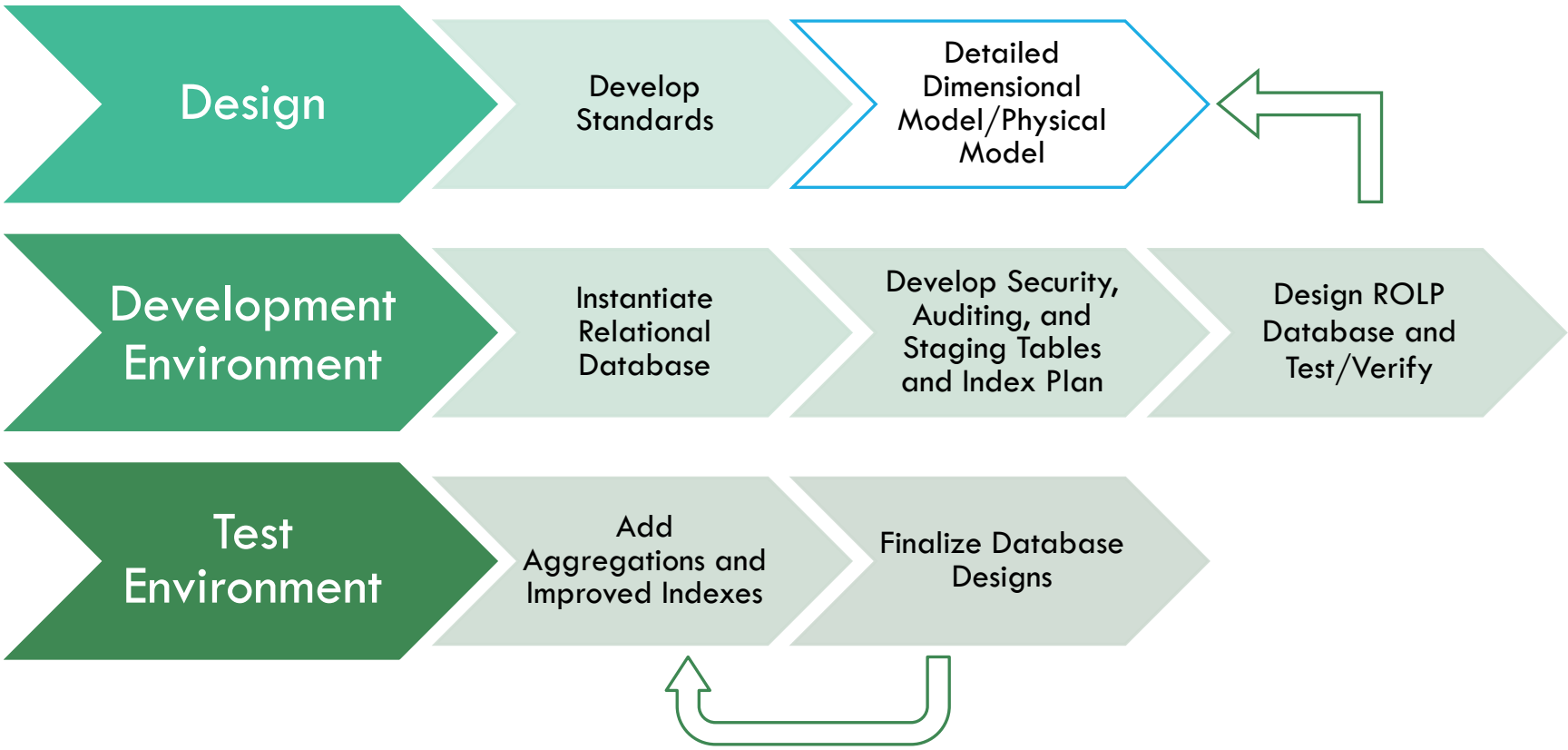
Data Type Selection

School of Information Studies
Syracuse University

The Physical Model



Every Byte Counts in a Fact Table!!!

	Total bytes of data types	88 Bytes
X	Total number of rows	5 Billion
	Total amount of disk storage for this fact table	**440 GB**

Data Types

Data Type	Description	Storage
smallint	Integers -32K to +32K	2 bytes
int	Integers -2B to +2B	4 bytes
bigint	Integers $-9B^2$ to $+9B^2$	8 bytes
decimal(p,s)	Decimal values from -10^{38} to $+10^{38}$	5-17 bytes
money	Monetary values -900T to + 900T	8 bytes
datetime	Date/times from 1/1/1753 to 12/31/9999	8 bytes
char(n)	Fixed-character string specific charset	n bytes
nchar(n)	Fixed-character string Unicode charset	2n bytes
varchar(n)	Variable length string specific charset	2 + n bytes
nvarchar(n)	Variable length string Unicode charset	2 + 2n bytes

Saving Space With Types

- Don't use money or decimal when int will suffice.
  - Savings: 1 to 13 bytes
- Don't use nvarchar unless you require Unicode.
  - Savings: 4 bytes
- Don’t use char(1) for Yes/No; use bit.
  - Savings: 0 to 7 bytes (depending on the number of bit fields)
- Use smallint instead of int.
  - Savings: 4 bytes
- Use int YYYYMMDD for date instead of datetime.
  - Savings: 4 bytes