误

在理性和证据面前，人们还会想起其他多种逻辑谬误，比如青少年在追求潮流物品时总爱说的“可是学校里人人都有啊”，或者归咎自然规律的“这就是违背常理嘛”。同样，在专业领域，类似谬误也层出不穷，比如“其他的市政

府/部门都有这个新设备，所以我们当然也该来一个。”当被问到为什么要保持现行的某种做法或者为什么不进行变革，官僚们最中意的答案就是“它本来就是如此嘛”，这已经取代了之前被滥用的“这是机构的本质所决定的呀”。而上述这些说法，在逻辑上都站不住脚。

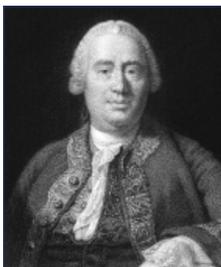
因果关系

我们之前尚未提到的谬误铺路石之一，用拉丁语叫做“post hoc, ergo propter hoc”，意思是“在它之后的，也正是由于它”。如果有学过统计学的读者，可能会发现这其实是“关联性不能证明因果性”类谬误的变形。

不能仅因为两个事物看起来有关联，就认为其中一个是另一个的诱因——或者说，就不能断言他们实际存在有任何形式的因果关系。“警车谬论”的故事可以作为一个根据两个事物的相关性就推断出因果性的滑稽例子：一个到访地球的火星来客，注意到一旦有犯罪发生，犯罪现场无一例外的就会出现警车。因此，这位火星来客就错误地断言，警车就是导致犯罪发生的原因。这一谬论也可以被放在救火车和火灾、或者救护车和伤亡上。

不能仅因为两个事物看起来有关联，就认为其中一个是另一个的诱因——或者说，就不能断言他们实际存在有任何形式的因果关系。

当然，从某种程度上来说，相关性或关联性和因果性是有联系的。问题是，我们如何才能确定或认为所面对的就是因果关系？这个问题非常重要，因为对因果的思考和联想已经成为我们日常世界中必不可少的一部分。当我们面对一些自己不太理解的事情时，我们的第一反应通常是问“怎么会这样呢？”，换句话说，就是“原因是什么？”



大卫·休谟(1711.4.26 - 1776.8.25)，苏格兰哲学家，历史学家，经济学家和评论家，他提出的经验主义哲学和怀疑论尤其为世人所知。他是苏格兰启蒙运动和西方哲学史上最为重要的角色之一。他被认为是“以英语进行写作的人中最为伟大的”哲学家。人们常把休谟和约翰·洛克(John Locke)、乔治·巴克利(George Berkeley)等人共同归为英国经验主义者。

从历史上来看，正规的因果关系论相对来说还是一个比较新的观点。很多学者都以大卫·休谟的著述为基础来解释什么是原因以及我们如何确定原因。休谟是一位生活于十八世纪早中期的苏格兰哲学家。对于因果关系，休谟没有过多的赘述，只是指出了构成此类关系最重要的三个条件。第一个条件是，原因和结果必须相吻合，或者按他说的叫做“结合”，即两件几乎总是同时出现的事物彼此关联的部分。

第二个条件是原因必须发生在结果之前。也就是说，一般情况下，只有犯罪或火险发生之后，警车或消防车才会出现。现实中，抢劫案的发生会导致警察出面应对；而警察的出现却不是引发抢劫案的原因。当然这个例子中也有例外，在当前的预测型警务模式下，我们会希望警车或者便衣小组在犯罪发生前就出现。不过，我们之后再谈这些例外情况。

第三个形成因果的条件是最难理解的，我们称之为“伪关系（non-spuriousness）”条件。伪关系是指“原因”还不足以或不充分导致“结果”的发生，但是它的出现却必然造成特定结果或效应的出现。这个条件有时从反向看更容易理解，即伪关系意味着没有其他第三方因素导致“原因”和“结果”同时出现。举个例子，辖区内的警察数量和发案量间可能密切相关，但是很难说这两者互为因果，而且人口密度的潜在增长或减少会同时对两者产生影响。

虚假关系（Spuriousness）指的是存在于两个或多个因素间的关系具有巧合性，真正的原因是某个潜在的第三方因素。

休谟的因果关系条件

1. 原因和结果必须相互吻合；
2. 原因必须发生在结果之前；
3. 没有其他潜在的第三方因素导致原因和结果同时出现。

这里我们面临的问题是，即使我们把一个显性“原因”移除，“结果”却依然存在。因此，在犯罪和警车关系中，如果一个人出于恶作剧而报警，而警车因此出动，那么我们很可能会看到出警地点并没有犯罪发生。从评估人员或科学家的角度出发，伪关系总的说来是最难掌控的一个因素。要观察两件基本上契合的事物并不难，要找出某件事物的前因性事件也不是很难。

真正的难点在于评估是否还存在其他潜在机制导致这两件事物同时发生。鉴于此，我们找到了应对虚假关系的两个办法。第一个办法是，构建解释性理论来阐明某件事是如何或为何会引发另外一件事的出现。用术语说，就是我们需要找到“因果机制（causal mechanism）”。从逻辑上看，即为为什么X会造成Y？而在业界，就是那句“怎么个来龙去脉”？

第二个应对伪关系的办法是，对各条件进行物理操控（physical manipulation），即考虑我们是否可以自行通过物理性的方法来重现某个结果。我们也把这种操控称为实验。

历经摸索，我们已经建立了一系列的试验设计或是操控条件的方法。通过它们，我们可以将被认为互为因果的两个因素和其他可能性或干扰因素相隔离。这些技术我们将在下一章节中加以强调。

所以，总的来说，这个环节上只要我们能肯定上述三个条件都具备，就有足够的理由相信某件事就是另外一件事的原因。这几个条件分别是契合性或关联性；时间先后顺序，即原因先于结果存在或发生；以及，不存在伪关系的情况，即没有其他潜在机制会导致明面上的原因和结果同时出现。

令人遗憾的是，我们在很多研究中并没有充分考虑到上述三个方面，比如我们常常听到一些医学调研宣称某些因素

（比如石榴）可能降低罹患癌症的风险，一般来说，这类研究的关联点在于某人进行了调查研究，并发现那些吃石榴的人中患癌症的比例比较低。这里我们比较容易确定的是，吃石榴这一行为是否发生在罹患癌症这件事之前。

但是这类型的研究一般容易忽略的就是是否存在虚假关系或混淆因素（confounding factor）。比如，相较于不吃石榴的人群，常吃石榴的人群可能抽烟更少、运动更多、饮食习惯更健康，而且整体上而言有着更为健康的生活方式，这些因素都可能是真正的原因；而饮食结构中有没有石榴这点并不相关。

当然了，一旦我们开始相信石榴和癌症相关，我们就会为这一“事实”（fact）找出各种可能的因果性解释（causal explanations），比如，我们或许会宣称，石榴中所含的丰富维生素C或者抗氧化剂能够预防癌细胞的滋生。

证据和解释相结合

人们常错误地认为，只要收集到足够多的证据，就可以“证明”一个假设或理论是正确的。实际上，如何将解释和相关证据相结合，是一个复杂的过程。

当需要验证某个关系时，我们通常需要采取以下两种方式来使用数据或证据。首先，如果要进行解释说明，那么我们必须确保这个解释符合当前所具备的大多数证据或所面临的实际情况。如果某

种解释无法说清楚我们已知的大部分情况，那么可能这并非一个理想答案。

其次，一旦我们从几个看似合理的解释中，找出从逻辑上说得通而且基本符合现有证据的那几个，那么我们就需要着手进行二次测试（secondary tests）以验证这些解释在紧急情况（critical circumstances）下是否还能站得住脚。

显然，我们所选择的解释是符合已知事实的，所以如果我们仅仅是不断采集相同情况下的数据，并不能作为进一步的确凿证据。

比如说，我们掌握的现实情况是，在主城区内的涂鸦要比其他地区多；但是单凭这点并不能证明“涂鸦导致犯罪”。正如我们之前讲的警车谬论例子，火星上总是看到犯罪和警车同时出现，但其实无法证明两者间的因果关系。换句话说，只要出现几次有犯罪发生但是却并没有警车到场的情况，那这个假设就马上可以被否定了。

这个或许是休谟在其关于因果关系的论述中最为重要的观点了。要就某事证真很难，但是要证伪却容易得多。

休谟就此举了一个例子：虽然从古至今太阳都是东升西落，但是并不能“证明”这一现象明天必然会发生。也就是说，我们只需要找到一个例子说明太阳没有从东边升起，就可以否定这个规律。这个例子虽然不合常理，但是却论证了这个观点，即可以对某种关系证真的证据和可以对某种关系证伪的证据，相对而言是不平衡的。

把该观点引入到评估人员和科学家的语境之中，我们就可以开始探讨工作假设（working hypothesis）这一概念了。

证明一个理论:

1. 我们必须确保某种解释符合当前所具备的至少是绝大多数证据；
2. 随后，我们必须进行二次测试来进一步验证这些解释能否站得住脚。

工作假设和虚无假设

假设检验（Hypothesis testing）是一种常用的验证知识真伪的科学方法。它要求通过一个正确（或错误的）说明，就特定问题给出一个最为可信的解释。对假设进行检验可以帮助我们获得某些结论³。一个工作假设可能是这样的：一个具有刑事司法从业经验（比如担任过警官或矫治官）的缓刑官，比起那些没有刑事司法从业经验的同事，会不那么支持囚犯改造项目。可能造成这样差别的原因是，这些曾经在一线工作过的缓刑官认为他们对于为什么有些罪犯不会改过自新这样的问题有深入了解，或者说带有倾向性偏见。但这个假设并不能作为切实证据，因为这类缓刑官中也有可能存在确实支持改造项目的。要提供证据证明这个假设是否成立，我们可以通过观察那些与假设相反的情况来进行假设检验。这会引导我们进入到虚无假设（null hypothesis）阶段：无论缓刑官是否具有在刑事司法领域的工作经验，从统计数据来看，他们对于犯人改造项目的态度并没有显著差异。

如果我们无法反驳或对这一虚无假设进行证伪（实际上，他们的态度没有差别），那么我们就可以从逻辑上排除之前那个工作假设，即认为有刑事司法从业经验的缓刑官会没有不具备这类经验的同事们那么支持改造项目。

这就是科学家们用来验证假设和理论的策略。我们不能对工作假设直接进行验证，而是创造出与之相反的虚无假设

（比如，我们发现先前的刑事司法工作经验对于结果没有任何影响），以驳倒工作假设，或者最起码我们需要严肃考虑它给出的结论。如果我们没有找到可以支撑起虚无假设的证据（我们没找到论点的可代支持），那么我们就有充分的理由相信之前的工作假设是有效的。我们能找到的可代支持越少，就说明我们工作理论的可信度越高。

工作假设和虚无假设的一个案例

在2010年11月到2011年2月间，在英国大曼彻斯特的奥尔德姆（Oldham）地区，在住宅区发生的盗窃案和前四个月相比上涨了18%，这意味着增加了91起盗窃案。有一种假设是，盗窃案的增多是由于近期监狱释放人员的增加导致本区域的盗窃案罪犯增加。

一些负责犯罪预防工作的警官却认为这两者之间并没有直接关联。那些获释的人员从监狱出来就马上受到严密监控，旨在将他们再犯罪的可能性最小化，同时尽可能保证他们能走上正轨。而导致他们再犯罪的原因，可能是体制的失败，或者社会保障的缺失，而不是因为他们刚从监狱里放出来。

鉴于此，奥尔德姆市的邻里警务小组，以及奥尔德姆社区安全合作处的一位高级警官、犯罪预防警官和犯罪分析师们共同协作，以找出两者之间的关联。

在研究工作的初始阶段，分析师们需要提出工作假设和虚无假设，在该案例中，可以进行如下陈述：

工作假设：

T 盗窃案的增加是因为盗窃案罪犯的增加，而后者增加的直接原因是近期监狱释放人数的增多。

虚无假设：

监狱释放人员的增加不太可能是奥尔德姆盗窃案增长的主要原因。

研究人员们对虚无假设进行了验证。如果证据都符合，他们就会酌情认同虚无假设为真，并且最终推翻工作假设。如果他们发现证据与该虚无假设不相符，他们就会推翻虚无假设，并且提供强有力的原因来证明工作假设为真。

搜集证据

环境扫描

我们为了有效设置工作重点所建立的计划，以及那些与我们履行机构使命的能力有关的情况，都会影响以证据为基础的决策。当问题发生并需要进行决策时，我们需要利用证据来确定这些决策可能带来的影响或是效果。警务机构可以通过采用相关方法路径，来提高绩效，进而探寻并建立与打击罪犯及犯罪行为相关的机制¹。

采集这些信息的常规做法是环境扫描。简单说来，环境扫描能帮助我们了解并掌握机构目前所面临的境况。它能让我们认识到那些对机构产生影响的内外在实际情况、重要问题以及趋势所向。这类信息能够帮助我们确定或否定所持观点，它能指引我们进行下一步的行动策划、战略重点设置以及资金预算。环境扫描还能够帮我们确定未来战略以及制定恰如其分的、有理有据的响应措施。

环境扫描可以给机构带来哪些益处呢？为什么我们需要投入时间和精力来进行扫描？以下是最为关键的一些答案：

环境扫描的好处：

- 不仅对机构内部的目标及职能进行全新的、客观的审视，而且还能够对存在问题进行最为有效地排序；有机会参考他人耗费心血所取得的研究成果、信息数据以及其他资料；
- 为社区利益相关方、机构、个人以及团体参与和他们息息相关的决策搭建平台，给予他们提出意见、观点和建议的机会；

环境扫描(Environmental Scans)能够帮助我们认识到对机构产生影响的内部及外部实际情况、重要问题以及趋势所向。

- 有助于较大的社区发现自己在应对相关问题上的优势及资源所在；
- 为了解你所在机构具备的资源和优势提供参考性框架或视点；
- 有助于了解你所在机构的项目和实践对其他机构、组织、个人或团体造成的影响，以及这些项目和实践从何种程度上对所在机构的职能履行发挥了效益。

进行环境扫描是一个序列性过程(sequential process)，包括从二级资源搜集信息，例如现有的研究报告、数据或其他信息。这还需要一手或首要信息资源的辅助，比如你亲自联络的个人或团体。对这些信息进行分析可以帮助你将在机构置于更为广阔的社会生态环境下进行考虑。

和其他很多管理步骤不同，并没有太多正规的指导方针说明环境扫描究竟该如何进行。而我们将做的，正是要为您提供一份整体程序以及一些有助于进行下一步发展的工具。

环境扫描的类型

环境扫描有两大基本类型。第一种类型是比较不正规的方法，即由您本人根据自身的知识能力，以及通过自己或助理坐在办公室里搜集来的信息进行环境扫描。第一步是先写下来据您所知，其他人在面临类似情况时所采取的做法有哪些；换句话说，就是了解相同的社会环境下，别人都是怎么做的。一般说来，多数人会选择联系那些在相关领域有深入研究、并经常参加相关会议和研讨的同事，因为他们会帮助这个过程进展得更加顺利。

第二步则涉及更为正规的检索。根据所面临的问题，您可能需要搜寻和该话题有关的期刊或新闻报道。一个非常有利的工具是互联网，诸如谷歌（Google）、必应（Bing）和网络爬虫（Webcrawler）这类的搜索引擎可以帮助您非常快速地获取到大量的信息。使用这些大众搜索引擎面临的巨大挑战之一是，有时候不太容易准确定位到您切实需要搜索的关键词，这就导致有一大堆搜索结果是毫无价值的。

利用互联网搜索引擎

对搜索引擎的使用是有诀窍的。如果您足够幸运，碰巧认识了市立图书馆的管理人员，或者可以进出于本地大学的图书馆，那么通常可以得到经验丰富的专业人员帮助。这里也提供给您一些小贴士，将使用谷歌搜索的内容进一步限定于您所期待的结果中。

无论是通过互联网还是亲自前往本地图书馆，您都可以对专业文献进行搜索。对于常见议题而言，通常能够从相关专业领域的杂志以及刊物找到论述。如果您能与一名就职于本地大学或学院的熟悉犯罪学或警察科学的图书馆管理员进行交流，那就是一个好的开端。

如果您还需要更为详细的信息资源，那么您就得查阅专业的研究报告或学术文献资料。这一步骤多少会让人有些畏难情绪，因为专业文献的种类可谓浩如烟海。其中一些是很容易就能查阅到，但也有一些是需要在对特定主题具备了较高的知识储备后才能搜索到的。这里的关键就是，永不气馁。

有时候，您需要浏览其他网站以查找进一步的信息。这种情况下，官方网站会带提供给您大量的信息，比如美国国家刑事司法文献服务（NCJRS）网站，或是联合国毒品和犯罪问题办公室（UNODC）网站。下一页的列表中，我们为您整理了一部分主流的警务和刑事司法期刊，以及几个关键站点的网址。很多本地的大学都和警方建有合作伙伴关系，因此也是非常好的信息资源，并且可能有助于开展联合研究项目。

如果您需要关于您所在社区或其他领域的的数据信息，那么请浏览加拿大统计局（Statistics Canada）网站，上面有大量信息可供使用。

利用谷歌网站进行有效搜索

1. 有的放矢.

如果要在特定网站中搜索相应主题的页面，那么请在谷歌搜索栏中键入：`site:网址 + 搜索关键词`；如果要搜索特定作者，则键入：`author:作者姓名`；如果要搜索包含特定关键词的标题，则键入：`intitle:关键词`。

2. 目标类型

通过在搜索栏中键入：`filetype:pdf` 或其他类型的文件后缀名，可以搜索到特定类型的文件（比如word文档和jpgs图片文件）。

3. 延展搜索.

使用星号键(*)作为搜索主题词中的占位符，比如输入“`polic*`”，那么返回的搜索结果将包括“`police`”（警察）、“`policing`”（警务）等。

4. 通过设定过滤关键词来限定搜索结果

在您希望排除的关键词前面加上减号(-)，就可以不显示和该词组有关的搜索结果。比如，输入 `alarms -burglar`，搜索结果中就不会出现和“`burglar`”（盗窃）相关的信息页面。如果您希望通过数字来限定搜索结果，那么可以使用两个点号(..)来进行范围限定，比如，如果输入“`2010 .. 2014 二手装甲车`”，那么返回的搜索结果将仅限于2010年至2014年之间。

5. 使用特定搜索引擎

比如说，使用谷歌学术页面进行搜索，可以找到选定主题的学术论文或其他文献资料。而网络爬虫(Webcrawler)则可以从一系列的搜索引擎上获取信息。您还可以通过Amazines网站(www.amazines.com)上的分类网站列表进行数据库检索获取免费文献资料。

信息资源

部分专业期刊名称

Aggression and Violent Behavior (《侵略性及暴力行为》)
 Alcoholism Treatment Quarterly (《酒瘾戒除季刊》)
 American Criminal Law Review (《美国刑法研究》)
 American Journal of Criminal Justice (《美国形式司法杂志》)
 British Journal of Criminology (《英国犯罪学杂志》)
 Campus Law Enforcement Journal (《校园执法杂志》)
 Canadian Journal of Criminology and Criminal Justice (《加拿大犯罪学及刑事司法杂志》)
 Canadian Society of Forensic Science (《加拿大法庭科学协会》)
 Child Abuse and Neglect (《虐待及忽视儿童》)
 Child Abuse Review (《儿童虐待研究》)
 Child and Youth Services (《儿童及青少年服务》)
 Correctional Compendium (《惩矫纲要》)
 Corrections Today (《今日惩矫》)
 Crime and Delinquency (《犯罪与违法行为》)
 Crime Prevention and Community Safety (《犯罪预防及社区安全》)
 Criminal Behavior and Mental Health (《犯罪行为及心理健康》)
 Criminal Justice and Behavior (《刑事司法及行为》)
 Criminal Justice Policy Review (《刑事司法政策研究》)
 Criminal Justice Review (《刑事司法研究》)
 Criminology and Public Policy (《犯罪学及公共政策》)
 Drug and Alcohol Review (《毒品及酒精研究》)
 European Journal on Criminal Policy and Research (《欧洲刑事政策及研究杂志》)
 Federal Probation (《联邦缓刑》)
 Global Crime (《全球犯罪》)
 International Criminal Justice Review (《国际刑事司法研究》)
 International Journal of Comparative and Applied Criminal Justice (《国际刑事司法比较及运用杂志》)
 International Journal of Offender Therapy and Comparative Criminology (《国际罪犯治疗及比较犯罪学杂志》)
 International Journal of Police Science and Management (《国际警察科学及管理杂志》)
 Journal of Alcohol and Drug Education (《酒精及毒品教育杂志》)
 Journal of Child and Adolescent Substance Abuse (《儿童及青少年药品滥用杂志》)
 Journal of Child Sexual Abuse (《儿童性虐待杂志》)
 Journal of Contemporary Criminal Justice (《当代刑事司法杂志》)

Journal of Crime and Justice (《犯罪及司法杂志》)
Journal of Criminal Law and Criminology (《刑法及犯罪学杂志》)
Journal of Drug Issues (《毒品问题杂志》)
Journal of Elder Abuse and Neglect (《虐待及忽视老年人杂志》)
Journal of Ethnicity in Substance Abuse (《种族及药品滥用杂志》)
Journal of Experimental Criminology (《实验犯罪学杂志》)
Journal of Family Violence (《家庭暴力杂志》)
Journal of Forensic Identification (《法医学鉴定杂志》)
Journal of Forensic Psychology Practice (《法医心理学实践杂志》)
Journal of Forensic Sciences (《法医学杂志》)
Journal of Gang Research (《犯罪团伙研究杂志》)
Journal of Interpersonal Violence (《人际暴力杂志》)
Journal of Offender Rehabilitation (《罪犯改造杂志》)
Journal of Police Crisis and Negotiations (《警察危机及谈判杂志》)
Journal of Quantitative Criminology (《定量犯罪学杂志》)
Journal of Research in Crime and Delinquency (《犯罪及违法行为研究杂志》)
Justice Quarterly (《司法季刊》)
Juvenile and Family Court Journal (《青少年及家庭法院杂志》)
Law and Order (《法律及秩序》)
Law and Policy (《法律及政策》)
Law Enforcement Technology (《执法技术》)
Legal and Criminological Psychology (《法律和犯罪心理学》)
Police Practice and Research: An International Journal (《国际警务实践及研究杂志》)
Residential Treatment for Children and Youth (《儿童及青少年入院治疗》)
Security Journal (《安保杂志》)
Studies in Conflict and Terrorism (《冲突及恐怖主义研究》)
Substance Use and Misuse (《药物使用及滥用》)
Trauma, Violence, and Abuse (《创伤, 暴力, 以及虐待》)
Violence Against Women (《针对女性的暴力》)
Western Criminology (《西方犯罪学》)
Women and Criminal Justice (《妇女及刑事司法》)

在线资源

加拿大国家司法局(National Institute of Justice): <http://www.crimesolutions.gov/>
加拿大统计局(Statistics Canada): www.statcan.gc.ca
兰德警务质量研究中心(RAND Center on Quality Policing):
<http://www.rand.org/jie/centers/quality-policing.html>
美国司法部社区警务服务办公室(US Department of Justice Community Oriented Policing Services): <http://www.cops.usdoj.gov/>

在加拿大统计局的网站上²，您可以找到关于警务工作及刑事司法系统的大量信息。

为了最为有效地浏览网站信息，请确保您已经从多个渠道收集了信息。这么做可以使您能够进行查证和相互对照，以获知相同的议题和观点是否出现在不同的信息渠道中。

有时候，您需要召集本部门的其他人员进行正规的信息分析。在这样的情况下，您可以考虑让机构外的相关人员加入，并且进行正规信息扫描。在SWOT分析章节的第二部分，我们介绍了进行正规扫描的步骤。环境扫描以及SWOT分析间的根本区别在于，整体上而言，环境扫描的关注点或纳入考虑的议题范畴要更加广泛。一般而言，SWOT分析针对的是和本机构所面临的挑战和机遇密切相关的问题。

进行环境扫描的程序

针对某一问题，您所需要收集的信息，取决于您试图找寻的答案。要组织进行一次环境扫描，我们可以先从重点回答以下问题入手：

- 面临的关键问题是什么？
- 关于该问题我们需要知道些什么？
- 会影响这些因素的趋势以及驱动力是哪些？

一旦对这些问题进行了认真细致的解答，并且通过直接和间接渠道收集到了研究文献资料，那么您就可以着手进行分析了。

首先，我们需要考虑在次级研究（Secondary research）中，出现了哪些观点、理论、问题或关注点。换句话说，即该议题对其他的团体、组织、社区或是警察部门带来了哪些影响？他们对于该问题的关注究竟适合程度？

将您的调研结果和此前目标小组（focus groups）讨论中得出的量化数据进行比较。回忆一下在一对一的访谈中，人们都说过些什么。有哪些观点是普遍存在的？这些结果是否具有关联性和重复性？

尽可能确定人们对问题各个方面的关注程度是否也体现在了您的次级研究结果之中。他们的看法是否与您一致？还是说他们有不同的观点、思考或担忧？

一旦您或您的小组确定了关注问题的先后次序，即将最严重、最紧迫的问题列于首位，那么您就可以开始考虑相关策略、项目活动，以及能够帮助您解决问题的各种方法了。同时，您还需要考虑应对这些紧要问题所涉及的资金预算。

正如我们所提到的，环境扫描的过程中，通常需要同步进行SWOT分析（SWOT Analysis），因为它能够帮助您确定内部和外部的优势、劣势、机遇以及威胁，这些因素都直接影响到机构履行其职能的能力。我们将在本章节的第二部分展开讲解SWOT分析。

案例：警察部门委员会环境扫描³

在2009年，警察部门委员会（Police Sector Council，简称PSC）发表了一份报告，对加拿大的社会发展变化进行了更为深入的说明，旨在提供一份与警方工作息息相关的加拿大社会以及全球性议题的概述。进行该报告的初衷是由于当今社会上的各种趋势日益复杂，而且交织在人口因素、社会因素和经济因素之间的关系也更加错综。在社会变革的大背景下，加拿大警方需要认识到只有和其他力量联手，才能应对好各类问题。该报告就加拿大警务工作进行了更加全面系统的总结，尝试得出世人对警务内涵的意见和反馈。

在PSC进行环境扫描前，特地征询了各相关警务机构的意见，即他们希望在报告中看到哪些方面的内容。以他们的意见为基础，环境扫描报告强调了人口学和公共安全与秩序，并在内容列表中纳入了社会学和人力资源管理学。这份环境扫描是文献综述和2008年一份警方内部环境扫描调研的成果。超过300所机构，从小规模的警察局到设置有警务课程的高等教育机构，以及其他各个层级的相关部门，都提供了自己的反馈意见。

请假设您是某个社区内一所中等规模警察局的局长，而这个警局的警务工作正面临着一系列挑战。您的某个员工建议说，应当了解一下其他社区的警察局都面临着那些挑战。除了对本机构进行环境扫描，您还可以对各类社区、地区警局网站进行扫描，不过可能在搜索信息的时候面临困难，因为多数部门都把数据作为内部资料。因此，另一个选择是，您可以利用PSC报告，因为它已经包含了超过300所机构的信息。这不仅节约了大量时间，而且还将向您提供其他机构面临问题的一份综述。

PSC报告所研究的重要趋势概述

和加拿大一样，全球人口呈老龄化趋势。 加拿大越来越依赖于国际移民来增加人口，补充劳动力。加拿大已经进入人口老龄化，每三个加拿大人中，就有一个出生于生育潮时期⁴，这导致国家面临着退休大潮来临的现实。由于当今的年轻人对于投身警务工作的热忱有限，因此在老一辈退休后，警察队伍也面临着挑战。

虽然加拿大被视为经济发达国家，但是也要看到各个次族群(sub-population)中面临的贫困问题。当前经济上的不确定因素也导致一些人出现了较为严重的不安全感。无家可归者和贫困人口，以及一些少数民族人口，一直以来在司法体系中占据了较高比例，无论是作为受害者还是施暴者。警务工作需要接受近在眼前的新现实，特别是正如一些人所讨论的，社会经济的发展失衡和高犯罪率息息相关。

加拿大具有高等教育水平的人口比例不断上升，同时会使用计算机以及拥有技术素养的人也在增加。对警方来说，科技是把双刃剑。一方面，通过收集和存储信息，科技能够在解决和预防犯罪方面发挥功效，但是另一方面，科技也越来越多的被用在高智商犯罪上，包括诈骗、身份盗窃以及相关的掠夺性行为。

大型综合商场、封闭式社区以及其他新模式物业弱化了公共和私人区域的界限，这使得警方需要进一步加强与私营保安部门的合作。网络犯罪也导致案件归属地和司法辖区上的混淆不清，这也是警方面临的挑战之一，需要更进一步的警务国际合作。

文化背景、生活方式和价值观念的差异给政治家和警方都带来了更大挑战，因为加拿大社会各界一直在努力寻求一个平衡点来保证多元文化共存，同时也在寻求对法律的公正解读。

根据所得来的各类信息，您或许会重新确定本部门的服务组件（service components）。不难看出，有几个可供选择的选项。根据您工作部门的地址（比如一个主流大都市，而非一个小型的市郊社区），您可能需要酌情拓宽活动范围。另一个选择是确定4-5个关键区

域，这些区域同时也是受到其他部门的关注，而您将重点关注本机构的核心使命所在。

再次强调，您能从环境扫描中得出那些结论，取决于您先期确定的亟待解决的首要问题。

SWOT 分析

SWOT分析是指对一所机构的优势(Strengths)、劣势(Weaknesses)、机遇(Opportunities)以及威胁(Threats)进行评估。请注意，通常情况下，优势和劣势是针对机构本身，而机遇和威胁则是受到外部环境影响。要进行SWOT分析并不困难，它也有助于进一步完成环境扫描。通过进行一次SWOT分析，不仅能够得到非常有用的信息，并且可以在一个大框架下获取相关资讯，为进一步分析奠定良好的基础，可谓事半功倍。它也是非常出色的决策辅助工具，能够帮助我们敲定重要决策——特别是正确决策。正如我们早前在本章节里提到的，环境扫描能够为您提供原始资料和次级资料，并能据此确定您所研究的问题中，哪些是最为紧迫的议题和隐忧。当这些信息与SWOT分析的结果相结合，您将能够更好地确定战略重点和未来方向。

同时，在探讨目前您所在机构具备的优势和劣势过程中，能让不同岗位的工作人员参与进来，因此SWOT分析也有助于得到更为全面的环境扫描结果。而对机构的外部环境予以考虑，您可以抓住对本机构有利的机会。

SWOT分析还有助于您发现威胁所在：即那些可能导致本机构当前的职能履行面临挑战的外在因素、现实情况或发展趋势。

SWOT分析有时可以通过引导式小组讨论来进行。在讨论前，可以先发放调查表，让每个成员都有机会先对问题进行考虑，提出自己的意见。

更简单的方法是，您可以给每位参与讨论的成员发一张空白的SWOT分析表，这样他们就能先写下自己的想法，然后携带该表格来参会。

进行 SWOT 分析的步骤

这可以使参与者们从不同视角出发，特别是从外在影响因素出发考虑，对本机构进行评估。而且，即使在最终的成果中没有体现出某些参与者的观点，他们至少也感到自己在参与过程中得到了表达意见的机会，一般情况下，这有助于鼓舞士气，并经常能帮助领导层在进行选择和执行不同方法时，得到更多的“支持者”。

在组织小组讨论时，首当其冲的问题是：哪些人来参与？选择职责不同但有彼此交叉的人员参与，将有利于形成最为全面的评估报告。虽然不敢说万无一失，但是这样的小组人员结构能够减少某些重要方面被忽视的可能性。SWOT分析的基本原则之一，是主体参与者至少得是中层管理人员，当然如果是更高领导层最好。此外，分析过程中，还应当有来自本机构不同部门的员工代表，一线的监管人员也应当参与。还有，虽然不是必然，不过机构领导人员们通常对

本机构机构着实需要考虑的外在和内在问题都有着更为深入的认识，这不仅是源自他们的经验，还来自于他们在机构内外的广泛人际网络。

在开始分析和填写表格前，也就是在SWOT分析会议召开前，通常需要给小组成员提供环境扫描的结果先行阅读。您所提出的引导性问题应当是基于环境扫描和SWOT过程的，这样才能确保您构建了研究的框架，为小组讨论奠定良好的基础。

案例: 完整的 SWOT 分析

优势 (Strengths)

- 社区及“问题解决”理念
- 提供的社区服务
- 财政责任
- 创新

劣势 (Weaknesses)

- 员工老龄化和即将退休人员
- 人员录用面临的挑战
- 缺乏多样性

机遇 (Opportunities)

- 和邻近社区开展关于灾难管理规划的合作
- 申请高级政府资助以进一步开发灾难应对策略

威胁 (Threats)

- 对科学技术依赖性的增加
- 人口老龄化 / 衰退
- 增加的财政预算
- 需要增加文职辅助人员

SWOT小组讨论的基本规则

- 一个时间段内只关注一个特定问题。
- 侧耳倾听，确保您理解了别人所表达的观点。避免打断或指责别人的发言。
- 对每个问题设置合理的讨论时间，以保证谈论能够有序推进。尊重他人——能够接受不同的观点和态度。
- 对于如何应对导致分心的情况提前达成共识，比如手机或是进入会场报告事情的员工造成的干扰。建议关闭手机，并告诉员工只有在紧急情况下才能进入会场打断讨论。
- 保密：哪些信息是会议室之外也可以传播的？信息最终应当如何处理？怎样保护匿名性？
- 所有的小组成员都应当参与讨论。

由于小组所讨论的问题是根据环境扫描的结果所提出的，因此应当让成员们充分考虑每一个问题，进而确定哪些问题和因素是需要更进一步强调或深入了解的。

您在进行SWOT分析的过程中，请时刻提醒自己：

- SWOT分析是一种主观过程，不是一门“科学”。不过，从环境扫描中获取的量化和质化数据，将有助于参与者们相信所得到的结论是有理有据的。
- 化繁为简，每次只关注少数议题。如果有其他情况发生，可以在稍后通过后续流程加以解决。如果不设这些限定，整个分析流程可能会因为一时间面临太多的数据和信息而陷入停滞。
- 对于本机构的优势和劣势要有客观的认识。建立相关的安全和透明条款，这样才与者们才能坦诚相对。

总而言之，SWOT分析将环境扫描和战略规划相结合，有助于得出客观真实、研究透彻并且受到本机构内部工作人员和外部的利益相关者们支持的结论。基于证据的决策正是得益于采取类似的工具，以事实和研究成果为基础，从不同的视角出发考虑并博采众长，进而做出有理有据的计划和决定。

注释及参考文献

1. 见《警务战略计划 第二部分》<http://mennozacharias.com/tag/environmental-scanning/>
2. <http://www5.statcan.gc.ca/subject-sujet/theme-theme.action?pid=2693&lang=eng&more=0&MM>
3. “Environmental Scanning Labour Market Information Project 2009.” Environmental Scan prepared for the Police Sector Council. <http://www.policecouncil.ca/wp-content/uploads/2013/03/PSC-Environmental-Scan-2009.pdf>
4. “生育潮婴儿” (Baby boomer) 是指在二战后1946-1964年期间出生的人。

统计学

一种有效的决策工具

统计学可能是最被人误解的学科之一。多数大学生惧怕但又不得不学习它，而多数教统计学的教授也往往极不情愿去研究它。此外，人们经常把统计学看成是骗子的工具。就像马克吐温曾经说的那样，“世上有三种谎言，分别是谎言，该死的谎言和统计学。”然而，只要使用得当，统计数据可以成为决策者决策方法中最有用和最有力的工具之一。

我们怀疑统计学的坏名声有两个来源。首先，许多人把它看作一种较为抽象的数学，关于它大多数人在中学时期都没有一个愉快的学习经历。其次，多数教统计学的人并不是统计学家，虽然他们多少能够掌握一些技术上的细节，但却很少能够领会其隐藏的逻辑。统计数据必然牵扯到一些数学运算，但是其中大部分的运算并不比平衡某人的收支更为复杂。理解统计学的关键因素就是把它看做一条有效的途径，使主宰我们这个世界的不确定因素变得有序和有意义。事实上，统计学其中一个定义就是把它看作一种能在不确定性条件下做出决策的科学，它不应作为一种最后的手段来支持或使一个基于直觉而匆忙做出的决策合理化。

理解统计学的关键因素就是把它看做一条有效的途径，使主宰我们这个世界的不确定因素变得有序和有意义。

对于大多数决策者来说，关键是不要纠缠于细节和统计分析，而是要理解基本的原则和事件背后的逻辑。那些基本原则很少，而且一般都十分简单。一旦理解了这些，那么，即使一个对基本数学和技术方面知之甚少的人，也能得心应手地运用好统计原则。

统计学包括两个基本分支。第一个，是用来描述事物的数据集合。第二个，是使用数据来帮助我们做出决定或推论。我们把第一个称作描述统计学（descriptive statistics），第二个称作推断统计学（inferential statistics）。

关于测量的讨论

我们经常把观察和记录数据的过程作为一种测量方法。与大多数人相比，统计学家对于测量有着不一样的看法，他们假定所有测量方法都包含误差的元素。换句话说，在统计学的世界里，如果想要使被测量的事物具备极高的精确性，那么运气成分比管理水准重要多了。从统计学的角度来看，测量误差有两个来源：固有误差（inherent error）或不稳定性，以及操作误差（Operational error）。

当我们提到固有误差或者不稳定性，我们指的是被测量事物的属性。比如，你问某人对其工作的满意程度，范围为0--100（0代表完全不满意，100代表完全满意），他们的答复可能会是71。如果你在其他不同的场合问他们同样的问题，他们的答案则可能在接近71附近的一个范围。

事实上，多数人对自已工作的满意程度有一个大致的想法，但很难给出一个精确的数字。另外，尽管他们可能大多满意现有的工作，其确切的满意程度也会随着每天的时段，或他们与上司是否有过争执，或天气等多种因素而改变。

所以，即使在相对稳定的范围内，多数人对其工作满意度的实际水平也会出现固有的不稳定。

这同样适用于酒精测试。到目前为止，呼吸测试法是在酒驾案件中最常用的测试血液酒精含量（BAC）的方法。假设一名警官从刚刚靠边的驾驶员身上进行两次单独的测试。

固有误差和我们测量的事物有关—例如，你嘴里是否有酒味会对酒精测试造成影响。

操作误差和我们如何进行测量有关—例如，测量设备的问题或我们计量的方式方法。

那么，他很可能会在两次不同的测试中得到不同的BAC数据，这要看被测试的驾驶员是否刚刚打过饱嗝或呕吐；或者来自手机和警用无线电的电子干扰；或者现场有没有烟草的烟雾，灰尘和湿度的干扰。

因此，从统计学的角度看，BAC数据是充满变数的。

除了固有变量的概念，还有操作上的失误。或许这名警官忘了对这个设备进行手动标准检验，或者电池没有完全充好电，或者设备使用不当，或者体内酒精测试表格填写错误。警官在誊写结果的时候也会出现错误，例如把0.08写成0.8，或者在测试BAC并记录的过程中忘记真实的数据。

关键在于，只要能够办到，我们都在尽量使测量完全准确，虽然通常来说很困难。相信自己能做到这点只不过是自欺欺人。对于大多数情况而言，“接近”已经足够好了。

BAC到底是0.08还是0.085有什么差别么？统计数据之所以有效，原因之一，就是统计学假设我们在实际测量中会出现各种误差。

统计学另外一个优点在于，当我们正确使用它时，能够估计出在测量过程中有多少误差存在。

根据统计学家的观点，那些相信测量能够完全精准的人无异于把头埋进沙地里的鸵鸟。既然不得不承认测量中的误差无处不在，那为什么不正视它，然后试着评估一下这些误差的范围呢？那我们应该怎样做？答案是我们既需要对同一种事物进行多次测量，也需要测量相同属性的多种事物。



根据统计学家的观点，那些相信测量能够完全精准的人无异于把头埋进沙地里的鸵鸟。

描述统计学

记住一种单一事物的特征相对比较容易，不管这种事物是一个人，一件事（比如月食），或一辆机动车的颜色。同样，多数人能够轻易回忆起几种事物的特性。然而，事物的数量越多，记住群组里面单一事物的特性就越困难。例如，我们或许能回忆起警局里所有同事的年龄。但想要记住这个地区所有警官的年龄事实上是不太可能的；如果我们想要能够说出这一地区所有警官的年龄，就需要聚合或汇总数据。这就是描述统计学起到的作用。

描述统计学就是总结一组对象的特性，从而使我们能够在庞大杂乱的信息中理清头绪。单单把600位男女警官的年龄列举出来是没有任何意义的，即使我们能够记住它。描述统计学让我们能够找

出信息列表中确定有用的那部分。通常来说，在对所研究对象的信息列表或是大量的观察中，我们最想知道两件事，即他们的典型性，以及存在的可变性是多少？

我们假设现在计算出这个地区警官的平均年龄是38岁。其他测量典型性的方法还包括中位数和众数的计算。中位数就是在年龄分布中的一个数值，其中小于它和大于它的年龄各占一半。在这里，中位数可能是35岁。换句话说，在我们地区一半警官大于35岁，另一半则小于35岁。众数则是指研究对象中最普遍的年龄。平均数、中位数和众数是测量典型性的最常用的几种方法。我们也可以把这些方法看作是信息列表或年龄分布的主要定位点。

最常见的测量典型性的方法是计算平均数或平均值。

我们假设现在计算出这个地区警官的平均年龄是38岁。其他测量典型性的方法还包括中位数和众数的计算。中位数就是在年龄分布中的一个数值，其中小于它和大于它的年龄各占一半。在这里，中位数可能是35岁。换句话说，在我们地区一半警官大于35岁，另一半则小于35岁。众数则是指研究对象中最普遍的年龄。平均数、中位数和众数是测量典型性的最常用的几种方法。我们也可以把这些方法看作是信息列表或年龄分布的主要定位点。

差异量数可以告诉我们测量的范围或差异是怎样的广泛。我们知道这个地区警官的平均年龄是35岁；此外，我们还应该知道其中大多数人年龄在30岁至40岁之间而不是25岁至50岁。最常见的测量差异性的方法是区域统计（range statistics）和方差统计（variance statistics）。

区域统计是一种针对两点间距的简单的测量方法。例如，在我们的警官中，最年轻的可能是24岁，而年纪最大的是58岁。那么这个区域简单来说就是58-24，或者说是34年。此范围的测量基于区域分布中最小值和最大值之间的差异。从最小到最大范围的计算是比较有趣的，因为它有时会具有误导性。例如，在一个区域中最老的人可能是65岁，而其他“年长”雇员大部分都小于55岁。在这里，我们就把65岁这个年龄叫做异常值。

对于一些奇怪的极端的分布案例，我们使用一种叫做四分位距（interquartile range）的统计方法。

描述统计学就是总结一组对象的特性，从而使我们能够在大杂乱的信息中理清头绪。

我们可以通过确定研究对象的平均值或是中间值来测量典型性。

我们可以通过确定研究对象中最年轻和最年长的年龄段，年龄扩展的范围，或结果偏离平均值的程度来测量可变性。

要计算四分位距，我们需要先算出位于年龄分布区四分之一这个位置的人的年龄，然后再算出四分之三这个位置的人的年龄。那么，简单来说四分位距就是这两个数据之间的差异。因此，像从最小到最大的范围一样，四分位距让我们明白了年龄扩展范围的概念。

除了区域，我们经常使用方差统计来界定在测量集中趋势时数据的扩散和分配程度。在方差统计中，两个最常见的数据是方差和标准偏差。乍一看，这些统计数据可能会让人有点儿望而生畏，但从概念上来说，它们还是十分简单的。理解它们的关键不在于把焦点放在数学计算上，而是要思考它们潜在的真正含义。

一个关于典型性的案例

即使是比较简单的描述统计学对于决策来说也是很有用的。让我们看下两个相邻城镇一周内（7天）劫案的数量。每天上报的违法行为在下表中都列了出来。

	日期	甲镇	乙镇
	星期一	10	9
	星期二	11	16
	星期三	8	7
中间值	星期四	9	7
	星期五	9	8
	星期六	9	7
	星期日	5	7
合计		61	61
平均数		9	9
中位数		9	7
众数		9	7

为了测量典型性，我们可以采取计算平均值或者等差中项，以及中位数和众数等方法。计算平均值或者等差中项，简单来说就是合计劫案总数，然后除以天数得到的数值。中位数就是指在一组数据中介于上下各50%中间的那一个数值。众数就是在一组数据中出现次数最多的数值。

表中的数据描述了劫案实际的数量。即使是从这些有限的信息量中，我们也可以得到一些有趣的数据。首先，两个城镇一周劫案的总数都是61件。所以平均劫案率的计算结果是每天9件。然而，在仔细观察数据后，我们发现乙镇其中一天的劫案数高达16件。从统计学的观点来看，这一类异常值是非常有价值的。异常值对于测量算术平均值影响较大。简单假设一下，如果我们把16换成30，而另外的数值保持不变，那么劫案的平均值将会激增至10.7件。

然而，异常值对中位数的测量影响较小（有时候也称作中值）。我们已经知道，中位数就是将区间内的数值区分成上下各50%的中间值。表中，A镇中位数或者中值是9，巧合的也是周四，一周中的中间那一天。对于乙镇来说，中位数则是7。

甲镇中位数比平均数较小是由以下一些原因造成的，除了有一个16件劫案的异常值外，乙镇劫案数量普遍低于甲镇。因为这里我们只用到少量的数据，所以看起来比较浅显易懂。如果数据量较大，那结果将不会如此明显。但是不管如何，原则是不会变的。

一个关于变异数的案例

在这个案例中，我们仍然使用上一个表格的劫案数据。从上表可以看出两个派出所劫案发生的典型性和平均值差不多一致。然而，从原始数据我们发现B镇劫案发生的变异性相较于A镇更为明显。从数学上看，平均数和中位数的稍有不同也印证了这个观点。

	甲镇		
	抢劫案	抢劫案	平方偏差
星期一	10	1	1
星期二	11	2	2
星期三	8	-1	1
星期四	9	0	0
星期五	9	0	0
星期六	9	0	0
星期日	5	-4	16
平均值	9	0	2.7
	乙镇		
星期一	9	0	0
星期二	16	7	49
星期三	7	-2	4
星期四	7	-2	4
星期五	8	-1	1
星期六	7	-2	4
星期日	7	-2	4
平均值	9	0	9.4

测量变异性的一种方法是计算极差。甲镇的劫案数从最少的5到最多的11，差值为6；乙镇的劫案数从最少的7到最多的16，差值为9。

另外两种普遍用来测量变异性的方法是计算方差和标准差。虽然看似复杂，这些方法在概念上是很简单的。第二列的数字，是我们用每一天的劫案数减去平均劫案数得到的结果。例如，在甲镇，第一天的平均差计算公式为 $10-9=1$ ，以此类推。

在第三列，我们简单的对平均差进行平方计算（就是用它的值乘以本身）。对所有数值进行计算后，我们可以得出两个结论。第一，平均值的平均差为0。这是恒定的，因为平均值在区间的中间，正负偏差将相互抵消。这就是我们计算第三列数值的原因：平方偏差。

平方偏差的平均值即方差。甲镇的方差是2.7，乙镇的是9.4。这个结果说明乙镇劫案数的变化远大于甲镇。方差在统计学中被大量的应用。在更为先进一些的统计学中，我们试图解释为什么一组数据的变异性会比另一组数据的更多。或许，两个城镇的经济发展水平有着根本的不同。又或许警务工作的效率影响了劫案的数量。这些都是我们想要进行测试和验证的假设。

由于平方值会生成较大的数字，我们通常比较方差的平方根。这将数值带回到原始测量的大小（非平方的原始数据）。方差的算术平方根叫做标准差。甲镇的标准差是1.6，乙镇则为3.1。这个结果说明乙镇的劫案变化率几乎是甲镇的两倍。

题外话

我们还可以利用方差统计来估算测量中存在多少误差。例如，两个人平均花25分钟来完成一项工作。假设第一个人的方差为8分钟，而第二个人的方差为3分钟。假定平均来说两个人的业绩近似一样，那么根据方差测量的结果，第二个人的变异性较小且更加稳定，实际意义上来说也就是更优秀。从管理的角度来看，问题的有趣性在于为什么这个人的表现比其他人更加恒定。

随后进行的调查可能会显示，第一个人必须在各种严格的条件下完成工作，而第二个则面临较少的环境挑战。也有可能第一个人开始的时候未尽全力，到后面才加紧时间工作，结果使平均值看起来处在一个可以接受的范围。

不管怎么说，了解方差的变化能够使我们掌握更多原因，而不是简单的知道平均值和集中趋势的不同。

推论统计学

数理统计学的第二个主要分支即推论统计学。推论统计学可以帮助我们得出结论，并作出决定。和描述统计学不同，推论统计学的数学计算将会比较复杂。因此，我们将把重点集中在推论统计学的逻辑推理中，以此来检验如何使用它帮助我们作出决定。从书上自学推论统计学通常来说是一件比较困难的事情。

对于那些没有相关专业背景的读者，去当地大学选读一个学期的统计学课程是一个很不错的投资。否则，了解基本的概念也就足够了，把细节留给专家们吧。

推论统计学有多种用途。然而，其最主要的用途有两个，一是从样本数据去估计或推断母体数量特征，二是估计两个或更多的母体或样品之间是否存在显著差异。

母体推算

首先我们对母体中的样本进行推断。如果我们想要知道一个城市里使用防盗警报系统的人口比例，我们可以联系每个家庭并提出相应的问题。对一个司法管辖区内的每个人都采集信息被称为人口普查。在一个拥有30万家庭的城市里，这绝对是一个既昂贵又耗时的提议。这就是为什么人口普查很少进行，同时必须在有限的条件下开展。幸运的是，早在二十世纪，统计学家们就想出如何从子组或样品的特点来估计整个人口的特点。

不管怎样，做到这一点的关键就是从母体数据中绘制或收集样品的方法。本质上，“任何旧的样本”都用其利用价值。从母体中选取样本也需要特定的方法。

在选取时会用到很多不同的方法，但让我们尽量把它简化，考虑基本情况就好。这里我们需要了解的是统计学家称作简单随机抽样的概念。简单随机抽样就是指每一个家庭在母体人口中都有同等机会被选中，且被选中的样本彼此间完全独立。让我们把它分解成三个组成部分：随机抽样，机会均等和独立性。

机抽样

这意味着被选取的样本家庭都使用一种机会机制——就像投掷硬币和计算机随机数生成器。换句话说，我们不能根据有效性或门的颜色去选择样本家庭。随机抽样意味着从一个目录上编号1到30万的家庭中选取一系列家庭（例如一个城市里家庭的名录）。假设要抽取1200个家庭样本，我们可以用随机数生成器在1到30万之间生成1200个随机数。一旦我们得到这些随机数，就可以按照目录上所对应的数字编号来抽取这些家庭样本。

机会均等

这意味着每个家庭有相同的机会或被选择可能性。对于目录上具有双重地址或被遗漏的某一特定类型的家庭（例如某些公寓或某一类住户处于某一特殊的环境），要么他们都有很大的几率被选中，要么都没有被选中的机会。

独立性

这意味着一个家庭的选取不能够决定或影响到另一个样本的选取。例如，选择样本的人可能会注意到在同一个街区的两个房子或彼此相邻的两栋房子出现在列表中。因此，研究员或许认为样本过于雷同，出于偏好去掉其中一家而选择其他家庭。这是不可能被接受的。因为不管什么原因，出现的选项必须包括在内。

如果我们遵循这些规则，那么从样品的特点来估计整个人口的特点就是可行的。当然，一些其他的事项也需要考虑到，例如样品的大小，但这些都属于细节，最好能进一步和相关的专家进行深入讨论。如果我们遵循上文所述的基本规则，就可以估计出在给定的可能范围内装有防盗警报系统的家庭人口比例。

换句话说，利用样本进行估计只能够接近实际数据，但可能不会得出精确的数字。然而，统计抽样和其他程序的不同，是它可以估计人口数据有可能属于的范围。因此，我们可以得出结论，安装防盗警报器的家庭的比是X%，范围为正负Y个百分点，例如，20个家庭中的19个。

新手往往贬低不能够提供精确数据的统计估算。但是，正如我们先前讨论的，统计范畴的基本假定是所有测量都会有误差，所以我们能做到最好的就是得出一个估计值，并使它达到一个合理的精确度。没有其他程序可以做到这一点。利用统计估算，你可以掌握你的判断是否足够精确来帮助你完成实际任务，或因变量太多需要重新考虑。

有很多不同的方法可以用来进行估算，但你无法知道他们是更接近母体数据中的实际指标，还是已经跑到火星附近的某个地方去了。

显著差异

推论统计学的另一个主要用途是能够估计两个样本是否相似或不同。例如，经过一年时间，一名警察局长可能会想了解各个派出所在出警时间上是否有差异。通常情况下，类似于出警时间这样的数据都是由自动调度系统进行统计。在一个阶段后，就可以计算出平均出警时间。如前文所述，平均值将基于误差数据进行估计，而实际数值将围绕这一估计值分布。因此，问题出现了，如果一个部门的出警时间是8分钟，另一个是9分钟，那么这1分钟差异是不是反映了真正的差异，还是它只是简单反映了可测量误差的范围？

在一些差异较大的情况下，实质意义上并不需要统计数据帮助我们作出决定。例如，如果出警时间的差异是10分钟，这样的差异就是一个真实且重要的存在。然而，当我们得到的是1分钟，那我们就无法确定这个差异是真实的或是处于正常变异的范围内。

统计学可以让我们知道这种差异是否是在正常变异范围以内或以外。如果在正常范围外，就说明这个差异是统计学意义上的显著差异。

不过，我们必须注意到，统计学上的显著差异，并不一定意味着它实质上的不同。另一方面，如果统计学上没有显著差异，那这种差异我们评估为正常变化范围的差异，因此，也不存在实质上的重要性。

推论统计学更多的使用在我们需要进行多重比较的时候。通常情况下，一座大城市可能有10个或更多的派出所。这10个派出所的差异是否明显不同？更先进的技术可以帮助我们弄清什么因素会与这些差异相关。这给我们带来我们这一章的最后一个主题，即统计建模的作用。

统计建模

对于大多数决策者来说，统计学真正强大的地方在于社会、自然和机械流程建模的能力。统计模型让我们能够研究那些多重因素影响下而产生特定结果的复杂问题。例如，统计模型经常被用于模拟受交通和天气影响下的出警时间。基于这些模型得到的结果，我们就可以确定如何优化停放警车以及设置巡逻路线。这些模型还可以用于确认是否改变警员派出的数量，或者某一巡逻线路是否更有效。

同样，居民区的房屋地理位置和社会人口特征可用于建模，来研究什么类型的住宅或住户最有可能遇到入室盗窃或非法闯入。目前，走在最前沿的建模领域之一是预测警务。预测警务是分析技术的应用，特别是定量分析技术，用于鉴别最有可能的目标，使警察能够提前干预，防止或解决犯罪。预测警务可以向执法机构提供几个优势¹，但它并非取代传统的警务工作；相反，它强化了现有的一些手段，例如问题导向警务、社区警务、情报主导警务和热点警务²。

在统计建模中，我们的重点是试图解释变化。对此，我们可以回顾下我们的基本统计概念——方差。例如，我们可能会问：对于来自支队1和支队2的同样类别的呼叫，究竟什么因素会影响到出警时间的变化？基于这些和其他模拟练习的结果，可以找出什么样的介入方式有效，什么样的无效。提高决策者的资源配置和效率也是有可能的。无论模型或基本过程复杂与否，统计分析有助于我们得到许多精确的估计结果。

在统计建模中需要特别注意的关键问题如下：

- 整体模型是否准确地反映了我们正在试图描述或仿效的过程？也就是说，有统计学意义吗？
- 在模型中解释影响结果因素的变量有多大？
- 在模型中哪些元素有统计学意义，哪些没有？

- 模型中各组成部分对结果因素的相对影响或排序是什么？
- 从实质性或政策性的角度来看，这些影响是有意义的吗？
- 在模型中的各子组件如何相互作用，并对结果产生影响？

正如我们所指出的，统计学并不是所有决策的法宝。然而，使用适当的统计技术，可以让我们对正在研究的问题有着更加深刻的认识。

决策是一个复杂的过程，最好的方法是在处理过程中使用尽量多的工具来帮助你得到答案。通常，我们需要权衡取舍。一些方法可能有统计学意义，但没有实质意义。同样，仅仅因为一个方法比另一个更有效，并不意味着它在社会层面或经济层面是合理的。不管怎样，知道某种方法是否有“真正”的影响就是一个很好的开始。

注释及参考文献

1. “Predictive Policing: Forecasting Crime and Law Enforcement.” Research Brief RAND Corporation. http://www.rand.org/content/dam/rand/pubs/research_briefs/RB9700/RB9735/RAND_RB9735.pdf
2. “Predictive Policing, National Justice Institute” <http://www.nij.gov/topics/law-enforcement/strategies/predictive-policing/Pages/welcome.aspx>

实验设计

如何发现事件的意义？

这本书的一个基本概念，是基于证据的决策比其他形式的决策更有优势。通过证据，我们指的是可观察的和可测量的“事实”或数据。虽然我们认为是事实依据是一个很好的事情，但实际上一个单一的事实或数据或一点点信息是没有任何意义的，这些事实只有在和其他事实相比较的情况下才会有意义。

举个例子，让我们假设您去印度旅行，看到一双鞋子卖2859卢比。如果您不熟悉印度的价格，您可能会问自己这值不值。“事实上”这双鞋子是否值2859卢比对于您是没有意义的，除非您把它和其他东西进行比较，这种比较可能与另一个产品或印度的每小时平均工资或与另一种等价货币有关。目前，2900卢比相当于50美元。只有通过比较，鞋子的相对价值才有意义。

同样，您所在部门可能整体破案率为40%，而暴力犯罪破案率为70%。在市议会会议上，议员可能会对这样的效率是否可以被接受提出疑问，而普通市民可能会期望至少有90%的罪行能够受到起诉或者得以结案。加拿大国家统计局的数据显示，加拿大的整体暴力犯罪破案率约为60%，而总体破案率为40%左右。通过比较，很明显，您的部门与其他地区整体破案率大致在同一水平，并且在处理暴力犯罪时表现的更好。

要了解某一事件的意义，
我们需要一个适当的比较点

这里要强调的重点是，要了解某一事件的意义，我们需要一个适当的比较点。

在以证据为基础的决策框架内，我们必须问自己的一个关键的问题：什么是最恰当的比较？一个互补性的问题也可能是：比较时最好的方式是什么？答案是使用标准的框架，方案评估者和应用科学家称之为实验设计。实验设计能简单地选用不同方法帮助我们作出适当的比较。

本章的其余部分将重点放在一些基本的实验设计上，我们常常考虑需要针对什么做出决策来评估信息或数据的价值。在应用研究中，设计也许会变得非常复杂。但是，不管设计的复杂性有多大，每一个设计所拥有的价值或优点都包含了几个基本原则。

“反事实”思维

当我们观察某一事情时，可能会问：如果这件事不曾发生，结果会如何？如果轴心国赢得第二次世界大战会怎么样？如果执政党没有赢得最后的选举会怎么样？如果由保险公司提供警察服务，而不是市政当局，会发生什么？与一些理论模型相比，它不能给我们提供什么证据，但作为一种意识训练，它迫使我们去认真思考某项政策或计划的重要内容是什么。什么相关或起作用的部分会使结果变的不同，或像我们预期一样产生一定影响？这些想法，与现有的结果或事实相反，被称为“反事实”（counterfactuals）。

爱因斯坦把对反设事实的深思熟虑称为思想实验（thought experiment）。思想实验包括在我们的头脑中进行分析，彻底想清楚一个特定事件或结果的潜在影响和后果。区分爱因斯坦的思想实验和那些简单的幻想或理论，关键就是要利用实际情况和观察到的数据来进行测试。

我们来看一个思想实验的例子。我们可能考虑的问题是，警察的效率会受到工作内容和性质所造成的压力大小的影响；例如，对风险的感知，长时间的工作，轮班制度和问责制度造成的影响。我们认识到这些元素可以是压力的来源，但它们实际上是否能影响警察工作效率的高低？在我们的思想实验中，我们可能会考虑其他因素，如组织压力。

经验告诉我们，其他因素也会影响工作压力水平。对警察来说最大的压力或许可能不是警务工作的性质造成的，相反，可能是由于人们组织和行为的特点而产生的压力。对于完成的很棒的工作

缺乏奖励或认可，绝对会对工作效率造成影响。我们也应该考虑其他工作环境的压力，在警察工作环境中可能造成压力的有一组织结构和各方面的组织生活（例如同事关系、培训、资源、领导和监督，以及内部事务）¹。通过这个思想实验我们得出的结论是，工作内容不是工作压力水平的唯一因果关系，像工作环境这些其他的压力反而可能是更大的原因。

认真思考后，我们得出了一个有意义的结论。然而，就它本身来说，在逻辑上有意义的事实在我们观察到的现实世界中并不总是行得通。我们所需要的是可以重复观测的确凿证据，不仅在想象中，也要能够看的见摸的着，可以和别人一起共享并进行评估。

好的证据由什么构成？

当我们基于证据作出决策时，最根本的问题是：什么构成了适当的证据？如果我们把科学当作一种途径来寻找某一事物的“真实”解释，那么在法庭案件上使用是有意义的。在法院，需要用科学的方法提供各种大量的证据。

即使它们从根本上来说都是真实的，我们也会认为一些证据比另一些证据更有效、更可靠和更有相关性。所以这就是科学性。好的证据经得起严格的交叉检查。那么，好的证据由什么构成？

好证据的一个特征就是经得起严格检验。在科学的框架内，测试的基本机制是实验设计。实验设计是逻辑的物理应用，所以我们需要检查实验设计潜在的逻辑性。

假设在一个阶段，我们想评估防盗警报器对住宅入室盗窃率的影响，方法之一是对一个在所有住宅都安装了防盗警报器的小区进行采样。我们可以分析住宅入室盗窃率在引入报警器前后存在哪些差别。不幸的是，任何差异都可能来自其他因素（回顾我们之前讨论过的虚假关系），住宅入室盗窃率可能由于一些偶然因素导致下降，例如，小区附近年轻人数量的减少，或一个更高效的就业市场，或由于警察的巡逻，或因为更有效的邻里守望计划，我们知道所有这些因素都与住宅入室盗窃案的发生高度相关。

理想情况下，我们希望能够观察到同一个社区在有防盗报警器和没有防盗报警器的情况下同时发生的情形。换句话说，我们希望根据同一个社区有或没有参与防盗报警器项目所产生的结果差异来评估这个项目的效果。尽管如此，我们知道这是不可能的。事物不可能在同一时间呈现出两种状态。在任何时刻，社区要么参加了该计划，要么没有参加。在同一时刻，我们无法观察到同一实体的两种不同情形，这在科学上被称为“反事实问题”。那么，在一种情形已经发生的情况下，我们如何衡量其他情形发生了什么？

如果我们无法评估在相反的情况下或反事实情形下发生了什么，那我们如何判断是防盗警报产生的影响而不是别的什么东西？科学家和项目评估人员采取的方法是找到一个和实验组情况尽可能接近的对照组。对照组和实验组的相似程度直接决定了我们的实验结果有多大的可信度。

寻找或创建对照组的方法有很多，其中一些会比另一些更好。评估人员甚至会花费大量的时间和精力去找到一个合适的对照组。

在逻辑上有意义的事实
在我们观察到的现实世界
中并不总是行得通。
我们需要的是可以重复
观测的确凿证据。

例如，我们可能会发现一个“姐妹”社区距离目标社区不远，将其拿来进行比较。另一方面，我们可以向每户住宅发放防盗报警器，或在社区南部发放，但忽略北部住宅。我们甚至可以考虑把我们的目标群体与该省或该地区的其他社区进行比较。所有这些方法都可以提供一个比较值，据此我们可以判断防盗报警器在目标社区的潜在影响。

然而，问题是所有这些选项都可能具有局限性。某些条件或环境使目标组和对照组有着不可避免的差异。有时我们可以发现这些差异。例如，在选择一个“姐妹”社区时，它可能是该镇老年人较为集中的住宅区，同时有一个更落后的整体安全设计。甚至对一个临时观察员来说，这都是一个明显的差异。然而，现实中的差异通常来说并不明显。

本章的其余部分将重点关注如何用不同的方法来确定有效的对照组，从而准确重现或模拟反事实情形的发生。无论我们正在评估的项目属于什么类型，确定这样的对照组都是影响评估的关键。简单地说，没有一个有效的反事实估计，我们不能确立一个项目的影响的确信程度。

单组测试不考虑其他情况可能带来的结果

与目标相比较(单组测试)

我们最简单的设计之一就是把我们感兴趣的母体与一个特定的目标或标准进行比较。通常，政策条例基于法律标准或目标群体，是经过最佳实践后研究制定的。由于环境不同目标的变化很多。例如，某一社区可能会制定一个五年内减少20%交通事故的目标。一部分制造商可能会实施六西格玛管理策略，预期使生产出来的次品率不超过百万分之3.4。人力资源政策还规定，公司应该积极雇佣一定比例的弱势群体的成员。

因此，关键是把我们感兴趣的母体与理论上可行的、可实现的目标进行比较。一旦我们开始实施一个行动，问题就变成了我们是否已经达到了目标。如果我们达到目标，我们就有理由相信，行动（通常是一个政策或计划的实施）已经成功。当然，我们会用一个统计程序来帮助确定我们是否足够接近目标或是已达成目标。

方法论的文献中有时称这种方法为单组测试。这是对照一个标准来执行一个行动，政策或程序，如果结果符合标准，我们一般就认为该行动是成功的。证据似乎合理地令人信服。不幸的是，单组测试也有缺陷。在下面的例子中，我们可以看到它的一个主要缺陷。

案例：单组测试

假设一个社区报告其整体犯罪率为每10万人中有3000人犯罪，并希望三年内使这个比例降低20%。于是警察局长决定发起一个邻里守望计划，这是一个较为划算的策略，以公民自发参与为中心打击犯罪。三年后犯罪率为每10万人中约有2400人犯罪，相当于达到了预期的20%的下降。那么我们可以认为是守望计划促使了犯罪的减少吗？似乎是如此，证据似乎也是令人信服的。

事实上，比例下降也可能存在其他的原因。这项运动在当地媒体上引起了大量的宣传，警察和志愿者挨家挨户分发小册子；网上宣传活动也同时展开，建议各个家庭和企业主如何采取简单、高效的措施来保护他们的住宅和建筑。手机技术的成熟和社交媒体的大量应用使公民感到更安全，因为当他们看到罪案发生或其他可疑的东西时可以更迅速地联系警方。由于宣传的作用，社区中的住户们增强了警惕意识，并对他们的住宅和其他住地进行了一些改造。一些人安装了闭路电视，并对门、窗和其他入口的锁和安全措施进行加固和强化。

也就是说，通过提高社区居民对犯罪的认识，并促使他们采取行动，那么，不管他们是否参加了该计划，都会减少犯罪的可能性。

这里的重点不是去争论邻里守望计划在减少犯罪方面起不起作用，而是强调其他可替代的或巧合的原因同样能够说明为什么达成目标。这些其他原因有多大的可信度取决于不同的因素。首先，其他原因在逻辑上说的通吗？如果以前的宣传活动没有明显的影响，我们不妨把邻里守望计划作为一种有效的机制。

其次，如果在其他社区的宣传活动导致犯罪率大幅下降，那我们可能会更支持其他原因。再次，可能是多种原因导致整体犯罪率下降，如人口统计的变化，或社区保障中心的增加，整体意识的提高，吸毒成瘾率的降低，等等。显而易见，无论哪种因素都会导致犯罪率的下降。

前后比对设计

单测设计或目标设计的一种变化是前后比对设计。例如，我们希望在一个感兴趣的群体或社区中找到一个有影响的点。我们可以预先测量环境，然后进行干预处理，最后再测量结果。这里我们假设所得结果之前和之后的任何差异都是干预介入的影响。与单测设计的政策目标参照物不同，在这个设计中隐含的比较物是前测基础与后测结果。

前后比对设计的优势和劣势与单组测试相仿。具体来说，我们不能完全肯定是干预处理造成的影响或仅仅是一些巧合的作用。例如，一个司法管辖区可能要减少年轻司机的交通肇事率，他们采取的方法是将25岁以下司机的法定血液中酒精浓度限制从8%上升到5%。通过对该法令出台前三年及执行后三年的数据进行研究，一名评估工作人员发现年轻司机交通肇事率确实有所下降。

同样，我们可以得出结论说在立法方面的变化是决定因素。但另一方面，也有可能年轻司机中饮酒的比率和BAC值的比率并没有多大变化。这种差异仅仅是由于警察提高了警惕，他们针对年轻的司机，试图强制执行新的法规。

即使没有立法方面的变化，而只是类似警察的警惕这类因素，也会产生类似的结果。也就是说，重要的因素并不是立法，而是加强警察的监督，这对年轻司机来说是一种普遍的威慑力。

局限性回顾

这些设计的局限性并不意味着收集不到相关证据。我们有充分的理由相信，如果将这些干预措施强加在许多社区，即使在不同情况下也会得到类似的结果。另外，进行一项干预，然后撤销它，也能告诉我们很多。如果进行干预得到了我们预期的结果，而撤销干预导致恢复原状，那么我们就得到一个更有力的证据说明这次干预是真正的原因。我们需要记住的是，证据很少是绝对的。它有着不同程度的可靠性或可信度。就像在法庭上一样，一些形式的证据比其他的更可信。

鉴于这些设计的固有弱点，我们可能会问那应该采取什么办法来解决这个问题。到目前为止，评估者和科学家之间公认的黄金法则是经典实验设计。

经典设计

科学的经验法则是，没有什么完美的，确定性是一个难以实现的目标。另一方面，只有一个人死亡的不确定概率几乎不可能是玩俄罗斯轮盘的原因。同样地，在股票市场中，一个人如果只有1%的风险会失去所有资产，那他的投资方式和风险超过80%时的投资方式肯定是不同的。所以，如果我们做不到完美，什么是目前最理想或达到黄金标准的实验设计呢？

到目前为止，评估专家和科学家都喜欢采用两组前后比对实验设计，用于提供最有效、最可信的证据。我们开始的时候用到之前提过的前后比对设计，然后再辅以对照组或控制组作为反事实依据。换句话说，我们的其中一组将接受治疗，另一组则顺其自然。如果治疗组出现显著的变化而对照组没有，那么我们有非常强烈的理由相信干预措施的影响。

本设计的关键是确保对照组从一开始就近似于实验组或治疗组。这可以追溯到我们前面讨论的反事实，理想的情况是，我们希望看到干预的元素同时既出现又不出现。这种情况在物质上来说是不可能的。然而，我们可以保证，无论是干预组还是对照组，初始情况是尽可能一致的。我们怎么做呢？

增强经典设计实验有效性的关键是保证对照（控制）组与实验组等效

一种方法是选用一对完全相同的人（或社区，或其他您要比对的事物），并将其分为2组。然而，除非这两人是标准的克隆人，否则我们永远不能肯定他们所有相关特点都是相同的。幸运的是，虽然我们很少能找到克隆人或完全相同的事物，但我们仍可以将受试群体分成两个统计学上的等效组。正如我们以前指出的，统计上的等同并不意味着真正的相同，但这确实意味着，平均而言，两者之间没有统计学上的显著差异。换句话说，实际意义上，他们是接近于相同的。

确保统计学上等效的方法是选取一个初始组，并随机分配给干预组和对照组。

通过随机分配，意思是可以使用一个抛硬币的方式（一个公平的硬币）或随机数发生器来进行分配。在一个足够大的初始组里，所得的2个子组将在统计学意义上相等。也就是说，各小组之间的任何有重大差异的个体将会自我消除。需要注意的一点是，初始组数量越大，两个子组的等效性越强。

任何可能会影响结果的其他的因素（超出了干预的因素）都将分布在这2个组之外。因此，这两个分组将有着同样的所有相关特点，除了一组是准备接受干预或采取措施，而另一组则顺其自然。

避免样本选择偏差

使经典设计更加有效的关键，是研究人员如何引导实验组和对照组或控制组的随机分配。如果我们没有随机分配受试者到干预组和对照组，那就可能存在样本选择偏差，这意味着，干预组和对照组可能会有不同的相关因素。例如，我们可能会进行一项涉及到有或没有防盗报警器的住户之间的研究。

如果犯罪率在已经安装了防盗报警器的居民区较低，那防盗报警器可能不是大多数或所有较低犯罪率的差异所在，很可能安装防盗报警器的人比那些不安装的人更加小心谨慎。

换句话说，那些安装了警报的人也同样是那些已经谨慎地安装高质量锁或窗的人，并且是邻里守望计划中较为活跃的志愿者。

通常情况下，任何人或受试者自愿或选择进入治疗组应被视为可疑的。受试者通常会做一个项目的志愿者，因为他们更有动力，或者看到干预措施可能更为有利。解决样本选择偏差的办法就是让评估专家或研究者做一个满足条件的随机分配。不得不说，研究人员进行真实的随机分配可谓势在必行，一些研究者把那些他们认为最合作或最有可能成功的人分在干预组而非对照组，这样的行为不可能不为人所知。



差强人意的变量

有时我们因为政策、程序介入或其他原因而不能随机分配小组成员。这种情况往往出现在当政府决定立法和制定政策时。按其本质来说，社会政策是在特定领域内实施，而不是随机分配的。例如，当卑诗省希望引入一套新的反应标准将会如何？很明显，我们可以使用前后比对统计模式，但我们知道这有局限性。那我们是否能用经典设计的框架来克服这些缺陷呢？

匹配比较设计

答案是，一些方法虽然没有经典设计这么理想，但却比前测和后测模式更具说服力。因为我们没有能力随机让司法管辖区分配到不同的反应标准，一个方法是找到潜在的克隆区。即我们所知道的不同标准的司法管辖区（或者，更像是假设）在其他大多数方面，甚至所有方面和实验区是相似的。对卑诗省而言，我们可能会考虑选择华盛顿州，俄勒冈州和阿尔伯塔省作为对照方。

当然，这里的假设，是这些司法管辖区有不同的反应标准，但也有和卑诗省类似的地理和社会人口统计特征。

我们称这种方法为匹配比较程序 (matched comparisons procedure)。我们试图找到匹配的司法管辖区，使它尽可能和实验组相似，以提供有意义的对照组。

同样，样本选择偏差问题可能出现，因为有可能某一司法管辖区正决定实施一项其他地方不执行政策。正如之前的前测后测模式，与经典设计的黄金法则得到的结果相比，我们需要对这一方法的结果慎之又慎。

无论如何，匹配比较设计给出了令人信服的证据，证明某些做法是有效的。其中一个最好的例子就是早期关于在汽车上使用日间行车灯来减少交通事故发生的研究。另一方面，匹配比较研究也表明，一些政策没有带来预期的影响，对死刑和谋杀率两者关系的研究是一个很好的案例。

对跨辖区调查后的多数证据表明，死刑可以减轻我们的报复心理，但它并没有减少谋杀案的实际数量。

我们经常地做出同样的行为，即使有证据表明它不会产生我们想要的结果

当我们需要作决定时，证据越充分，决定的正确性越高。即使我们作出错误的决定，错误也可以告诉我们什么是不对的。

如果结果不改变，那不停地做同一件事就没有任何意义。当我们现在的做法明显达不到我们的预期效果时，逻辑上我们就应该尝试换不同的做法。最终，我们很可能会找到一些确实有效的做法。这里需要强调的是，面对相反的证据我们必须愿意改变我们的看法。

我们经常地做出同样的行为，即使有证据表明它不会产生我们想要的结果。几个世纪以来，医生都热衷于放血治疗，而不考虑其他因素，这似乎是当时的“常识”。事实上，许多患者被不必要地放血消耗了体能导致随后死亡，但没有人考虑过这些。

要点概述

本章节讨论的潜在重点是如何收集证据，使用程序或是设计生成的数据是帮助我们判断证据可信度的重要因素。其中的关键因素是我们之前提过的概念，就是一件单独的事物没有任何意义，除非它能与其他的東西进行比较。

换句话说，每一件事物都需要一个参照物，我们才能够理解它的意义。一种干预或行动只有和另一种行动或非行动（无行动）进行比较才有意义。这被称为反事实比较。

既然同一事物不能在同一时刻呈现两种不同的状态，我们就必须找到最合适的对照物。正如我们已经看到的，克隆是不现实的，因此，至今为止最佳的做法是随机实验，把被试者或感兴趣的对象随机分配到干预组，以及适当的对照组或控制组里。随机过程有助于确保没有系统性的样本选择偏差。

在某些情况下，随机分配到干预组和对照组是不可能的，所以我们试着创造出尽可能接近理想的情况。

通过这些方法产生的证据应该始终被认为是可疑的，但如果方法合理，同时对于干预效果逻辑上来说只有极少数的其他解释，那我们一般都愿意相信这些证据，除非我们找到其他更优越的证据。

即使对于最好的实验设计，结果也并不总是同样可信。设计只是我们考虑的一个因素，其效果产生的影响的大小和范围是另一个因素。显然，相比较小的影响或是毫无影响而言，干预措施产生较大的影响会更有助于我们作出决定。但这会导致我们对其他因素的思考，如成本收益或是成本效益分析，也就是我们下一章的主题。

注释及参考文献

1. Shane, J.M. “Organizational stressors and police performance (组织压力与警察效率)” *Journal of Criminal Justice* (刑事司法杂志) (2010) p. 2. doi:10.1016/j.jcrimjus.2010.05.008

成本分析

基本概念

成本分析 (Costing analysis) 有两种类型。第一种是分析做某件事情所产生的成本，举个例子，当您决定购买一辆车时，不仅要考虑购车的费用，还要考虑日常维护和损耗带来的开销，比如维修、加油、保险等等，而根据具体情况，可能还会产生其他费用，比如修建一个新车库或是租一个车位的花销。如果我们关注的是所有权的实际成本，我们还应当考虑车辆的贬值（但愿我们几年后卖车时还多少能收回些成本），以及购置车辆投入资金的利息。

另一种成本分析被我们称为“成本-收益” (cost-benefit) 或“成本-效果” (cost-effectiveness) 分析。在这一类型中，我们会比较某项决策带来的成本以及预期收益的价值。例如，一所警察局可能打算在今后的培训上追加投资，那么所带来的问题就是：这项投入带来的回报是什么？如果是关于如何打击网络犯罪的培训，然而这个社区却压根没有互联网存在，那可以预计此项投入的收益为零。

实际上，这就是个一目了然的成本核算。换个情景，如果培训是关于警官的健康和安全事宜，那么回报可能就是事故、病假和伤亡率的减少，保险赔付也会降低，员工们会更加积极高效地工作，而且士气也得到了提升。我们可以将这些收益的价值和培训课程成本做个比较，就能估算出投资的相应回报值。

对成本进行研究使我们能够确定一项决策的总花销以及由此带来的收益。

经济学里有一个基础性的概念叫做“机会成本” (opportunity cost)。假如您的预算有限，那决定做某件事很可能就意味着要放弃另一件事情。比如，根据警察局的资金预算，您决定要购置一辆轻型货车，一旦做出这一决策，就意味着无法再购置越野车、轿车、摩托或者别的车辆，也就是说，您已经不再拥有选择其他车辆的机会了。这可不仅仅是放弃了选择另一个最佳备选项的机会，同时也意味着我们得放弃做出其他选择会带来的收益。

对成本进行研究使我们能够确定一项决策的总花销以及由此带来的收益。而且，我们还可以考虑被认为是第二或第三最佳备选项的预期成本及收益，并将其和我们最为优先的选项进行比较。有时候，这样的做法会让我们发现，原来那个看起来“差一点”的选项其实比我们最初偏爱的选择还要好。

当我们要做出选择时，货币成本 (monetary cost) 并非而且也不应当成为唯一的考虑因素，政治和其他社会因素也会影响我们的选择。

不过，货币成本确实重要，而且一般来说也更容易量化，因为大多数产品和服务都有与之对应的货币或市场成本。同时，社会和政治成本通常也和经济决策紧密关联。和正规的货币成本评估一样，在评估非货币决策的影响时，也可以套用成本核算的基本框架，唯一区别在于，这些情况的成本和收益通常更加难以量化，因此，决策者们可以而且应当采用定性数据（qualitative data）来评估这一类型决策所带来的影响。

无论我们是进行直接成本分析、成本-效益分析或成本-收益分析，总的来说都需要考虑以下五个步骤。

需要考虑的步骤

1. 确定您所考虑的问题或进行的分析和本部门的运营或战略计划是相关联的；
2. 制定本次决策旨在达到的目标；
3. 确定哪些是可供选择的方案或策略；
4. 进行所选择方案或备选方案的财务分析（成本-收益分析或成本-效用分析）；
5. 准备一份会计报表对结果进行总结。

这些步骤似乎和我们之前提到的没啥太大区别。其实的确如此，只不过，我们需要把有效的基于证据的决策过程看作是整体大框架的一部分，也就是从一开始，就要考虑我们在做什么，原因是什么，备选项是什么，以及可以使用什么证据来帮助我们进行决策。只有在我们知道自己在做什么以及为什么这么做以后，才有可能找准信息。如果不知道这一点，那么我们可能会采集了大量的数据，却压根没怎么收集到证据。

成本分析

直接成本研究包括估算某一特定仪器或服务的总寿命周期 (Life cycle)。我们所说的寿命周期，指的是我们能够使用该仪器或服务的时间段，比如说，在被扔到废品站前，一辆机动车的平均寿命大约是12年；而某人或某所机构可能决定要买一辆车，准备使用5年然后就转手，这种情况下，对于车主而言，这辆车的寿命周期就是5年。

进行精准成本分析的关键在于，确保我们考虑到了所有恰当的开销。一般说来，对于仪器或有价物品而言，开销主要涉及以下几个类别：

- 折旧 (depreciation)；
- 资本利息；
- 维护费用 (耗材及维修)；
- 执照或监管成本；
- 运营成本。

虽然分析师们通常不会对运营者成本 (operator costs) 进行分析，不过，哪怕是在运营者成本对最终决策没有任何影响的情况下，我们还是应当对其予以考虑。如果某个器材是作为替换用品购置的，那么原有器材的运营者成本就会转移过来。不过，让我们假设，一所警察局决定要购置数辆新警车，或者是决定在原有基础上，再购置一辆战术车。

新增车辆可能也意味着要增配人手，而人力资源成本就成为了我们要纳入分析的内容之一。

某些读者可能在想，为什么我们在列表中仅仅列出了折旧，而不列出初始成本 (initial capital cost)。

进行精准成本分析的关键在于，确保我们考虑到了所有恰当的开销。

这里，我们的假设是，这件新购置的器材将在其使用寿命终期被出售，也就是说，资本投入其实是购入价和卖出价之间的差额，这也就是我们所说的“折旧”。

折旧率会根据某件器材的不同部件而有所区别，不过一般说来，每年大概是贬值20-30%。我们根据物品的实际购入价格来计算折旧，一件售价1万美元的器材，如果每年的折旧率是20%，那么一年后价值8000美元，第二年的折旧则是 $\$8,000 \times .2$ ，即贬值1,600美元；据此计算，两年的折旧总值为 $\$2,000 + \$1,600$ ，即折旧3600美元，届时，这件器材的剩余价值将是 $\$10,000 - \$3,600$ ，即6400元。

在成本核算中常常被忽略的一个方面是购置所带来的利息。利率有时候在文献中也被称为贴现率 (discount rate)。如果您需要通过借贷来筹资进行购买，那一般来说就自然而然要考虑利息的问题，因为银行或者融资机构会收取这笔费用。不过，即便是您可以一次性付清器材的购置费用，我们也应当考虑资本的“借贷”价值，原因是如果我们没有进行本次购置，那么我们还可以将这笔资金投入在某个有收益的项目上，或者作为其他用途。这一点，实际上也就是另一种形式的机会成本。

获取可靠的成本估算

无论您是准备雇佣人员还是购置器材，想要得到合理的成本核算结果，其关键在于确保我们把所有项目都纳入了计算，并且对它们进行了最为可靠的成本估算。由于各机构的工作环境不同，一般来说我们需要从自身经验中汲取最佳的信息。您可以回顾一下所在单位公布的资金记录，因为这些经费公示能够反映出真实的财务经验，您可以通过它们看到发生了哪些预算外开销（节流）。您可别把这些信息看作是特殊或偶然情况，而是应当将它们纳入您的成本分析列表进行横向比较。

在无法取得机构所发布的资金记录的情况下，您也可以通过其他途径获取相关信息。通常，供销商可以提供和同类产品竞争者的成本比较，不过，您也得留个心眼，这类型的比较分析通常会选择性地突出或回避“不太方便横向比较”的条目。因此，您得确保自己一是一、二是二地对相关条目进行了比较，如果您发现有些条目没有被列出，那么请务必联系供销商提供相应的信息。

很多第三方机构也会就不同种类的项目进行成本分析，特别是那些专业或行业协会。同时，政府和公共服务机构通常也会就预算和成本研究进行公示。很多这一类型的信息都可以在网上或者本地的图书馆查阅到，有时候甚至简单的一通电话就可以发掘到数据金矿。

下面的列表为我们展示一个关于直接成本核算的例子。这里，我们按5年以上的使用年限来计算了拥有一辆常规的轻型货车所需要支付的成本。不过，本案例中，我们并没有计算驾驶人员的成本。

进行直接成本核算旨在估算某一物品使用期限内的成本，以确定我们是否有能力负担购置该物件。在对不同物品进行比较时，直接成本核算同样很有用，比如，某一品牌的轻型货车可能购买的费用更高，但是日常维护成本却比较低廉，这时，我们所面临的问题就变成了：选择哪一辆会更好？

购买一辆轻型货车作为私人用途的成本

以五年为期限，购置并驾驶一辆轻型货车的成本：

项目	成本(美元)
车辆购置价格	23,500
二手车卖出价格	9,500
折旧	18,577
分期付款利息	3,387
加油	10,079
保险	3,471
税费及牌照费	3,650
养护	2,069
维修	821
总开销	42,054
每公里成本	0.47

假设：

- 每年行驶2万公里；
- 首付2,750.30美元，分期付款年利率2.7%；
- 油费每升1.25美元；
- 每百公里耗油10.46升。

对于采购一辆二手车是否真的比起购买新车整体开销要低，或是长期租赁车辆是否比完全购置要来得划算这样的问题，我们也可以用类似的成本分析来决定。不过，对于这些不同的情况，显然我们必须对寿命周期、运营成本和折旧等进行不同的假设。也许我们需要重申一下，在成本核算中使用到的数据基本上都是估算值。正如我们在统计学那一章节所讨论的，所有数值都是预估值，关键在于，在一定的研究和经验基础上，尽可能将错误减小到最低程度。但也要注意，很多需要核算的项目，比如车辆的售价以及运营的真实成本，作为依据的预估价值其实都不在我们的掌控之中。

我们对有价物品的成本进行了考虑，对于人力资源，其实我们也可以进行类似的成本核算，所遵循的基本原则都是一样的。一般来说，当决定聘用某个人时，我们关注的是他的工资；不过伴随而来的其他开支也可能不菲。当我们把退休金、税赋、保险、福利和其他补偿性款项纳入考虑后，通常会发现大概需要在工资基础上再增加15-30%，这还不算上对员工提供培训的开支和人员维持成本，维持成本包括办公空间以及他们完成本职工作所需要的装备和供给。在之前的例子中，我们也提出仪器设备一般需要配置专人使用，所以，同样的，人们通常也需要借助仪器设备来进行工作。

关于成本—效果分析的一点说明

在此前的分析里，我们的关注的是在寿命周期内拥有和使用一辆机动车所需要支付的总成本。了解了某一事物的总成本对于决策过程来说是一个重要环节。不过，很多情况下仅仅知道总成本还不够我们掌控全局。很多设备或其他物品会带来某种形式的成果或产品，以车辆为例，其成果就是利于交通，因此，相较于总成本，了解每公里的运营成本才是更加有意义的信息。

在上面例子的核算列表中，每公里的预计成本大约是0.47美元。我们把这种以某一特定计量单位计算成本的标价方式称为“成本—效益”。虽然成本—效益毫无疑问和总成本息息相关，但是我们在决策过程中应当把它看作独立的一环。通常，总成本上的差异很可能互不相干；这种情况下，某特定计量单位下的成本支出就显得尤为重要了。

成本—效果分析的关键因素之一是能够确定恰当的成果评估标准，并且能用它们恰到好处地进行测评。

单位成本不同于总成本的原因之一就是，总成本是由两个要素构成的，即固定或沉没成本，以及变量成本。固定成本是指那些采购时发生的一次性支出，而变量成本则一般指运营和维护开支¹。某件仪器很可能固定成本比同类产品高，但是如果它的效率也更高的话，那么一般说来它的计量单位成本就会比同类产品低。

这点在人力资源成本上同样适用。比起把钱付给看起来便宜一些的第三方服务承包商，为那些生产力更高、缺勤率更低以及服务质量更好的员工提供更高的薪酬要更加“划算”，这里需要考虑的重点在于提供了多少产品、服务了多少人以及服务质量有多高。

不过，成本—效果分析的关键因素之一是能够确定恰当的成果评估标准，并且能用它们恰到好处的进行测评。我们又一次需要审视本机构的运营或战略计划，因为这些计划中涵盖了本机构的愿景和行动目标。无论是直接的还是间接的，一所机构的效能高低和它是否能够提供预期的产品或服务密切相关。

成本—效益分析

一般说来，成本—收益分析（Cost-benefit analyses）是简单的成本—效果分析的延伸。成本—收益分析的首要特点是，对于投资所带来的回报，它所关注的范围更广。多数的成本—收益分析纳入考虑的是那些不容易被量化的效果（收益），或者是那些可能有更加广泛的社会影响的成果。

成本—效益分析的基础是福利经济学（welfare economics）。和多数经济学的分支学科不同，福利经济学关注的不仅仅是顾客和企业的决策，它还关注那些在更大的社区范围内对经济利益造成影响的公共决策。因此，成本—效益分析的关注点主要是类似生活品质或环境质量这样的问题。在行进此类成本—效益分析的过程中，最大的挑战就是如何来评测效益，使它们可以和相关情况进行交叉比较。在商品市场中，苹果和变压器是截然不同的物品，但是它们仍然可以共存于一个市场之中，并且都可以具备货币价值。货币是一种常见的交换单位，使种植苹果的商人可以购买到变压器——即使变压器的厂商压根就没兴趣用自己的产品来换取苹果。

而对很多公共产品和服务而言，难点在于不存在一个公开的交易市场来赋予它们相应的货币价值。

而且，出于意识形态领域的原因，很多人拒绝赋予公共产品货币价值，比如，常见的看法有“你不能对环境进行标价”或“人命无价”，但事实却并非如此。问题在于，我们缺乏一个独立或中立的市场来对这类事物进行定价，而这也正是成本—效益分析和直接成本分析相比所存在的主要劣势。

显示偏好和陈述偏好

虽然人命究竟能否被标价这样的哲学性问题难以回答，福利经济学家们却也有着两套高明的方法来进行工作，其中一套方法被称为“显示偏好”（revealed preferences）法。显示偏好和人们在面临某种定性现象时的真实表现相关联。比如，将某一特定房产和另一所相似房产进行比较，能够显示出人们对于屋前草地上的消防栓会赋予“眼中钉价值”；而家长们对自己孩子所接受教育赋予的价值也能够从一定程度上通过孩子学费占其工资收入的比例显示出来。

而另一套经济学家们惯用的工具则被我们称为“陈述偏好”（stated preferences），它指的单纯就是：人们陈述自己就某一事物愿意支付多少价值。

比如，通过看人们愿意增付多少税款来获得清新的空气或支持其他应对温室效应的项目，我们可以判断出他们对于环境因素所赋予的价值。一般情况下，陈述偏好是通过问卷调查和类似的程序来得出的。

虽然显示偏好和陈述偏好各有优势，但也存在它们的局限性。通过一个人购买的人寿保险数额来判断他对自己生命的定价似乎是一个挺好的显示偏好例子；然而，他购买保险的能力其实是受到自己收入制约的。而且，一个把自己的生命看得非常宝贵的人，可能并不愿意亲朋们从自己的死亡中“受益”，毕竟人寿保险的保金是支付给受益人而不是投保人。而社会学家们则从多个方面对陈述偏好进行了大量研究，并得出了惊人的结论，就是人们口中说的和实际上做的，其实完全大相径庭。

同时，在对各种政策和项目的更广泛影响进行评估时，成本—效益分析还是当前我们可以采用的为数不多的新技术之一。它能够帮助我们看清问题，确定哪些是关键因素，并且为应对这一问题提供一些依据。在公共领域，它已经得到了广泛的认可，而且被不少政府部门奉为必备方法，比如，加拿大财政部就要求所有由联邦政府制定的监管性框架都必须是在成本—效益分析的基础之上。

我们可以通过成本—效益分析来：

- 决定是否要采取某个项目方案；
- 决定是否要继续进行某个现行项目；
- 从几个备选项目中进行选择。

这一要求的目的在于，“各部门和机构必须对监管性和非监管性的方案进行评估，以从社会的整体层面上将净收益的最大化。因此，所有监管部门和机构都应当能够证明所提请的方案，相较其他任何类型的监管性或非监管性行为，从长远上对加拿大公民、商业和政府而言，在经济、环境和社会利益上实现了净收益的最大化²。”

总之，我们可以在很多事情上运用成本—效益分析，比如：

- 决定是否要采取某个项目方案；
- 决定是否要继续进行某个现行项目；
- 从几个备选项目或计划中进行选择。

成本—效益分析的步骤

这里所强调的，还是关于如何把所面临的情况和本机构的运营或战略计划相结合。

1. 界定问题

这里所强调的，还是关于如何把所面临的情况和本机构的运营或战略计划相结合。

2. 确定任何限制性或局限性因素

这里讨论的是您所面临的行政性要求和其他可能出现的挑战，包括各种财务上的限制、管理上或人事上的挑战、环境和其他规章制度，以及其他任何您可能需要解决的因素或“障碍物”。

3. 列出备选方案

每一个新项目都有其备选方案，包括顺其自然或保持原状。举个例子，假设我们面临的问题是应否关闭某一处警署，那还应当考虑的方案包括和其他部门合并，和消防或紧急医疗救护等服务部门联合办公，或者与其他警署合作，扩大行动范围。

4. 列出效益

在列出了备选方案后，就要考虑投资相应的回报是多少。是会得到经济上的回报呢，还是会提高效率或成效？也可能，问题的关键不在于产生未来收益，而是在于减少或避免支出。是否可以得到健康、安全或环境上的收益？问题也可能和整体的生活质量息息相关。是否可以节约设备、时间或人力资源？

5. 如何对成本和效益进行量化？

显然，商品和服务的市场或货币价值是最容易进行量化的。我们也已经对获得市场价值上所存在的挑战进行了说明，不过，即使要找的是影子价格或代理价格的特定成本或收益，也并非不可能的任务。社会学家们找到了各种途径来估算人类生命的价值。比如说，与一个社区内的噪音水平或较大的交通流量相关的成本，可以通过观察喧闹社区和安静社区的房价差异或高交通流量和低交通流量之间的差别来进行估算。

通常，我们能够通过浏览适合的文献资料来找到方法对那些难以凭借货币单位计算的问题进行评估。我们也已经讨论了如何利用技术方法来进行更为精准的在线信息检索。我们还可以求助于本地高等教育机构的经济学专家和其他社会科学领域专业人员。一旦遵循了上述步骤，我们就可以形成一份报告，把各个相关方面进行总结，并得出相应的成本和收益。

净现值

俗话说得好，“一鸟在手胜过双鸟在林”，这句话在钱的问题上更是如此。我们之所以要对借出的钱收取利息，原因就是当把钱借给别人时，我们就折损了机会成本，这笔钱不可以再被用于其他地方，而为了弥补放款人的机会成本，借款人就必须支付利息。举个例子，当您购买期限为5年的定期理财产品时，投资将会得到相应的回报，在年利率3%的情况下，一份1000美元的产品的增值情况是：

$$\begin{aligned} & \$1,000 \times 1.03 \times 1.03 \times 1.03 \times 1.03 \times 1.03 = \\ & \$1,000 \times 1.03^5 = \$1,159. \end{aligned}$$

我们也可以反过来考虑，如果您准备在五年内陆续借款2000美元，那么它的当前价值又是多少呢？换句话说，您愿意支付多少成本来即刻拥有这笔现款？

这就是隐藏在“反向抵押贷款”（reverse mortgages）后面的原则。某家银行或融资机构将会根据您的房屋当前的市值提供您一定比例的款额，前提是您同意他们在未来几年内管理房屋并按照市场价值进行出售。这一例子是前面说的问题的反例，我们把这类资本成本率称为折现率（discount rate）。如果折现率按3%计算，那么这笔未来价值2000美元的资金今天的市值是：
 $\$2,000 \times 1/1.03^5 = \$2000 \times .863 = \$1,725$ today.

我们把未来价值一定金额的钱在当前的价值称为“净现值”（net present value）或者简称为NPV。与NPV相对应的，就是“未来值”（future value）。由于项目和货物资本有预定的使用期限，因此我们也普遍采用标准化的净现值来衡量它们当前的价值。另外一个理解NPVs的方式是，当我们试图控制通货膨胀时，可以将它们看作是恒定常数而不是货币面值。

在这些例子里，我们讨论的内容也被经济学家称为“私人时间偏好率”（private time preference rate），因为都是从个人角度出发进行计算。在公共领域里，如果我们选择在某个特定项目上投入公共资金，那通常就意味着无法再在其他惠及公众的项目上投资了。在公共或社会福利领域，经济学家们将被延后的价值称为“社会机会成本”（social opportunity cost）。虽然使用的术语不同，不过其背后运用的原则是相似的。

收益—成本比率

对于那些长期开展的项目，我们对于成本和收益都需要进行分期偿付。根据情况，成本周期可能会与收益期间或收益有效期不同，因此，为了使二者具备可比性，分析师们会同时计算成本及收益的净现值。

我们把收益和成本的比值称为“收益—成本比率”（benefit-cost ratio）或简称为BCR。假设某项目收益的净现值是1350万美元，成本的净现值是1000万美元，那么根据计算公式即可得出：

$$BCR = \frac{(\text{收益NPV})}{(\text{成本NPV})} = \frac{13.5}{10.0} = 1.35$$

理想情况下，BCR的值应当要大于1；在任何小于1的情况下，就意味着成本大于收益；在没有特殊原因的情况下，这一备选方案就应当被否决。如果我们对几个备选方案进行评估，那么一般说来BCR值最高的方案就是最佳选择。如果一个收益—成本比率不是最高的方案获选，那很可能就是我们把其他方面可能获得的收益纳入了考虑范畴。

案例：对一架警用直升机的效用评估³

对北美的不少城市而言，社区是否应当配备警用直升机一直是个颇具争议的话题，人们对于这一议题的很多方面都存在较大的意见分歧，特别是在直升机是否成本过于昂贵、噪音过大、侵犯隐私等方面，或者简单说来，是否有确有必要。把资金投入配备更多警力、更多警车或者更多自行车“骑警”上是不是更为恰当呢？

是不是有证据说明启用直升机犯罪率将呈“滑雪式”下降？还是说地理区域、人口规模或密度说明有/无必要启用直升机？启用警用直升机的首要目的是震慑犯罪分子吗？它旨在提升警方的行动效率吗？警用直升机是作为日常巡逻之用呢还是仅仅为特殊或紧急情况而待命？在追捕过程中，警用直升机能够减少伤亡率吗？有人认为答案是肯定的，但也有人认为直升机根本没必要存在，因为本来就没有什么追捕大战。

1999年，伦敦警察局发布了一份《关于伦敦警务直升机的研究》，着重探讨了：1) 直升机巡逻对于不同类型犯罪的发生是否具有抑制作用（民宅盗窃、商店盗窃、机动车盗窃、从机动车内进行盗窃、抢劫、毁坏财物、夜间非法入侵、可疑人员及车辆）；2) 警用直升机是否可以提高警方行动的有效性或工作效率。伦敦警察局通过1年的时间，对警用直升机的使用情况进行了评估，成本细则分析如下：

伦敦警察局直升机成本分析细则

成本分析细则：

项目	成本
直升机租约：	\$256,328.55
<ul style="list-style-type: none"> • 直升机 • 保险 • 驾驶员 • 例行保养维护 	<ul style="list-style-type: none"> • 附加保养维护 • 零部件储备 • 管理
聚光灯租约：	\$6,026.56
附加器材：	\$11,155.67
<ul style="list-style-type: none"> • 航空电子设备 • 无线电通讯设备 • 警笛/标识 	<ul style="list-style-type: none"> • 液晶显示器 • 其他技术装备
安装器材设备的成本：	\$16,775.00
额外维修费用：	\$7,922.99
<ul style="list-style-type: none"> • 店铺配件 	<ul style="list-style-type: none"> • 其他零配件
燃油成本：	\$32,820.81
<ul style="list-style-type: none"> • 总共使用的燃油量（升）= 43,168.92 	
运营成本（美元）：	\$331,009.58 或 \$331/小时

备注：

1. 1999年，购置相同型号的直升机大约是268,000美元或400,000加币；
2. 由于油费根据时间和地域有所区别，因此提供的是耗油总量，以便今后进行比较；
3. 研究是根据1,000小时的飞行时间来计算，但没有计入飞行员（由警员担任）以及其他相关工作人员的工资。

假设：成本是根据一架直升机一年内1000小时使用时间的实际租赁费用和运营开销来计算。收益则是根据效能和效益来衡量，效能的计算是根据启用直升机后节约的警力成本和时间成本，而效益则是根据启用直升机后更短的犯罪侦查时间以及犯罪减少率来计算。

假设一位警员的工作成本是50.84美元 / 小时，侦查员的成本是55.17美元 / 小时，下面的表格展现了一定工作时长的高效能和效益成本。

下表提供的是672起案件中节约的 15,321分钟（255小时）带来的显著效能收益。

表格: 使用警用直升机在多种类型犯罪中显著节约的时间

案件类型	直升机巡逻		整体警务工作			直升机显著效能 (分钟)*
	总出警量	平均每位警官的处理时间 (分钟)	总出警量	平均每位警官的处理时间 (分钟)	涉及警官人数 (直升机)	
人身攻击	14	76.00	958	89.2	32	422
武器	37	69.11	367	104.5	198	7,007
家庭暴力	16	50.11	2,862	69.00	74	1,398
失踪人口	25	67.96	1,723	79.00	88	980
噪音	21	24.76	4,686	32.00	66	478
扰乱治安	77	47.33	2,194	46.00	359	-477
人际纠纷	85	44.82	8,5889	55.26	299	254
醉酒	12	41.37	1,800	41.00	35	-13
人身安全检查	16	72.08	4,422	61.00	58	-643
协助其他机构	19	54.12	1,337	80.00	37	958
协助消防机构	18	51.48	747	77.78	74	1,946
入室盗窃 (住宅)	35	52.20	540	59.22	207	1,453
偷盗	16	64.24	704	78.72	58	840
破坏财物	16	51.08	665	67.85	52	872
小偷	26	32.17	1,028	42.01	75	738
机动车肇事逃逸	22	64.03	2,208	64.65	67	42
不清醒状态驾驶	9	74.43	655	88.31	31	461
可疑人员	124	43.60	3,249	42.00	426	-682
可疑车辆	29	50.68	789	42.00	105	-911
警报器	55	24.88	2,805	26.00	177	198
总数	672				2,518	15,321 =255 小时

* 显著效能的计算方式为，将每位警官的平均处警时间（整体警务工作）减去每位直升机警官的处警时间，再乘以通过警用直升机处理案件的警官人数。计算结果会比实际效能值稍微偏低，因为在总体出警量的计算中，直升机出警数量同样被纳入了计算。

附表：运营价值：通过时间（小时）和资金（美元）展示效能价值和效益价值

案件类型 (“刚刚发生”及“正在 在进行”的案件)	效能		效益		运营价值
	小时	美元	小时	美元	美元
人身攻击	7.03	357	80	4,414	4,771
武器	116.78	5,932	480	26,482	32,414
家庭暴力	23.30	1,184			1,184
失踪人口	16.33	830	280	15,448	16,278
噪音	8.50	432			432
扰乱治安	1.25	64			64
人际纠纷	14.38	731			731
醉酒	1.83	93			93
人身安全检查	-10.72	-545			-545
协助其他机构	17.20	874			874
协助消防机构	36.38	1,848			1,848
入室盗窃(住宅)	24.22	1,230	320	17,654	18,884
偷盗	16.93	860	160	8,827	9,687
破坏财物	22.30	1,133			1,133
小偷	24.97	1,268			1,268
交通	4.73	250			240
不清醒状态驾驶	7.68	390			390
可疑人员	-2.97	-151			-151
可疑车辆	-11.68	-593			-593
警报器	9.80	498			498
其他*	10.12	514			514
总数	338.36	17,189	1,320	72,823	90,014

* “其他”案件是指在表格所列出的20项案件类型之外，涉及警方参与解决的事件。

警员们节约下来的338.36小时的效能价值17,189美元，而1,320小时效益的侦查价值为72,825美元，两项相加的总价值大约估算为90,014美元。按672件次的出警量来计算，平均下来警方在处理每种类型案件上耗费的时间总计为543小时，而剩余的457小时（1,000 - 543）价值23,334小时，这一数值得加到运营成本上。因此，本年度的运营总成本是354,344美元（\$331,010 + \$23,334）。而上述计算的收益为90,014美元，换句话说，占成本的25.4%。

因此，这里我们可以写入报告的结论是，显著实际收益为各项成本之和的25%。而且，我们还可以看到效能和效益的具体情况，将两者通过货币价值进行量化后，收益约为138,463美元，也就是占到成本的42%。我们需要将这些货币化的收益结合警用直升机的特点和特色进行考虑，包括它的速度、空中视角、对特殊地点进行搜索的方便性和安全性（比如房顶、铁路沿线、河岸等），以及强大的夜间照明系统对于提高能见度性和市民及警官安全保障的作用。

案例：关于委派专人使用车辆和整体调度使用车辆的成本核算⁴

马纳迪县（Manatee County）警察局就将车辆委派专人使用和整体调度使用车辆的成本效益进行核算，以下数据均由该警察局警长办公室提供。究竟是将车辆委派专人使用更为经济实惠还是整体调度使用车辆的理念更好，本项研究为我们提供了合理的成本核算。

下面的表格比较了两种车辆调配方式的成本。除了车辆的购置开销、油费、维护费用，还得计算其他衍生成本，最值得关注的就是一位警官在执勤始末交接车辆、查验仪器所耗费的时间，据估算，每位警官在交接车上耗费的时间约为40分钟，而这意味着每年要耗费11天的巡逻时间。本项分析将计入和不计入这项成本都纳入了考虑。

委派个人使用车辆项目		整体调度使用车辆	
机动车使用寿命	5 年	机动车使用寿命	1.8 年
使用寿命里程数	90,000	使用寿命里程数	90,000
使用周期成本		使用周期成本	
机动车成本	\$22,488	机动车成本	\$22,488
维修救援	\$3,750	维修救援	\$6,063
日常维护	\$12,285	日常维护	\$21,297
油费	\$19,731	油费	\$19,688
		每年耗费的警官时间成本*	\$7,899
总成本	\$50,754	总成本	\$65,309
每英里成本	\$0.56	每英里成本	\$0.73
		不计算耗费的警官时间成本	\$0.64
每英里的汽油及维护成本	\$0.36	每英里的汽油及维护成本	\$0.46

* 耗费的警官时间计算是假设每次接车要耗费25分钟，每次交车要耗费15分钟，以移交器材和查验车况。

这张成本核算表显示出，对于警队而言，将车辆委派给专人使用是最为经济有效的做法，具体到财政成本，专人用车每英里成本仅为0.56美元，而调度用车则需要0.64美元每英里，也就是说前者能节约14%的资金；而要是再算上损耗的工作时间成本，那么节约额则高达30%。对于大多数还没有开始采取委派专人用车做法的机构而言，执行这一项目最大的障碍在于添置车辆所带来的首要成本。

该研究还指出了采取委派专人用车做法的其他好处，具体说来如下：

- 由于警官可以开警车上下班或者驾车执行任务，因此专人专车计划可以提高道路上的见警率，这对于犯罪和潜在罪犯具有一定的威慑作用。
- 在紧急警情下，由于警官们不用先赶到指定区域去取车再前往现场，因此这种做法可以带来更快的响应速度。如果是严重事件，警官们不用驾驶私家车先赶往指定地点集合以取得装备和听取事件说明，而是可以直接从家里出发奔赴现场进行处理。如果没有该项目，那么刑事案件案件的侦查员需要先赶到警察局取得各类器材，然后才能赶赴犯罪现场，这很可能就浪费了宝贵的侦查时间。
- 可以将警车直接开回家使得巡逻警官、刑事犯罪侦查员以及指挥官们可以在任何时间、任何地点对必要情况进行响应。在比较严重的情况下，比如重大灾祸，警察局可以即刻调动所有在职警官，召集他们在接到通知后直接赶赴任务地点。
- 将警车开回家还可以加大执法力度。警官们在非工作时间驾驶警车时，如果发现严重违法犯罪活动，必须马上采取行动。
- 为值勤警官提供了更加坚固的后盾。在紧急情况下，如果距离案发地最近的执勤人员位于另一个警区或者亟需援助，而某位非工作时间的警官恰好就在附近，那么就可以被临时委派前往现场。
- 由于专人专车，警官们会较为精心地维护车辆并保持清洁。
- 实行专人专车项目会一定程度上提高警官们的士气。

小结

虽然成本核算并不是基于证据的决策中唯一获取数据的途径，但是一般情况下，它确实是更为普遍使用的一种工具。简单说来，成本核算能为我们提供三种必要信息。首先，如果使用得当，它能帮助我们把握所期待的结果和行动方案和战略计划的远景目标和近期目标相结合，从而使我们把关注点放在某项活动是否在本机构职责的大框架内这一问题上。

其次，成本核算能够帮助我们关注究竟哪些才是带来实际成本开支的条目。多数情况下，“粗略估算”或者基于惯例的成本核算会忽略了很多和我们活动相关的附加成本，比如，成本核算中常常会忽略了需要支付的利息或是增聘人手带来的开支。通过进行细致分析，我们才能进一步确定自己已经将所有这些项目都纳入了考虑。另外，我们应当把成本分析的结果提供给同事和其他相关人员，以便他们从旁观者的角度确定是否还有被遗漏了的成本消耗项目。

最后，由于成本核算是一个较为客观且重点明确的过程，因此它可以提供一种透明、公正的机制性方法以帮助我们做出最终选择。每项成本核算所立足的假设，以及每个项目所产生的价值，都经得住推敲。这一过程的透明性也让决策更经得住质疑：即使由独立第三方进行决策，也可以得出一样的结论。而且，和完全基于价值的决策不同，如果人们要对基于证据的决策提出质疑，首先得提出其他方案或分析，才能对公布的分析结论进行合理的驳斥。

即便是某个人能够提出其他证据，净利益（net benefit）依然存在，因为这个证据只是对某个情况提供了更为精准的评估，而最终，我们将得到一个更好的基础以做出决策。

注释及参考文献

1. 沉没成本 (sunk costs) 是指由于过去的决策已经发生了的, 而不能由现在或将来的任何决策改变的成本。——译者注
2. Treasury Board of Canada (2007) *Canadian Cost-Benefit Analysis Guide: Regulatory Proposals*. Ottawa: Government of Canada. Catalogue No. BT58-5/2007. <http://www.tbs-sct.gc.ca/rtrap-parfa/analys/analys-eng.pdf>
3. 案例来源: Whitehead, Paul C., “Operational Value of Police Helicopters: A Cost-Benefit Analysis,” *International Journal of Police Science and Management*. (2002) Vol 4, p. 233
4. 案例来源: 马纳迪县警长办公室备忘录, “re: Assigned Vehicle Costs”, November 30, 2007

进行决策

运用证据

无论是在私人生活里还是专业工作中，我们都需要做出各种决策。在大多数情况下，我们是根据所接受的教育、传统智慧或是常规做法来做决定。通常，对普遍性的做法提出质疑只会让我们绕个圈子又回到原点；然而，也有不少情况，传统做法和常识发挥不了作用，我们可能无法取得所希望的结果，或者我们的做法带来的效益差强人意。但是，出于某些原因，人类总是抗拒改变。我们是一种很保守的物种，对于不停重复的事情觉得习以为常，即使我们对结果并不那么满意；不过呢，正如《酒瘾戒除手册》里所陈述的，“所谓精神失常，就是不断重复同一件事情，却指望能够从中得到不同的结果。”

从历史的角度出发，我们可以原谅那些在热衷于进行各种耗时仪式的决策者们，毕竟，在求雨法师们看来，如果人们的舞蹈足够取悦于神，那么最终会雨从天降。



基于证据的决策整个过程是公开透明的——决策不再是暗箱操作，而是一个有目共睹的逻辑清晰、证据合理的过程。

但是，由于当今的天气预报已经足够准确，因此求雨这一专业可能不再吃香了。

由于气象学上已经积累了足够多的系统知识，从而使人们能够较为准确地预报本地气温、降水和其他天气现象。气象学家们通过进行科研和其他形式的系统性研究，完成了这一工作。

系统研究和数据采集是科学的基础所在，然而对它们的过度依赖却缚住了我们对其他学科进行探索的手脚，这点让人遗憾，因为其实我们可以利用当前的大量经验证据（empirical evidence）来做出更好的决策。而且，如果没有现成的分析结果，通常我们自己来进行分析以进行决策也不是那么困难的事情。一个人完全没必要成为科学家才能进行这类分析，而是可以从经验总结中收集到可靠的主要证据。

基于证据的决策会帮助您做出更好的选择。由于它的整个过程是公开透明的，因此外部人员也有机会获悉决策的依据是什么。决策不再是暗箱操作，而是一个有目共睹的逻辑清晰、证据合理的过程。

基于证据的决策是以当前就警务工作所开展的最好的研究和信息为依据，制定出指导方针并且对警察局、各支队以及个人进行评估。

我们并不是说对于所面临的任何问题，您都能找到一个最为理想的解决之道；但是，基于证据的决策能够帮助我们确定哪些方案和措施是起不了作用的，也就是说，您可以避免尝试一些会造成事态恶化的“新举措”。不过，多数情况下，查阅现成的证据或是自行收集数据，都有助于您找到更加见成效的方法路径。每个人都可以根据证据进行推论。

推理总结是一项基本的人类技能，进行分析式思考就像是画画或着骑自行车一样的技能，可以被传授或者被习得，并且会随着练习而不断熟练；不过，它也像类似空手道这样的技能，需要不断进行实地演练和运用。因此，本手册也配备了一本练习册，您可以通过里面提供的案例研究来练练手。

总之，您如何才能通过这本手册，归纳出自己的策略以做出基于证据的优秀决策呢？简要说来，可以通过以下几个步骤。

如果没有提出恰当的问题，那么即使数据成堆，也无法从中得出答案。

确认和限定问题

本书的前三章关注点在于如何确认实际问题。要是提出的问题不在点子上，那找再多的数据也无济于事。我们还反复强调，找准了问题，还得把它们限定在合适的框架内。理想情况下，您应当根据所在机构的方案或战略计划来提出问题，这有助于您紧扣本单位的主旨和目标来考虑具体事宜。很多机构失败的主要原因就在于他们没有把关注重点放在自己的主要职责上，而是希望自己能够三头六臂搞定一切，而这根本就是不现实的。

如果您没有机构或战略计划，那么下一步最好是对所面临的问题进行深入研究，并回答以下几个基本问题：

- 我们为什么计划做这件事？
- 可能带来哪些结果？
- 这一行动和本机构的职能有无联系？
- 这一行动会给我所在的单位或者所服务的人群带来哪些益处？
- 还有其他更好的成本—效益或者成本—效能方案吗？
- 这一行动有什么长期或者短期的影响吗？
- 如果我们要执行这一行动，那么还需要动用哪些资源？

如果您打算着手的事情超出了本机构的传统职能范围，那么请您制定一个重点突出的行动计划，以对您即将展开的活动提供支持或证明。

一旦您确定并能证实找准了症结，就罗列列出可行的行动方案，一般情况下两到三个为佳。不过，也有些情况是各个行动方案及其优势似乎并没有显著差异，这时您可以考虑进行环境扫描或SWOT分析。如果事情至关重要，那您应当考虑向第三方机构进行咨询。

收集证据

通常，最佳的证据来源就是您所在的机构本身，您可以从接处警纪录、财务报表等档案记录中，得到宝贵的内部资源。一般来说，内部数据是了解或衡量单位现状的基础或基准。

在机构外部，我们也能够通过各种信息渠道获得资讯，专业联盟或行业协会就是很好的例子；供销商也可以提供相关产品在功效、使用寿命以及维护成本等方面的比较数据；互联网虽然有很多垃圾信息，但是同样也有一座座的信息金矿等待您去发掘，您只要掌握如何利用您喜爱的搜索引擎来尽可能过滤掉无关内容就好了。另外，也别怯步于去找其他执法机构寻求帮助，在英国，国家数据办公室（Office for National Statistics）负责保管犯罪案件的信息，同时，它还提供了很多的年度报告及研究；与之相类似，美国联邦调查局（FBI）以及司法数据局（Bureau of Justice Statistics）也是信息宝库。

图书馆以及本地的大学或学院也是绝佳的信息来源。图书馆管理员可以登录网络数据库搜索学术文献和其他的专业资料，其中一些可能是我们自己无从获取的。请记住，图书馆管理员可以成为您的最佳帮手，因此，请和市立图书馆的管理员保持联系，或者到一所本地高等学府去寻求专业意见。

图书馆管理员还可以帮助您有效地进行数据库检索。很多省份和省级机关都会收集并提供区域性数据，不过，虽然很多数据都是公开的，但是也有一些数据仅供被授权的机构使用。如果您就职于一所公共服务机构，那么很有可能您所在的单位就是这些被授权的机构之一。加拿大统计局（Statistics Canada）网站同样提供了大量的宝贵数据。

请记住，图书馆
管理员可以成为
您的最佳帮手。

一些大学和学院具备犯罪学实验室和研究组，同样的，您通常也能够通过网络检索来获取相关信息，或者向图书馆管理员寻求帮助。

不过，您也要记住，并不是所有的证据都具备相同的价值，因此，不要害怕提出质疑或反对，特别是当您的想法和本部门或同事的经验相左时。还有，虽然不是必然，但是看起来像天上掉馅饼这样的好事通常确实也不会是真的，您要自问消息来源是否可信，比如，某机构在提供数据是开诚布公还是有所隐瞒，或者它是不是有自己的计划？某份研究或出版物是否经过了外部机构的核查程序？

整理证据

一旦您收集好了证据，那您就得把它们按照一定的顺序进行整理。我们很容易通过一份电子表格进行成本核算，而其他材料则可以通过列表来进行呈现。您得确保记录了信息的来源以便倒查您是在哪里找到它的，这样做的目的在于，如果有人质疑它的真实性，您可以告诉他们来源。

展现数据的法则之一是，要将其放在一定的情形下。请记住，单独的信息是毫无意义的，除非它和其他事情相关联，因此，您得问自己：“相较什么而言？”

对于一所都市警局而言，10分钟的平均响应时间是否足够了？我们是否应当集中精力处理优先案件？我们可以断言，市长和市政委员会、警察理事会以及本地社区的其他人员也会提出这些问题。

展现数据的法则之一在于
将它放在一定的情形下。

在一辆警务战术车辆上投资一百万美元合适吗？这究竟是一项一本万利的保险性投资抑或根本就是资金的巨额浪费？我们在培训上的投入够充足吗？在近期内培训要求会发生变化吗？如果会，是什么变化？只有在我们按一定标准进行了比较之后，才能回答这些问题。在市面上一定货物和服务的价格区间是什么？对于业绩的行业基准或规范有哪些？有哪些好的做法您可以用来和所在单位或部门的做法相比较？

您进行以证据为基础决策的频率越高，您就会觉得它越简单。知识在于积累。

回顾决策过程

一旦您完成了分析，那么就该对整个决策过程进行回顾了。您从中获悉了什么？整个过程是否逻辑清楚？是否可以更具效率？您进行以证据为基础的决策频率越高，您就会觉得它越简单。知识在于积累。

您很快就会确定信息的最佳来源，您还将懂得如何才能让整个流程更加高效，以及如何避免误入歧途。

虽然总的说来，以证据为基础的决策会比其他方法要更耗时，但是它有自身的优势，做出的决策以有力证据为支撑，因此在面对审视的时候，也更容易站得住脚。为了我们所服务的社区，在考虑问题和实施策略的过程中，我们确实有必要做到进一步以证据为基础。

如果您递交给市政委员会的申请有着强有力的外部证据为依托，那么将更有可能得到正面的决策结果；那些展示了现有的成功案例或可信的预期成果的呈述报告也更加具有说服力。最后，如果有人对您提出质疑，您可以利用手中的证据来为自己的提议辩护，如果对方持反对意见，那么您也可以要求他们用数据来说话。

其他人对于《正确决策》的评论

加拿大卑诗省萨尼治市警察局 鲍勃·唐宁 (Bob Downie)
局长

《正确决策》一书是在加拿大警察局长联盟 (CACP) 警务研究基金会的委托和资助下面世的。我们对于能与任·加里斯 (Len Garis) 先生和他的团队合作感到非常骄傲，我们所完成的这本手册、配套的练习册以及相关案例研究，将提供给整个加拿大的警务机构使用。我们也得知，这一作品的前身其实是供消防人员参考的决策手册，因此，对于它能够改编成为供警方参考的读物，我们感到非常高兴。

对警务研究重要性的重视从未像当今这么高。随着对于有限资源的需求增高，以及对我们所做决策、所立项目、服务质量等方面问责力度的加大，我们必须要有能力解释我们为什么做某事以及做事方法，同时还要能说服各方我们所做出的决策是正确的。要做到这些，我们需要依靠信实的证据而不是依靠直觉来进行决策。

《正确决策》一书揭开了笼罩在以证据为基础决策上的神秘面纱，向我们展示了寻找和依靠证据进行决策的重要性，以及不要畏惧打开本机构的大门迎接各类研究成果。曾几何时，警务机构对于研究方法有时会持不信任态度，因为我们总觉得自己身处他人无法衡量和理解的特殊环境，但是这样的日子，也已经过去了。本书向您展示了哪些是您在找寻证据的过程中应当提出和解答的问题，哪些是需要采取的步骤，以及您可以通过哪些途径来找到可信且可靠的证据。

您所在机构的任何人都可以使用本手册，它提供了一个框架来帮助大家制定可靠的方案来做出以信息为依托的决策，它还强调了将项目计划和本机构战略或工作方案以及相关远景目标和近期目标相结合的重要性。这一方法可以在您任职机构的任何层级上得到使用，并能够帮助您的员工进一步拓展工作思路。

在确定加拿大警务的重点研究方向上，CACP研究基金会做了大量的工作。感谢全国各地不同层面的警务、学术以及政府机构所做的努力，为我们做出正确决策创造了机会。这一手册将帮助我们把握机会，最为充分地利用这些研究成果，作为回报，我们也将为人民提供最高品质的警务服务。

加拿大卑诗省阿博兹福市警察局鲍勃·里奇 (Bob Rich)
局长，法学硕士，警队荣誉勋章

如何才能做一个够格的警察局长？一个重要的因素就是具备做出正确决策的能力。我还意识到一点，随着职级的提升，你所在的位置越高，你的职责范围越难定义。比如，如果你在街上拦下了一个违法人员，你需要考虑的是应当按照那条法律为依据来逮捕他，你几时行动、如何行动都是清清楚楚的；但是当你面对的是管理一个警务机构的时候，那连应当做哪一类决策这样的问题似乎都是很不明朗的。

本手册列出了一系列步骤以帮助我们做出缜密且可靠的决策。通过阅读本书，您将获知或重温如何循序渐进地执行一个决策程序以增加成功的几率，不止是使您的决策“更加英明”，而且还能提高周边相关人员对您的支持度。看起来，花点时间来阅读《正确决策》将会是您能即刻做出的最佳决策！

美国德州阿灵顿市副市长长官，前警察局长，塞隆·鲍曼 (Theron L. Bowman) 博士

《正确决策》一书通过逻辑清晰、简明易懂、有理有据的方式向读者展示了决策过程。保罗·马克西姆教授、任·加里斯局长兼教授、戴雷·普利卡斯荣誉教授以及法律专家莫娜·戴维斯携手，以简洁明了的语言，向读者展示了如何提出问题，以及如何获取数据。本书也阐述了如何制定基础性和战略性的方案，并且还进一步讨论了研究方法、数据分析以及实验设计，在每一个章节，它仍以实际警务工作为案例，因此，即使是没有理论基础的读者，也能够理解相关理论；而那些重要的术语，也得到了清晰的定义。

那些以本书为基础，了解并应用以证据为基础的决策的读者们将很快获取较强的能力来更好地执行方方面面的警务工作程序并取得预期成果，这些工作程序包括进行初步和专业的犯罪及行政调查、目击者确认、问话及讯问，以及犯罪现场调查等等。这本书不仅可以为警务工作的新手提供多样化的资讯，而且也适用于有经验的警官、侦查员以及分析师，帮助他们在当今对“以证据为基础”呼声高涨的大环境下，能够更上一层楼。本书开具的温和经济学药方，包括成本—效益分析以及机会成本等，让我作为市政长官很想确保警察局长以及其他部门官员人手一本。

这本《正确决策》确实是很好的以证据为基础决策的入门读物，它应当和所配套的练习册一起，成为司法领域工作者的基础读物。

做出正确决策

作为警界或其他公共服务领域的专业人员，您可能每天都需要在平衡需求和可用资源的基础上做出关键的决策。那么，您该如何做出这类决策呢？同时，您该如何证明这些决策是正确的呢？

在本书中，保罗·马克西姆教授、消防局长任·加里斯教授、戴雷·普利卡斯荣誉教授以及法律分析师梦娜·戴维斯，将和您一道探讨什么是基于证据的决策，我们为什么需要做到以及如何做到基于证据的决策。

其他人对于《正确决策》的评论

请翻阅封底里页以查看完整评论。

对于任何需要进行决策的人员，我都推荐把《正确决策》作为一本非常好的参考读物。本书把各类材料按照简明易懂的方式进行了归纳，就决策过程所涉及的各个不同方面提供了辅助资料。

皇家骑警“E”分部指挥官 克雷格·盖雷斯 (Craig J. Callens) 副警监，警队指挥官荣誉勋章

《正确决策》一书揭开了笼罩在以证据为基础决策上的神秘面纱，向我们展示了寻找和依靠证据进行决策的重要性，以及不要畏惧打开本机构的大门迎接各类研究成果。本书向您展示了哪些是您在找寻证据的过程中应当提出和解答的问题，哪些是需要采取的步骤，以及您可以通过哪些途径来找到可信且可靠的证据。

加拿大卑诗省萨尼治市警察局 鲍勃·唐宁 (Bob Downie) 局长

本手册列出了一系列步骤以帮助我们做出缜密且可靠的决策。通过阅读本书，您将获知或重温如何循序渐进地执行一个决策程序以增加成功的几率，不止是使您的决策“更加英明”，而且还能提高周边相关人员对您的支持度。看起来，花点时间来阅读《正确决策》将会是您能即刻做出的最佳决策！

加拿大卑诗省阿博兹福市警察局鲍勃·里奇 (Bob Rich) 局长，法学硕士，警队荣誉勋章

那些以本书为基础，了解并应用以证据为基础的决策的读者们将很快获取较强的能力来更好地执行方方面面的警务工作程序并取得预期成果。这本书不仅可以为警务工作的新手提供多样化的资讯，而且也适用于有经验的警官、侦查员以及分析师，帮助他们在当今对“以证据为基础”呼声高涨的大环境下，能够更上一层楼。

美国德州阿灵顿市副市长长官，前警察局长，塞隆·鲍曼 (Theron L. Bowman) 博士

这是一本对于警务工作的未来发展很有帮助的工具书，可谓为那些思考着明天的警察们量身定制。任何读者都可能会得出“基于意见”的警务将淡出历史这样的结论，而基于证据的警务实践将成为一种全新的、公正透明的基础性原则。

劳伦斯·谢尔曼 (Lawrence Sherman) 教授，《基于证据的警务》(1998) 作者

<http://www.policefoundation.org/content/evidence-based-policing>