



Metadata and Data visualization in Clinical Trials

Metadata in the Pharma Industry

Presented by
Bharath

Agenda

- What is Metadata?
- Types of Metadata
- Advantages of Metadata
- Changes to be done in meta data to facilitate clinical programming.

Introduction

Metadata is simply data about data, which is a way of establishing meaningful standards applicable to electronic data submitted for FDA review:

What information about the data can Metadata give?

- Description of the content.
- Context
- Structure.
- Purpose of a database

Types of Metadata

- **Dataset Level Metadata** – Data about datasets - Contains information about all the datasets present in the study.
- **Variable Level Metadata** – Data about variables- Contains information about all the variables present in the dataset.

Dataset Level Metadata

- Dataset Name
- Description
- Location
- Structure
- Purpose
- Key fields

Variable Level Metadata

- Variable Name
- Variable Label
- Type
- Decodes/ Formats
- Origin
- Role
- Comments

Is Our Metadata Robust

Is our Metadata Robust ?

- Yes

How ?

- Traceability

eg : We can trace back a variable and its derivation at any given point of time.

Changes to Metadata to facilitate Clinical Programming

Though our Metadata is robust in terms of submissions , Did we ever think of doing any changes to the metadata to facilitate programming ?

Vardef – Variable definition sheet

I) All variables in a single sheet ?

All variables in a single sheet helps find out the issues in a simple go.

Eg :Attributes related issues like length, Spell issues

Variable definition sheet - Example

DOMA	VARNAME	VARLABEL	TYPE	CTFORM	ROLE	ORIGIN	VARLEN	DEFCOM	INTCOM	STYLPRI	NOTES	CORE
EG	EGTESTCD	ECG Test or Examination	text	EGTESTCD	Topic	eDT, CRF Page 28				Uppcase test	Short name	Req
	IETESTCD	Inclusion/Exclusion Criteria	text		Topic	Assigned	200			Uppcase test	Short name	Req
LB	LBTESTCD	Lab Test or Examination	text	LBTESTCD	Topic	CRF Page 13, 14	8			Uppcase test	Short name	Req
RP	RPTESTCD	Repro System Findings	text	RPTESTCD	Topic	CRF Page 6	8			Uppcase test	Short name	Req

Custom AutoFilter

Show rows where:

VARNAME

ends with testcd

☒ And ☐ Or

Use ? to represent any single character
Use * to represent any series of characters

Cancel

Length Inconsistency

Vardef – Variable definition sheet - Example

- I) In the above example TESTCD variable in different domains have inconsistent attributes- i.e., Length which was not found at domain level sheet is found at variable level sheet when applied a filter on similar suffix variables.

Computational Algorithms sheet

2) All Computational Algorithms in a single sheet?

All Computational Algorithms in a single sheet serves as a one stop destination for a programmer to look for regular derivations.

Computational Algorithms :

Example

DOMAIN	VARNAME	VARLABEL	TYPE	CTFORMAT	ROLE	ORIGIN	VARLEN	INTCOMM
LB	LBBFL	Baseline Flag	text	NY	Record Qualifier	Derived	1	LB should be Sorted by USUBJID LBTESTCD LBCAT LBDTC; if last.LBTESTCD of LBDTC=<RFXSTDTC then LBBFL='Y' else LBBFL= NULL. Only non-missing results should be considered. Note: if RFXSTDTC is missing then consider RFSTDTC.
EG	EGBFL	Baseline Flag	text	NY	Record Qualifier	Derived	1	EG should be Sorted by USUBJID EGTESTCD EGDTC; if last.EGTESTCD of EGDTC=<RFXSTDTC then EGBFL='Y' else EGBFL= NULL. (Only non-missing results should be considered)

Computational Algorithms sheet - Example

In the above example:

RFSTDTC is used when RFXSTDTC is missing when LBLFL is being derived.

Unlike which RFSTDTC is not used in EGBLFL derivation which is clearly a manual error (Specification writer) which leads to the inconsistency.

Computational Algorithms sheet - Solution

Hyperlinking the general derivations/ computational algorithms to a single sheet from individual domain level sheets will serve the purpose of meeting integrity and achieving consistency.

Merging sheets to standardize the values?

3) Merging sheets to standardize the values?

eg : ROUTE

FREQ

DOSU etc

Introducing a sheet which contains all the study specific raw values and standardizations against it will facilitate programming, avoids rework on the source and validation mismatches.

Merging sheets to standardize values- Example

Name of Format	Code Value (From SDTM Dataset)	Code Text	source
UNIT	U	Unit	units
UNIT	U	Unit	u,
UNIT	U	Unit	unit
UNIT	mg	Milligram	mg
UNIT	mg	Milligram	1 mg
UNIT	mL	Milliliter	ml
UNIT	mL	Milliliter	milliliters
UNIT	mL	Milliliter	cc
UNIT	UNKNOWN	Unknown	unknown
UNIT	UNKNOWN	Unknown	unk
UNIT	TABLET	Tablet Dosing Unit	TABLETS
UNIT	TABLET	Tablet Dosing Unit	Tablet
UNIT	TABLET	Tablet Dosing Unit	tab
UNIT	TABLET	Tablet Dosing Unit	tabs
UNIT	TABLET	Tablet Dosing Unit	tablets
UNIT	TABLET	Tablet Dosing Unit	1 tab

SUPPQUAL Variables in Value level sheet

- 3) Including SuppQual variables in the value level sheet?
- In General we submit all our value level information in the value level sheet.
 - Including SuppQual variables in value level sheet gives the information about the supp level value information all together.

SUPPQUAL Variables in Value level sheet- Example

Source Variable	Variable Value	Value Label	Variable tpye	Controlled Terminology Codelist Name	Origin
QNAM	CMROUOTH	Route, Other	Float		CRF Page 6
QNAM	CMFRQOTH	Frequency, Other	Float		CRF Page 6
QNAM	RACEOTH	Race Other	Text		CRF Page 3
QNAM	RANDNUM	Randomization Number	Float		CRF Page 21
QNAM	DSABNAME	Additional Antibiotics	Text		CRF Page 24
QNAM	LBCLSIG	Clinically Significant	Text		CRF Page 5, 13, 24, 27
QNAM	ORNRIND	Original Referencel Range Indicator	Text		CRF Page 5, 13, 24, 27
QNAM	LBTYPE	Test Performed	Text		CRF Page 20
QNAM	ADVEVENT	Reason Not Seen For Evaluation	Text		CRF Page 25
QNAM	ABNEEDED	Specify Antibiotics Needed	Text		CRF Page 25
QNAM	NOTSNDT	Date of Withdraw Consent	Text		CRF Page 25
QNAM	NOTSNCO	Reason for Not Seen	Text		CRF Page 25

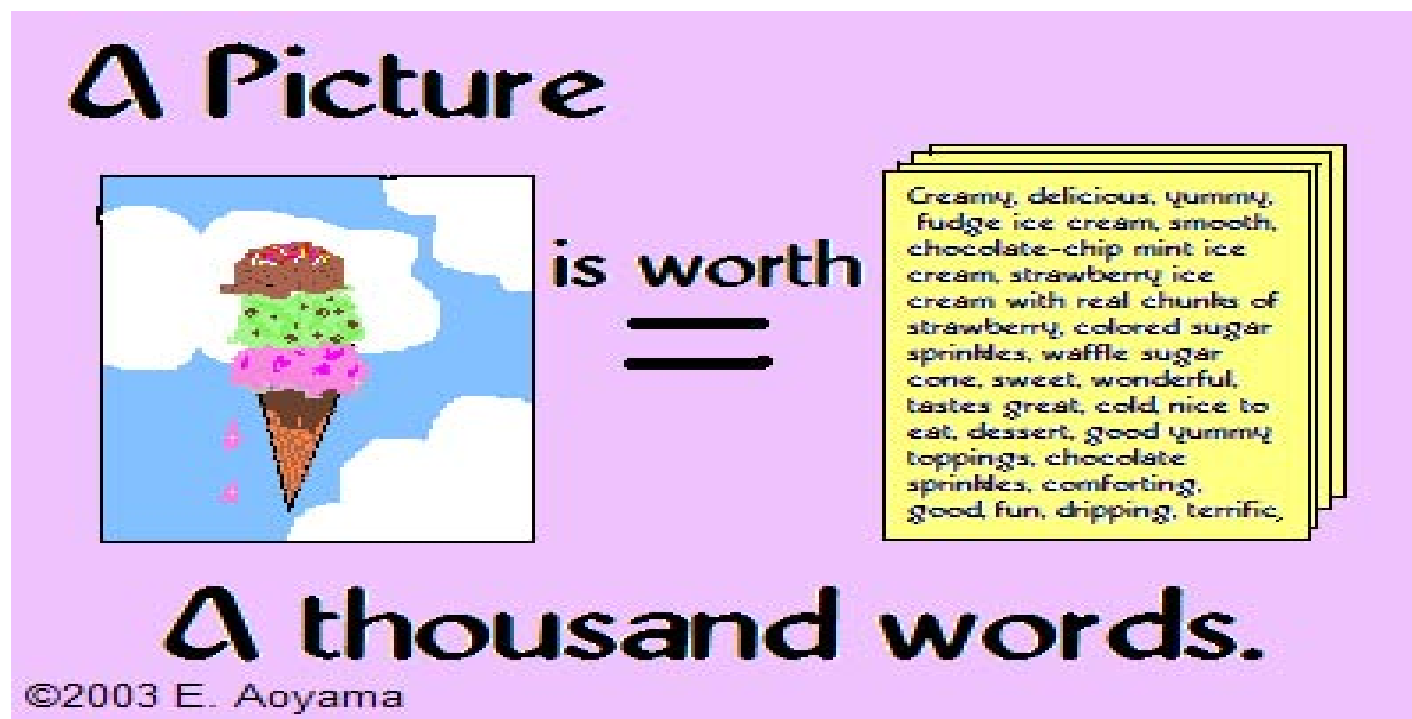
Data Visualization In Clinical Trials

Presented by

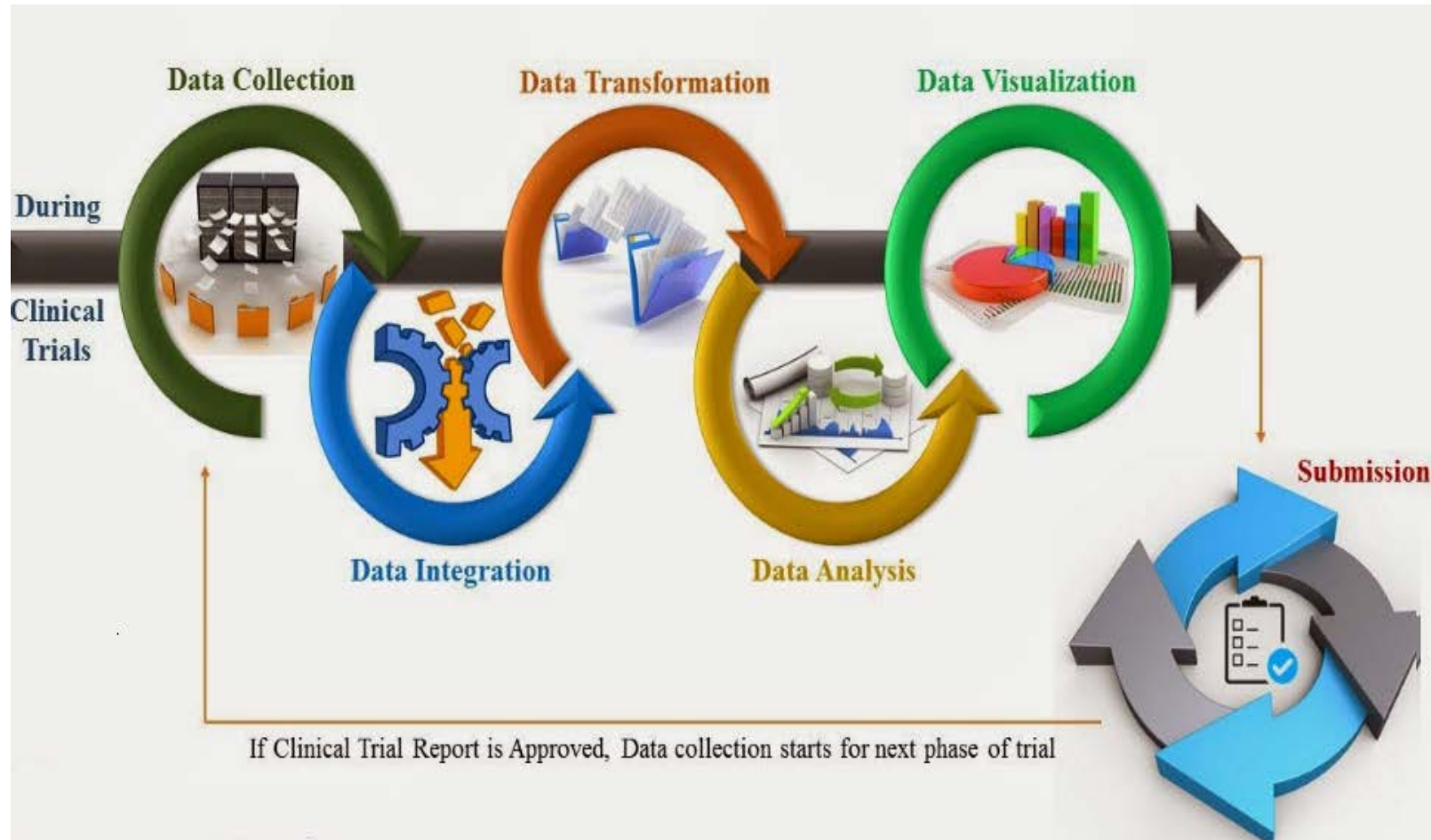
Raghava Reddy

What is Data Visualization?

- Data visualization is the process of displaying vast amount of data can be structured and viewed through a variety of charts and graphs to facilitate analysis.



Data Visualization flow chart in Clinical Trials.



Ex:AE Summary Table

Table 1. AE Summary Table

BODY SYSTEM/PREFERRED TERM	INTENSITY							
	Unknown		Mild		Moderate		Severe	
	N	(%)	N	(%)	N	(%)	N	(%)
PATIENTS WITH ADVERSE EXPERIENCE			xxx	(xx)	xxx	(xx)	xxx	(xx)
AUTONOMIC NERVOUS SYSTEM			xxx	(xx)	xxx	(xx)	xxx	(xx)
MOUTH DRY			xxx	(xx)	xxx	(xx)	xxx	(xx)
BODY AS A WHOLE GENERAL			xxx	(xx)	xxx	(xx)	xxx	(xx)
EDEMA DEPENDENT			xxx	(xx)	xxx	(xx)	xxx	(xx)
FEVER								
PAIN								

Ex:AE Summary Visual Analytics

AE During On – therapy by Intensity

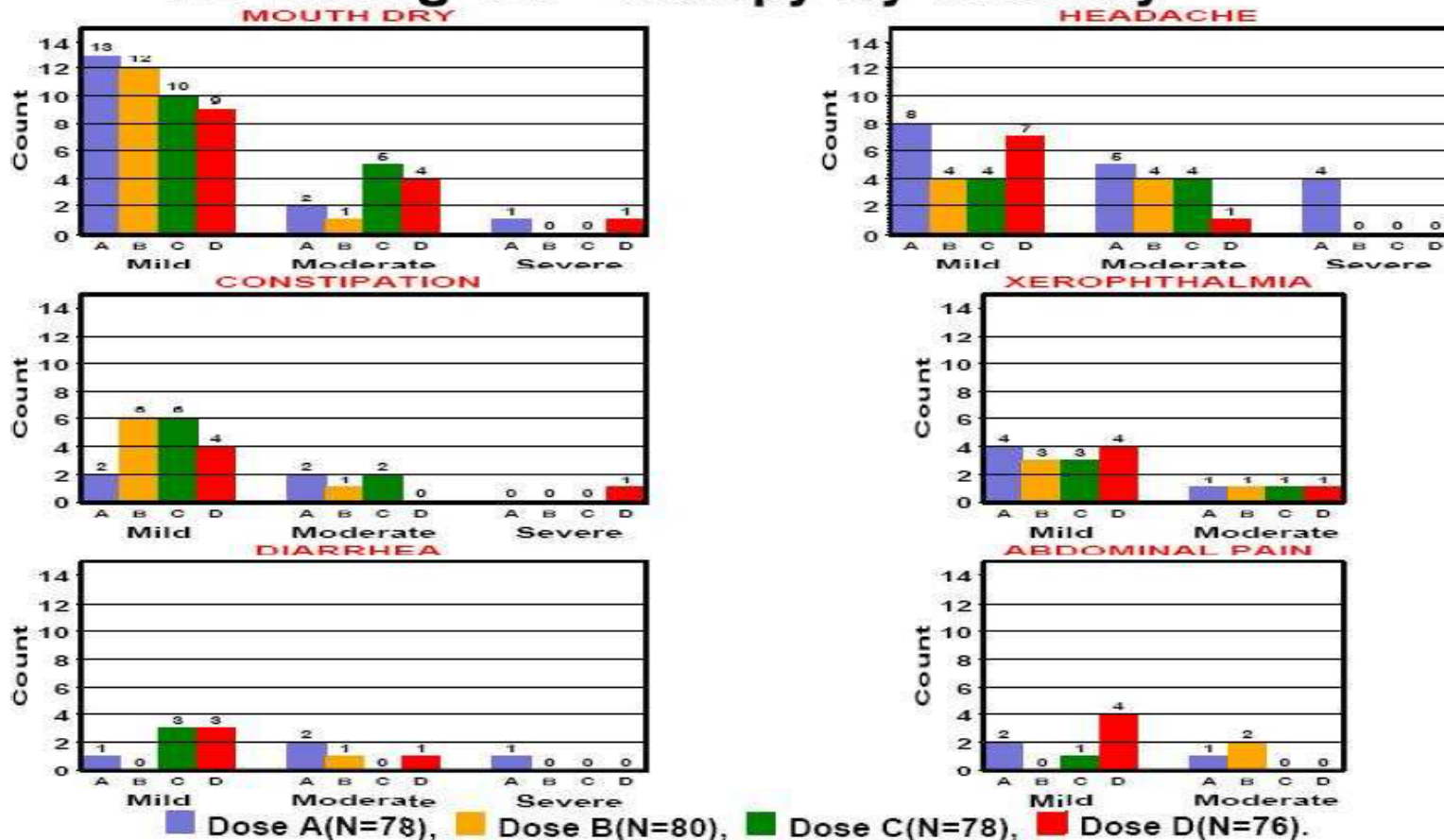


Figure 1. Bar Chart Display of AE by Intensity

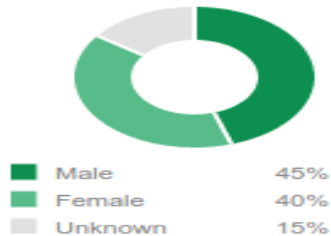
Ex: Demographics Summary Table

Table DM 1
Demographics
(Format 1)
All Subjects

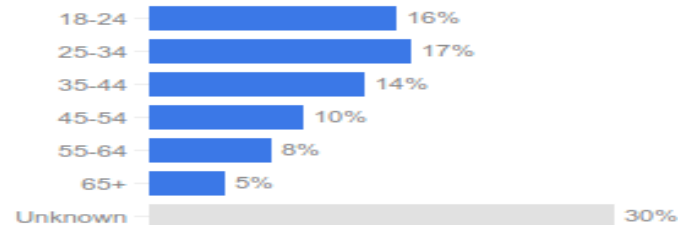
Parameter / Summary	Placebo (N=xxx)	Trt 1 (N=xxx)	Total (N=xxx)
Age (year)			
n	xxx	xxx	xxx
mean (SD)	xx (x)	xx (x)	xx (x)
median	xxx	xxx	xxx
range	(xxx, xx)	(xxx, xx)	(xxx, xx)
Sex			
Male	xxx (xx)	xxx (xx)	xxx (xx)
Female	xxx (xx)	xxx (xx)	xxx (xx)
Race			
Caucasian	xxx (xx)	xxx (xx)	xxx (xx)
Black	xxx (xx)	xxx (xx)	xxx (xx)
Hispanic	xxx (xx)	xxx (xx)	xxx (xx)
Asian	xxx (xx)	xxx (xx)	xxx (xx)
Other	xxx (xx)	xxx (xx)	xxx (xx)

Ex: Demographic Data Visualization

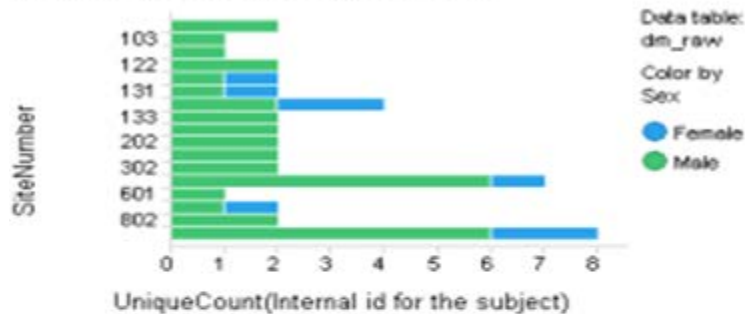
Gender



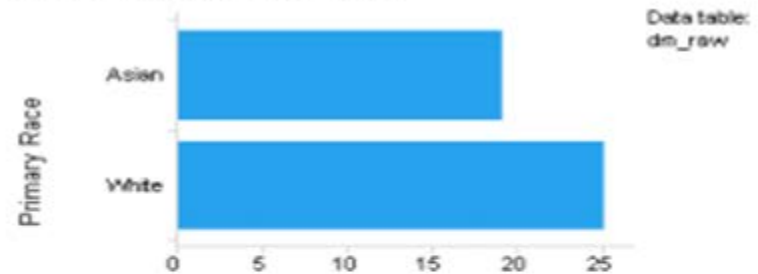
Age



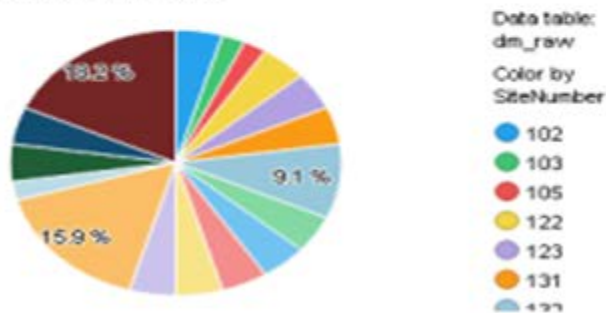
Unique Subjects per Site and Sex



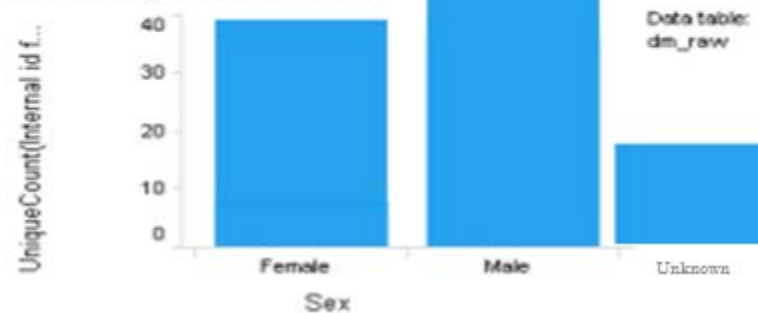
Unique Subjects per Race



Unique Subjects per Site



Unique Subjects per Sex



Ex: Summary of Demographic VA



Why Data Visualization?

- Data Visualization is powerful way to display large data coherently.
- Reviewers can quickly understand high-level overview information.
- Compared to a large table with lots of rows and columns, visualization is a easier way to Understand.
- Save reviewer time to make important decisions.
- Dynamic and Interactive presentation of reports.
- Present many numbers in a small space.

Tips to Get Started:

- The same data can be displayed using different types of graphs.
- Each graph type enhances the different aspects of the data.
- Understand the data you are trying to visualize, including its size and Type.
- Determine what you are trying to visualize and what kind of information you want to communicate.
- Use a visual that conveys the information in the best and simplest form.

Common Techniques

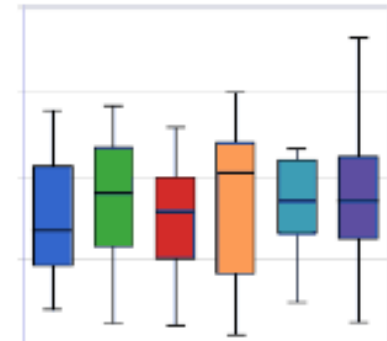
- **Charts:** bar or pie
- **Graphs:** good for structure, relationships
- **Plots:** 1- to n-dimensional
- **Maps:** one of most effective
- **Images:** use colour/intensity instead of distance (surfaces)
- 3-D surfaces and solids

Which chart type (or graph) to use for your data

- **Box Plot:**

Box plots are great for succinctly displaying the distribution of the data as well as the summary statistics such as mean and quartiles.

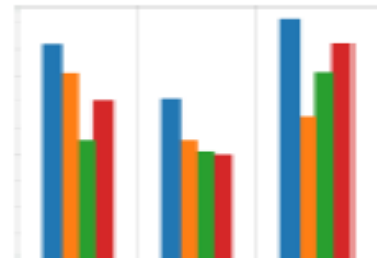
Box Plot



- **Bar Charts:**

Bar Charts are great for visualizing Comparisons (e.g. frequencies, means, count) across groups.

Bar Chart

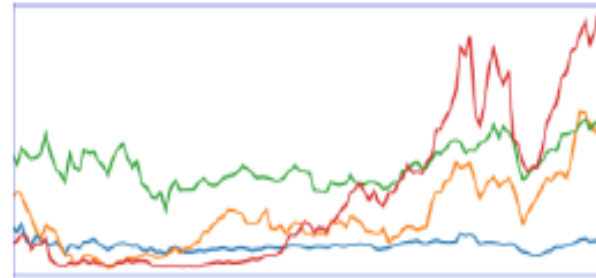


Which chart type (or graph) to use for your data

- **Line Charts:**

Line graphs are good for showing trends over time and show the relationship between two or more variables.

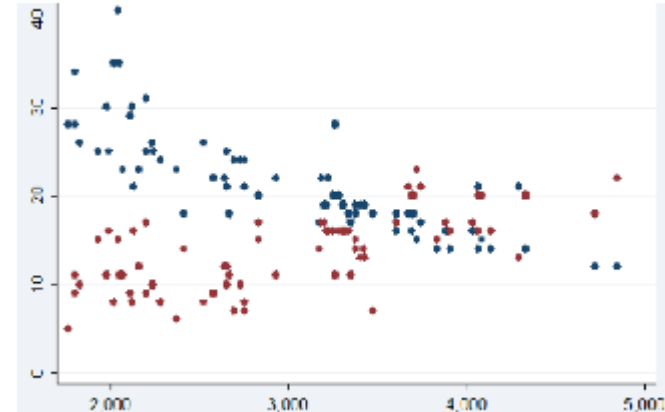
Line Graph



- **Scatter Plots:**

Scatter Plots show joint variation of two data items.

Scatter Plot



References:

- https://www.cdisc.org/system/files/all/reference_material_category/application/pdf/submissionmetadatamodelv2.pdf
- http://oels.byu.edu/student/idioms/proverbs/a_pic_thousand_words.html
- <https://support.google.com/adwords/answer/7072127?hl=en>
- <https://pharmasug.org/proceedings/2017/BB/PharmaSUG-2017-BB18.pdf>

Thank You!

Name: Bharath Kumar

Email: p.bharath2039@gmail.com

Contact: +91-9989812030

Name: Raghava Reddy Konda

Email: raghavareddykonda@gmail.com

Contact: +91-9492425267