**Программа повышения квалификации**

**«Анализ данных с использованием IBM SPSS Statistics»**

**На кого рассчитана программа:**

- для научных и педагогических работников,

- специалистов компаний,

- государственных служащих,

для всех, кто осуществляет профессиональную деятельность в сферах, связанных с анализом данных, обработкой результатов опросов общественного мнения.

**Результаты освоения программы**

Слушатель, освоивший программу, будет способен формировать массивы данных на основе проведения обследований населения и проводить их анализ:

* иметь представление о том, как проводятся обследования различных целевых групп и какое внимание надо уделить формулировке вопросов с учетом их последующей обработки;
* знатьметоды, применяемые для анализа результатов проводимых выборочных исследований (проверки гипотез, наличия связи, выявление однородных групп, построение моделей);
* уметь применять полученные знания для формирования, корректировки и обработки данных исследования, анализа различных проблем жизнедеятельности населения и социально-демографических групп, факторов и детерминант социальной и экономической политики.
* получить навыки анализа и интерпретации результатов выборочных обследований с применением IBM SPSS Statistics.

**Программа обучения** (**6 тем**)

**Тема 1. Сущность и основные направления выборочных обследований населения. Возможности использования специальных ППП для обработки данных выборочных обследований**

 Методы сбора количественной информации. Выборочные исследования. Выборочные социально-демографические обследования в России. Основные пакеты статистических прикладных программ для социальных исследований. Функции специальных ППП (Statistica, SPSS) в обработке данных выборочных исследований. Структура, модули SPSS. Направления обработки данных. Подготовка данных. Ввод и сохранение данных. Шкалы измерений (количественная, порядковая, номинальная). Свойства шкал и их допустимые преобразования. Типы категоризации данных.

**Тема 2. Подготовка данных. Отбор и модификация данных.**

Выбор наблюдений. Сортировка наблюдений. Разделение наблюдений на группы. Модификация данных. Вычисление новых переменных. Вычисление новых переменных в соответствии с определенными условиями. Формулировка условий. Агрегирование данных. Ранговые преобразования. Веса случаев. Причины и механизмы порождения пропусков данных. Возможности игнорирования пропусков. Методы заполнения пропущенных значений. Методы выявления аномальных значений. Применение устойчивых процедур оценивания. Анализ множественных ответов.

**Тема 3. Дескриптивная статистика. Таблицы сопряженности.**

Роль статистики в обработке результатов выборочных обследований. Микро- и метаданные. Области применения и границы применимости математико-статистических методов. Сводка наблюдений. Описательная статистика. Одномерные распределения. Показатели вариации. Дисперсия, вариационный размах, среднее абсолютное отклонение, квантильные размахи. Построение таблиц сопряженности. Графическое представление таблиц сопряженности.

**Тема 4. Параметрические и непараметрические критерии.**

Анализ взаимосвязи между признаками. Независимость переменных. Основные характеристики связи. Непараметрические и параметрические критерии. Критерий независимости (критерий согласия χ2). Сравнение двух и нескольких выборок (зависимых и независимых). t-критерий. Статистические критерии для таблиц сопряженности. Коэффициенты корреляции (для номинальных и ранговых шкал). Меры тесноты связи между переменными. Простейшие меры тесноты связи (для дихотомических переменных). Меры связи для таблиц с порядковыми данными. *t*-меры Кендэла и их свойства. *d*-меры Сомерса. Мера Гудмена-Краскала и ее свойства. Дисперсионный анализ.

**Тема 5. Корреляционно-регрессионный анализ.**

Сущность и задачи корреляционного анализа. Диаграммы рассеяния. Парные коэффициенты корреляции. Измерение степени тесноты статистической связи, «очищенной» от влияния посторонних признаков с помощью частных коэффициентов корреляции. Проверка существенности связи признаков. Доверительные интервалы для коэффициентов корреляции. Множественный коэффициент корреляции. Коэффициент детерминации. Двухмерная модель регрессионного анализа: линейная и нелинейная модели регрессии. Кривые роста в задачах прогнозирования, «фиктивные» переменные и их применение. Множественная линейная модель регрессии. Нелинейная регрессия (бинарная логистическая регрессия, мультиноминальная логистическая регрессия, порядковая регрессия, пробит-анализ, приближение с помощью кривых).

**Тема 6. Методы снижения размерности.**

Статистический подход в методе главных компонентов. Вычисление главных компонентов и их графическая интерпретация. Информативность редуцированного признакового пространства. Регрессия на главные компоненты. Роль и место непараметрических методов в структурном моделировании. Иерархический кластерный анализ. Метрики признакового пространства. Принципы измерения расстояния между группами объектов. Алгоритмы быстрого кластерного анализа, метод k-средних. Двухэтапный кластерный анализ. Построение дерева целей.

**Продолжительность программы**

72 часа

**Форма обучения**

Дистанционная

**Стоимость программы**

42 тыс. рублей.

**Документ по окончании программы**

удостоверение о повышении квалификации МГУ

**Контакты руководителя и менеджера**

Руководитель программы – Кучмаева Оксана Викторовна, доктор экономических наук, профессор Экономического факультета МГУ, специалист по количественным исследованиям в социальной сфере, руководитель более 30 исследовательских проектов, имеет опыт преподавания аналитических курсов в НИУ ВШЭ, РЭУ им. В.Г.Плеханова. Данный курс апробирован в открытом университете Фонда Е. Гайдара.

e-mail: kuchmaeva@econ.msu.ru

Менеджер программы – \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

e-mail: @\_\_\_\_\_\_\_.ru

контактный тел:

сайт программы (если есть отдельный):