The Joy of Data: *Levels of Abstraction*

wayne.smith@csun.edu

[updated: Thursday, November 28, 2024]

"One should have a devotion to abstraction and a passion for details."
--Alfred North Whitehead (1861-1947)

Introduction:
Using computing technology to analyze data involves making design decisions about the best "level of abstraction". A "too-low" level of abstraction makes the analysis difficult to generalize broadly and apply appropriate statistical methods. A "too-high" level of abstraction makes too many assumptions on the part of many stakeholders and is difficult to justify or reproduce. Finding the best level of abstraction takes careful thought and practice. Sometimes, an analysis that incorporates the best elements of both levels of abstraction is the shrewd strategy.

The following are examples of levels of abstraction common in business analytics.

Design Choice	**Low-level**	**High-level**	**Key Question(s)**
Data	You have access to an entire dataset.	You have access to a summary of a dataset.	Is an original analysis needed? Will the dataset fit into a single Worksheet?
Analysis (calc. mean)	Use =sum()/count()	Use =average() or Analysis Addin	What does the reader or listener know or use?
Business Context	The company or organizational unit is a very technological firm.	The company or organizational unit is <u>not</u> a very technological firm.	Do you know the company and industry culture well enough?
Written	The company norm is to always "show your work."	Individuals make the assumption that you have double-checked your work.	Can you expect your written to be circulated widely in the organizational unit, company, or industry?
Presentation	You have yet to establish that you can use evidence well.	You have established your general credibility.	Do you know who will be in the audience on the day of the business meeting?