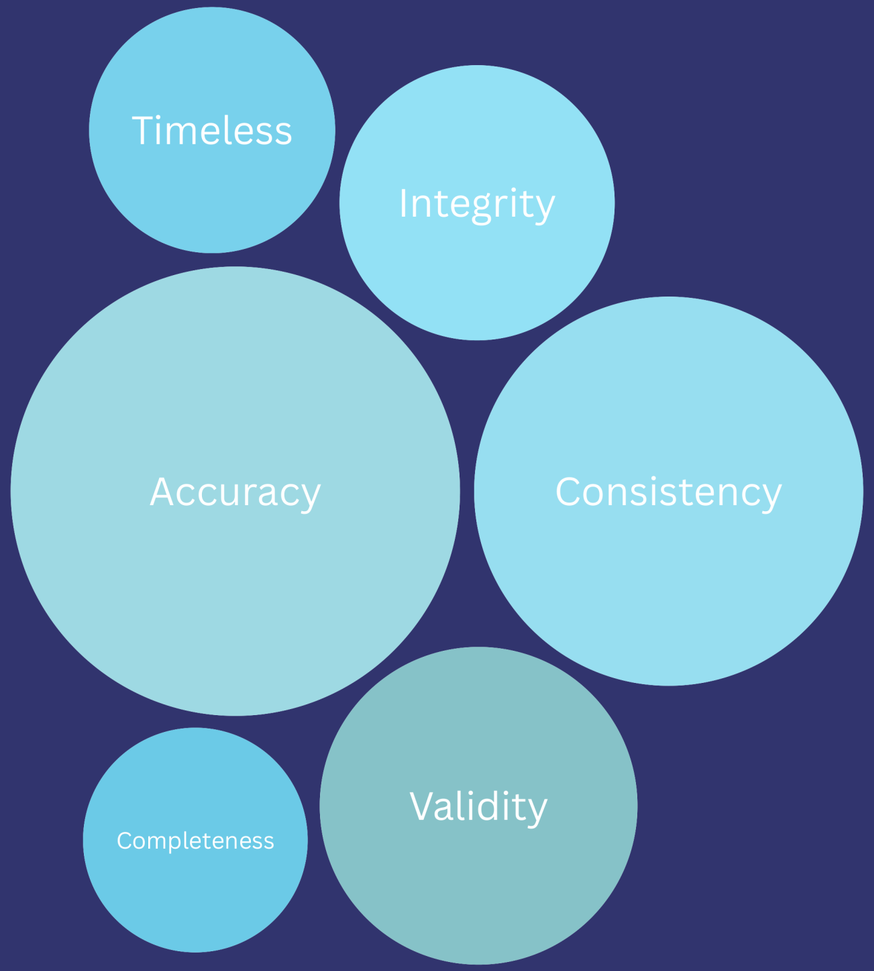
TESTING DATA QUALITY

Ensuring Reliable and Accurate Data

By: John Emmanuel A. Reyes



WHAT IS DATA QUALITY?

The measure of how well a data set serves its purpose, based on characteristics like accuracy, completeness, consistency, validity, uniqueness, and timeliness.

High-quality data is fit for its intended purpose and adds value. Conversely, a low data quality indicates poor, misleading data that can lead to poor decision-making.

IMPORTANCE OF TESTING DATA QUALITY

- Organizations use data to drive decisions in marketing, product development, communication, and more.
- High-quality data simplifies processing and analysis, leading to faster insights and better decision-making.
- It reduces risk and costs by identifying and mitigating potential issues related to misinformation or incorrect data outputs.
- Producing high-quality data enhances customer satisfaction, especially when focusing on improving customer experience

GARBAGE IN, GARBAGE OUT

GARBAGE DATA

Non-compliant Data

Data Entry Errors

Wrong Data Format

Inconsistencies

Nulls

Data Misinterpretation

Transposition Error

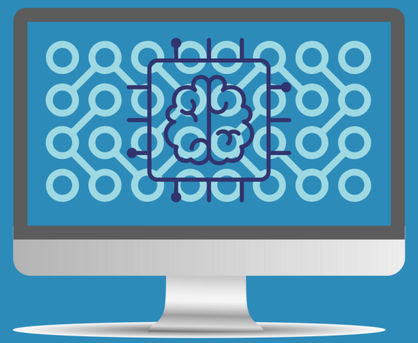
Duplicates

Incomplete Data

Lack of informative Data

Biased Data

PERFECT MODEL



GARBAGE RESULTS

Inaccurate Information

Misleading insights

Bad Decision-Making

Reputation Risk

Increased Risk

Financial Risk

Duplicated Efforts

Misallocation

Decreased Trust

Poor data quality destroys business value.

Recent Gartner research has found that organizations believe poor data quality to be responsible for an average of $15 million per year in losses.[4]

Testing Data Quality

1. Data Profiling
2. Data Validation & Integrity
3. Data Standardization
4. Data Monitoring

CODE REPOSITORY LINK

TESTING DATA QUALITY

Ensuring Reliable and Accurate Data

By: John Emmanuel A. Reyes

CODE REPOSITORY LINK

REFERENCES

[1] HEAVY.AI. (n.d.). Data Quality. Retrieved July 24, 2024, from https://www.heavy.ai/technical-glossary/data-quality

[2] Atlan. (2023, December). Data quality testing. Atlan. Retrieved July 24, 2024, from https://atlan.com/data-quality-testing/

[3] Profisee. (2023, October). Garbage in, garbage out: Understanding data quality and its impact. Profisee. Retrieved July 24, 2024, from https://profisee.com/blog/garbage-in-garbage-out/

[4] Gartner. (2018). How to create a business case for data quality improvement. Gartner. Retrieved from https://www.gartner.com/smarterwithgartner/how-to-create-a-business-case-for-data-quality-improvement#:~:text=Poor%20data%20quality%20destroys%20business,million%20per%20year%20in%20losses.