**Steps for IDELA Data Cleaning and Basic Analysis**

1. **Clean child data**
	1. SAVE an original version of the data for your records.
	2. Check student codes
		1. Check child ID codes and names against the Team Leader records to see if student codes were entered correctly for each center/community
		2. Use =CONCATENATE formula to create unique IDs for all children
	3. Check for missing data
		1. Insert 1 row above data labels in Excel sheet and use =COUNTIF formula to count 999 values for each item. Make note of any variables that have more than 20% missing data, and include this information in your in report. This may indicate inadequate translation, adaptation or scoring of particular items.
		2. Replace to replace all 999 values in data sheet with 0.
	4. If data was collected using paper and pencil, check for outlying values (values above or below what is possible in IDELA scoring guidance).
2. **Create IDELA summary variables**
	1. Use =SUM command to create total scores for each subtest. Subtests are defined in the IDELA Tool documentation.
		1. Example: Add the scores from all items in the shape identification item together (shapeid1 – shapeid5).
	2. Create percentage correct score using the formula Total score achieved/Total possible score.
		1. Example: Total score of shape identification item/5.
	3. To create total domain scores, add percentage correct for all domain items and divide by total number of items.
		1. Example: Add percentage correct columns for all numeracy items and divide by 7 =(measurement+sorting+shape id+number id+one-to-one correspondence+simple operations+puzzle)/7.
		2. See IDELA Tool documentation or the appendix below for guidance on which items fall into which domain.
	4. To create total IDELA score, add percentage correct scores all core domain scores EXCEPT executive function and learning approaches and divide by the total number of domains.
		1. Example: (motor+literacy+numeracy+social-emotional)/4.
	5. SAVE a clean version of the data for your records.
3. **Test inter-rater reliability (Optional)**
	1. Sort data by IRR column (largest to smallest), then select ‘Add a level’ and add sorting by Assessor Status column
	2. Use =CORREL command to test correlation between assessors who spoke with children and those who observed.
	3. Record IRR correlation values for each item in a separate tab and then deleted rows with the ‘Assessor status’ = 2 (assessor was observing the assessment).
		1. We don’t want duplicate records in the final dataset and choose to use the responses from the assessor who was speaking to the child for the final analysis.
4. **Analyze data (Optional)**
	1. Use =AVERAGE and =AVERAGEIF command to generate averages for different groups of children
		1. Remember to lock (F4) column that references the groups you are interested in before dragging formula across columns. (Example: Lock the column with boy/girl data when looking at gender differences.)
	2. Sort data by variable of interest (Example: Boy/girl), then use =TTEST command to test for significant differences between scores
		1. Use conditional formatting to easily see differences that are significant (highlight if < .05)
5. **Generate summary variables for caregiver questionnaire (Optional)**
	1. If data was collected using paper and pencil, check for outlying values (values above or below what is possible in scoring guidance).
	2. If data was collected using paper and pencil OR tablets, review variable coding and ensure you aware of how response values have been coded. Recode variables to have more meaningful coding if needed.
		1. Example: If variables were coded 1=Yes and 2=No, recode variables so that 1=Yes and 0=No.
		2. Example: If caregiver attitude items are coded such that 1=Strongly agree and 5=Strongly disagree, recode responses so that higher values are positively associated with more positive responses.
	3. Use =SUM command to create summary variables for: Total types reading materials, total types of toys, total home learning activities, total home possessions, total caregiver attitudes, etc.
6. **Merge child and caregiver data (Optional)**
	1. Copy and paste unique ID and child names from IDELA sheet into a new sheet and sort by unique ID
	2. Copy and paste unique ID and child names from caregiver sheet into a new sheet and sort by unique ID
	3. Use =IF command to test whether child and caregiver codes match, and fix or remove any IDs that do not match. Make sure changes are recoded in full data sheet as well as on smaller matching sheet
	4. Copy child data into new sheet and sort by unique ID (if not already done). Then sort caregiver data by unique ID if not done already
	5. Copy caregiver data and add it to new child data sheet
	6. Check that child and caregiver records match correctly.
7. **Create Pivot Table (Optional)**
	1. Select all data in new merged child/caregiver sheet and choose the option ‘Insert Pivot Table’ with table created in new sheet.
	2. Choose variables of interest for analysis. Note that all associations will be correlational only.
8. **Share data (Optional)**
	1. Remember to delete all identifying information (names, specific locations) before sharing data with anyone not involved with your study.

**Appendix A. IDELA Domains**

|  |  |  |  |
| --- | --- | --- | --- |
| **Gross and Fine Motor Development**  | **Emergent Literacy and Language**  | **Emergent Numeracy** | **Social-emotional Development** |
| Hopping on one foot  | Print awareness | Measurement and comparison  | Peer relationships |
| Copying a shape | Expressive vocabulary | Classification/Sorting | Emotional awareness & regulation |
| Drawing a human figure | Letter identification | Number identification | Empathy  |
| Folding Paper  | Emergent writing | Shape identification | Self-awareness  |
|  | Phonemic awareness (First sound discrimination)  | One-to-one correspondence |  Conflict resolution |
|  | Listening comprehension | Simple operations (addition & subtraction) |  |
|  |   | Simple problem solving (puzzle) |   |
| **Executive Function (optional):** Short-term memory and inhibitory control  |
| **Approaches to Learning (optional)**: Persistence, motivation and engagement |

**Note**: Executive function and approaches to learning items are NOT included in the total IDELA score.