UNICEF Database Manager Assessment

Task 1: Maternal and Child Health Indicators Analysis

Your task is to calculate population-weighted coverage of health services (antenatal care and skilled birth attendance) for countries categorized as on-track and off-track in achieving under-5 mortality targets as of 2023.

The bar chart provides a snapshot of the relationship between maternal health indicators, specifically Skilled Birth Attendant and Antenatal Care (4+ visits) for women aged 15-49 years worldwide, based on household survey data from 2018-2022. It displays population-weighted averages for these indicators according to the SDG progress status of Under-5 Mortality Rate (U5MR), categorized as "Achieved," "On Track," and "Acceleration Needed." The findings suggest an association between maternal health indicators and U5MR status. However, these relationships are not statistically significant everywhere. Overall, the pattern indicates that countries with higher rates of skilled birth attendants and antenatal care visits tend to have lower U5MR rates. Specifically, countries that have achieved under-five mortality targets exhibit significantly higher rates of skilled birth attendance and antenatal care. While this analysis does not allow for causal claims, it does suggest that maternal health indicators play a crucial role in achieving under-five mortality targets.

The particular focus of this task is the distinction between the "Acceleration Needed" (off track) and "On Track" categories. For the "Antenatal Care 4+ visits" indicator, countries in the "On Track" category (56.5%) exhibit slightly higher coverage than those in the "Acceleration Needed" category (54.4%), but the difference is not significant, as the confidence intervals overlap.

In contrast, a different pattern emerges with the "Skilled Birth Attendant" indicator. Countries that are "On Track" have a mean coverage of 85.7%, while those requiring "Acceleration" lag behind at 63.8%. Since the confidence intervals do not overlap, these differences are statistically significant. This suggests that the "Skilled Birth Attendant" indicator is more closely related to achieving the U5MR target than the "Antenatal Care 4+ visits" indicator. However, as mentioned earlier, further research is needed to establish a causal relationship between these indicators and U5MR to make well-grounded policy recommendations.
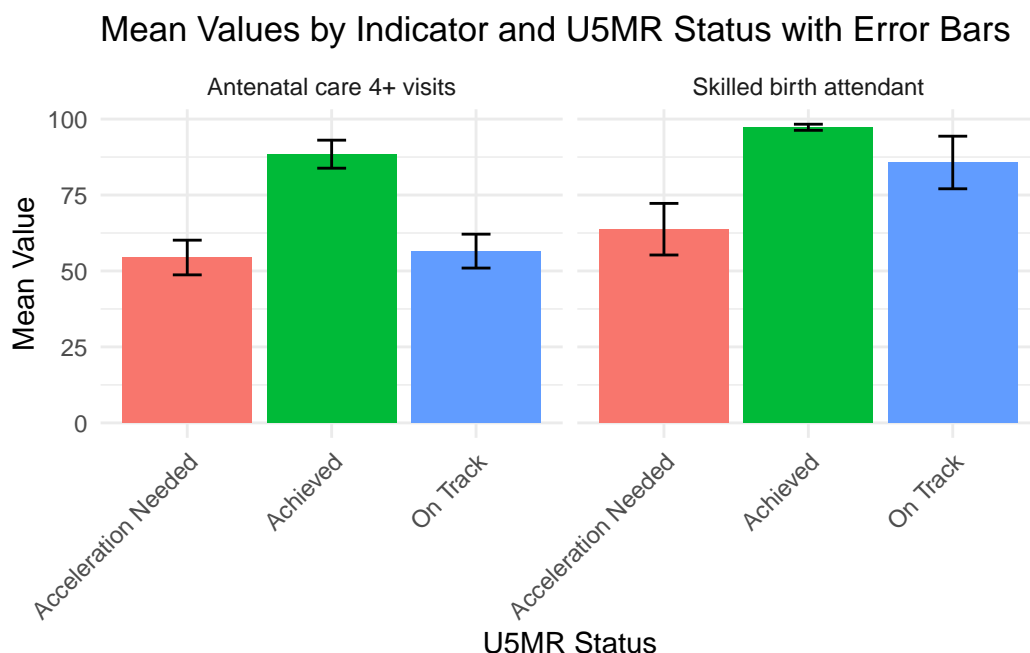
Mean Values by Indicator and U5MR Status with Error Bars

Figure 1: Mean Values by Indicator and U5MR Status with Error Bars

Task 2. Early Childhood Development Index (ECDI) Analysis

Your task is to produce a Data Perspective on the evolution of education for 4- to 5-year-old children. Particularly interesting for the Perspective is understanding how educational performance evolves month by month at these critical ages, considering both general education and specific subjects (e.g., literature and math, physical education). Methods should be sound and well-argued if necessary.

The variables outlined in the task for Zimbabwe MICS 2019 represent the old, not revised measure of Early Childhood Development Index (ECDI). The ECDI is a composite index that measures the development of children aged 36-59 months in four key domains: literacy-numeracy, physical, socio-emotional, and learning. The ECDI is calculated based on the presence or absence of specific skills in each domain. The ECDI is a binary variable, with a value of 1 indicating that a child has reached the expected level of development in all four domains, and 0 indicating that they have not. However, recently ECDI has been replaced by ECDI 2030, a new measure with quite a different methodology. While the features of the new measure are beyond the scope of this assessment, I will focus on producing the ECDI based on the old methodology and present it's main findings in accordance with the outlined task.

How Many Children Achieved the Expected Level of Development?

Below are the findings of the ECDI and its key components for Zimbabwe MICS 2019:

Descriptive Statistics
Early Childhood Development Index and Subcomponents

Variable	Number of Observations	Mean	Minimum	Maximum
ECDI	2491	0.71	0	1
Literacy and Numeracy	2491	0.10	0	1
Physical	2491	0.96	0	1
Socio-Emotional	2491	0.77	0	1
Learning	2491	0.91	0	1

The estimates are unweighted as the weights were not provided in the dataset. It can be seen that the overall percentage of children developmentally on track accounts for 71%. The value matches the MICS Survey Findings Report. While the highest share of population is on track in physical development (96%), the lowest share of those on track refers to literacy and numeracy and accounts for only 10%.

Correlation of index and its subcomponents

We wanna measure how the final score of ECDI correlates with its subcomponents. Since all the measures are binary, we used Phi coefficient to measure the correlations as this method suits best for dichotomous measures. The correlation matrix below shows weak correlations with literacy/numeracy domain (0.21) and physical domain (0.31), and moderate correlations with learning (0.48) domains. However, the most problematic domain is socio-emotional, which has a strong positive correlation with the ECDI (0.77). While detailed psychometric analysis is beyond the scope of this work, it suggests that socio-emotional domain is the major driver of the ECDI score in Zimbabwe. Ideally, no index should express the strong correlation between the domains and finakl score, as it 1) may indicate that the index is not a composite measure but rather a measure of a single domain 2) the domains are not fully independent from each other.
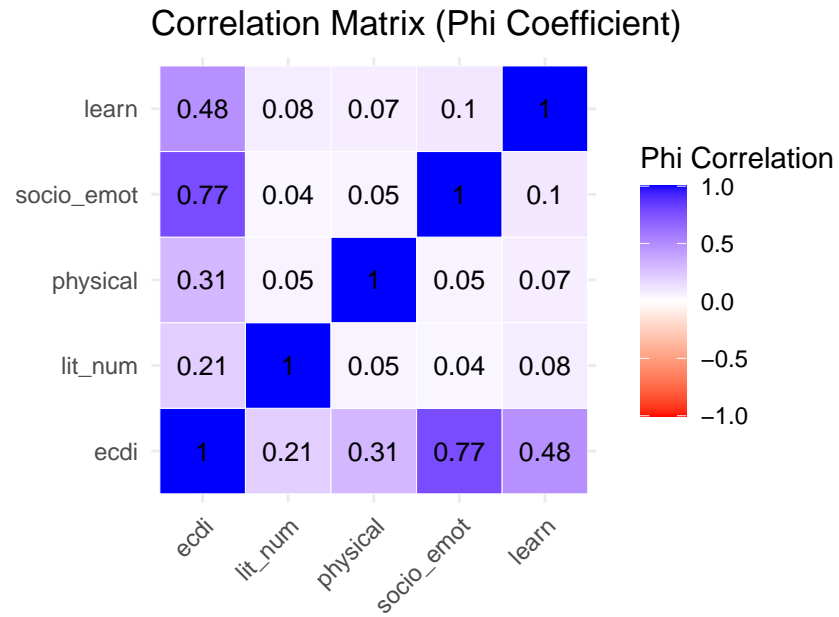
Correlation Matrix (Phi Coefficient)

Figure 2: Corelation Matrix of ECDI and subcomponents

Quality check: how many observations per month?

In order to make sure our further analysis on addressing the changes in ECDI overe time by month is valid, it makes sense to check the number of observations per month.
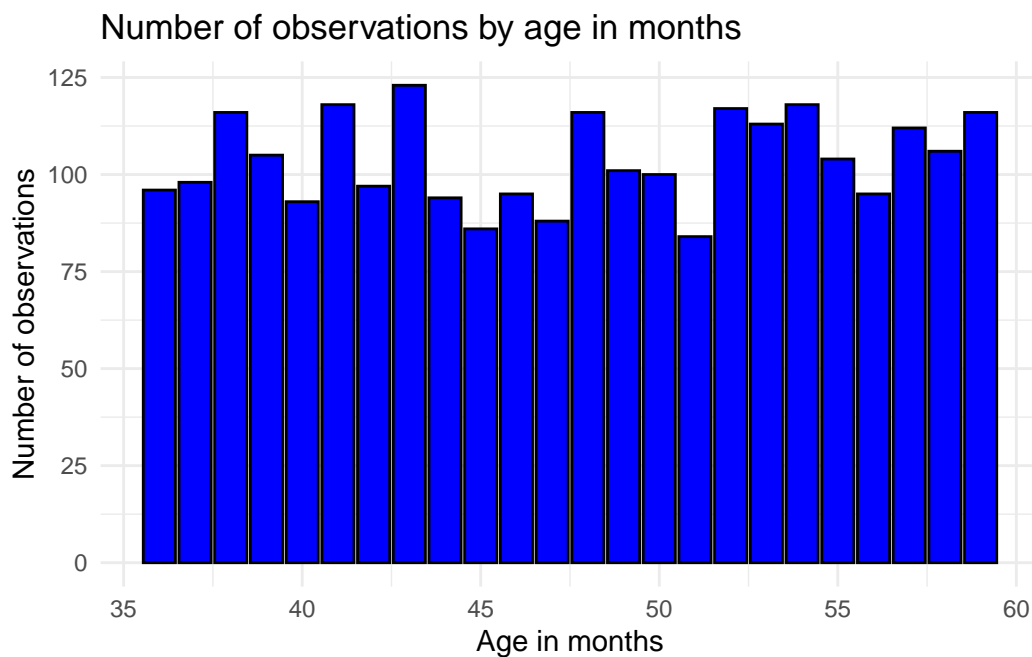
4

Figure 3: Number of ECDI Observations by Age in Months

The lowest number of obs is 86 and refers to 51st month, while the highest number of obs is 123 and refers to 43th month. This suggests that the data is sufficient to analyze the changes in ECDI over time by month.

Descriptive exploration: distribution of ECDI and its subcomponents by age in months

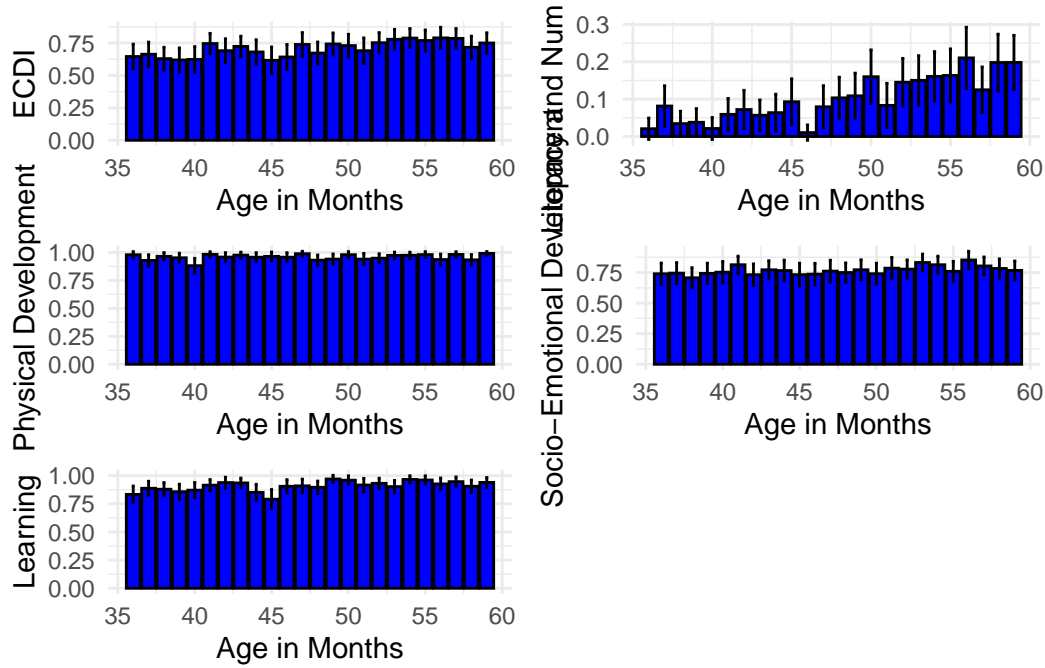The chart below checks the distribution of ECDI score and its main components by age in months.

Figure 4: Distribution of ECDI and its subcomponents by Age in Months

The figure suggests that the changes in ECDI by month, as well as in most of it components, are not very substantial. Moreover, the overlapping 95% confidence intervals suggest that in predominant cases they are not even statistically significant. This may indicate that the ECDI and its components are not sensitive enough to detect small changes in child development over time. However, further research is needed to confirm this hypothesis.

Regression analysis

We run 5 logit regressions, each predicts ECDI or its 4 subcomponents by age in mobths to see if the changes in time either in index or in one of it's main components are significant over time. The results of the GLM are presented in the table below.

GLM Model Summary

Effect of Age in Months on Various Outcomes

Outcome Variable	Term	Estimate (Logit)	Std. Error	z-value	p-value
ecdi	(Intercept)	−0.515	0.303	−1.697	0.090
ecdi	age_in_months	0.030	0.006	4.672	0.000
lit_num	(Intercept)	−6.391	0.532	−12.005	0.000
lit_num	age_in_months	0.086	0.010	8.254	0.000

physical	(Intercept)	2.343	0.691	3.390	0.001
physical	age_in_months	0.017	0.015	1.190	0.234
socio_emot	(Intercept)	0.449	0.326	1.378	0.168
socio_emot	age_in_months	0.016	0.007	2.332	0.020
learn	(Intercept)	0.272	0.476	0.571	0.568
learn	age_in_months	0.044	0.010	4.255	0.000

The table shows that the age has a significant effect on the ECDI and such subcomponents as literacy-numeracy, socio-emotional, and learning. The physical domain, however, does not show a significant relationship with age. However, to understand the magnitude of these changes, logit coefficients are not useful for the interpretation. Instead, based on the built models, predicted probabilities are calculated and visualized.
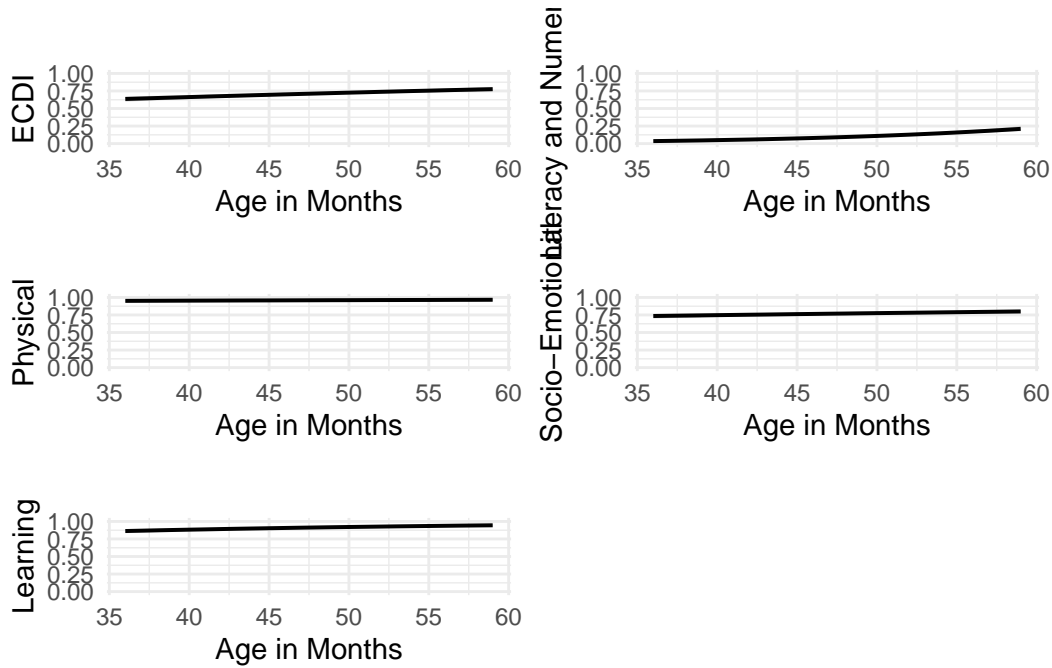

Figure 5: Predicted probabilities of age in months on ECDI and its subcomponents, logit model

The figure suggests that while some significant effects are present, their magnitude is very small, which may indicate that the ECDI and its components are not sensitive enough to detect small changes in child development over time. However, further research is needed to confirm this hypothesis. It needs to be indicated that that the current results are confounded by the lack of such variables as sex of a chilld, wealth quintile, parental education, area of residence, and others which are necessary to control for in the analysis.