

DATA QUALITY AND MANAGEMENT

Lecturer: Mark Gillan

DATA QUALITY

- Quality data: the degree to which data is fit for its intended purpose.
- All analysis is only as good as the data it is carried out on.
- Therefore, the quality of the underlying data is critical to any analysis.



DATA QUALITY — IMPACT - BENEFITS

Impacts of poor data quality

- Analysis rework
- Organisational inefficiencies
- Customer dissatisfaction
- Opportunity cost of missed sales
- Reputational costs from loss of trust
- Compliance costs/fines from incorrect reporting

Benefits of high-quality data

- Improved customer experience
- Reduced risk
- Competitive advantage from accurate insights
- Increased revenue
- Higher staff productivity



MEASURING DATA QUALITY

- Quality data does not mean perfect data. High quality data is data that is good enough for the activity it is being used for.
- There are many ways data can be incorrect.
- It could be missing, out of date, duplicated, in the wrong format, doesn't match the same information elsewhere, or just plain wrong.
- The different ways that data can be incorrect are organised into six different areas called the **Dimensions of Data Quality**.
- It is important to check quality against as many of the different dimensions as necessary when reporting data quality.



DIMENSIONS OF DATA QUALITY - CUTVAC



- **Completeness:** how non-blank or populated the data is.
- **Uniqueness:** data is not recorded more than once - identifies duplicates.
- **Timeliness:** how up to date the data is.
- **Validity:** that data is in the correct format, type and range.
- **Accuracy:** data represents the real world.
- **Consistency:** data matches if two copies of the same information are compared.



REMINDER



Define what is meant by 'Data Quality'



Name some impacts of poor-quality data



Name some benefits of high-quality data



What are the Dimensions of Data Quality?



Does CUTVAC help?

