ЧАСТНОЕ ПРОФЕССИОНАЛЬНОЕ ОБРАЗОВАТЕДЛЬНОЕ УЧРЕЖДЕНИЕ «СЕВЕРО-КАВКАЗСКИЙ КОЛЛЕДЖ ИННОВАЦИОННЫХ ТЕХНОЛОГИЙ»

ЛЕКЦИЯ 5. ТЕМА 2.3. ТЕХНОЛОГИЯ ОБРАБОТКИ ЧИСЛОВОЙ ИНФОРМАЦИИ

1. Общие сведения об обработке числовой информации

Обработка числовой информации — это совокупность методов и процедур преобразования, анализа, хранения и визуализации количественных данных с применением программно-аппаратных средств. В современной ИТ-среде она реализуется преимущественно в средах электронных таблиц, СУБД, специализированных математических и статистических пакетов, а также в средах программирования.

Числовая информация характеризуется дискретностью значений, возможностью применения арифметических и логических операций, а также статистической обработки. Её ключевые свойства: измеримость, упорядоченность, возможность агрегации (суммирование, усреднение, дисперсия и др.), масштабируемость и совместимость с алгоритмами машинного обучения и прогнозирования.

В профессиональной деятельности обработка числовой информации применяется для учёта, планирования, бюджетирования, контроля качества, анализа эффективности, моделирования и принятия управленческих решений.

2. Технологии обработки числовой информации

Современные технологии обработки числовой информации классифицируются по способу реализации и назначению:

- **Локальные табличные процессоры** (Microsoft Excel, LibreOffice Calc, Google Sheets) — для оперативного ввода, расчётов, построения диаграмм и фильтрации небольших и средних массивов данных.
- **Системы управления базами данных** (СУБД: PostgreSQL, MySQL, MS Access) — для хранения, выборки, агрегации и анализа больших массивов числовых данных с обеспечением целостности и многопользовательского доступа.
- **Математические и статистические пакеты** (Mathcad, MATLAB, R, SPSS) — для сложных вычислений, моделирования, регрессионного и факторного анализа.
- **Языки программирования** (Python, C++, Java) с библиотеками для работы с данными (pandas, NumPy) — для автоматизации, интеграции и масштабируемых вычислений.
- **BI-системы и средства визуализации** (Power BI, Tableau) — для построения дашбордов, многомерного анализа и отчётности.

Выбор технологии определяется объёмом данных, требованиями к точности, скоростью обработки, необходимостью совместной работы и уровнем автоматизации.

3. Электронные таблицы: основные понятия и способы организации

Электронная таблица — это программное средство, представляющее данные в виде прямоугольной матрицы, состоящей из строк и столбцов, на пересечении которых располагаются ячейки. Каждая ячейка может содержать число, текст или формулу.

Способы организации данных в электронной таблице:

- **Линейная структура** — данные располагаются в строках (записи) и столбцах (поля), что соответствует табличному представлению реляционной модели.
- **Иерархическая группировка** — с помощью функций группировки и сводных таблиц данные агрегируются по уровням (например, по периодам, подразделениям, категориям).
- **Сводные таблицы** — автоматизируют построение сводных отчётов, группировку, фильтрацию и расчёт итоговых показателей.
- **Динамические диапазоны и именованные области** — позволяют создавать гибкие ссылки, используемые в формулах и диаграммах.

Организация данных должна соответствовать принципам нормализации: отсутствие дублирования, атомарность значений, однозначность интерпретации полей.

4. Структура электронных таблиц: ячейка, строка, столбец. Адреса ячеек

Основные структурные элементы электронной таблицы:

- **Ячейка** — минимальная единица хранения данных, расположенная на пересечении строки и столбца.
- **Строка** — горизонтальный ряд ячеек, обычно нумеруется числами (1, 2, 3…).
- **Столбец** — вертикальный ряд ячеек, обозначается буквами латинского алфавита (A, B, C…; при превышении — AA, AB и т. д.).

Адреса ячеек формируются как комбинация буквенного обозначения столбца и номера строки: A1, B5, Z100. Существуют два типа адресации:

- **Относительная** (A1) — адрес изменяется при копировании формулы в другую ячейку.
- **Абсолютная** (AAA1) — адрес фиксируется и не изменяется при копировании.
- **Смешанная** (A1илиAA1 или AA1илиA1) — фиксируется только строка или только столбец.

Адресация используется в формулах для ссылок на данные, что позволяет автоматизировать расчёты и обеспечить гибкость моделей.

5. Строка меню. Панели инструментов

Строка меню содержит основные команды управления документом: «Файл», «Правка», «Вид», «Вставка», «Данные», «Формат», «Сервис», «Окно», «Справка». Каждая команда открывает подменю с набором операций.

Панели инструментов предоставляют быстрый доступ к наиболее часто используемым функциям:

- **Стандартная панель** — содержит кнопки для создания, открытия, сохранения, печати, отмены и повтора действий.
- **Панель форматирования** — позволяет изменять шрифт, выравнивание, границы, заливку ячеек.
- **Панель формул** — отображает и позволяет редактировать формулы, вводимые в активную ячейку.
- **Панели диаграмм, сводных таблиц, сортировки и фильтрации** — обеспечивают специализированные операции по анализу данных.

Настройка панелей инструментов позволяет адаптировать интерфейс под конкретные задачи пользователя.

6. Ввод данных в таблицу. Типы и формат данных: числа, формулы, текст

Ввод данных осуществляется непосредственно в активную ячейку или через строку формул. Данные могут быть введены вручную, импортированы из внешних источников (CSV, базы данных) или сгенерированы с помощью функций.

Типы данных:

- **Числа** — целые и дробные значения, используемые для расчётов. Поддерживают арифметические операции и статистические функции.
- **Текст** — строковые значения, применяемые для заголовков, описаний, меток. Не участвуют в вычислениях, но могут использоваться в логических и текстовых функциях.
- **Формулы** — выражения, начинающиеся со знака «=», которые выполняют вычисления на основе значений других ячеек и встроенных функций.

Форматы данных:

- Числовой, денежный, процентный, дата/время, текстовый, пользовательский.
- Форматирование не изменяет значение данных, а только способ их отображения.

7. Алгоритмы действий при работе с данными

Алгоритм ввода и форматирования данных:

1. Выделить диапазон ячеек для ввода данных.
2. Ввести значения вручную или импортировать из внешнего источника.
3. Применить соответствующий формат к столбцам (числовой, дата, текст и т. д.).
4. Проверить корректность данных (отсутствие ошибок, пустых ячеек, несоответствий).
5. При необходимости использовать условное форматирование для выделения аномальных значений.

Алгоритм создания формул:

1. Выбрать целевую ячейку для результата.
2. Ввести знак «=» для начала формулы.
3. Указать ссылки на ячейки или ввести константы.
4. Применить встроенные функции (СУММ, СРЗНАЧ, ЕСЛИ и др.) при необходимости.
5. Нажать Enter для подтверждения формулы.
6. Скопировать формулу в другие ячейки с учётом типа адресации.

Алгоритм построения диаграммы:

1. Выделить диапазон данных, включая заголовки столбцов и строк.
2. Выбрать тип диаграммы (гистограмма, график, круговая и др.) в меню «Вставка».
3. Настроить элементы диаграммы: заголовок, подписи осей, легенду.
4. При необходимости изменить формат отдельных элементов (цвета, шрифты, маркеры).
5. Разместить диаграмму на текущем или отдельном листе.

Алгоритм фильтрации и поиска данных:

1. Выделить заголовок таблицы или весь диапазон данных.
2. Активировать фильтр (меню «Данные» → «Фильтр»).
3. Установить критерии фильтрации в выпадающих списках столбцов.
4. Для сложных условий использовать расширенный фильтр с отдельной областью критериев.
5. Для поиска конкретного значения использовать функцию «Найти и заменить».

8. Построение диаграмм и графиков. Форматирование готовых диаграмм

Диаграммы и графики служат для визуализации числовых данных и выявления закономерностей. Основные типы:

- **Гистограммы** — для сравнения величин по категориям.
- **Графики** — для отображения динамики показателей во времени.
- **Круговые диаграммы** — для представления долей в общем объёме.
- **Точечные диаграммы** — для анализа корреляции между двумя переменными.
- **Линейчатые диаграммы** — альтернатива гистограммам при большом количестве категорий.

Форматирование диаграмм включает:

- Изменение типа диаграммы.
- Редактирование заголовков, подписей осей, легенды.
- Настройка цветов, стилей линий, маркеров.
- Добавление вспомогательных элементов: сетки, трендовых линий, подписей данных.

Корректное форматирование повышает читаемость и информативность визуализации.

9. Обеспечение поиска и фильтрации данных. Типы критериев

Поиск и фильтрация позволяют оперативно извлекать нужные данные из больших массивов.

Методы поиска:

- Быстрый поиск по значению (Ctrl+F).
- Поиск по формату или примечаниям.
- Использование функций поиска (ВПР, ГПР, ПОИСКПОЗ) для извлечения данных по условию.

Фильтрация данных:

- **Автофильтр** — простая фильтрация по одному или нескольким столбцам с использованием условий «равно», «не равно», «больше», «меньше», «содержит» и др.
- **Расширенный фильтр** — позволяет задавать сложные условия в отдельной области критериев, включая логические операторы И/ИЛИ.

Типы критериев фильтрации:

- **Текстовые** — по совпадению, началу, окончанию или содержанию подстроки.
- **Числовые** — по диапазону значений, равенству, неравенству.
- **Дата/время** — по периодам, конкретным датам, диапазонам.
- **Пользовательские** — на основе формул, возвращающих логическое значение (ИСТИНА/ЛОЖЬ).

Эффективное применение поиска и фильтрации существенно повышает производительность работы с данными.