# Central Tendency White Paper Requirements Specification

## Scripts for Output defined in: 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
* QUESTION TO THE GROUP: Is this supposed to a configurable or hardcoded script? Example: should the 25% / 75% be configurable or hardcoded?
* *Basic script functionality (user options)*
* parameterize core variables required for analysis & display:
* notched box plot
* outliers: IQR outliers as boxes, reference range outliers as red dots
* reference ranges: options to draw reference lines for various scenarios (based on values found in data)
* location and name of each input data sets
* output location (user-specified path)
* AVAL (measured outcome)
* ANRLO (lower limit of reference range)
* ANRHI (upper limit of reference range)
* SAFFL (population flag)
* ANL01FL (analysis flag)
* maximum number of boxes to display on a page (basic paging control)
* on/off jitter option (to reduce overlap of reference range outliers)
* Footnote:
* Box plot type=schematic, the box shows median, interquartile range (IQR, edge of the bar), min and max within 1.5 IQR below 25% and above 75% (ends of the whisker). Values outside the 1.5 IQR below 25% and above 75% are shown as outliers. Means plotted as different symbols by treatments

#### General Output and Formatting Requirements

NOTE: These might or might not apply to the specific output or feature of the script.

**[DDT] Suggestion:**

* **Move all general requirements to a "CSS Standard Analyses General Output and Formatting Requirements" document.**
* **Move all "Central Tendency" requirements to a "Output and Formatting Requirements" section of the CT white paper**
* **Keep only target-specific requirements in this document**
* The page margins used in this document mirror the margins required for US regulatory submissions. TFL fits within 9.25” x 6” common text area.
* This document contains generic titles and footnotes to aid review.
* Titles must be centered.
* Footnotes must be left justified
* Placeholders for the program, output, date generated, and source data footnotes do not need to be displayed for each shell but will be produced for each TFL consistent with the standard macro output.
* P-values should follow the formats below (these guidelines align with medical journal requirements):
* For p > 0.1, report the p-value to 2 decimal places
* For 0.1 > p > 0.001, report the p-value to 3 decimal places
* For p < 0.001, report p < 0.001.
* Decimal places for continuous data summaries should follow these guidelines:
* The mean and estimates of precision (e.g., variance, SD, SE, Confidence Intervals) should be displayed to 1 more decimal place than the collected data as should distributional parameters (eg, Q1, Q3) except the Minimum and Maximum which are to presented to the same decimal places as the collected data.
* Percentages are to be displayed to at least one decimal place.
* All tables summarizing baseline data must include a “total” column for studies with more than 1 arm.

### Usage Requirements

#### Required Input

* Dataset-1: ADSL
* Variables: STUDYID, USUBJID, SAFFL, TRT01P, TRT01PN

##### Vital Signs Domain

* Dataset-2: 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-2: 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-2: ???
* Variables: ???
* Record selection: ???

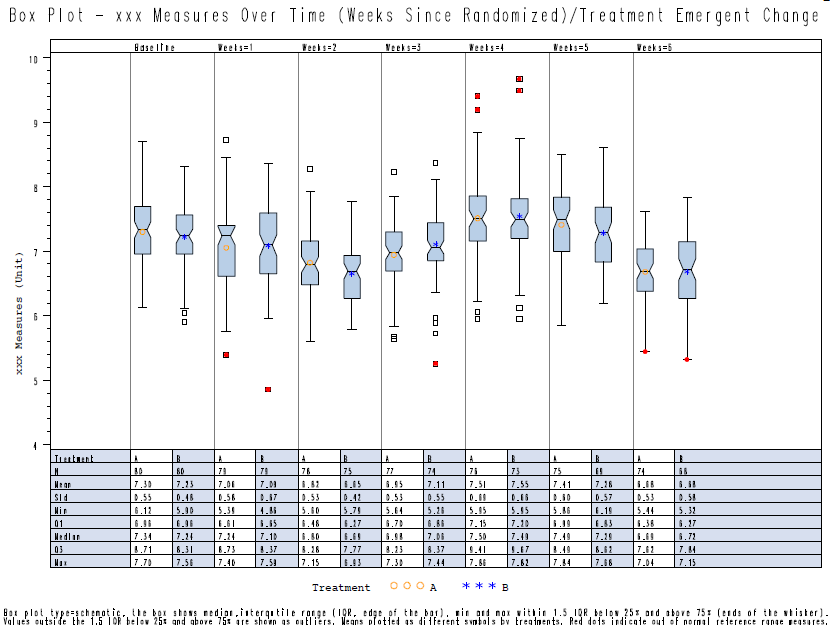
#### Macros

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

### Test Data

* Data sets from <https://github.com/phuse-org/phuse-scripts/tree/master/scriptathon2014/data>
* ADSL – Subject-level data
* 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>