

Эксклюзивные курсы по анализу данных,

подготовленные специалистами компании StatSoft Russia на основе уникального опыта решения реальных задач

Введение в современный анализ данных в программе *STATISTICA* для задач бизнеса, маркетинга, промышленности

16 акад. часов

Углубленный курс статистического анализа данных с использованием программного пакета *STATISTICA*

36 акад. часов

Реализуется на базе МИЭМ НИУ ВШЭ

- Курсы по анализу данных – это грамотное сочетание качественного теоретического материала и увлекательной практической работы, обеспечивающее максимально эффективное применение приобретенных знаний.
- По завершении курса слушатели получают сертификат Академии Анализа Данных StatSoft, который станет отличным дополнением к личному портфолио.

Менеджер учебных программ

Валентина Калиникова

Телефон: 8 (916) 316-85-59

E-mail: vkalinikova@hse.ru

Сайт программы:

<http://miem.hse.ru/edu/pm/dpo>

Адрес: ул. Таллинская, д. 34, здание МИЭМ



Милков Максим Леонидович

О лекторе

Является выпускником одного из ведущих вузов России – Московского физико-технического института (МФТИ) по специальности "Прикладные математика и физика". Имеет высокую квалификацию и многолетний опыт работы в области статистического анализа и обработки данных для крупных российских и зарубежных компаний, таких как ЦБ, ИнтерРАО ЕЭС, ВСМПО-АВИСМА, ТНК-ВР, Норильскгеология, Такеда Фармасьютикалс, Полисан, Днепроспецсталь, Арнест и мн. др. Регулярно проходит стажировки и тренинги у ведущих разработчиков аналитического программного обеспечения.

О программе

Программа дает уникальные навыки работы с данными от постановки задачи до интерпретации результатов, что позволяет специалистам реализовать себя в различных областях.

Полученные на наших курсах знания помогут новичкам успешно начать свою карьеру, а знатокам анализа данных значительно повысить свой уровень и квалификацию, освоив современные инструменты анализа данных.

Практическая работа осуществляется с использованием программного продукта *STATISTICA* - универсального пакета для всестороннего анализа данных, занимающего лидирующие позиции на рынке аналитического программного обеспечения, а также его приложений для современных методов анализа, таких как нейронные сети, Data Mining и др.

Курсы рассчитаны на студентов, выпускников, а также всех желающих связать свою будущую профессию с аналитикой и решением прикладных задач в бизнесе, маркетинге, промышленности.

Курсы охватывают как классические, так и современные методы анализа данных, а также инструменты, позволяющие как новичку в области аналитики, так и требовательному пользователю всесторонне их исследовать, проводить оперативный мониторинг регулярно обновляемых данных, строить прогнозы с высокой точностью, создавать шаблоны анализа и готовить аналитические отчеты презентационного качества, используя средства автоматизации и широкий спектр графических возможностей.

Слушателям курсов открываются новые перспективы профессионального роста, т.к. современные инструменты анализа данных серии **STATISTICA** используются для широкого спектра задач.

Курс «Введение в современный анализ данных в программе **STATISTICA** для задач бизнеса, маркетинга, промышленности»

16 академических часов

1. Возможности интерфейса **STATISTICA**
 - Настройки программы с учетом требований пользователя
 - Взаимодействие с Excel
 - Интеграция с MS Office
2. Работа с данными
 - Импорт и экспорт данных из любых других приложений на примере из бизнеса
 - Подготовка данных для исследования
 - Все этапы проведения анализа – от сбора данных до интерпретации результатов
 - Подготовка аналитических отчетов
 - Преобразование данных любого вида к нужному формату
3. Вероятностные распределения
 - Вероятностные распределения и их свойства
 - Подгонка наблюдаемых распределений под известные теоретические
 - Вероятностный калькулятор и его применение на практике
4. Элементарные понятия статистики
 - Описательные статистики – получение обобщенной информации о данных на примере маркетингового исследования
 - О чем молчит коэффициент корреляции – исследование связей между парой признаков на примере данных из промышленности
 - Проверка равенства средних на примере маркетингового исследования (t-критерии для разных случаев)
 - Таблицы частот и таблицы сопряженностей – работа с дискретными величинами на примере данных социологического опроса
 - Группировка и однофакторный дисперсионный анализ – t-критерии для трёх и более групп
5. Графический модуль для подготовки материалов презентационного качества
 - Типы графического представления с учетом исходных данных и требований пользователя
 - Связь графиков различного типа с другими видами анализов
 - Построение графиков любой сложности и их точечная настройка
 - Интерактивный визуальный анализ с использованием средства «Кисть»
 - Интерактивная подгонка кривых под известные распределения
 - Импорт графиков в отчеты и презентации
 - Интерпретация результатов графического анализа
6. Основные методы анализа данных
 - Построение предсказательных моделей с помощью парной и множественной регрессии на примере маркетинговых и промышленных данных:
 - Анализ рядов продаж многономенклатурной продукции
 - Создание материалов с заданными свойствами на примере многокомпонентной стали
 - Нелинейное оценивание
 - Логит и пробит модели
 - Классификация маркетинговых данных с помощью кластерного анализа
 - Дискриминантный анализ
 - Классификация данных с помощью деревьев классификации
 - ABC и XYZ анализ на примере реальных задач из маркетинга и бизнеса
 - Простейшие методы прогнозирования – выделения тренда и сезонных составляющих
 - Знакомство с нейронными сетями
 - Знакомство с технологиями добычи данных (Data Mining)

Углубленный курс статистического анализа данных с использованием программного пакета **STATISTICA**

36 академических часов

1. Обзор курса: анализ данных на компьютере и его применение для решения практических задач в маркетинге, бизнесе, экономике, промышленности
2. Программа **STATISTICA** – универсальный инструмент анализа данных
3. Обзор программы **STATISTICA**: интерфейс, управление данными, аналитические методы, визуализация
4. Обзор многомерных методов анализа данных:
 - Дисперсионный анализ
 - Факторный анализ
 - Многомерное шкалирование
 - Кластерный анализ
 - Дискриминантный анализ
 - Деревья классификации
 - Множественная регрессия
 - Общие линейные модели
 - Обобщенные линейные и нелинейные модели
 - Общие регрессионные модели
 - Временные ряды и прогнозирование
5. Дисперсионный анализ (ДА):
 - Однофакторный ДА
 - Многофакторный ДА
 - ДА с повторными измерениями
 - ДА главных эффектов
 - Примеры исследований
 - Упражнения
6. Факторный анализ и метод главных компонент – обзор методов, систематическая работа в модуле, разбор практических примеров
7. Многомерное шкалирование - обзор методов, систематическая работа в модуле, разбор практических примеров
8. Современные методы классификации данных: кластерный анализ, дискриминантный анализ, деревья классификации:
 - Иерархическая кластеризация
 - Метод k-средних
 - Двухходовое объединение
 - Дискриминантный анализ стандартный и пошаговый
 - Деревья классификации
 - Сравнение методов
 - Примеры
 - Упражнения
9. Прогнозирование временных рядов:
 - Реализм в прогнозировании
 - Типы моделей
 - Горизонт и точность прогноза
 - Сравнительный обзор методов прогнозирования
 - Графический анализ и преобразование данных
 - Графическое представление
 - Как наглядно и эффективно представить результаты исследования
 - Заполнение пропусков
 - Методы сглаживания временных рядов
 - Простейшие методы прогнозирования
 - Модели экспоненциального сглаживания
 - Экспоненциальное сглаживание
 - Линейное экспоненциальное сглаживание
 - Методы декомпозиции временных рядов
 - Тренд и сезонность
 - Случайная компонента
 - Выделение циклической компоненты
 - Выделение скрытых периодичностей – спектральный анализ
 - Введение в спектральный анализ
 - Анализ периодограммы
 - Линейные модели прогнозирования: классическая ARIMA (АРСС) модель
 - Условия стационарности линейного процесса
 - Методология Бокса-Дженкинса идентификации ARIMA модели
 - Прогнозирование на основе ARIMA модели
 - Анализ остатков
 - Автоматический перебор моделей
 - Анализ распределенных лагов
 - Нейросетевой подход к решению задач
 - Основные парадигмы нейронных сетей
 - Прогнозирование временных рядов
 - Сравнение нейросетевых и классических методов
 - Прогнозирование временных рядов с короткой историей
10. Нейронные сети
11. Data Mining
12. Big Data
 - Что такое Big Data?
 - Платформы и технологии
 - Примеры решений