**Методы обработки социологических данных**

Заключительный этап эмпирического социологического исследования предполагает обработку, анализ и интерпретацию данных, получение эмпирически обоснованных обобщений, выводов и рекомендаций.

*Обработка данных*включает в себя следующие компоненты:

Редактирование и кодирование информации. Основное назначение этого шага состоит в унификации и формализации той информации, которая была получена в ходе исследования.

Создание переменных. Собранная на основании анкет информация в ряде случаев прямо отвечает на те вопросы, которые необходимо решить в исследовании. Вопросы получили форму индикаторов в процессе операционализации. Сейчас же необходимо провести обратную процедуру, то есть перевести данные в форму, которая бы отвечала на вопросы исследования.

Статистический анализ. Этот шаг является ключевым в процессе анализа социологических данных. В ходе статистического анализа выявляются некоторые статистические закономерности и зависимости, которые позволяют социологу сделать определенные обобщения и выводы. Для проведения статистического анализа социологи используют большое число различных математических методов, позволяющих полно и всесторонне анализировать собранную информацию. В современной социологии для этой цели активно применяются ЭВМ, дополненные программами математико-статистической обработки.

В зависимости от методов получения первичной информации возможно применение различных приемов обработки и анализа данных. Так, если социолог определенную часть информации извлекает из документальных источников, то он использует два основных метода анализа документов. *Неформализованный (традиционный) и формализованный (контент-анализ)*. Традиционный анализ основан на восприятии, понимании, осмыслении и интерпретации содержания документов в соответствии с целью исследования. Формализованный анализ документальных источников (контент-анализ - анализ содержания) рассчитан на извлечение социологических информации из больших массивов документальных источников, недоступных традиционному интуитивному анализу. Он основан на выявлении некоторых количественных статистических характеристик текстов (или сообщений). При этом предпола­гается, что количественные характеристики содержания документов отражают некоторые существенные черты изучаемых социальных явлений и процессов.

Формализованный анализ документов основан на стандартизации процедур поиска, определения в содержании документа единиц счета, которыми могут быть отдельные слова (термины, географические названия, имена политических деятелей и т.д.), суждения, выраженные и виде предложений, абзацев, фрагментов текстов и т.д., а также различные виды публикаций (по жанру, типу авторов, темам и т.п.). Единицы счета определяются в зависимости от целей социологического исследования.

При обработке и анализе данных, полученных методом опроса, широко применяются методы ранжирования, шкалирования, корреляции и др. Так, *ранжирование* - это процедура установления относительной значимости (предпочтительности) исследуемых объектов на основе их упорядочивания. *Ранг* - это показатель, характеризующий порядковое место оцениваемого объекта в группе и других объектов, обладающих существенными для оценки свойствами. Для каждого объекта вычисляют сумму рангов, полученную от всех экспертов, затем упорядочивают эту сумму. Ранг 1 присваивают объекту, получившему наименьшую сумму, самый низкий ранг - объекту с наивысшей суммой. Ранжирование дополняется, как правило, другими методами экспертных оценок. Шкалирование – это  ***процедура, с помощью которого устанавливается соответствие между свойствами социального объекта и свойствами сопоставляемой им математической модели (например, степень удовлетворенности респондента сопоставляется ряду натуральных чисел от 0 до 100).***  В социологии обычно выделяют следующие виды шкал:

 номинальные,

порядковые (ранговые),

метрические (шкалы отношений, интервальные).

***Номинальная шкала*** это шкала соответствия социальных свойств равенства и неравенства соответствующим числовым отношениям с целью отличия одного социального объекта от другого, то есть выявление непересекающихся классов - каждый класс соответствует определенной позиции на шкале. Например, пол: мужской - 1, женский - 2; мужчины: холостой - 1, женатый - 2, разведенный - 3, вдовец - 4. Это*закрытые вопросы*, то есть все возможные ответы предлагаются респонденту. При этом решаются следующие задачи: во-первых, измеряется количество каждого признака; во-вторых, определяется различие между ними. Для выполнения второй задачи достаточно решить математическое уравнение сравнения равенства или неравенства, при этом первая задача решается методом выбора той или иной категории или числового балла на шкале или применением *открытых вопросов*, то есть вопросов, в которых дается респонденту возможность дать свой ответ, например, “Ваш возраст?”, “Ваше образование?”, “Ваше отношение к коллективному отдыху? - 1. Мне нравится отдыхать коллективом, 2. Предпочитаю отдыхать один, 3. Предпочитаю незнакомую компанию, 4. Другое, напишите свой ответ” и т.п.

***Порядковая*** (ранговая) ***шкала***это шкала соответствия социальных свойства равенства или неравенства и их степени выраженности с соответствующими числовыми отношениями. Например, удовлетворенность чем-либо: полностью - 5, скорее да чем нет - 4, трудно сказать - 3, скорее нет - 2, полностью не удовлетворен - 1. Для этого помимо задач, приведенных в номинальных шкалах, необходимо измерить интенсивность выраженности признаков. Чтобы решить эту задачу, достаточно выполнить математического сравнение "больше", "меньше" и расположить признаки в порядке возрастания или убывания, то есть проранжировать.

***Метрическая шкала*** это шкала соответствия социальных свойств равенства или неравенства и пропорций внутри них с соответствующими числовыми отношениями. Например, равных пропорций - определение возрастных групп: 16-20 лет - 1, 21-25 лет - 2, 26-30 лет - 3; неравных пропорций - срок службы: 1-2 года - 1, 2-5 лет - 2, 5-20 лет - 3, 20-25 лет - 4, более 25 лет - 5. Если на шкале ноль условный, то это интервальная шкала, если естественный - шкала отношений (возраст, заработанная плата и т.д.). Метрические шкалы помимо задач, решаемых с помощью номинальных и порядковых шкал, позволяют ответить на вопрос: на сколько или во сколько раз  интенсивность выраженности одного признака больше или меньше другого.

Основная задача корреляционного анализа – это установление характера и тесноты связи между результативными (зависимыми) и факторными (независимыми) показателями (признаками) в данном явлении или процессе. Корреляционную связь можно обнаружить только при массовом сопоставлении фактов.

Характер связи между показателями определяется по корреляционному полю. Если Y - зависимый признак, а Х - независимый, то отметив каждый случай X(i) с координатами xi и yi получим корреляционное поле.

Теснота связи определяется с помощью коэффициента корреляции, который рассчитывается специальным образом и лежит в интервалах от минус единицы до плюс единицы. Если значение коэффициента корреляции лежит в интервале от 1 до 0,9 по модулю, то отмечается очень сильная корреляционная зависимость. В случае, если значение коэффициента корреляции лежит в интервале от 0,9 до 0,6, то говорят, что имеет место слабая корреляционная зависимость. Наконец, если значение коэффициента корреляции  находится в интервале от - 0,6 до 0,6, то говорят об очень слабой корреляционной зависимости или  полной ее отсутствии.

Таким образом, корреляционный анализ применяется для нахождения характера и тесноты связи между случайными величинами.

Регрессионный анализ своей целью имеет вывод, определение (идентификацию) уравнения регрессии, включая статистическую оценку его параметров. Уравнение регрессии позволяет найти значение зависимой переменной, если величина независимой или независимых переменных известна.