



# 线下调查实战- 以CGSS为例

胡以松

中国人民大学中国调查与数据中心  
中国人民大学统计学院



# 内容

---

1. 线下调查简介
2. CGSS项目概述
3. 调查组织实施
  - 组织模式
  - 人员管理
  - 进度管理
  - 成本管理
4. 调查质量控制

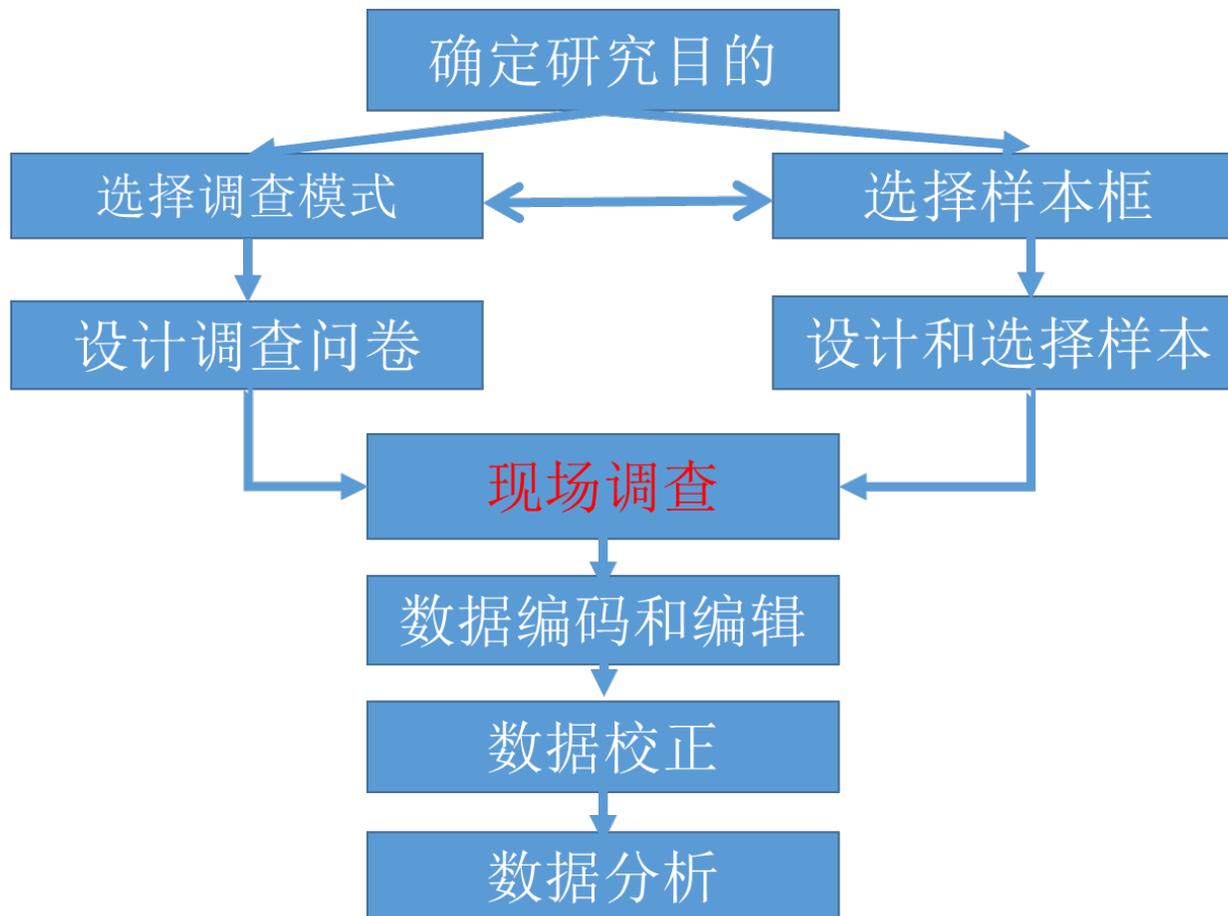


# 1. 线下调查简介

1.1 线下调查定义

1.2 调查模式举例

## 抽样调查的一般流程





## 1.1 线下调查定义

---

- **线下调查**：又称面对面访问（**Face to face interviewing**），是相对于线上调查（网络调查）而言的，需要访问员和受访者见面调查。一般情况下，访问员通过念读问卷询问问题，受访者回答后，访问员在纸质或电脑问卷上记录问题的答案。
- **应用情况**：国内目前大型全国性学术性调查主要的调查方法，如：
  - ✓ 中国综合社会调查（**CGSS**）、中国教育追踪调查（**CEPS**）
  - ✓ 中国健康与养老追踪调查（**CHARLS**）、中国家庭动态跟踪调查（**CFPS**）
  - ✓ 中国社会状况综合调查（**CSS**）
  - ✓ 中国劳动力动态调查（**CLDS**）
  - ✓ 中国家庭金融调查（**CHFS**）
  - ✓ 中国健康与营养调查（**CHNS**）



## 1.2 调查模式举例

举例	共同优点	共同缺点	特点
电脑辅助的个人面访 (CAPI)	<ul style="list-style-type: none"><li>• 可以实时获取数据</li><li>• 可以实时监控调查进度</li></ul>	<ul style="list-style-type: none"><li>• 软硬件准备和风险</li><li>• 招募和培训压力</li></ul>	<ul style="list-style-type: none"><li>• 应答率高</li><li>• 费用高</li><li>• 复杂问卷</li></ul>
电脑辅助的电话访问 (CATI)	<ul style="list-style-type: none"><li>• 增加实时质量控制手段</li><li>• 方便沟通管理</li></ul>		<ul style="list-style-type: none"><li>• 应答率低</li><li>• 费用低</li><li>• 短问卷</li></ul>
电脑辅助的网络访问 (CAWI)			<ul style="list-style-type: none"><li>• 代表性差</li><li>• 质控难</li><li>• 费用低</li><li>• 短问卷</li></ul>



## 2. 中国综合社会调查(CGSS)概述

---

- 中国历史最长的的大型连续性学术性社会调查项目之一
- 目的：通过收集追踪性数据、总结社会变迁的长期趋势
- 创始人：李路路教授和边燕杰教授
- 两期：第一期2003-2009年，第二期2010-2021年
- 问卷：社区问卷和个人问卷
  - ✓核心模块：调查全部样本，年度调查，固定不变
  - ✓主题模块：调查全部样本，5年重复一次，两次调查内容重合率>80%
  - ✓附加模块：调查1/3或1/4随机样本，不确保重复周期和内容



# CGSS问卷内容

---

- 社区问卷

- ✓ 基本信息、基础设施及公共设备、社会经济情况、劳动力迁移、养老及医疗保险覆盖情况、物价水平、选举情况、社区组织、社区历史

- 个人问卷

- ✓ 核心模块：家庭成员、社会人口属性、健康、迁移、生活方式、社会态度、阶级认同、政治参与行为与态度、个体认知能力、劳动力市场、社会保障、家庭、创新与创业

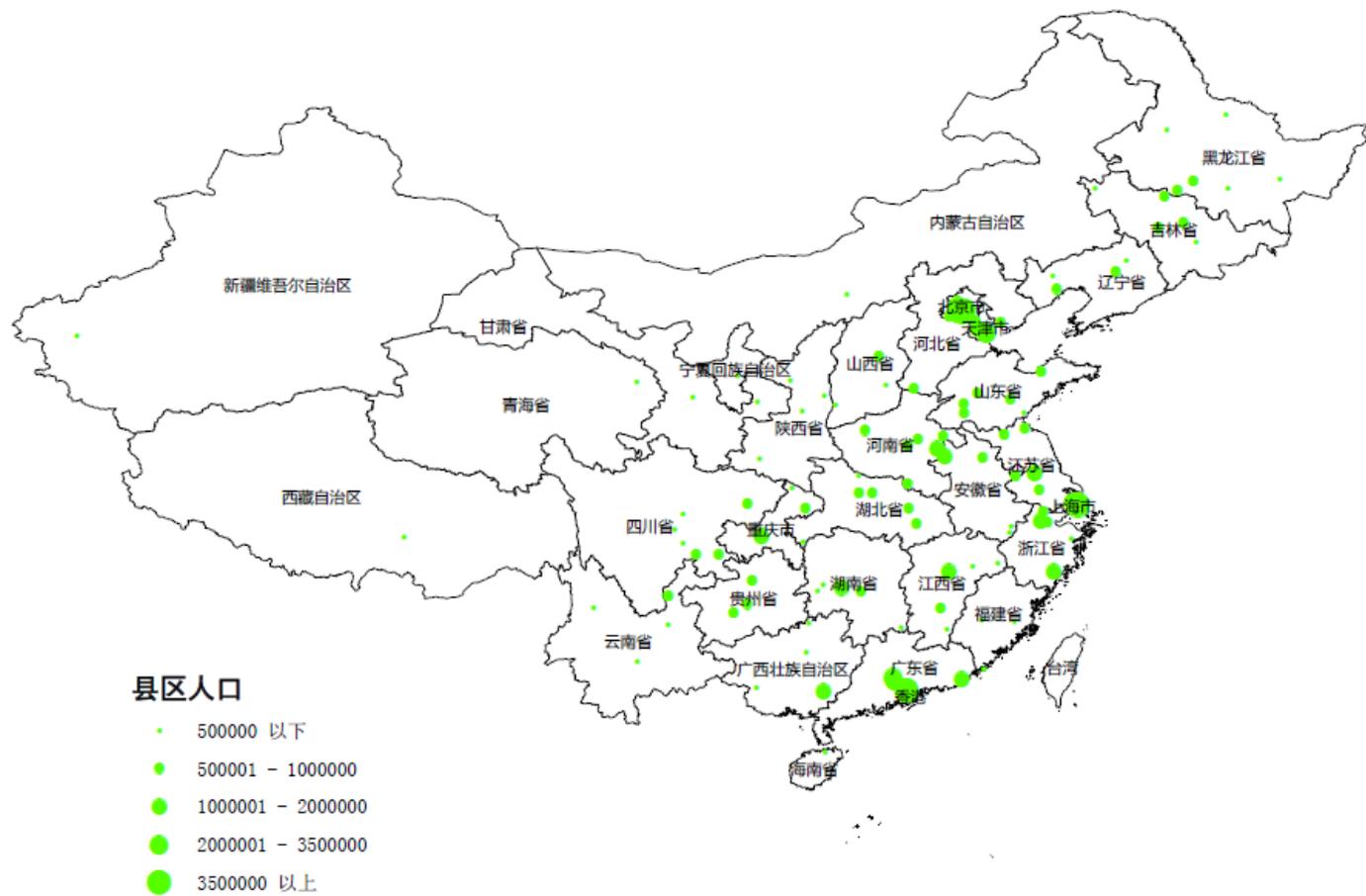
- ✓ 附加模块：

- 2011：住房；2012：主观幸福感，社会慈善；2013：公共服务满意度，2015：家庭能源消费、法制文明；

- 2017：社会网络、政府角色；2018：家庭能源消费、宗教、全球化；2019：体育消费；2021：健康、环境以及疫情相关



# CGSS抽样



- 多阶段PPS随机抽样
- 全国5个大都市+100个县区
- 480个村居
- 12000 家庭
- 12000个人





# 住户清单列表示例

CLASS 村/居委会住户清单及抽样名单

序号	建筑物编号	所在楼层	楼层住宅数量	住户编号	对应分图	实际门牌号	抽样标识	有无老人*	户主姓名	住宅地址	备注
1	1	3	2	1	3-1	302					
2	1	3	2	2	3-1	301					
3	1	2	2	1	3-1	202	√	1	李明	人民大学林园 5 号楼 2 单元 202	门口墙上写着“热爱毛主席”
4	1	2	2	2	3-1	201	○ <sub>3</sub>	2	张敏	人民大学林园 5 号楼 2 单元 201	白色防盗门
5	1	1	2	1	3-1	102					
6	1	1	2	2	3-1	101					
7	2	2	2	1	3-1	202					
8	2	2	2	2	3-1	201					
9	2	1	1	1	3-1	101	√	3	赵康	人民大学林园 39 号楼 3 单元 101	横批是“合家欢乐”
10	3	6	3	1	3-1	无					

\* 有无老人一栏选择以下情况中的一种，只需要填写序号即可：1. 有 60 周岁及以上的老年人长住；2. 无 60 周岁及以上的老年人长住；3. 无法进入；4. 拒绝回答；5. 无法联系或空户；6. 其他情况。

制表人签名： 李明 共 91 页 第 6 页



# CGSS数据开放共享

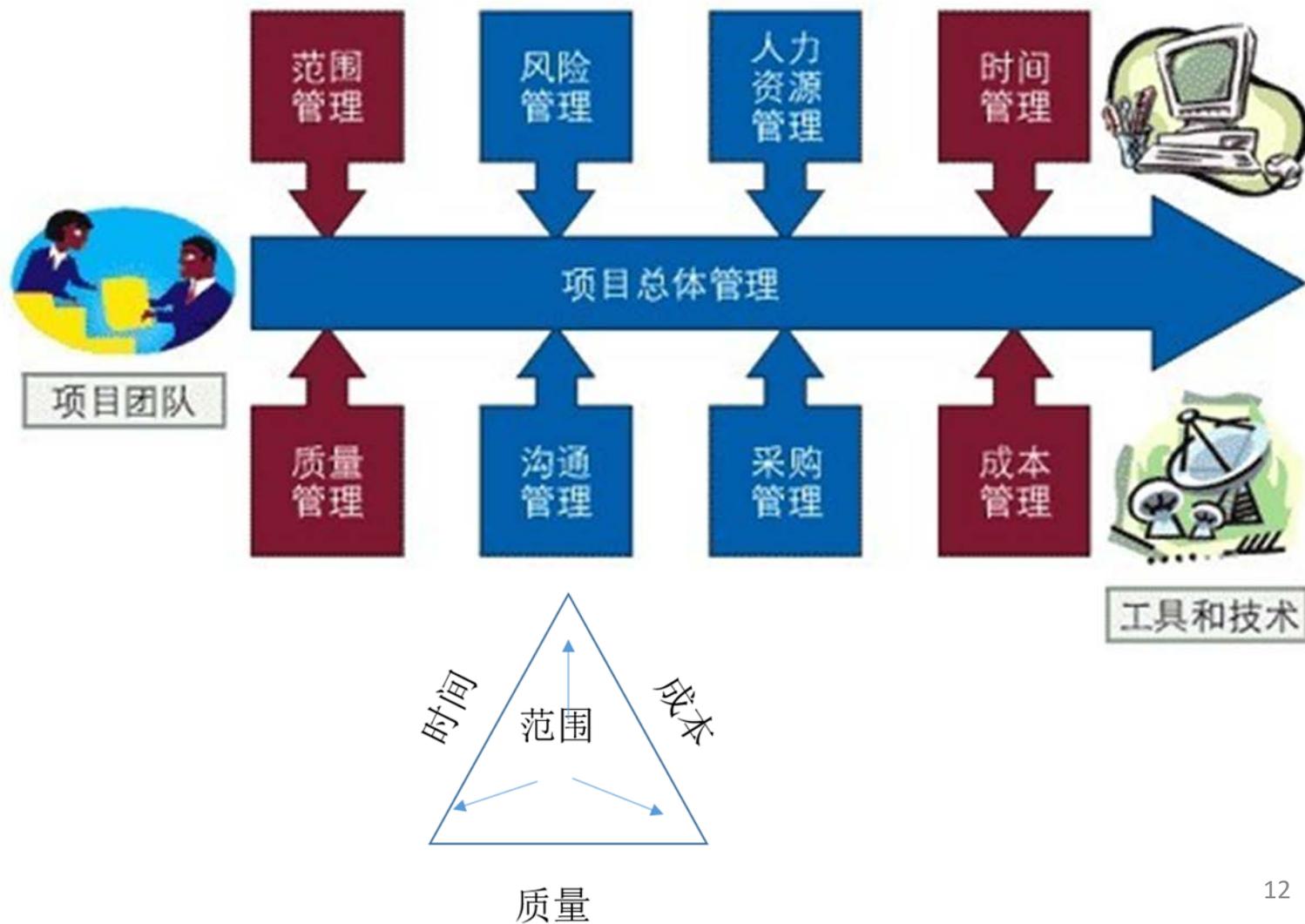
---

- CGSS开创了我国社会调查数据开放共享的先河，在CGSS的推动下，建成了中国第一家向全社会开放的社会科学调查数据库——中国社会调查开放数据库（CSSOD），并在其基础上建设国家自然科学基金重点项目”中国国家调查数据库（CNSDA）”  
<http://cnsda.ruc.edu.cn/>

- 至2019年底，CGSS数据用户超过了40000人，基于CGSS数据发表期刊论文超过3000篇，博硕士学位论文705篇，学术著作27本。除了科研外，并被广泛应用于教学、政府决策、商业等领域。

## 3. 调查组织实施

- 3.1 调查组织模式
- 3.2 调查人员管理
- 3.3 调查进度管理
- 3.4 调查成本管理





## 3.1 调查组织模式

---

### 3.1.1 不同的组织模式

- ✓ 行政工作网络
- ✓ 社会招募组织团队
- ✓ 建立合作性的工作网络
- ✓ 外包方式

### 3.1.2 调查中的行政资源



# 建立合作性的工作网络

- 如：中国综合社会调查项目（CGSS），2009年建立的中国社会调查网络（CSSN）。



- 中国人民大学
- 复旦大学
- 中山大学
- 南开大学
- 浙江大学
- 吉林大学
- 东南大学
- 厦门大学
- 西安交通大学
- 云南大学
- 兰州大学
- 河海大学
- 上海大学
- 河北大学
- 山西大学
- 山东大学
- 贵州民族学院
- 云南师范大学
- 沈阳师范大学
- 西南大学
- 郑州大学
- 中南大学
- 华中师范大学
- 安徽师范大学
- 广西大学
- 黑龙江社科院
- 四川社科院
- 江西社科院
- 新疆财经大学
- 广州大学
- 宁夏大学
- 上海财经大学
- 上海社科院
- 华东师范大学
- 内蒙古大学
- 四川大学
- 太原理工大学
- 山西师范大学
- 三亚学院
- 贵阳学院
- 天津社科院
- 南京大学
- 江西财经大学
- 西北民族大学
- 西南财经大学
- 河北师范大学
- 福州大学



## 3.1.2 调查中的行政资源

---

- 目前情况
  - ✓ 高校执行的面对面调查项目，也都借助行政资源，才能够有效开展。
  - ✓ 举例：CGSS（各地高校联系村居委会）、CHARLS（CDC系统、北大团委）、CHFS（中国人民银行系统）
- 行政资源在面访抽样调查中的作用
  - ✓ 协助抽样过程，如样本框的制作。
  - ✓ 协助访问过程中的受访者动员和解释。
- 行政资源在调查中可能存在的问题
  - ✓ 每个村、居委会配合力度会有差异。
  - ✓ 村、居委会可能会影响调查质量。
  - ✓ 村、居委会可能会影响调查进度。



## 3.2 调查人员管理

---

### 3.2.1 团队组建（学术性机构开展的调查）

- ✓ 调查团队构成

- 研究团队

- 执行团队

- ✓ 实地调查人员招募途径

- 校园招募

- 网站招募

- 自媒体招募

### 3.2.2 团队建设（访员培训）

- 实地调查人员招募要求

- ✓ 一般要求

- 诚实与认真
- 优秀的学习能力
- 规则意识
- 兼职人员

- ✓ 访问员特殊要求

- 良好的读写能力
- 与陌生人良好沟通的能力
- 较强的抗压能力(能吃苦)
- 弹性工作时间

人口学特征：年龄、性别、民族和教育水平等。

其他影响因素：调查经验、调查主题相关的专业培训。



## 3.2.2 团队建设

---

- 培训方式
  - ✓ 讲解和问题解答
  - ✓ 角色扮演
  - ✓ 模拟（试）调查
  - ✓ 调查员评价
- 培训内容
  - ✓ 调查员的职业道德要求
  - ✓ 调查员的知识及技术要求



# 调查员评价

---

- 培训时评价
  - 态度和能力评价
  - 培训考试
- 现场调查评价
  - 客观数据评价（合格问卷率等，**过程数据**）。
  - 督导主观评价



# 调查员的职业道德要求

---

- 1、尊重受访者的权利，包括知情权、隐私权和拒绝权。
- 2、认识到调查工作的重要性和意义所在。
- 3、诚实负责，不弄虚作假。不因个人问题影响全局。
- 4、严格执行每项调查的技术规范，不得擅自更改。
- 5、文明礼貌，耐心细致，和气待人。
- 6、不得在调查中进行与调查不相干的活动。



# 调查员的知识及技术要求

---

- 1、主持者：何人、什么单位主持这项调查。
- 2、目的：为什么要做这次调查？需要调动调查员的积极性。
- 3、总体：了解调查什么人，以及选择这些人的主要依据。
- 4、抽样程序：调查过程中的哪些行为会破坏样本的随机性。
- 5、工作流程：掌握问卷调查的整个流程。
- 6、调查管理：时间要求、质控标准、工作待遇、安全考虑等。
- 7、**一般访问技术(GIT)**：从容消除受访者各种顾虑及拒访等。



# 编制参考话术

## ✓核心思想

- 你是专业的调查员：身份的真实（有发文、函件、之前的联系人等）
- 你代表的是一个权威或知名的机构
- 受访者的参与是十分重要的，收集的信息是有价值的。

## ✓如何回应受访者顾虑

- 首先，调查员应当简洁的提一下最初受访者所提及的顾虑。例如“您讲的非常对！”，“我理解您的心情！”、“我知道您着急有事！”等等。
- 当调查员在消除受访者的顾虑时，参考培训话术，**针对顾虑进行解释说明。**
- 向受访者强调“**希望他们参与调研，他们参加这个调研非常重要**”，向受访者展示他们的参与是有价值的。
- 当然，也可能出现**中途拒访**，一般的中途拒访出现在询问到敏感或隐私的问题，或者由于时间太长，受访者不愿意接受调查。



## 如何消除受访者顾虑（以CGSS项目为例）

拒访情况分类		受访者的顾虑	如何应对
入户前的拒访	时间原因的拒访	把忙碌，没有时间作借口	介绍项目的目的、内容等进行劝说
		确实工作忙碌	改其他时间段或者约定合适时间
		访问时间过长	把访问分成几个时间段来进行
		到访时间不对	选择受访者极有可能在家的时间到访
	信任问题的拒访	对项目组的怀疑	熟知项目意义，让联系人配合说服
		对访员身份的怀疑	出示相关证件和介绍信，或联系社区联系人
		对社会不信任或者对项目的影响力不相信	表示体谅与理解，描述项目的意义和影响力
		经历过推销，上当受骗过，不愿意信任	社区联系人证明以及表明我们访问后给予的一点礼品，强调不做任何推销。
	各种排斥心理拒访	认为跟自己 and 自身利益无关，不愿意参加	描述项目可能对未来政策的影响，表明受益的就是受访者和他的家人
		社会、社区治安不好，与陌生人保持距离	找社区联系人以及对受访者有影响力的人劝说
		对于社会改革带来的影响不满，比如下岗、失业，拒绝参加	表示对受访者处境的理解，告诉他参与项目是为了更多像他一样的人可以得到保障
		由于个人原因拒绝访问（家庭吵架、心情问题、爱打麻将、身体不适、需要看孩子等）	根据不同情况随机应变
访谈中的拒访	个人信息保密以及隐私问题	问及关于子女信息时	表示数据使用时会清除受访者姓名信息
		问及收入及支出情况时	表示感兴趣的是和受访者一样的人的整体情况而非个人
		采集家户联系信息时	为了后续访问可以找到他们
	时间原因中途拒访	访问途中，因为时间过长，拒绝继续接受访问	耐心劝说，或者再约时间



# 标准化访问

- 概述：标准化访问一种**高度控制的访问形式**，即通过实现设计好的问题和程序来完成访问，以达到最大限度实施**科学**测量的目的。
- 含义及目的：
  - ✓控制体现在：
    - 访问是按照预先设计好的问卷进行的。
    - 访问时，只能按照问卷上的题目、指导语和顺序进行提问。
    - 当受访者不清楚题目含义时，访问员只能重复题目，或按统一说明进行解释。
  - ✓目的：
    - 每个访员使用相同的访谈方法。
    - 不同的访员访问，受访者的回答仍一致。
    - 收集的指标只反映受访者的差异，不能反映访员的差异。



# 标准化访问的原则

---

- 完全按题目提问
- 适当的追问
- 完整记录答案
- 保持立场中立
- 培训受访者



## 3.3 调查进度管理

---

- 包括为完成调查项目按时完成所需的各个过程，更侧重的是过程管理，简单来说就是到什么时候完成哪些工作，最后按照预期的计划完成整个项目过程。
- 一般进度管理包括如下几个过程：
  - 1) 估计工作量和时间
  - 2) 项目进度计划
  - 3) 控制进度



# CAPI控制进度举例

- 可以随时通过数据了解现场调查的工作进度。
  - 每天计算工作进度，同时设定**不同维度**的进度，例如不同地区（省市、县区或村居）的进度、不同队伍的进度、不同队伍中每个访员的进度。
  - 具体的进度还可以有多种**进度指标**，例如**发放样本的进度**、完成问卷的进度、回收电话的进度等。
- 进度和质量控制有关
  - 访员进度过快、过慢
  - 队伍进度过快、过慢
  - 队伍进度发生变化



## 3.4 调查成本管理

---

- 成本是调查项目最重要的问题之一。一项调查的经费往往是相对不足的。成本管理的目的利用最少的花费获得最高质量的调查结果。
- 不重视成本管理的结果：
  - 数据收集后质量很差而无法使用或解释
  - 为了达到项目目的成本大大超支
- 调查费用分类
  - 固定费用：调查设计、数据分析等和样本量大小无关的费用
  - 非固定费用：和样本量大小有关的费用



- $C = F + SM(S) + S(S) + M(S) + I(S) + A(S) + D(S)$

- ✓C: 总费用
- ✓F: 固定费用
- ✓SM: 抽样费用
- ✓S: 督导费用
- ✓M: 核查费用
- ✓I: 访员费用
- ✓A: 管理费用
- ✓D: 设备费用



- 权衡费用

- ✓ 很多时候，面临的是费用、质量和进度的权衡。
- ✓ 费用几乎关系到调查的每一个环节设计和执行，如：
  - 调查方式
  - 问卷长短
  - 内容复杂性
  - 抽样设计
  - 调查时效性
  - 。。。

- CGSS项目费用介绍:

- CGSS项目每年执行时，除北京外的其他省份的现场调查工作依靠中国社会调查网络的合作单位具体负责，**相对节省费用**。
- 各合作单位的费用是按照项目委托的方式，具体费用执行**兼顾中国人民大学的财务执行要求**，由各合作单位负责实施。
- 在具体项目预算过程中，需要注意对整体和调查每个部分工作量的估计。
- 好的预算在调查开始前，就已经对整个项目的执行过程做了一次**预演**。



## 4. 调查质量控制

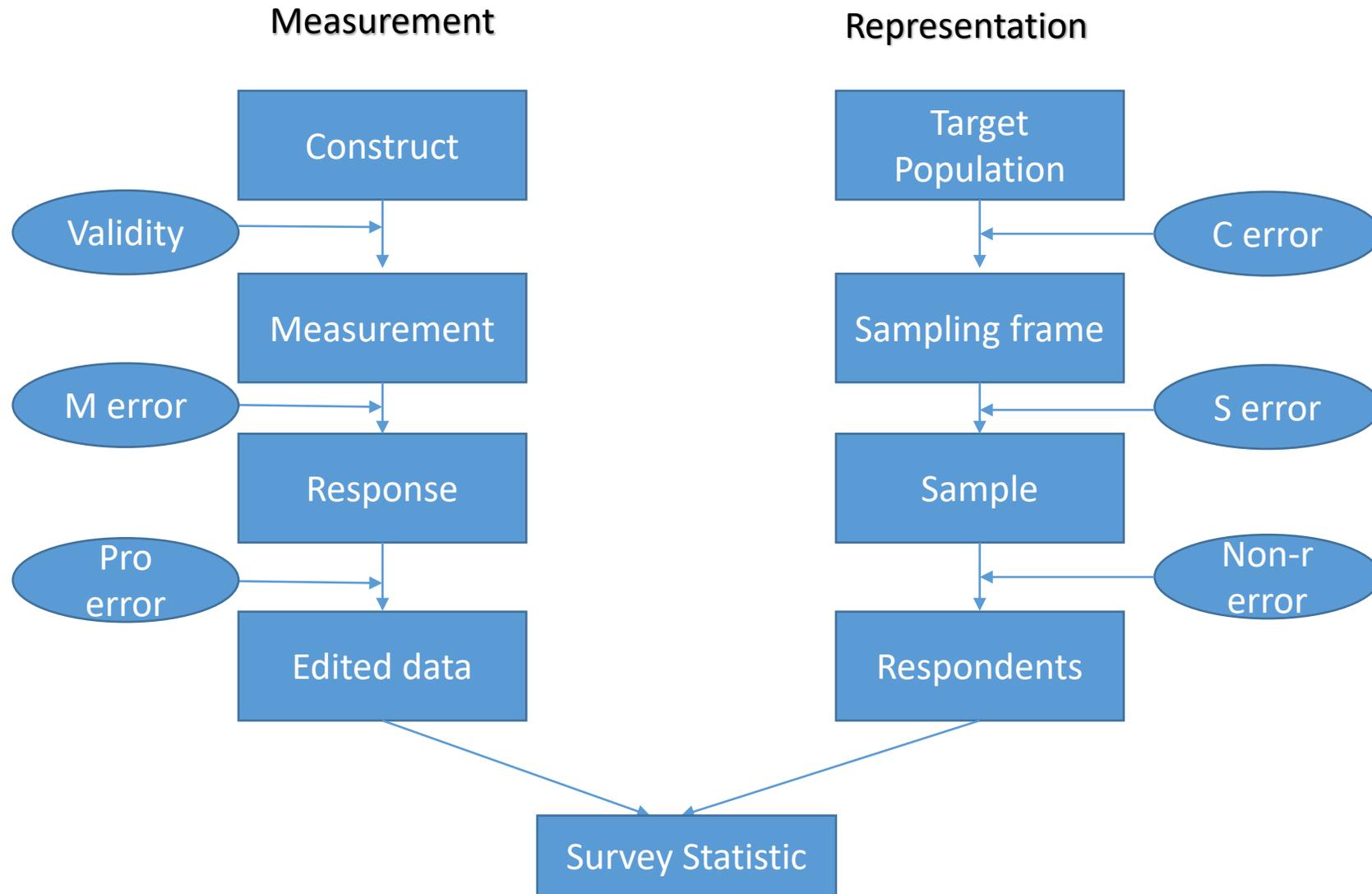
---

4.1 调查误差来源

4.2 调查质量控制



# 4.1 调查误差来源





## 4.2 调查质量控制

---

4.2.1 调查准备阶段

4.2.2 调查实施阶段



## 4.2.1 准备阶段

---

- 1、组织和协调保障
- 2、抽取有代表性的样本
- 3、认真设计调查问卷
- 4、调查员的招募与培训
- 5、组织试调查
- 6、建立现场调查质量的核查方法



# 建立现场调查质量的核查方法

- 现场核查

- 调查员自查：封面、调查对象、逻辑关系、填写，**核查清单**。**电脑调查可以实现必填和漏填提醒功能。**
- 督导审核：初审、复审
- 督导陪访
- 现场复核

- 远程核查

- 数据核查：应答率、电话收集率、录音采集率、问卷缺失、逻辑错误等
- 电话复核：专门的电话核查问卷（效果较差）
- 录音核查：还原现场访问情况，关注访员访问行为（标准化访问）。



## 4.2.2 调查实施阶段质量控制方法

---

- 1、设法提高应答率
- 2、尽量询问了解情况的受访者，减少回答误差。
- 3、调查开始的第一天，及时总结、讨论，统一标准，解决难题。
- 4、减少环境因素带来的影响。
- 5、做好实时现场和远程质控，及时反馈、改正和补充培训。



## 4.2.3 质控结果分析

---

- 1、应答率计算：单位和项目无应答
- 2、调查员培训考试成绩
- 3、抽样复查的复核率
- 4、计算机逻辑检错率
- 5、样本代表性分析



更多信息:

- 中国调查与数据中心, <http://nsrc.ruc.edu.cn>

**预祝同学们市调大赛取得好成绩!**