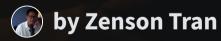# Descriptive Analytics – Understanding Historical Data

This handout introduces descriptive analytics, focusing on summarizing historical data to identify patterns and trends. By understanding past data, you can make informed decisions that guide future strategies.

by Zenson Tran

What is Descriptive Analytics?

Definition: Descriptive analytics is the process of analyzing historical data to understand what happened in the past. It provides a high-level summary and a foundational understanding of data patterns.

Importance: By analyzing past performance, businesses can identify patterns, assess performance, and uncover insights to help guide future decisions.

Example: A retail company analyzing monthly sales data to discover seasonal trends.

Basic Descriptive Statistics

- Mean (Average)

- Median

- Mode

- Standard Deviation

The mean is the sum of all values divided by the number of values. Example: Calculating the mean sales amount for a quarter to find the "typical" sales figure. Note: Outliers (extremely high or low values) can skew the mean.

The median is the middle value in a sorted dataset, offering a "center" value less impacted by outliers. Example: Finding the median customer age in a diverse dataset to understand the core customer demographic.

The mode is the most frequently occurring value in a dataset, revealing popular choices or categories. Example: Identifying the most purchased product type to understand customer preferences.

Standard deviation measures data variability around the mean, indicating data consistency or volatility. Example: A high standard deviation in monthly sales suggests fluctuating performance, while a low standard deviation indicates stability.

Practical Data Summarization Techniques

1 **Summing and Counting**

Calculate totals and counts to get an overall view of the data. Example: Counting the number of transactions to understand purchase frequency.

2 **Grouping and Aggregating**

Group data by categories (e.g., product type or region) and calculate summary statistics (like averages or totals). Example: Summing sales by region to see which areas perform best.

3 **Filtering for Insights**

Narrow down data by specific criteria, like time periods or customer segments, to focus on relevant sections. Example: Filtering sales data to analyze performance in specific quarters or regions.

Practical Application: Descriptive Analysis Walkthrough

Objective: Practice descriptive analysis on a sample dataset.

Steps:

1. Calculate Descriptive Statistics: Apply mean, median, mode, and standard deviation to understand key metrics.

2. Aggregate Data: Group data by categories (e.g., product, month) and calculate summary statistics.

3. Interpret Results: Identify patterns, such as high-performing products or monthly trends.

Outcome: By completing this exercise, you'll gain a foundational understanding of summarizing historical data to derive insights.

Key Takeaways

Descriptive Analytics for Historical Insight: Provides a solid foundation for understanding past performance and identifying trends.

Essential Descriptive Statistics: Mean, median, mode, and standard deviation help summarize and interpret data effectively.

Data Summarization Techniques: Summing, counting, grouping, and filtering are fundamental methods for exploring large datasets.

Reflective Questions

How can the mean and median give different insights when analyzing a dataset with outliers?

Why might standard deviation be helpful for understanding the consistency of monthly sales figures?

In what scenarios would mode be more insightful than mean or median?

Tips for Applying Descriptive Analytics

Double-check outliers: They can skew data, so consider median or mode if outliers are significant.

Use standard deviation to assess stability: Low standard deviation means data points are close to the mean, while high values indicate spread and variability.

Group data to see trends: By categorizing data, you can identify key insights for specific products, regions, or time periods.