# Central Tendency White Paper Requirements Specification

## Scripts for Figure 7.1 Box Plot – Observed Values of xxx Over Time

### Script Specification

#### Specific Output Requirements

* Generates a box plot of AVAL by AVISITN and TRTPN. See domain variations in Usage Requirements, below
* Basic script functionality (user options)
* ***data display features***
* Schematic or Tukey notched box plots, as described in [SAS/STAT 9.4 Graph Template Language: Reference (Boxplot Statement)](http://support.sas.com/documentation/cdl/en/grstatgraph/67882/HTML/default/p0vuh82v39fsasn1vqhzmhdl8y16.htm#n0g9wrnpg9zsryn1dk5csdrdfnqy)
* analyze by treatment
* group boxes by time point, and label clearly
* outliers:
* **IQR outliers** as box outlines, and
* **reference range outliers** as red dots
* reference ranges - options to draw reference lines for various scenarios (based on values found in data)
* **uniform:** DEFAULT. If all reference range values are uniform for a set of measures, draw those uniform low/high reference range line.
* **narrow:** If reference range values are NOT uniform for a set of measures, draw the only the *narrow* reference range lines (highest of the lows, lowest of the highs)
* **all:** Discouraged by white paper authors – draw all reference lines relevant to the set of measures.
* **none**
* **numeric value(s):** draw exactly these specific reference lines (e.g., a zero (0) line for change from baseline boxplots.
* ***User variability***
* location and name of each input data sets
* output location (user-specified path)
* parameterize core variables required for analysis & display:
* "Treatment name" which could be planned or actual
* "Treatment number" used for display order of Treatment Names
* AVAL (measured outcome)
* ANRLO (lower limit of reference range)
* ANRHI (upper limit of reference range)
* SAFFL (population flag)
* ANL01FL (analysis flag)
* ***basic readability features***
* maximum number of boxes to display on a page (basic paging control)
* Restrictions
* Footnote:
* Box plot type is schematic: the box shows median and interquartile range (IQR, the box height); the whiskers extend to the minimum and maximum data points within 1.5 IQR of the lower and upper quartiles, respectively. Values outside the whiskers are shown as outliers. Means are marked with a different symbol for each treatment. Red dots indicate measures outside the normal reference range *(indicated by red lines ...)*.
* The parenthetical clause *(indicated by red lines ...)* should be dynamic, based on whether lines appear and which lines the user specified.

#### General Output and Formatting Requirements

See the specification document **CS\_GeneralOutputandFormattingRequirements.docx**.

### Usage Requirements

#### Required Input

##### Vital Signs Domain

* Dataset: ADVS
* Variables: STUDYID, USUBJID, SAFFL, ANL01FL, TRTP, TRTPN, PARAM, PARAMCD, AVAL, ANRLO, ANRHI, AVISIT, AVISITN, ATPT, ATPTN
* Record selection: SAFFL='Y' and ANL01FL='Y'

##### Laboratory Domain

* Dataset: ADLBC or ADLBH or ADLBHY?
* Variables: STUDYID, USUBJID, SAFFL, ANL01FL, TRTP, TRTPN, PARAM, PARAMCD, AVAL, A1LO, A1HI, AVISIT, AVISITN~~, ATPT, ATPTN~~
* Record selection: SAFFL='Y' and ANL01FL='Y'

##### ECG Domain

* Dataset: ???
* Variables: ???
* Record selection: ???

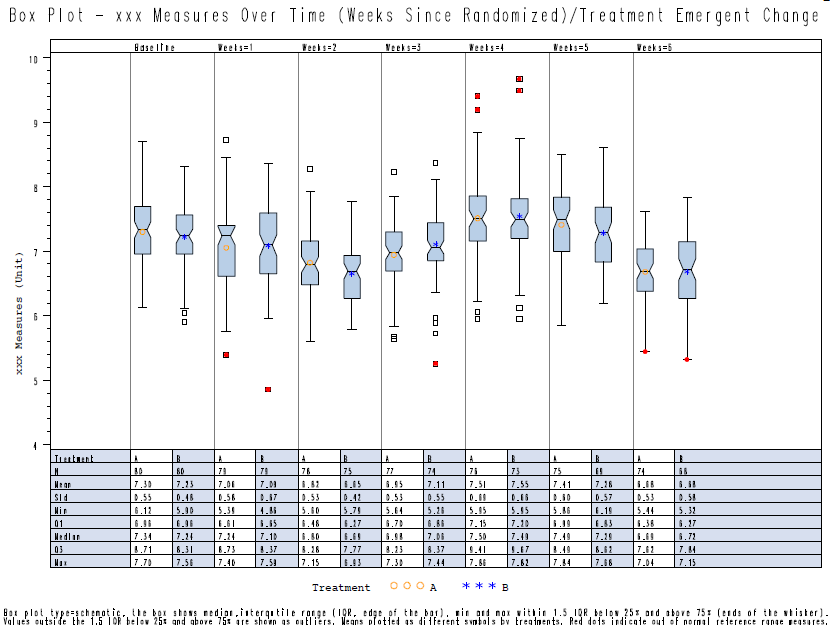
#### Macros

* Requires the PhUSE CS macro library:   
   <https://github.com/phuse-org/phuse-scripts/tree/master/whitepapers/utilities>
* User must ensure that SAS can find PhUSE CS macros in the SASAUTOS path

### Test Data

* Data sets from <https://github.com/phuse-org/phuse-scripts/tree/master/data/adam/cdisc>
* Domain data
  + Vital Signs Domain
  + ADVS – Vital signs
  + ADLBC, ADLBH, ADLBHY – Laboratory measures
  + ??? – ECG measures

### Sample Output



### Reference Documents:

White paper: <http://www.phusewiki.org/wiki/images/4/48/CSS_WhitePaper_CentralTendency_v1.0.pdf>

Programming Guidelines: <http://www.phusewiki.org/wiki/index.php?title=WG5_P02_Programming_Guidelines>