

Словарь начинающего аналитика

Полезен тем, что:

- ❖ Содержит расшифровку основных терминов, которые будут встречаться в курсе
- ❖ Всегда будет под рукой, и к нему можно обратиться в любой момент при обучении/работе

Для удобства термины расположены по порядку, исходя из последовательности лекций

Лекция 3

Визуальный анализ данных — это процесс, в котором исследование данных происходит с помощью наглядного отображения информации.

Временная метка (Timestamp) — это последовательность символов или закодированной информации, показывающей, когда произошло определённое событие. Обычно показывает дату и время (иногда с точностью до долей секунды).

Гистограмма (столбчатая диаграмма) — это визуальное изображение данных, отражающее изменение нескольких видов данных за какой-то период времени. Представляет собой набор столбиков, каждый из которых отвечает за своё значение. Применяется для сравнения каких-либо параметров.

График — это визуальное изображение зависимости какой-либо величины от другой, дающее наглядное представление о характере изменения. Применяется, чтобы отразить изменения одного и того же показателя во времени.

Диаграмма — это графическое представление данных, позволяющее быстро оценить соотношение нескольких величин. Позволяет проанализировать динамику изменения показателя во времени (например, объём продаж по месяцам); сравнить одни показатели с другими (например, объём продаж в разных регионах); сравнить доли между собой (например, какую долю занимают каналы продаж в процентах); сравнить, как соотносятся между собой данные (например, как изменяется объём продаж с ростом затрат на маркетинг).

Именованный диапазон — это область ячеек, имеющая определённое название, присвоенное пользователем, которое расценивается как адрес указанной области.

Круговая диаграмма — это визуальное изображение данных, позволяющее отразить пропорции или доли показателя относительно целого.

Парсинг данных — это автоматизированный сбор открытой информации в интернете по заданным условиям. Для этого используются специальные программы — парсеры сайтов. Они собирают неструктурированную информацию, преобразуют её и выдают в удобном виде.

Точечный график (Scatter plot) — это визуальное представление данных, изображающее значения двух переменных в виде точек на декартовой плоскости (координаты x и y). Применяется для иллюстрации и сравнения числовых значений, например, научных, статистических или технических данных.

Условное форматирование — это инструмент для наглядного представления результатов, который позволяет выделять цветом значения таблиц по заданным критериям, искать дубликаты, а также графически «подсвечивать» важную информацию.

Условный оператор — это оператор, обеспечивающий выполнение определённой команды (набора команд) только при условии истинности некоторого логического выражения либо выполнение одной из нескольких команд (наборов команд) в зависимости от значения некоторого выражения. Применяется, когда выполнение или невыполнение некоторого набора команд должно зависеть от выполнения или невыполнения некоторого условия.

Функции в анализе данных — это инструмент для анализа данных, позволяющий производить вычисления с данными и обрабатывать их. Применяется, когда данных много и формулу нужно применить для большого количества значений (например, посчитать общую сумму множества ячеек). Например, СУММ, СРЗНАЧ.

Функция ВПР (VLOOKUP) — это функция для поиска и извлечения данных из определённого столбца в таблице. Позволяет работать с большими таблицами и базами данных.

Часовой пояс (Time zone) — это географическая область, в которой действует одно и то же время.