

Приказ № 38/1 от 28.04. 2022г.

Об утверждении положения репрезентативности выборки

Во исполнение муниципальной программы «Организация системы профессиональной ориентации учащихся в образовательных организациях Сергокалинского района»

**ПРИКАЗЫВАЮ:**

1. Утвердить положение репрезентативности выборки
2. Руководителям образовательных организаций района в проводимой в образовательных организациях использовать положение репрезентативность выборки
3. Методическому центру управления образования разместить настоящий приказ на официальном сайте управления образования и в социальных сетях.
4. Контроль за исполнением настоящего приказа возложить на заместителя начальника управления образования Мусаева М.И.



  
Х. Исаева.

## Репрезентативность выборки

Свойство выборки, благодаря которому результаты выборочного исследования позволяют делать о генеральной совокупности и эмпирическом объекте в целом, называется *репрезентативностью*.

*Репрезентативность* выборки – это способность выборки воспроизводить определенные характеристики генеральной совокупности в пределах допустимых погрешностей. Выборку называют репрезентативной, если результат измерения определенного параметра для данной выборки совпадает с учетом допустимой погрешностью известным результатом измерения генеральной совокупности. Если выборочное измерение отклоняется от известного параметра генеральной совокупности больше выбранного уровня погрешности, то такая выборка считается не репрезентативной.

Предложенное определение прежде всего устанавливает связь между выборочной и генеральной совокупностью исследования. Именно генеральную совокупность представляет выборка, и только на генеральную совокупность могут быть распространены выявленные в выборочном исследовании тенденции.

Выборка не может представлять иную совокупность, отличную от той из которой фактически проводится отбор единиц для измерения. Если исследователь заблуждается относительно фактических границ генеральной совокупности, то его выводы будут некорректны. Если он по ошибке или намеренно расширяет или искажает границы генеральной совокупности в отчетных материалах, публикациях, презентациях по результатам исследования, то вводит в заблуждение пользователей и может рассматриваться в качестве фальсификации результатов.

Проверка репрезентативности осуществляется посредством сравнения отдельных параметров выборки и генеральной совокупности.

Репрезентативность или не репрезентативность выборки может быть установлена исключительно в отношении отдельных переменных. Более того, одна и та же выборка может быть репрезентативна одним параметрам и нерепрезентативна – по другим.

В выборке может быть одни параметры генеральной совокупности воспроизводить более точно, а другие – менее точно. Поэтому корректнее вести речь о мере репрезентативности конкретной выборки. Как и в случае с выборкой в целом, ключевым моментом в определении репрезентативности выборки является обоснование погрешности, в пределах которой выборка признается репрезентативной для целей исследования. Возможно и обратное — фиксация размера фактических ошибок и констатация факта, что выборка представляет генеральную совокупность с определенными погрешностями. И опять-таки ключевую роль в этом играет характер использования результатов исследования. Следовательно, одна и та же выборка может быть признана достаточно репрезентативной для одних целей (например, для прогноза явки избирателей на предстоящих выборах), но недостаточно репрезентативной для других (например, для определения рейтингов кандидатов и прогноза результатов голосования).

По каким параметрам следует проверять репрезентативность выборки? Во-первых, таких параметров в большинстве исследовательских ситуаций немного. Ведь сопоставить результаты выборочного измерения с данными о генеральной совокупности можно только при наличии последних. А исследования проводятся потому, что таких данных как раз не хватает. Поэтому еще на этапе моделирования объекта и последующей разработки инструментария целесообразно предусмотреть измерение одного или нескольких контрольных параметров, по которым доступны данные, характеризующие генеральную совокупность. Это создаст необходимую эмпирическую базу для проверки репрезентативности.

Во-вторых, нужно стремиться проверять репрезентативность выборки по параметрам, имеющим существенное значение для предметной области исследования. В современной практике широкое распространение получил контроль репрезентативности по основным демографическим параметрам — полу, возрасту, образованию и пр. Эти данные, как правило, доступны для любого территориального объекта, так как фиксируются во время переписей населения и впоследствии пересчитываются статистическими учреждениями с применением обоснованных математических моделей. По этой причине обязательное включение нескольких демографических переменных в «паспортчику» стало общепринятой профессиональной нормой. Однако такая практика может быть отнесена к разряду наивных и подвергнута обоснованной критике. Дело в том, что основные и общедоступные для сравнения демографические параметры далеко не всегда играют роль структурирующих факторов в отношении предметов социологических исследований. Их природа сама по себе не является социальной, а влияние на объекты исследований часто достаточно опосредованно. Поэтому репрезентативные по демографическим параметрам выборки на деле могут скрывать значительные проблемы в виде системных ошибок и неконтролируемых смещений. Напротив, демографическая репрезентативность эффективных с точки зрения целей и задач исследования выборок может оказаться невысокой.

Дело в том, что основные и общедоступные для сравнения демографические параметры далеко не всегда играют роль структурирующих факторов в отношении предметов социологических исследований. Их природа сама по себе не является социальной, а влияние на объекты исследований часто достаточно опосредованно. Поэтому репрезентативные по демографическим параметрам выборки на деле могут скрывать значительные проблемы в виде системных ошибок и неконтролируемых смещений. Напротив, демографическая репрезентативность эффективных с точки зрения целей и задач исследования выборок может оказаться невысокой.

Интересно, что в исследовательских отчетах и публикациях практически всегда упоминаются именно репрезентативные выборки. Неужели нерепрезентативные выборки настолько редки? Конечно нет. Выборок, которые являются проблемными с точки зрения репрезентативности по тем или иным параметрам, в исследовательской практике встречается достаточно. Скорее их даже больше, чем выборок, репрезентативность которых может быть оценена не формально (по демографическим параметрам), а по существу. Однако их публичное упоминание в профессиональной социологической среде, к сожалению, табуировано. И никто из исследователей не готов признать, что репрезентативность его выборки по существенным для предметной области измерения параметрам проблематична либо не проверяема.

На самом деле обнаружение признаков не репрезентативности выборки не является катастрофой. Во-первых, существующие технологии «ремонта» (перевзвешивания) выборки во многих случаях позволяют полностью исключить эффект не репрезентативности относительно беспокоящего социолога или его клиента параметра. Суть метода перевзвешивания состоит в

присваивании различным категориям наблюдений (в случае опроса — респондентов) определенных весовых коэффициентов, компенсирующих недостаточное или избыточное фактическое представительство данных категорий в выборке. В дальнейшем эти веса учитываются при проведении всех расчетных операций с массивом данных, что позволяет получать распределения, полностью соответствующие сбалансированному (соответствующему расчетным квотам) массиву данных. Современные статистические программы, такие как БРВВ, позволяют производить расчеты с учетом весовых коэффициентов в автоматическом режиме, что делает эту процедуру достаточно легкой для исполнения.

Во-вторых, даже если получить «хорошую» репрезентативную выборку не получается, «умеренной» репрезентативности может оказаться достаточно для решения многих исследовательских задач. Напомним, что репрезентативность — это скорее мера соответствия, чем дихотомический маркер. И только отдельные исследовательские задачи — в основном связанные с точным прогнозированием определенных событий — требуют от выборок действительно высокой (статистически подтвержденной) репрезентативности.

Например, для того, чтобы прогнозировать рыночную долю нового продукта в маркетинговом исследовании требуется выборка, охватывающая и репрезентирующая потенциальных клиентов. Однако чаще всего маркетологи не имеют достаточных данных о том, кто на самом деле образует круг их клиентов, тем более — потенциальный. В этой ситуации проверить репрезентативность выборки вообще невозможно — ведь неизвестно, какие параметры она должна воспроизводить. Тем не менее многие маркетинговые задачи успешно решаются, так как для выявления клиентских предпочтений, реакции на рекламные материалы, анализа отзывов на новый продукт статистически репрезентативные выборки не нужны — достаточно обеспечить охват типичной клиентуры, которую легко найти прямо в магазинах. Нерепрезентативные выборки вполне подходят для решения поисковых задач, выявления сильно выраженных тенденций, анализа специфики отдельных категорий (представленных маленькими самостоятельными подвыборками), сравнения таких категорий между собой (двумерный анализ), анализа взаимосвязей между переменными и других задач, в которых точность полученных статистических распределений имеет второстепенное значение.