



УНИВЕРСИТЕТ ИТМО

Обработка статистических данных

Методы работы с собранной информацией. Анализ.
Обработка. Ошибки.

НИР
Лекция 7



С помощью обработки и систематизации полученных материалов наблюдений анализируются особенности и закономерности исследуемых процессов, проверяется, насколько правильными были исходные предположения. На основе полученных результатов формулируется решение более частных задач, делаются выводы относительно общих задач дальнейших исследований.

Основной инструмент анализа – сформулированные в программе гипотезы, а также предположения, которые возникают в ходе их проверки. Возможно и опровержение первоначальной гипотезы. Для проверки описательных гипотез используются процедуры группировки данных и их первичной классификации. Для подтверждения объяснительных гипотез необходимо использовать более сложные аналитические методы



Методы статистического описания направлены на получение количественных характеристик данных, полученных в конкретном исследовании. **Цели статистического описания** во многом совпадают с целями качественного анализа данных: стремление к сжато и наглядному их представлению. Математический язык позволяет достигнуть этой цели.

Методы статистического вывода позволяют корректно распространять результаты, полученные в конкретном исследовании, на всё явление как таковое, делать заключения общего характера.

Эти методы являются мощным инструментом построения научной теории, с их помощью осуществляется переход от изучения локальных явлений к познанию универсальных закономерностей, а тем самым и к научному прогнозированию.



Провести полноценный анализ первичной социологической информации можно только после ее обработки, то есть после приведения в состояние, удобное для сравнений, обобщений, интерпретаций. Чтобы подготовить собранную информацию к анализу, необходимо пройти семь этапов, выполняемых в определенной последовательности.



Этап 1-й – проверка документов (бланки социологического инструментария) на точность, полноту и качество заполнения.

Говоря о точности заполнения, имеют в виду правильность (адекватность) ответов респондентов на вопросы. Например, если на вопрос анкеты «Были ли у Вас взыскания за последние 12 месяцев?» дан ответ «да», а на вопрос «Если «да», то какие?» – ответа не было, значит, ответ «да» в предыдущем вопросе следует исправить на вариант «затрудняюсь ответить».

Под полнотой заполнения понимают наличие ответов на все или большинство вопросов. Например, если нет ответов более чем на 20 % вопросов или отсутствуют ответы на вопросы о себе, то анкета бракуется.

Очень важно обращать внимание на **качество** заполнения. Если, например, ответы на открытые вопросы сделаны неразборчивым почерком, если контрольные вопросы противоречат основным, если из фильтрующих вопросов следует вывод о некомпетентности респондента – во всех этих случаях анкета исключается из обработки



Этап 2-й – добор пропущенных данных, то есть проведение дополнительных опросов взамен забракованных.

При проверке анкеты или других форм социологического инструментария на точность, полноту и качество заполнения следует не упускать из виду вопрос о репрезентативности выборки. Например, после исключения некоторого количества анкет из обработки по причине их низкого качества необходимо провести дополнительные опросы, чтобы выборка оставалась в пределах расчетной.

Этап 3-й – уточнение программы обработки и анализа информации.

Разработка первичных схем обработки и анализа информации проводится на начальном этапе исследования. Однако в процессе подготовки к анализу собранных данных необходима корректировка выбранных методов обработки, коэффициентов, показателей, компьютерных программ и т. д. Важно уточнить не только общие направления обработки, но и составить конкретные задания на выполнение этой работы.



Этап 4-й – кодирование (кодировка) количественных данных в соответствии с заранее разработанной программой (заданием, инструкцией).

Кодирование, то есть присвоение числового кода каждой единице информации, необходимо, в первую очередь, для компьютерной обработки данных. Однако и при ручной обработке информации применение кодов может значительно упростить выполняемые процедуры. Ручную обработку анкет легче проводить, используя специальные бланки обработки, где соответствующие строки и графы нумеруются кодом.

Способы кодирования вопросников могут быть различными:

а) **сплошная нумерация** – кодируется порядковым номером каждый вариант ответов, например:

1. Введена ли на Вашем предприятии СМК?

001. Да.

002. Нет.

003. Я не знаю.

004. Другое.

2. С какими проблемами в деятельности предприятия Вам приходилось сталкиваться?

005. Отсутствие чёткого распределения ответственности.

006. ...



б) **позиционное кодирование** – когда числовой код состоит из порядкового номера вопроса и номера варианта ответа, например соединив номер вопроса «04» и вариант ответа «1», получим код «041»:

04. Как часто на вашем предприятии проводятся обучающие семинары?

1. Раз в месяц (041)
2. Раз в квартал (042)
3. Раз в полгода (043)
4. Раз в год (044)
5. Другой вариант (045)

05. Какое обучение принято?

1. Внутреннее (051)
2. Внешнее (052)
3. ...

С особыми трудностями приходится сталкиваться при кодировании открытых вопросов анкеты. При выполнении такой работы необходимо вначале обработать часть анкет, выявить повторяющиеся варианты, затем классифицировать их, наделяя числовыми кодами.



Этап 5-й – ввод первичной информации на машинные носители.

Особое внимание здесь обращается на контроль за качеством ввода и исправления допущенных ошибок.

Этап 6-й – систематизация однородных величин, определение процентных значений, группировка, ранжирование данных по возрастанию или убыванию признака, вычисление средних арифметических, средних взвешенных и других значений, необходимых для «сжатия» информации, приведение ее в состояние, удобное для анализа.

Этап 7-й – оформление данных в виде аналитических таблиц, графиков, диаграмм с использованием различных способов, приемов, повышающих наглядность информации.

Одним из основных критериев достоверности информации можно назвать ее **адекватность** изучаемым процессам, т.е. собранная информация должна соответствовать действительному состоянию исследуемых явлений, их объективным параметрам.

Достоверная информация должна отличаться **надежностью**, то есть быть способной устойчиво воспроизводить именно те характеристики объекта исследования, отражать те его стороны, которые важны для исследователя. Собранная информация, как правило, не может противоречить изложенным в программе исследования основным теоретическим принципам и методическим подходам. Она должна быть **валидной**, то есть измерять те качества или свойства изучаемых явлений и процессов, которые задуманы исследователем.

Основы достоверности информации формулируются на методологическом уровне в рамках подготовительного этапа исследований. Это могут быть и недостаточная изученность теоретического материала, а также и то, что выдвинутая проблема не отражает противоречий реальной ситуации.



К наиболее распространенным **просчетам**, влияющим на достоверность информации, можно отнести следующие:

- программа исследований составлена при недостаточном объеме исходных данных, а потому в ней не выражена специфика изучаемого объекта;
- методика сбора информации ориентирована на описание или объяснение исследуемых процессов, а не на получение практических рекомендаций;
- составленный социологический инструментарий по своему объему и содержанию выходит за рамки действительных материально-технических ресурсов исследовательской группы;
- преувеличены или не учтены профессиональные возможности сотрудников, необходимые для успешного проведения обработки и анализа информации;
- сроки выполнения отдельных видов работ не соответствуют нормативным и недостаточны для качественного завершения исследования;
- календарный план и график работ не согласован;
- не предусмотрены меры по профилактике негативного отношения респондентов к сбору информации;
- расчетная выборка отличается от фактической после исключения забракованных документов;
- слабый контроль организаторов работ за ходом исследования.



Среди **способов непосредственной верификации (проверки) информации** выделяют следующие:

- 1) контроль за адекватностью ситуаций опроса, благоприятно или негативно влияющих на искренность респондентов;
- 2) проверку полноты ответов респондентов, наличие ответов на открытые вопросы;
- 3) использование психологических приемов контроля за поведением опрашиваемых и на этой основе оценка степени их заинтересованности в представлении достоверной информации;
- 4) применение логического анализа ответов респондентов, сопоставление с ответами на контрольные вопросы или ответами других респондентов;
- 5) проведение параллельно экспертных опросов, бесед с руководителями или другими лицами, компетентными в исследуемых проблемах;
- 6) сопоставление данных, приобретенных в результате опроса, с информацией, полученной другими методами;
- 7) предварительный анализ собранной информации с целью выяснения возможности ее использования для подтверждения или опровержения гипотез исследования.

Важнейшим условием сбора достоверной информации является также *стандартизация* исследовательской ситуации, выработка единого подхода к процедурам сбора данных. Стандартизация осуществляется за счет детального инструктирования всех исполнителей и проведения тренировочных работ под наблюдением организаторов с последующим обсуждением результатов в исследовательском коллективе.

Существенный фактор, обеспечивающий достоверность информации, – *контроль* организаторов исследования за сбором данных. Руководители работ должны определить, в какой степени тот или иной сотрудник социологической группы владеет навыками сбора информации и насколько эффективно, умело и добросовестно использует свои возможности.



Особое внимание обращается на следующие критерии:

- достаточно ли подготовлен исследователь к проводимым работам;
- насколько правильно и быстро устанавливает необходимый контакт с респондентами;
- обеспечивает ли условия для гарантии анонимности опросов;
- устраняет ли в ходе проведения бесед внешние помехи (шум, присутствие посторонних лиц);
- не производит ли замену единиц выборочной совокупности на более удобные для опроса;
- правильно ли регистрирует собранную информацию;
- соблюдает ли требования единого подхода к исследовательским ситуациям.



Для контроля и самоконтроля можно разработать инструкцию, в которой указывается:

- полное наименование организации, в которой проводится опрос;
- общая характеристика объекта исследования, его структурных подразделений, с перечислением фамилий, имен и отчеств руководителей, их телефонных номеров;
- формулировки целей, основных задач и гипотез исследования;
- сведения о применяемых технических средствах и рекомендации по их использованию;
- описание используемого в процессе сбора информации социологического инструментария;
- правила заполнения бланков инструментария, перечень требований, обязательных при выполнении соответствующих процедур;
- способы подбора единиц наблюдения;
- сроки начала и завершения работы и др.



Многие недостатки в методике сбора информации можно выявить с помощью **пилотажных** исследований. Для этого используют небольшую выборку, опрашивая 10–20 человек или более в зависимости от общего объема работ и сложности социологического инструментария. В состав пилотажной выборки включают респондентов, намеченных для основного опроса. Подобное анкетирование нередко заменяют пилотажным интервью, когда в качестве плана интервью применяется бланк анкеты. Такой прием позволяет, если использовать наблюдение за реакцией (поведением) респондента, собрать больше сведений не только о качестве инструментария, но и о возможных психологических трудностях опроса. Одновременно этот метод помогает сэкономить средства на копировании двух вариантов бланков анкет, предварительного и окончательного.

Проведение **опосредованной верификации (средств проверки) информации** осуществляется следующими способами:

- 1) сравнением собранной информации с результатами ранее выполненных аналогичных исследований;
- 2) сопоставлением информации, полученной различными методами;
- 3) логическим анализом содержания полученных данных;
- 4) дополнительными опросами экспертов, компетентными в исследуемых проблемах;
- 5) последующим (дополнительным) наблюдением за объектом исследования;
- 6) экспериментальной проверкой собранных сведений.



При осуществлении **контроля** за достоверностью информации необходимо особое внимание обращать на наличие следующих критериев:

- на адекватность социологического инструментария измеряемым свойствам, признакам исследуемого явления;
- соответствие шкал измерения поставленным задачам исследования;
- систематичность ошибок в использовании инструментария;
- добросовестность сотрудников, их точное следование инструкциям;
- уровень компетентности (высокий/низкий) исследователей, участвующих в сборе данных



Способы контроля за качеством информации можно представить в виде двух основных форм:

- 1) внешний контроль – проверка добротности инструментария, компетентности исполнителей, их стремления выполнять обязательные требования, предъявляемые к научному исследованию;
- 2) внутренний контроль – выяснение причин несовпадения Контрольных вопросов с основными, анализ последствий «неработающих» блоков инструментария, выявление ошибок путем сопоставления данных, собранных в различных (аналогичных) подразделениях объекта исследования, и др.

Аналитическое описание текстов. Этот процесс имеет циклический характер. На этапе первичного прочтения полученный текст внимательно изучается, составляются авторские комментарии, определяется место документа в применяемой системе кодирования. В дальнейшем, при многократном повторном чтении текста, полученные предварительные выводы и оценки корректируются и уточняются, дополняются новыми результатами. Иногда требуется вернуться к собеседованию с клиентом для уточнения содержания некоторого события или явления. Данная работа продолжается до получения четкого и конкретного представления об исследуемом социальном явлении.



Большое внимание следует уделять **структурированию документов**. При анализе текст можно рассматривать не только как единый объект, но и как совокупность отдельных фрагментов, ситуаций, эпизодов, представляющих собой внутренне законченные сюжеты. В процессе структурирования зарегистрированные факты и события обобщаются по различным признакам: темам, времени и т. д., при этом анализируется, как они взаимосвязаны между собой, подтверждают ли друг друга. Происходящие изменения рассматриваются в контексте ключевых событий или этапов исследуемого процесса, оценивается степень влияния на них некоторого комплекса взаимодействующих скрытых факторов.



Следующим этапом качественного исследования, являющимся первым шагом концептуального анализа, является **первичная классификация** полученной информации. Она предусматривает новое построение анализируемого текста уже на основе обобщенных понятий.

Выводы и результаты, полученные в рамках исследований, позволяют уточнить структуру предварительного классификатора. Этот динамический процесс может предусматривать либо детализацию классификатора с вводом дополнительных классов, либо, напротив, обобщение классов с их объединением в другие, более крупные. При этом следует иметь в виду, что идентифицированные классы не должны пересекаться между собой. Решение проблемы оптимизации классификатора должно обуславливаться потребностями выполняемого исследования.



В некоторых качественных исследованиях бывают востребованы ***кластеризация и метод аналитической индукции***. При кластеризации, на основе поиска отклонений от первоначального обобщения, применяется объединение внутренне тождественных и внешне не связанных между собой категорий. В результате анализа сформированных кластеров с помощью метода аналитической индукции могут быть получены выводы и концепции, не следующие в явном виде из материалов, обобщенных обычными способами.



Конечной целью качественного исследования в социальной работе является формулирование **концепции** и построение **мини-теории** некоторого явления, наблюдаемого на практике. Следует иметь в виду, что создание мини-теории является характерной особенностью качественных исследований, не ориентированных на проверку и верификацию более общих теорий.

Создание мини-теории предполагает изучение уже существующих теорий и исследований в анализируемой области, совокупности материалов, имеющих по рассматриваемой проблеме. Как указывалось ранее, очень важной является идентификация событий и фактов, противоречащих уже выявленным закономерностям. При этом следует иметь в виду, что привлечение для анализа более представительного массива информации может привести не только к уточнению созданной теории, а и к ее опровержению.

Результаты качественного исследования во многом зависят от эрудиции исследователя, его научной интуиции, способности систематизировать и обобщать противоречивые, на первый взгляд, данные.

Типичными **ошибками** при проведении качественных исследований в научно-исследовательской работе являются:

- 1) субъективизм интерпретации обрабатываемых данных, что может невольно привести к получению желаемых, а не фактических результатов;
- 2) обобщение недостаточно представительного информационного массива;
- 3) отсутствие технологий и средств проверки (верификации) полученной информации;
- 4) отсутствие критического подхода к полученным данным.

Перечисленные ошибки неизбежно будут возникать в процессе работы начинающего исследователя, их можно избежать только при накоплении достаточного практического опыта.

1. Разработать анкету (если по теме исследования, то замечательно, но можно и по любой другой теме).
2. Провести опрос (не менее 10-ти человек).
3. Систематизировать данные, используя графические методы.
4. Сделать выводы, опираясь на материалы лекции.

Спасибо за внимание!

IT'sMO^{re} than a
UNIVERSITY