

README FOR THE REPRODUCIBILITY PACKAGE for TURNING PROMISE INTO PRACTICE: BUILDING A COHERENT ECEC WORKFORCE SYSTEM IN VIET NAM

Table of Contents

Overview.....	1
Data Availability Statement.....	1
Statement about Rights.....	6
Instructions for Replicators	6
Analysis Teacher-Training Institutions' Course descriptions.....	32
Software Requirements.....	34
Folder Structure.....	34
Annex 1: Data received from the Ministry Of Education And Training.....	35
Annex 2: ECEC Data Stocktaking.....	37

Overview

This reproducibility package supports verification of the empirical findings presented in the World Bank working paper Turning Promise into Practice: Building a Coherent ECEC Workforce System in Viet Nam.

The package documents the data sources, analytical routines, and processing steps used to move from cleaned data inputs to the tables, figures, and diagnostic findings reported in the paper.

The analysis relies primarily on descriptive quantitative analysis and structured qualitative coding undertaken in Microsoft Excel. No statistical programming software was used by the authors. Where inferential statistics (including ANOVA results) are reported, these were provided as summary outputs by the research team from Vietnam Institute of Education Studies (VNIES) as part of the original study materials and were used for contextual interpretation only.

Materials included:

- Cleaned data from two surveys conducted by the World Bank team: Learning in Early Years (LEY) survey and Teaching and Learning in Early Childhood (TLEC) survey.
- An output Excel file include reproducible tables and figures

Data Availability Statement

Dataset Name	File Name(s)	Link/Access Instructions	Access Date	Data Access
Teaching and Learning in Early Childhood	TLEC_Cleaned_Response_Level_Data.xls	Access instruction: The TLEC survey was conducted by the World Bank team. The dataset was provided to the World Bank team in cleaned, structured format by a Vietnamese research group collaborating on the study (VNIES) The data will be available in the MicroData library in the future	2025	Accessible
Learning in Early Years (LEY)- Teachers	LEY_Teacher_responses.xls	Access instruction: The LEY survey was conducted by the World Bank team. The cleaned and structured LEY datasets were provided by the contractor–Faculty of Early Childhood Education, Hanoi National University of Education (HNUE). Disaggregation tables for the LEY survey were produced using Excel PivotTables; the authors used these outputs for integration and interpretation. The data will be available in the MicroData library in the future	2025	Accessible
Learning in Early Years	LEY_Manager	Access instruction: The LEY survey was conducted by the World Bank team. The cleaned and structured LEY datasets were provided by the	2025	Accessible

(LEY)-Managers	responses.xls	contractor—Faculty of Early Childhood Education, Hanoi National University of Education (HNUE). Disaggregation tables for the LEY survey were produced using Excel PivotTables; the authors used these outputs for integration and interpretation. The data will be available in the MicroData library in the future		
MOET Administrative Data – National Level	ECEC_National_level_Data.xls	Access instruction: Data was manually compiled from early childhood education statistics published on Ministry of Education and Training’s website. The data will be available in the MicroData library in the future Source: Viet Nam Ministry of Education and Training. 2024. Early Childhood Education Statistics. Available at: https://moet.gov.vn/thong-ke School Year (SY) 2014-2015: https://moet.gov.vn/thong-ke/Pages/thong-ke-giao-duc-tieu-hoc.aspx%3FItemID=4031 SY 2015-2016: https://moet.gov.vn/thong-ke/Pages/thong-ke-giao-duc-tieu-hoc.aspx%3FItemID=4042 SY 2016-2017: https://moet.gov.vn/thong-ke/Pages/thong-ke-giao-duc-tieu-hoc.aspx%3FItemID=5132 SY 2017-2018: https://moet.gov.vn/thong-ke/Pages/thong-ke-giao-duc-tieu-hoc.aspx%3FItemID=5391 SY 2018-2019: https://moet.gov.vn/thong-ke/Pages/thong-ke-giao-duc-tieu-hoc.aspx%3FItemID=6631 SY 2019-2020: https://moet.gov.vn/thong-ke/Pages/thong-ke-giao-duc-tieu-hoc.aspx%3FItemID=7385 SY 2020-2021: https://moet.gov.vn/thong-ke/Pages/thong-ke-giao-duc-tieu-hoc.aspx%3FItemID=8821 SY 2021-2022: https://moet.gov.vn/thong-ke/Pages/thong-ke-giao-duc-tieu-hoc.aspx%3FItemID=8822 SY 2022-2023: https://moet.gov.vn/thong-ke/Pages/thong-ke-giao-duc-tieu-hoc.aspx%3FItemID=10545 SY 2023-2024: https://moet.gov.vn/thong-ke/Pages/thong-ke-giao-duc-tieu-hoc.aspx%3FItemID=10548	2024	Accessible
MOET Administrative Data – Provincial Level	ECEC_Provincial_level_data.xls	Access instruction: The original data was shared with the World Bank Task Team by the Department of Early Childhood Education of Ministry of Education and Training (MOET) during the World Bank’s technical assistance to the Department in 2022-2023. The data is not publicly available. The original datasets were compiled by the Ministry of Education and Training (MOET) for the 2020–2021 school year to support the drafting of (i) a Government Decision on early childhood education (ECE) development in disadvantaged areas and (ii) a National Assembly resolution on the universalization of early childhood education and care (ECEC) for children aged 3–5. These original datasets were shared exclusively with the World Bank Task Team and, due to confidentiality restrictions, cannot be made publicly available. As a result, only an intermediate dataset constructed by the World Bank Task Team is included in this reproductivity package. Assembly Process: The Excel files received from MOET consist of multiple worksheets, with each worksheet corresponding to a specific indicator and containing several layers of disaggregation (for example, by province and by public versus non-public provision). To facilitate analysis, the World Bank Task Team manually consolidated these data into a new Excel worksheet. In this intermediate dataset, indicators are organized by rows and provinces by columns.	2024	Limited

		Only data for the 2020–2021 school year were included, and solely at the total level and the province-level disaggregation. A detailed description of the original MOET datasets is provided in Annex 1, while the intermediate dataset used for analysis is documented in Annex 2. Source: Viet Nam Ministry of Education and Training. 2023. Early Childhood Education Administrative Data.		
MOET Administrative Data – Regional Level	ECEC_Regional_level_data.xls	Access instruction: The data was compiled from the appendix of the Draft National Assembly Resolution of Early Childhood Education Program Renovation posted on chinhphu.vn website on January 10, 2025. The dataset compiled by the World Bank Task Team is available. Source: Government of Viet Nam. January 10, 2025. Appendix of Draft National Assembly Resolution of Early Childhood Education Program Renovation. https://chinhphu.vn/du-thao-vbqppl/du-thao-nghi-quyet-cua-quoc-hoi-ve-doi-moi-chuong-trinh-giao-duc-mam-non-7211	2025	Accessible
MOET Administrative Data – District Level	ECEC_District_level_data.xls	Access instruction: The original data was shared with the World Bank Task Team by the Department of Early Childhood Education of Ministry of Education and Training during the World Bank’s technical assistance to the Department in 2022-2023. The data is not publicly available. The original dataset was compiled by the Ministry of Education and Training (MOET) for the 2020–2021 school year to support the drafting of a National Assembly resolution on the universalization of early childhood education and care (ECEC) for children aged 3–5. The original dataset was shared exclusively with the World Bank Task Team and, due to confidentiality restrictions, cannot be made publicly available. As a result, only an intermediate dataset constructed by the World Bank Task Team is included in this reproductivity package. Assembly Process: Two excel files for 15 provinces received from MOET consist of multiple worksheets, with each worksheet corresponding to ECEC indicators of a province and its districts, as described in Annex 2. To facilitate analysis, the World Bank Task Team manually consolidated these data into a new Excel worksheet. In this intermediate dataset, indicators are organized by rows and provinces and districts by columns. Only data for the 2020–2021 school year were included and sole the district-level disaggregation. A detailed description of the original MOET datasets is provided in Annex 1, while the intermediate dataset used for analysis is documented in Annex 2. https://worldbankgroup-my.sharepoint.com/:x/r/personal/ntran6_worldbank_org/Documents/VIETNAM-PCACL-91Q9M53/EDU Projects/Vietnam FY24 ECEC - Ethnic/3. ECEC REPORT/Chapter 3 - WFD/Reproductivity package/WFD-Reproductivity/1. Data/2. ecec_province_level_data.xlsx?d=w82b47ea331c14d3682700c5db53ca4bf&csf=1&web=1&e=jMPRcm Source: Viet Nam Ministry of Education and Training. 2023. Early Childhood Education Administrative Data.	2024	Limited
Provincial GDP per capital	2022_GDP per capita.xls	Access instruction: The data was retrieved from the open database of Viet Nam General Statistics Office at National Accounts and Finance (URL: https://www.nso.gov.vn/en/px-web/?pxid=E0312&theme=National%20Accounts%20and%20State%20budget)	2022	Accessible

		A data table can be generated by choosing 69 provinces in the 'Cities, provinces' box and 2022 in the 'Year' box and click Continue. Source: Viet Nam General Statistics Office. National account. 2022		
Total and Urban population by provinces	Urbanization rate.xls	Access instruction: The data was retrieved from the open database of Viet Nam General Statistics Office at Population and Employment (URL: https://www.nso.gov.vn/en/px-web/?pxid=E0203-07&theme=Population%20and%20Employment) A data table can be generated by choosing 69 provinces and regions in the 'Cities, provinces' box, Total and Urban in 'Average Population' box, and 2022 in the 'Year' box and click Continue. Source: Viet Nam General Statistics Office. Population and Employment. 2022		
Training targets of selected teacher training colleges and universities 2023 (Figure 7)	2023 TTIs training target.xls	Access instruction: Training targets for Early childhood education major of teacher training universities and colleges in 2023 were provided to the WBTT by Department of Higher Education during a technical mission in November 2022. Since the training targets are published by universities and colleges on their websites, this dataset can be available in the MicroData library in the future. Source: Department of Higher Education. Ministry of Education and Training. 2023.	2023	Accessible
Training targets of selected teacher training colleges and universities 2018-2022 (Figure 7)	Figure 7.xls	Access instruction: Training targets for Early childhood education major of teacher training universities and colleges from 2018 to 2022 was manually collected by the WBTT from the institutions' Admission Plans posted on their official websites of the analysed universities and colleges. URL: ❖ Hanoi National University of Education (HNUE): https://tuyensinh.hnue.edu.vn/tin-tuc/185 ❖ Viet Nam National University, University of Education (VNU-UED): https://education.vnu.edu.vn/index.php/WebControl/viewnews/1557 ❖ National College of Education (NCE): https://cdsptw.edu.vn/document.aspx?sitepageid=665&id=1060 ❖ Vinh University, School of Education (VUED): https://old.vinhuni.edu.vn/DATA/0/upload/1/documents/2018/07/dean_dh_chinh_ngay_0472018.pdf ❖ Hue University, School of Education (HUED): https://tuyensinh.dhsphue.edu.vn/Modules/nganhhoc/front_list_nganh.aspx?idmenu=117 ; ❖ Thai Nguyen University, School of Education (TNUED): https://tuyensinh.tnue.edu.vn/Data/Sites/120/media/thangtm/files/De%20an%20tuyen%20sinh%202020_6.pdf ❖ Hanoi Metropolitan University (HMU): https://drive.google.com/file/d/1uyIR9YyzIV01V7YoYbPTiNamdDmMWv9v/view	2024	Accessible
Ethnic minority population (Figure 8)	Population-Ethnicity by province-2019.xlsx	Access instruction: The data was retrieved from the open database of Viet Nam General Statistics Office at Censuses/Population and Houses census (URL: https://gis.gso.gov.vn/printed_material/) Source: General Statistics Office. 2020. 2019 Viet Nam Population and Housing Census. Population by province is in Table 1, pages 9-43 and Ethnic minority by province is in Table 2, pages 44-210.	2025	Accessible
Ethnic minority	Ethnic ECE teachers.xlsx	Access instruction: The original data was shared with the World Bank Task Team by the Department of Early Childhood Education of Ministry	2025	Limited

teachers (Figure 8)		of Education and Training during the World Bank’s technical assistance to the Department in 2022-2023. The data is not publicly available. Only an intermediate dataset compiled by the World Bank Task Team is available. Assembly Process: The datatable was extracted from the Annex of the draft Scheme on ECEC development for disadvantaged areas. Source: Viet Nam Ministry of Education and Training. 2023. Early Childhood Education Administrative Data.		
Children: teacher ratio by countries (Appendix B5)	B5.xlsx	Access instruction: Statutory children:teacher ratio by countries were looked up using Google using key words including “Statutory children to early childhood teacher” AND “[country]”. Detailed sources are below. Source: - Vietnam: Ministry of Education and Training (Vietnam). <i>Circular No. 19/2023/TT-BGDĐT: Regulations on the organizational structure of public early childhood education professional titles</i>, Article 3. Hanoi, 2023. https://thuvienphapluat.vn/van-ban/Lao-dong-Tien-luong/Thong-tu-19-2023-TT-BGDDT-co-cau-vien-chuc-theo-chuc-danh-nghe-nghiep-giao-duc-mam-non-cong-lap-508175.aspx - Malaysia: SpringerLink. “Table 1,” in <i>Early Childhood Education in Malaysia</i>, by Kimberly Kong, in <i>International Handbook on Education in South East Asia</i>. Singapore: Springer, [year]. https://link.springer.com/chapter/10.1007/978-981-16-8136-3_13-1/tables/1#ref-CR19. - Singapore: Early Childhood Development Agency (ECDA). “Educator-to-Child Ratios.” Government of Singapore. https://ask.gov.sg/ecda/questions/cm4mgs6r9001cptnr5ez88gro - Finland: Eurydice. “Organisation of Centre-Based ECEC in Finland.” European Education and Culture Executive Agency. http://eurydice.eacea.ec.europa.eu/euryperia/finland/organisation-centre-based-ecec#:~:text=The%20maximum%20number%20of%20children,than%205%20hours%20a%20day. - Australia: Australian Children’s Education and Care Quality Authority (ACECQA). Educator to child ratios. URL: https://www.acecqa.gov.au/nqf/educator-to-child-ratios	2024	Accessible
ECE Primary and Secondary school Teachers by grade (Appendix B6)	B6.xlsx	Access instruction: The number of ECE, primary- and secondary-school teachers by grades are published by MOET on a Vietnamese online newspaper named VnExpress. Source: Ministry of Education. URL: https://vnexpress.net/giao-vien-mam-non-khap-khoi-cho-tien-luong-cao-nhat-4959000.html	2025	Accessible
Teacher Training Course Outlines	Folder: Course outlines	Access instruction: The course outlines were shared by selected teacher training institutions (TTIs) upon request from the World Bank Task Team. Since most TTIs do not publish their detailed course outlines on their websites. The data will not publicly available in the MicroData library in the future. Only the data tables compiled by the World Bank Task Team is available. Source:	2024	Limited

		1. Bachelor of Early Childhood Education Course Outline. Hoa Lu University. 2024. 2. Bachelor of Early Childhood Education Course Outline. Vietnam National University – University of Education. 2023. 3. Bachelor of Early Childhood Education Course Outline. Hanoi National University of Education. 2023. 4. Bachelor of Early Childhood Education Course Outline. Ho Chi Minh City University of Education. 2024. 5. Bachelor of Early Childhood Education Course Outline. Quang Nam University. 2022. 6. Bachelor of Early Childhood Education Course Outline. Quy Nhon University. 2020. 7. Bachelor of Early Childhood Education Course Outline. University of Danang-University of Education. 2021. 8. Diploma program of Early Childhood Education Course Outline. Tan Trao University. 2020. 9. Diploma program of Early Childhood Education Course Outline. Hanoi National College of Education. 2023. 10. Diploma program of Early Childhood Education Course Outline. Ho Chi Minh City College of Education. 2018 11. Diploma program of Early Childhood Education Course Outline. Kontum Community College. 2022.		
--	--	---	--	--

Statement about Rights

I certify that the author(s) of the manuscript have legitimate access to and permission to use the data used in this manuscript.

Instructions for Replicators

1. Instructions for creating figures and tables in Excel

Figure	Description	Instruction for replicator
Output file: MOET administrative data_tables and figures		
Figure 1	Regional teacher shortages (as a percentage of quotas) in public institutions	<i>Teacher shortage (Source data_Province/column AL) by region (Source data_Province/column E):</i> Sum total of teacher shortage by provinces (<i>Teacher quota by region:</i> Sum total of teacher quota by provinces (Source data_Province/ column AK) <i>Shortage as % of Quota</i> = Teacher shortage by region/Teacher quota by region Figure 1 was generated using data in sheet Figure 1/Cells C7:C12 (the column) and E7:E12 (the line)
Figure 2	Regional teacher–class ratios in public institutions	Teacher-class ratio = Number of teachers in public schools (Source data_Regional/column AC)/ Number of classes (Source data_Regional/column AC) Figure 2 was generated using data in sheet Figure 2/Cells E8:E13
Figure 3	Provincial teacher–class ratios in public institutions	Teacher-per-class = Number of teachers (Source data_Province/column AI)/Number of public 0-3 year old classes (Source data_Province/column R) + Number of public 3-6 year old classes (Source data_Province/column U) Steps to create Figure 3 in QGIS: please see section B below.
Figure 4	Provincial teacher shortages (as a percentage of	<i>Shortage as % of Quota by province</i> = Teacher shortage by province (Source data_Province AL column) / Teacher quota by province (Source data_Province AK column)

	quota) in public institutions	Steps to create Figure 3 in QGIS: please see section B below.
Figure 5	Percentage of public teachers who are nontenured or on contract	Number of untenured or contract teachers = Number of public teachers (<i>Source data_Regional/column AC</i>) – Tenured teachers <i>Source data_Regional/column AD</i> Figure 5 was generated using data in sheet Figure 5/Cells F10:F15
Figure 6	Percentage of tenured teachers in selected provinces	Total teachers by district is calculated by summing number of teachers of the district by age group (<i>Source data_District/Column AK</i>) Total number of tenured teachers by district is calculated by summing the number of tenured teachers of the district by age group (<i>Source data_District/Column AL</i>) Steps to create Figure 3 in QGIS: please see section B below.
Figure 9	Trends in teacher–class ratios in public institutions 2016–2024	Teacher-class ratios for 2014–2022: <i>Source data_National/cells AJ2:AJ9</i> Teacher-class ratios for 2023–2024: <i>Source data_National</i> Figure 9 was generated using data in sheet Figure 9/Cells C7:C15
Figure 10	Total preschool teacher workforce, 2018–2024	Total number of preschool teachers: <i>Source data_National/cells U26:U31</i> Figure 10 was generated using data in sheet Figure 10/Cells C7:C12
Table 1	Teacher shortages and ratios by geographical and economic status	Urbanization rate = Average Urban population/ Average population Teacher shortage (as proportion of teacher quota) of 10 least urbanized provinces = Total teacher shortages of 10 least urbanized provinces (<i>Source data_province/Column AL</i>) / Total teacher quota of 10 least urbanized provinces (<i>Source data_province/Column AK</i>) Average teacher per class of 10 least urbanized provinces = Total teachers of 10 least urbanized provinces (<i>Source data_province/Column AI</i>) / Total classes of 10 least urbanized provinces (<i>Source data_province/Column R plus U</i>) Average child-teacher ratio of 10 least urbanized provinces = [Total classes of 10 least urbanized provinces (<i>Source data_province/Column R plus U</i>) * Child per class (<i>Source data_province/Column AB</i>)]/ Total teachers of 10 least urbanized provinces (<i>Source data_province/Column AI</i>) Similar calculation are performed for: - Teacher shortage (as proportion of teacher quota) of 10 most urbanized provinces and provinces having GDP per capita of less than 50% of national average/ 51-75% of national average/75-100% of national average/over 100% of national average - Average teacher per class of 10 most urbanized provinces and provinces having GDP per capita of less than 50% of national average/ 51-75% of national average/75-100% of national average/over 100% of national average - Average child-teacher ratio of 10 most urbanized provinces and provinces having GDP per capita of less than 50% of national average/ 51-75% of national average/75-100% of national average/over 100% of national average.
Output file: Figure 7		
Figure 7	Training targets of selected pedagogical colleges and universities	Figure 7 was generated using data in cells B7:H18 in the Source Data worksheet.
Output file: Figure 8		
Figure 8	Ethnic minority representation in the preschool teacher workforce	Percentage of Ethnic minority (EM) population by province = Number of EM population (<i>Source data_pop/Column C</i>)/Total population (<i>Source data_pop/Column B</i>) Percentage of EM ECE teachers by province = Number of EM ECE teachers (<i>Source data_teachers/Column C</i>)/Total ECE teachers (<i>Source data_teachers/Column B</i>) Figure 8 was graphed by data in worksheet 'Figure 8' in cells D6 to E22.
Output file: LEY and TLEC figures		
Figure 11	Mean Frequency of Practicum Activities Reported by Student	The data is in columns AN to AU in the 'TLEC data' worksheet. It was extracted to 'Figure 11' worksheet.

	Teachers (TLEC Survey, n = 449)	The mean frequency of each practicum activity is calculated by averaging all responses given by 449 teachers and 150 managers for question number P2.4 in the TLEC survey.
Figure 12	Unmet training needs (LEY survey) and effectiveness by professional title grade (TLEC survey)	Effectiveness by professional title grade (TLEC survey) The ratings teachers in the TLEC survey gave to their training courses are recorded in columns EG to FB of the “<i>TLEC data</i>” worksheet. These data were extracted to the “<i>Figure 12</i>” worksheet and sorted by professional title grade (TLEC data, Column R). The average effectiveness is calculated by averaging all training ratings for teachers within each professional title grade (Grades 1, 2, 3, and 4). Unmet training needs (LEY survey) 1. Pivot tables summarizing responses to the 15 options in Question P3.8.2 of the LEY survey were created in the worksheet “<i>Pivot table Fig 12</i>”, with teachers’ professional title used as a filter 2. In the pivot tables, for each training topic (Literacy, Mathematics, Special Needs, and Classroom Management), filters were applied to determine how many teachers in each professional title group (Grades 1, 2, 3, and 4) reported both having participated in training on that topic and still needing further training. The number of teachers responding “Yes” to both questions is entered in cells P23 to W29. 3. In the “<i>Figure 12</i>” worksheet, the percentage of teachers who participated in training is calculated by dividing the number of participating teachers by the total number of teachers in each grade. The percentage of teachers with unmet training needs is calculated by dividing the number of teachers reporting low, medium, or high need by the total number of teachers in each grade. 4. The training gap was calculated as follows: The training gap = (Percentage of teachers having need for further training on one topic) – (percentage of teachers participating in the training topic). Figure 12 was graphed by data in worksheet ‘Figure 12’ in cells K10 to M 13 (columns) and cells K14 to M14 (line).
Figure 13	Access to induction and mentoring by sector (LEY and TLEC surveys)	Access to induction by sector (TLEC survey) First, answers of teachers to the question P3.1 “Did you receive induction or not?” are in column BJ in the ‘TLEC data’ worksheet. The data was extracted to ‘Figure 13’ worksheet and sorted by Type of school (TLEC data/column Q). Second, Count function was used to calculate the number of public- and private-school teachers who received induction or not. The results are in cells C12 to CE. Third, the percentage of teachers receiving induction by school type = (number of teachers in public or private schools receiving induction)/(total number of sampled teachers in public and private schools) Access to induction and mentoring by sector (LEY survey) Column S in ‘LEY data’ contains answers of teachers to the question “14. When you started your job as a teacher, did you participate in an induction and/or mentorship program?”. 1. All values in Column S in ‘LEY data’ were copied to Column M in ‘Figure 13’. Since some cells contain more than one value, function IF(ISNUMBER(SEARCH(findtext,within_text,1)),1,0) was used to find of if choice Q14-1 (Yes, an induction program), Q14-2 (Yes, a mentor) or Q14-3 (No) was chosen. Data on school where teachers work at in Column G of ‘LEY data’ was also copied to Column L in ‘Figure 13’. 2. A pivot table using source data in cells L2:P1082 was created.

		3. Sum values of Q14-1, Q14-2, Q14-3 were calculated and school setting was chosen as a filter. Choosing number 1 (public school) in school setting will show the number of public-school teachers choosing Q14-1, Q14-2, and Q14-3. The numbers were recorded manually to the output table in cells R6:T11. Repeat this process for number 4 (private school) in school setting. 4. The percentage of teachers receiving induction by school type = (number of teachers in public or private schools receiving induction)/(total number of sampled teachers in public and private schools) Figure 13 was generated using Excel's graph function using data in cells B7:E11.
Figure 14	Participation in collaborative activities by role	The level of participation in professional exchange and learning activities – in the form of a practice community or professional group or network as answered by teachers in TLEC survey are in columns from FD to FI in the 'TLEC data' worksheet. It was extracted to 'Figure 14' worksheet and sorted by professional title (Core teacher and teacher as in TLEC data/column J) and grade (TLEC data/column R). Teacher participation in collaborative activities was calculated by averaging all ratings of teachers in three categories (Core teachers, Grade 1 and 2, Grade 3 and 4). Figure 14 was generated by Excel's graph function using data in worksheet 'figure 14', cells B7:H9.
Output file: Appendix		
B3	Weighted average working hours per week (LEY survey)	The input data come from individual-level responses to the LEY survey, including: - Participant identification code - Demographic variables (e.g., gender, age) - Work context variables (province, location type, school type, contract type, rank) - Self-reported working hours: Q28: Total hours worked in the most recent complete calendar week Q29: Hours spent working directly with children in that same week Data Screening and Exclusion Criteria To ensure data quality and internal consistency, the following exclusion rules were applied to Q28 and Q29 responses: Logical consistency check - Responses were excluded if hours working with children (Q29) exceeded total working hours (Q28). Plausibility threshold - Responses were excluded if total working hours (Q28) exceeded 80 hours per week, as these were considered reporting errors. As a result of these rules: 144 entries were removed due to unrealistic values in Q28. 75 entries were removed due to unrealistic values in Q29. Only records passing all checks were retained for further analysis. After exclusions, two validated variables were created: Adjusted Q28: Cleaned weekly total working hours Adjusted Q29: Cleaned weekly hours working with children These adjusted variables form the analytical base for all subsequent calculations. Weighting and Aggregation The cleaned data were then used to compute weighted averages, reflecting the survey design and ensuring representativeness. Weighted average values were calculated for total weekly working hours (values between 1 and 80)

		Pivot Table Construction (Table B.3) A pivot table was generated to summarize results across key dimensions: • Overall sample • Province (e.g., Hai Phong, Quang Binh, Can Tho) • Location type (Urban, Rural) • School type (Public, Private) • Contract type ○ Indefinite term ○ Definite term (over 1 year) ○ Definite term (less than 1 year) • Rank (Rank 1 to Rank 4) For each category, the table reports: • Weighted average total weekly working hours • Weighted average weekly hours working with children • Number of valid observations contributing to the estimate
B4	Teacher-reported stressors (LEY survey)	Source Data The analysis is based on individual responses to Question 39 (Q39) of the LEY survey: <i>“Thinking about your job at this ECEC centre, to what extent are the following sources of stress in your work?”</i> Respondents rated each listed stressor using a 1–4 Likert scale where higher scores therefore indicate greater perceived stress. The pivot table in Worksheet B4 was constructed by summarising each stressor choice using its average score across respondents, following these steps: 1. Selection of values ○ For each Q39 sub-item (e.g. <i>“Managing classroom behaviour”</i>, <i>“Too much preparation work”</i>), the numeric response values (1–4) were selected as the pivot Values field. 2. Aggregation method ○ The mean (average) of the responses was calculated for each stressor. ○ This resulted in one average score per stressor, reflecting the typical intensity of stress reported by teachers for that item. 3. Independent summarization ○ Each stressor was summarized separately, not combined into a scale or composite index. ○ All valid respondents who answered a given item contributed to that item’s average. After the mean scores were calculated: • Stressors were ordered from the highest to the lowest mean value. • A rank was assigned to each stressor based on this ordering, where: ○ Rank 1 indicates the highest average stress level. ○ Lower-ranked items indicate relatively lower average stress.
B5	Teacher investment priorities: LEY versus TLEC	Source Data Two survey instruments are used: • LEY survey – Question 31 (Q31) Teachers were asked to select and rank their top three investment priorities from a predefined list of spending options. • TLEC survey – Question P4.2 Teachers were asked: <i>“If the ECEC budget increased by 5%, please select three spending</i>

		<i>priorities from the list below and rank them from highest to lowest priority.”</i> Data Preparation For each investment option: - Only rankings 1, 2, or 3 were retained for analysis. - Values indicating “not selected” (e.g. 0 or 4–8) were excluded from calculations. - Each investment option was treated as a separate variable, not combined into a composite scale. This ensures that averages reflect priority among those who selected the item, rather than diluting results with non-selections. Pivot Table Construction: Averaging Priority Rankings The pivot tables in Worksheet B5 were constructed by summarizing each investment option using its average ranking, following these steps: Selection of values - For each investment area (e.g. <i>Improving teacher salaries, Supporting children with special needs</i>), the numeric rank values (1–3) were assigned to the pivot Values field. Aggregation method - The mean (average) ranking was calculated for each investment option, using only valid ranks (1–3). - Each average reflects the typical priority level assigned to that option by respondents who chose it. Independent summarization - Each investment option was averaged independently. - A lower mean value indicates that an option was more frequently ranked as a higher priority. The same averaging approach was applied separately to: - LEY survey data, and - TLEC survey data. Ranking of Investment Priorities After average values were calculated: - Investment options were ordered from lowest to highest mean score. - Ranks were assigned accordingly: Rank 1 = highest priority (lowest mean score) - Higher numeric ranks indicate lower relative priority The ranking is therefore a direct consequence of the average ranking values, not a separate weighting or scoring procedure.												
B5a	Children per teacher and per group by setting (2022)	Source Data The following columns in Source data_National were used: <table border="1"> <thead> <tr> <th>Concept</th><th>Nursery</th><th>Kindergarten</th></tr> </thead> <tbody> <tr> <td>Number of classes/groups</td><td>class03</td><td>class36</td></tr> <tr> <td>Number of children</td><td>child03</td><td>child36</td></tr> <tr> <td>Number of teachers</td><td>teacher03</td><td>teacher36</td></tr> </tbody> </table> (Suffix 03 = Nursery, 36 = Kindergarten) Data were first filtered to school year 2022–2023 and grouped by sector (Public, Non-public, Total). Calculation of Indicators For each sector and stage, indicators were calculated using simple division of aggregated totals, not averages of pre-computed ratios.	Concept	Nursery	Kindergarten	Number of classes/groups	class03	class36	Number of children	child03	child36	Number of teachers	teacher03	teacher36
Concept	Nursery	Kindergarten												
Number of classes/groups	class03	class36												
Number of children	child03	child36												
Number of teachers	teacher03	teacher36												

		Children per group (Nursery) = Total number of nursery children / Total number of nursery classes = child03 / class03 Source columns: Numerator: child03; Denominator: class03 Children per teacher (Nursery) = Total number of nursery children / Total number of nursery teachers = child03 / teacher03 Teachers per group (Nursery) = Total number of nursery teachers / Total number of nursery classes = teacher03 / class03 Children per group (Kindergarten) = Total number of kindergarten children / Total number of kindergarten classes = child36 / class36 Children per teacher (Kindergarten) = Total number of kindergarten children / Total number of kindergarten teachers = child36 / teacher36 Teachers per group (Kindergarten) = Total number of kindergarten teachers / Total number of kindergarten classes = teacher36 / class36 Sector-Level Aggregation Logic For each indicator: Counts were first summed within: ○ Sector (Public / Non-public / Total) ○ Stage (Nursery / Kindergarten) Ratios were then calculated using aggregated totals, not by averaging school-level ratios.
B10	Teacher participation in centrally provided and other CPD modalities, by group and province (LEY survey) (%)	The calculations are based on LEY survey Question 16 (Q16): <i>“During the last 12 months, did you participate in any of the following professional development activities?”</i> Subgroup variables, including: • Q7 – Contract type (Permanent / Contract) • Q4 – School type (Public / Non-public) • Q5 – Location (Urban / Rural) • Province – Tỉnh/Thành Pivot Table Setup: Counting Participation For each CPD modality, a pivot table was built using the following structure: • Rows ○ CPD modality (each Q16 sub-item analysed separately) • Columns ○ One subgroup at a time (e.g. Contract type, Sector, Location, Province) • Values ○ Sum of the Q16 binary variable for that modality Because the variables are binary, the sum equals the number of teachers who participated in that CPD modality within each subgroup. Example: Sum of Q16_InPerson = number of teachers who joined in-person training Pivot Table Setup: Subgroup Totals (Denominators) A second pivot (or a subtotals row within the same pivot) was used to calculate subgroup totals, which serve as the denominator in percentage calculations: • Rows: Subgroup categories

		• Values: Count of respondents This produces the total number of teachers in each subgroup, regardless of CPD participation. Example: Total Permanent teachers Total Contract teachers Total teachers in Hai Phong Calculation Formula (General) For any CPD modality <i>M</i> and subgroup <i>S</i>: Participation rate for modality <i>M</i> in subgroup <i>S</i> = Number of teachers in subgroup <i>S</i> who selected <i>M</i> ÷ Total number of teachers in subgroup <i>S</i> Subgroup-Specific Calculations The same formula was applied separately for each breakdown shown in Table B.10: • Contract type ○ Denominator = total teachers with that contract type • Public / Non-public ○ Denominator = total teachers in that sector • Urban / Rural ○ Denominator = total teachers in that location • Province ○ Denominator = total teachers in that province • Total ○ Denominator = all responding teachers Link between Pivot Tables and Final Percentages The final percentages shown in Table B.10 were calculated by combining outputs from the pivot tables: Participation percentage = Pivot table count for CPD modality and subgroup ÷ Pivot table total respondents in that subgroup The pivot tables therefore provided: • Numerators → participation counts • Denominators → subgroup totals while the final table reports derived participation rates.
B12 B13	Barriers to CPD by province (LEY and TLEC surveys, % of respondents) Barriers to CPD by teacher subgroup (LEY and TLEC surveys)	Pivot Table Setup For each survey (LEY and TLEC), pivot tables were created from the survey microdata using the following structure: • Rows ○ Barrier category (each Q27 or P3.7 sub-item analysed separately) • Columns ○ Province (Table B.12), or ○ Teacher subgroup (Table B.13): ▪ Public / Non-public ▪ Permanent / Contract • Values ○ Sum of the binary barrier variable Because the barrier variables are binary, the sum equals the number of respondents who selected that barrier. Example (TLEC, Thai Nguyen):

		Sum(P3.7_PD too expensive) = number of teachers in Thai Nguyen who selected “PD too expensive” Percentage Calculation Logic For each barrier B and subgroup S, percentages were calculated as: Percentage reporting barrier B in subgroup S = Number of respondents in subgroup S who selected B ÷ Total number of respondents in subgroup S Because these are multiple-response questions: - Percentages do not sum to 100% - Each barrier is calculated independently Calculations for Table B.12 (By Province) General Formula (Province Level) Barrier percentage (Province) = Barrier count in province ÷ Total respondents in province Examples A. PD too expensive – Thai Nguyen (TLEC) $48 \div 150 = 0.32 \rightarrow 32.0\%$ - Numerator: Teachers in Thai Nguyen selecting “PD too expensive” - Denominator: Total teachers in Thai Nguyen (TLEC) Calculations for Table B.13 (By Teacher Subgroup) General Formula (Subgroup Level) Barrier percentage (Subgroup) = Barrier count in subgroup ÷ Total respondents in subgroup Examples A. PD too expensive – Public teachers (TLEC) $106 \div 361 \approx 0.294 \rightarrow 29.36\%$
B14 B15	Supports for CPD access by province (LEY and TLEC surveys, % of respondents) Supports for CPD access by teacher subgroup (LEY and TLEC surveys, % of respondents)	Pivot Table Setup For each survey (LEY and TLEC), pivot tables were created from the survey microdata using the following configuration: Rows - Support type (each Q26 or P3.6 sub-item analysed separately) Columns - Province (Table B.14), or - Teacher subgroup (Table B.15), including: - Public / Non-public - Permanent / Contract - Core teachers / Regular teachers (TLEC) Values Sum of the binary support variable Because each support item is coded as 0/1, the sum equals the number of respondents who reported receiving that support. Example (LEY, Hai Phong): Sum(Q26_Released during working hours) = number of teachers in Hai Phong who received release time Pivot Table Construction – Subgroup Totals (Denominators) For each table, a second pivot (or subtotal row) was used to calculate subgroup totals, representing: - Total teachers in each province - Total teachers in each subgroup (e.g. Public, Contract, Core teachers) These totals serve as the denominators in percentage calculations. Examples:

		• Total teachers in Hai Phong (LEY) = 630 • Total teachers in Thai Nguyen (TLEC) = 150 • Total Public teachers (LEY) = 794 Percentage Calculation Logic (General Rule) For each support type S and subgroup G, percentages were calculated as: Percentage receiving support S in subgroup G = Number of respondents in subgroup G who received support S ÷ Total number of respondents in subgroup G Because the questions are multiple-response: • Percentages do not sum to 100% • Each support type is calculated independently Calculations for Table B.14 (By Province) ❖ General Formula Support percentage (Province) = Support count in province ÷ Total respondents in province Examples A. Released during working hours – Hai Phong (LEY) $98 \div 630 \approx 0.156 \rightarrow 15.56\%$ Calculations for Table B.15 (By Teacher Subgroup) General Formula Support percentage (Subgroup) = Support count in subgroup ÷ Total respondents in subgroup Examples A. Released during working hours – Public teachers (LEY) $118 \div 794 \approx 0.1486 \rightarrow 14.86\%$
B16	Frequency analysis of teacher-reported CPD topics provided by MOET/DOET/BOET (LEY survey; total mentions = 574)	Co-pilot was used to count mentions by most-mentioned topics The analysis of professional development topics was conducted using <i>mentions</i> as the unit of analysis rather than individual teachers. Open-ended responses were standardized and coded into thematic groups, with technology and digital transformation–related topics prioritized over generic descriptions. A total of 574 valid mentions were identified. More than half concerned technology and digital innovation, while child-centered pedagogy and social topics accounted for comparatively smaller shares. PROMPT 1. Core Unit of Analysis: <i>Mentions</i> (NOT Individuals) Key principle:  The unit of analysis is “mentions”, not teachers. • Each teacher was allowed to list up to two CPD topics (Q17-1 and Q17-2). • If a teacher wrote: “STEAM, AI” → this was counted as 2 mentions, not 1. Therefore: • Number of teachers ≠ number of mentions • Total mentions = total number of topics listed, after data cleaning For Table B16: •  Total valid mentions after cleaning = 574
		2. Text Standardization Prior to Counting (Crucial Step) Responses were open-ended text, so all answers were standardized before counting.

	2.1. Mapping similar expressions to a standard keyword Each mention was mapped to a standard keyword, as illustrated below: <table><tr><th>Original response text</th><th>Standardized as</th></tr><tr><td>Steam, STEAM, stem, GD Steam, pp STEAM, steam training</td><td>STEAM/STEM</td></tr><tr><td>CNTT, information technology, UDCNTT</td><td>IT Application</td></tr><tr><td>digital transformation, CDS, digital technology</td><td>Digital Transformation</td></tr><tr><td>AI, Al, artificial intelligence, AI technology</td><td>AI</td></tr><tr><td>Canva, Can va, Canva software</td><td>Canva Software</td></tr><tr><td>Child-centered approach</td><td>Child-centered education</td></tr><tr><td>Educational environment</td><td>Building educational environments</td></tr><tr><td>Child rights, human rights</td><td>Child rights</td></tr><tr><td>Multicultural music</td><td>Multicultural music</td></tr><tr><td>Traffic safety, ATGT</td><td>Traffic safety</td></tr></table>	Original response text	Standardized as	Steam, STEAM, stem, GD Steam, pp STEAM, steam training	STEAM/STEM	CNTT, information technology, UDCNTT	IT Application	digital transformation, CDS, digital technology	Digital Transformation	AI, Al, artificial intelligence, AI technology	AI	Canva, Can va, Canva software	Canva Software	Child-centered approach	Child-centered education	Educational environment	Building educational environments	Child rights, human rights	Child rights	Multicultural music	Multicultural music	Traffic safety, ATGT	Traffic safety
Original response text	Standardized as																						
Steam, STEAM, stem, GD Steam, pp STEAM, steam training	STEAM/STEM																						
CNTT, information technology, UDCNTT	IT Application																						
digital transformation, CDS, digital technology	Digital Transformation																						
AI, Al, artificial intelligence, AI technology	AI																						
Canva, Can va, Canva software	Canva Software																						
Child-centered approach	Child-centered education																						
Educational environment	Building educational environments																						
Child rights, human rights	Child rights																						
Multicultural music	Multicultural music																						
Traffic safety, ATGT	Traffic safety																						
	3. Topic Group Assignment Rules (Priority-based) After standardization, each mention was assigned to one and only one topic group, following a fixed order of priority. Group 1. Generic / Non-Specific Training Assigned if the response was general and non-specific, with no clear content area. Keyword rule: • “professional training” • “regular training” • “specialized training” • “in-service training” • “professional development” ➡ Even if the wording sounds professional, it was classified as Generic if no specific theme was mentioned.																						
	Group 2. Technology and Digital Innovation Assigned if any technology-related keyword appeared, even when combined with generic wording. Keyword rule (OR): • STEAM / STEM • IT / ICT • Digital transformation • AI • Canva • Software / digital signatures / electronic records ➡ Higher priority than Generic Example: “Training on applying AI” → Technology and Digital Innovation, not Generic.																						
	Group 3. Child-Centered Pedagogy Assigned when the focus was on pedagogical approaches for children. Keyword rule: • Child-centered education • Building educational environments																						

		<table><tr><td>Technology and Digital Innovation</td><td>STEAM/STEM, IT Application, Digital Transformation, AI, Canva Software</td><td>301</td><td>52.4%</td></tr><tr><td>Generic / Non-Specific Training</td><td>Professional training, Specialized training, Regular training</td><td>190</td><td>33.1%</td></tr><tr><td>Child-Centered Pedagogy</td><td>Child-centered education, Building educational environments</td><td>48</td><td>8.4%</td></tr><tr><td>Other Specific Topics</td><td>Child rights, Multicultural music, Traffic safety</td><td>35</td><td>6.1%</td></tr><tr><td>Total</td><td></td><td>574</td><td>100%</td></tr></table>	Technology and Digital Innovation	STEAM/STEM, IT Application, Digital Transformation, AI, Canva Software	301	52.4%	Generic / Non-Specific Training	Professional training, Specialized training, Regular training	190	33.1%	Child-Centered Pedagogy	Child-centered education, Building educational environments	48	8.4%	Other Specific Topics	Child rights, Multicultural music, Traffic safety	35	6.1%	Total		574	100%
Technology and Digital Innovation	STEAM/STEM, IT Application, Digital Transformation, AI, Canva Software	301	52.4%																			
Generic / Non-Specific Training	Professional training, Specialized training, Regular training	190	33.1%																			
Child-Centered Pedagogy	Child-centered education, Building educational environments	48	8.4%																			
Other Specific Topics	Child rights, Multicultural music, Traffic safety	35	6.1%																			
Total		574	100%																			
B17	Percentage of teachers reporting each feature of centrally provided CPD (LEY survey)	Pivot Table Setup A pivot table was created from the LEY survey microdata using the following structure: • Rows ◦ CPD feature (each Q20 sub-item analysed separately) • Columns ◦ Province ◦ Location ◦ School type • Values ◦ Sum of the binary Q20 variable Because each Q20 feature is coded as 0/1, the sum equals the number of teachers who selected that feature within each subgroup. Example: Sum(Q20_1 = Relevant and useful) = number of teachers who reported the CPD was relevant and useful A second pivot (or subtotal row within the same pivot table) was used to calculate subgroup totals, based on: • Count of respondents, grouped by: ◦ Province ◦ Location ◦ School type These counts provide the denominators used to calculate percentages. Example subgroup totals shown in the worksheet: • Total respondents = 1080 • Hai Phong = 630 • Quang Binh = 252 • Can Tho = 198 • Urban = 696 • Rural = 327 • Public = 794 • Private = 274 Percentage Calculation Logic (General Rule) For each feature F and subgroup G, the percentage was calculated as: Percentage reporting feature F in subgroup G = Number of respondents in subgroup G who selected F ÷ Total number of respondents in subgroup G Example - Relevant and useful – Total 862 ÷ 1080 ≈ 0.798 → 79.8% • Numerator: teachers who selected “Relevant and useful” • Denominator: all LEY respondents																				

B18	CPD participation and training needs by professional title grade (reporting participation in past 12 months, LEY survey) (%)	Source Question and Variables 2.1 LEY Survey Question Question Q21: <i>“In the past 12 months, were the following topics addressed in your professional development?”</i> (with an additional question on need for additional PD) 2.2 Response Structure For each topic listed in Q21, respondents reported: Whether they had been trained, and Their level of need for additional training. Each topic is represented by binary indicator variables: 1 = yes / selected 0 = no / not selected Training-need intensity was recorded in four ordered categories: - No need - Low need - Medium need - High need ❖ CPD Participation Participation refers to the percentage of teachers who reported having received CPD on the topic in the past 12 months. ❖ Training Need Need refers to the percentage of teachers who reported a <i>medium or high</i> need for further training on the topic. In the table, Total need = Medium need (3) + High need (4) Pivot Table Construction – Counts (Numerators) ❖ CPD Participation Pivot tables were created from the LEY survey microdata using the following structure: Rows - CPD topic (each Q21 item analysed separately) Columns - Professional title grade (Grade 1, Grade 2, Grade 3&4, Total) Values Sum of the binary “Have been trained” variable for the topic Because the variable is binary, the sum equals the number of teachers who reported CPD participation for that topic within each grade. Example: Sum(Q21_1 = Have been trained on child development) = number of teachers trained on child development ❖ Training Need A second pivot table was constructed for training need, using: Rows - CPD topic Columns - Professional title grade Values Sum of responses classified as Medium need (3) or High need (4) This yields the count of teachers reporting a medium or high need for each topic and grade. Separate tables (or subtotal rows) were used to calculate:
-----	--	---

		- Total number of respondents in each professional grade These totals serve as the denominators for both participation and need indicators. Example subgroup totals: - Total respondents = 1080 - Grade 1 = 733 - Grade 2 = 49 - Grade 3 & 4 = 161 Percentage Calculation Logic ❖ CPD Participation For each topic T and grade G: Participation rate (T, G) $= \frac{\text{Number of teachers in grade G trained on topic T}}{\text{Total number of teachers in grade G}}$ ❖ Training Need (Medium + High) For each topic T and grade G: Training need rate (T, G) $= \frac{\text{Number of teachers in grade G reporting (medium need OR high need) for topic T}}{\text{Total number of teachers in grade G}}$
B19	CPD effectiveness ratings by theme, topic, and teacher group (TLEC survey; mean score, 1–4 scale)	Source Question and Variables - TLEC Survey Question P3.8.2.* (multiple sub-items): <i>"Please rate the quality of the CPD courses you participated in."</i> Each CPD topic under P3.8.2 is rated on a four-point Likert scale (1–4). Only teachers who reported participating in CPD on the given topic provided a rating. Pivot Table Setup in Excel (Topic Level) For each CPD topic (P3.8.2.x), a PivotTable was created using the following configuration: 1. Rows CPD Topic (One topic per pivot row, or separate pivots for each topic) 2. Columns Teacher group, defined separately as: - ALL teachers - Grade I - Grade II - Grade III & IV - Regular teachers - Core teachers These groupings were created using the corresponding classification variables placed in the Columns area of the PivotTable. 3. Values Rating variable (P3.8.2.x) placed in the Values area - Value Field Settings set to: Summarise Values By → Average This instructs Excel to compute the arithmetic mean of all valid ratings in each group. Resulting Pivot Table Output (Topic Level)

		For each CPD topic and teacher group, the pivot table produces: Average rating == (Sum of individual ratings for the topic) ÷ (Number of teachers who provided a rating) ❖ Consolidation of Pivot Outputs Results from two above pivot tables were manually paste to the crosstab table on the right. Theme-Level Pivot Calculation For each theme, a separate pivot (or calculated field) was used to compute the theme average, defined as: Theme average = Arithmetic mean of topic-level mean scores within the same theme This was done separately for each teacher group. Example Assessment and Inclusive Practice – ALL teachers This theme includes multiple topics, such as: • Tracking child development • Adjusting learning goals • Needs of children from diverse backgrounds The theme average is: Theme average = (Mean Topic 1 + Mean Topic 2 + Mean Topic 3) ÷ Number of topics = 3.0172 Calculation of the Overall Average An overall CPD effectiveness score was calculated as: Total Average = Arithmetic mean of all topic-level mean scores This represents the average perceived effectiveness across all CPD topics, by teacher group.
B20	Participation in induction and mentoring (percentage of teachers reporting), LEY versus TLEC surveys	Pivot Table Construction – Counts (Numerators) ❖ LEY Survey Pivot Tables Pivot tables were created from the LEY survey microdata using the following structure: • Rows ○ Induction/mentoring response category: ▪ Yes, an induction program ▪ Yes, a mentor ▪ No • Columns ○ One disaggregation at a time: ▪ School type (Public / Non-public) ▪ Location (Urban / Rural) ▪ Total • Values ○ Sum of the corresponding binary variable (Q14-1, Q14-2, Q14-3) Because the variables are binary, the sum equals the number of teachers reporting that response in each subgroup. Example: Sum(Q14-1 = Yes, an induction program) = number of teachers who received an induction program

		❖ TLEC Survey Pivot Tables For the TLEC survey, pivot tables were created with: • Rows ○ Induction status: ▪ Received induction ▪ Did not receive induction • Columns ○ School type (Public / Non-public) ○ Location (Urban / Rural) ○ Total • Values ○ Count of respondents (or sum of the binary indicator) ❖ Consolidation of Pivot Outputs Results from two above pivot tables were manually paste to the crosstab table on the right. Percentage Calculation Logic (General Rule) For each category C and subgroup G, percentages were calculated as: Percentage (C, G) $= \frac{\text{Number of respondents in subgroup G reporting category C}}{\text{Total population}}$ The same calculation logic was applied consistently across: • Surveys (LEY and TLEC) • School type • Location • Total population
B21		Pivot Table Construction – Counts and Mean Values (Numerators) ❖ LEY Survey Pivot Tables Pivot tables were created from the LEY survey microdata using the following structure: • Rows ○ Training content reported as learned during induction: ▪ Educational planning ▪ Pedagogical skills ▪ Creating and maintaining a safe learning environment ▪ Communication and collaboration with school, families, and community ▪ Inclusive education methods ▪ Other • Columns ○ Total (no further disaggregation applied in Table B.21) • Values ○ Sum of the corresponding binary variable (Q15-1 to Q15-6) Because the variables are binary, the sum equals the number of teachers selecting that training content. Example: Sum(Q15-1 = Educational planning) = Number of teachers reporting knowledge gains in educational planning A separate pivot (or subtotal) was used to calculate the total number of LEY respondents, which serves as the denominator. ❖ TLEC Survey Pivot Tables

		For the TLEC survey, pivot tables were created using the following structure: • Rows ○ Training content gained during induction: ▪ Educational planning ▪ Pedagogical skills ▪ Safe and supportive learning environments ▪ Communication with families/community ▪ Inclusive education methods ▪ Other • Columns ○ Teacher group: ▪ Total ▪ Public ▪ Non-public • Values ○ Average of the corresponding P3.2 rating variable The Value Field Settings in Excel were set to: Summarise Values By → Average This produces the mean knowledge-gain rating on a 1–4 scale, calculated among teachers who participated in induction training and provided a valid rating. ❖ Consolidation of Pivot Outputs • Results from the LEY pivot table (percent selecting each option) and the TLEC pivot table (mean ratings) were manually copied and aligned into the final crosstab used to produce Table B.21. • LEY and TLEC results are presented side-by-side to reflect differences in survey design. Percentage and Mean Calculation Logic (General Rules) LEY Survey (Percent Selecting) For each training content C, percentages were calculated as: Percentage (C) = Number of respondents selecting content C ÷ Total number of LEY respondents TLEC Survey (Mean Rating) For each training content C and teacher group G, mean scores were calculated as: Mean rating (C, G) = Average of P3.2 ratings for content C among teachers in group G • Ratings use a 1–4 scale: ○ 1 = No knowledge gained ○ 4 = Substantial knowledge gained • Blank or non-applicable responses were automatically excluded by Excel PivotTables. Consistency of Application The same calculation logic was applied consistently across: • Surveys (LEY and TLEC) • Training content areas • School type (Public / Non-public) in the TLEC survey • Total population
B22	Participation in and perceived value of school-based professional support, by sector,	Pivot Table Construction – Counts and Mean Values (Numerators) ❖ TLEC Survey Pivot Tables Pivot tables were created from the TLEC survey microdata using the following structure.

	role, and title (TLEC survey)	A. Participation in School-Based Professional Support Source question: • P3.10 – “<i>Mức độ tham gia của bản thân vào bồi dưỡng chuyên môn tại trường theo hình thức cố vấn, tư vấn chuyên môn?</i>” (Scale: 1 = Ít / very low to 4 = Nhiều / very high) Pivot table structure • Rows ○ Teacher groups: ▪ Public schools ▪ Non-public schools ▪ Grade I teachers ▪ Grade II teachers ▪ Grade III/IV teachers ▪ Core teachers ▪ Regular teachers ▪ All teachers • Columns ○ None (each pivot produced one column per indicator; disaggregation handled by row grouping) • Values ○ Average of P3.10 → mean participation score ○ Mode of P3.10 → most common response (using PivotTable “Mode” function) ○ Count of “Yes” responses to P3.11 (see below, used for participation rate) Because P3.10 is a numeric 1–4 scale, the <i>Average</i> value equals the mean participation rating for each teacher group. B. Receipt of Professional Support (Yes/No) Source question: • P3.11 – “<i>Thầy/cô có nhận được sự hỗ trợ phát triển chuyên môn thường xuyên từ giáo viên cốt cán hoặc cố vấn chuyên môn có trình độ không?</i>” Pivot table structure • Rows ○ Same teacher groups as above • Columns ○ None • Values ○ Sum of the binary variable P3.11 (Yes = 1, No = 0) ○ Sub-group total count Because P3.11 is binary, the sum equals the number of teachers who reported receiving support in each subgroup. Example: Sum(P3.11 = Yes) = Number of teachers who reported receiving school-based professional support C. Value of School-Based Professional Support Source question: • P3.12 – “<i>Mức độ giá trị của những hỗ trợ phát triển chuyên môn thường xuyên từ giáo viên cốt cán hoặc cố vấn chuyên môn có trình độ đối với Thầy/Cô?</i>” (Scale: 1 = Ít / very low to 4 = Nhiều / very high) Pivot table structure • Rows ○ Teacher groups (same as above)
--	--------------------------------------	--

		• Columns ○ None • Values ○ Average of P3.12 → mean value rating ○ Mode of P3.12 → most common response ❖ Consolidation of Pivot Outputs • Results from the P3.10 (participation level), P3.11 (receipt of support), and P3.12 (value of support) pivot tables were manually copied and combined into the crosstab used to produce Table B.22. • Participation level and value are presented as mean scores (with mode shown for reference). • Receipt of support is presented as counts and percentages. Percentage and Mean Calculation Logic (General Rules) Participation Rate (Receipt of Support) For each teacher group G, the participation rate was calculated as: % receiving support (G) = Number of respondents in group G reporting “Yes” (P3.11) ÷ Total number of respondents in group G Mean Participation Level For each teacher group G: Mean participation score (G) = Average of P3.10 ratings for group G Mean Value of Support For each teacher group G: Mean value score (G) = Average of P3.12 ratings for group G • All averages are based on a 1–4 scale • Blank or non-applicable responses were automatically excluded by Excel PivotTables Consistency of Application The same calculation logic was applied consistently across: • Teacher role (Core / Regular) • Professional title grade (Grade I, II, III/IV) • School sector (Public / Non-public) • Total TLEC respondent population
B23	Mean ratings of teacher engagement in collaborative professional learning activities, by role and school type (TLEC survey, 1–4 scale)	Pivot Table Construction – Mean Values (Numerators) ❖ TLEC Survey Pivot Tables Pivot tables were created from the TLEC survey microdata using the following structure. A. Participation in Collaborative Professional Learning Activities Source question: • P3.9 – “<i>Mức độ tham gia của bản thân vào các hoạt động trao đổi, học hỏi chuyên môn – hình thức cộng đồng thực hành hoặc nhóm, mạng lưới chuyên môn?</i>” (Scale: 1 = Very low; 4 = Very high) Each sub-item under P3.9 corresponds to one collaborative learning activity, including: • Peer observation and feedback

		• Discussing holistic child development • Collaboration across age groups • Exchanging teaching materials • Joint planning of activities • Discussing individual children's needs B. Pivot Table Structure • Rows • Collaborative professional learning activities: ○ Peer observation / feedback ○ Discussing holistic development ○ Age-group collaboration ○ Exchanging materials ○ Planning activities together ○ Discussing child needs ○ Average (across all listed activities) • Columns • Teacher grouping variables: ○ Grades I and II ○ Grades III and IV ○ Core teachers ○ Regular teachers ○ Public schools ○ Non-public schools ○ Overall average Each grouping was applied one dimension at a time within the PivotTable columns. • Values • Average of the corresponding P3.9 sub-item The Value Field Settings in Excel were set to: Summarise Values By → Average Because P3.9 uses a numeric 1–4 Likert scale, the resulting value represents the mean level of engagement in each activity for each subgroup. ❖ Consolidation of Pivot Outputs • Pivot tables were created separately for each P3.9 sub-item. • The resulting mean values were manually copied and aligned into the final crosstab layout used to produce Table B.23. • An additional average row ("Average") was calculated as the arithmetic mean of the activity-specific mean scores. Mean Calculation Logic (General Rule) For each activity A and subgroup G, mean participation scores were calculated as: Mean participation score (A, G) = Average of P3.9 ratings for activity A among respondents in subgroup G • Ratings are on a 1–4 scale • Only valid responses were included • Blank or non-applicable responses were automatically excluded by Excel PivotTables Consistency of Application The same calculation logic was applied consistently across:
--	--	---

		• Teacher role (Core / Regular) • Professional title grade (Grades I–II vs Grades III–IV) • School sector (Public / Non-public) • Overall TLEC respondent population
B24	Teacher self-ratings on professional practice domain (LEY and TLEC surveys)	Pivot Table Construction – Mean Values (Numerators) ❖ TLEC Survey Pivot Tables Pivot tables were created from the TLEC survey microdata using the following structure. A. Participation in Collaborative Professional Learning Activities Source question: • P3.9 – “<i>Mức độ tham gia của bản thân vào các hoạt động trao đổi, học hỏi chuyên môn – hình thức cộng đồng thực hành hoặc nhóm, mạng lưới chuyên môn?</i>” (Scale: 1 = Very low; 4 = Very high) Each sub-item under P3.9 corresponds to one collaborative learning activity, including: • Peer observation and feedback • Discussing holistic child development • Collaboration across age groups • Exchanging teaching materials • Joint planning of activities • Discussing individual children’s needs B. Pivot Table Structure • Rows • Collaborative professional learning activities: ○ Peer observation / feedback ○ Discussing holistic development ○ Age-group collaboration ○ Exchanging materials ○ Planning activities together ○ Discussing child needs ○ Average (across all listed activities) • Columns • Teacher grouping variables: ○ Grades I and II ○ Grades III and IV ○ Core teachers ○ Regular teachers ○ Public schools ○ Non-public schools ○ Overall average Each grouping was applied one dimension at a time within the PivotTable columns. • Values • Average of the corresponding P3.9 sub-item The Value Field Settings in Excel were set to: Summarise Values By → Average Because P3.9 uses a numeric 1–4 Likert scale, the resulting value represents the mean level of engagement in each activity for each subgroup. ❖ Consolidation of Pivot Outputs • Pivot tables were created separately for each P3.9 sub-item.

		- The resulting mean values were manually copied and aligned into the final crosstab layout used to produce Table B.23. - An additional average row (“Average”) was calculated as the arithmetic mean of the activity-specific mean scores. Mean Calculation Logic (General Rule) For each activity A and subgroup G, mean participation scores were calculated as: Mean participation score (A, G) = Average of P3.9 ratings for activity A among respondents in subgroup G - Ratings are on a 1–4 scale - Only valid responses were included - Blank or non-applicable responses were automatically excluded by Excel PivotTables Consistency of Application The same calculation logic was applied consistently across: - Teacher role (Core / Regular) - Professional title grade (Grades I–II vs Grades III–IV) - School sector (Public / Non-public) - Overall TLEC respondent population
B25	Self-reported participation in targeted professional development (TLEC survey)	Pivot Table Construction – Counts and Mean Values (Numerators) ❖ TLEC Survey Pivot Tables Pivot tables were created from the TLEC survey microdata to produce Table B.25: Self-reported participation in targeted professional development. A. Participation in Targeted Professional Development Source question: P3.19 – “<i>Mức độ tham gia của bản thân vào bồi dưỡng chuyên môn tại trường theo hình thức huấn luyện có mục tiêu (như làm mẫu, hướng dẫn, luyện tập để phát triển kỹ năng nào đó)?</i>” Responses were recorded on a 1–4 scale: 1 = None 2 = Not much 3 = Some 4 = A lot B. Pivot Table Structure Rows - Response options: 1 – None 2 – Not much 3 – Some 4 – A lot Columns - Total (no further disaggregation applied in Table B.25) Values Count of respondents selecting each response option Percentage of respondents Average of P3.19 (mean score) Because P3.19 is a numeric 1–4 scale, the <i>Average</i> value represents the mean level of participation in targeted professional development. ❖ Consolidation of Pivot Outputs

		• A frequency pivot table was used to calculate counts and percentages for each response option. • A second pivot table was used to calculate the mean score for P3.19. • Results from both pivot tables were manually aligned and presented together in the final crosstab used to produce Table B.25. Percentage and Mean Calculation Logic (General Rules) Distribution of Responses For each response option R, percentages were calculated as: Percentage (R) = Number of respondents selecting option R ÷ Total number of TLEC respondents Mean Participation Score Mean participation score = Average of P3.19 ratings • Calculated on a 1–4 scale • Only valid responses were included • Blank or non-applicable responses were automatically excluded by Excel PivotTables Consistency of Application The same calculation logic was applied consistently across: • All TLEC respondents • Response-category distributions • Mean score calculation
B26	Average time spent on professional tasks, by professional rank (LEY survey; 1–4 scale)	Pivot Table Construction – Mean Values (Numerators) ◆ LEY Survey Pivot Tables Pivot tables were created from the LEY survey microdata to produce Table B.26: Average time spent on professional tasks, by professional rank. A. Time Spent on Professional and Non-Professional Tasks Source question: • Q30 – “<i>Trong tuần làm việc đầy đủ gần đây nhất, thầy/cô đã dành bao nhiêu thời gian cho các nhiệm vụ sau tại cơ sở mầm non?</i>” Each task was rated using a 1–4 categorical scale indicating the share of working time: • 1 = No time • 2 = 1–10% • 3 = 11–30% • 4 = 31–60% B. Task Categories Individual Q30 items were grouped into the following task categories for reporting: • Core Professional Planning (e.g. planning and preparing educational activities) • Collaboration and Communication (e.g. communicating with colleagues and parents/guardians, excluding direct contact with children) • Documentation (e.g. recording children’s development, wellbeing, and learning) • Administration (e.g. management activities, staff meetings, administrative tasks)

		• Professional Development (e.g. participating in professional development activities) • Reflection (e.g. reflecting on one's work individually or with colleagues) • Non-Professional Chores (e.g. laundry, cleaning, shopping, cooking) C. Pivot Table Structure • Rows • Task categories (listed above) • Columns • Professional rank: ○ Overall average ○ Grade I ○ Grade II ○ Grade III ○ Grade IV • Values • Average of the corresponding Q30 item The Value Field Settings in Excel were set to: Summarise Values By → Average Because Q30 responses use a numeric 1–4 scale, the resulting value represents the average share of working time spent on each task category. ❖ Consolidation of Pivot Outputs • Pivot tables were created separately for each Q30 item. • Item-level averages were then grouped and labelled into broader task categories. • Mean values were manually copied and aligned into the final crosstab used to produce Table B.26. • All results reflect LEY respondents only. Mean Calculation Logic (General Rule) For each task category T and professional rank R, mean values were calculated as: Mean time score (T, R) = Average of Q30 ratings for tasks in category T among respondents in rank R • Scores are interpreted using the 1–4 time-share scale • Only valid responses were included • Blank or non-applicable responses were automatically excluded by Excel PivotTables Consistency of Application The same calculation logic was applied consistently across: • All task categories • Professional ranks (Grade I–IV) • Overall LEY respondent population
--	--	---

2. Instructions for creating maps in QGIS

- Source of Shapefiles:

Open Development Mekong (ODM) Datahub. Vietnam provinces.

<https://data.vietnam.opendevlopmentmekong.net/vi/dataset/a-phn-tnh>

Open Development Mekong (ODM) Datahub. Vietnam districts.

<https://data.opendevopmentmekong.net/vi/dataset/a-phn-huyn>

- The «QGIS » folder contains the files needed to produce the maps.

Figures 3&4

After obtaining the 'ece_teacher_province_data.csv' file from the 'Source Data' folder (file 2. ECEC_province_level_data), save it in the 'QGIS' folder. Then open QGIS (version 3.38) and load both the 'ece_teacher_province_data.csv' and 'vnm_province.shp' files.

Step 1: Load the files

In the Explorer window, open QGIS. Open a new project by clicking 'Project' > 'New'. Select the files one by one and drag them into the Layers window.

Step 2: Join the files

In the 'Layers' window, right-click 'vnm_province'. In the new window that opens, click on 'Properties', then 'Join'. Click on '+' at the bottom left. Then select the layer to join (the 'ece_teacher_province_data.csv' file), the join field (province) and the target field (Name). Click OK, then Apply, then OK.

Step 3: Mapping

In the 'Layers' window, right-click 'vnm_province'. Click 'Properties' > 'Symbology'. On the right, choose 'Rule-based'. Create the following rules by clicking the '+' button in the bottom-left corner each time:

- Label: ≥ 2.0

Filter: "ece_teacher_province_data_teacher_per_class" ≥ 2.0

- Label: 1.8 - 2.0

Filter: "ece_teacher_province_data_teacher_per_class" ≥ 1.8 AND

"ece_teacher_province_data_teacher_per_class" < 2.0

- Label: 1.6 - 1.8

Filter: "ece_teacher_province_data_teacher_per_class" ≥ 1.6 AND

"ece_teacher_province_data_teacher_per_class" < 1.8

- Label: 1.4 - 1.6

Filter: "ece_teacher_province_data_teacher_per_class" ≥ 1.4 AND

"ece_teacher_province_data_teacher_per_class" < 1.6

- Label: < 1.4

- Filter: "ece_teacher_province_data_teacher_per_class" < 1.4

Click 'Apply', then 'OK'.

Step 4: Save and export the map

In the 'Layers' window, right-click 'vnm_province', click 'Zoom to Layer(s)'.

In the top-left menu, click 'Project' > 'New Print Layout'. Name the layout 'Map 1'. In the new window, click 'Add Item' > 'Add Map'. An arrow appears. Use the arrow to draw a square that takes up all the space on the white sheet. Click 'Add Item' > 'Add Legend'. An arrow will appear again, draw a box where you want the legend to appear.

Then adjust the legend. On the right, click 'Item Properties'. In the 'Legend Items' tab, untick 'Auto update'. Right click 'ece_teacher_province_data' and choose 'Hidden'. Right click 'vnm_province' and choose 'Hidden'.

To export the map, in the top-left menu, click on 'Layout' > 'Export as Image'.

NB: Follow the same procedure to create Figure 4. The difference is in Step 3, where the rules are:

- Label: $< 5\%$

Filter: "ece_teacher_province_data_teacher_shortage_quota" < 5

- Label: 5 – 10%

Filter: "ece_teacher_province_data_teacher_shortage_quota" ≥ 5 AND

"ece_teacher_province_data_teacher_shortage_quota" < 10

- Label: 10 – 20%

Filter: "ece_teacher_province_data_teacher_shortage_quota" >= 10 AND
 "ece_teacher_province_data_teacher_shortage_quota" < 20
 - Label: 20 – 30%

Filter: "ece_teacher_province_data_teacher_shortage_quota" >= 20 AND
 "ece_teacher_province_data_teacher_shortage_quota" < 30
 - Label: >= 30%

Filter: "ece_teacher_province_data_teacher_shortage_quota" >= 30

Figure 6

After obtaining the 'ece_teacher_district_data.csv' file from the 'Source Data' folder (file 6. ECEC_district_level_data), save it in the 'QGIS' folder. Then open QGIS (version 3.38) and load both the 'ece_teacher_district_data.csv' and 'vnm_district.shp' files.

Step 1: Load the files

In the Explorer window, open QGIS. Open a new project by clicking 'Project' > 'New'. Select the files one by one and drag them into the Layers window.

Step 2: Join the files

In the 'Layers' window, right-click 'vnm_district'. In the new window that opens click on 'Properties', then 'Join'. Click on '+' at the bottom left. Then select the layer to join (the 'ece_teacher_district_data.csv' file), the join field (dist) and the target field (District). Click OK, then Apply, then OK.

Step 3: Mapping

In the 'Layers' window, right-click 'vnm_district'. Click 'Properties' > 'Symbology'. On the right, choose 'Rule-based'. Create the following rules by clicking the '+' button in the bottom-left corner each time:

- Label: 100%

Filter: "ece_teacher_district_data_pct_tenured_teachers" >= 100
 - Label: 90% - 100%

Filter: "ece_teacher_district_data_pct_tenured_teachers" >= 90 AND
 "ece_teacher_district_data_pct_tenured_teachers" < 100
 - Label: 80% - 90%

Filter: "ece_teacher_district_data_pct_tenured_teachers" >= 80 AND
 "ece_teacher_district_data_pct_tenured_teachers" < 90
 - Label: 70% - 80%

Filter: "ece_teacher_district_data_pct_tenured_teachers" >= 70 AND
 "ece_teacher_district_data_pct_tenured_teachers" < 80
 - Label: < 70%

- Filter: "ece_teacher_district_data_pct_tenured_teachers" < 70

Click 'Apply', then 'OK'.

Step 4: Save and export the map

Zoom in on different provinces before exporting.

In the top-left menu, click 'Project' > 'New Print Layout'. Name the layout 'Map 3'. In the new window, click 'Add Item' > 'Add Map'. An arrow appears. Use the arrow to draw a square that takes up all the space on the white sheet. Click 'Add Item' > 'Add Legend'. An arrow will appear again, draw a box where you want the legend to appear.

Then adjust the legend. On the right, click 'Item Properties'. In the 'Legend Items' tab, untick 'Auto update'. Right click 'ece_teacher_district_data' and choose 'Hidden'. Right click 'vnm_district' and choose 'Hidden'.

To export the map, in the top-left menu, click on 'Layout' > 'Export as Image'.

Analysis Teacher-Training Institutions' Course descriptions

The analysis of teacher preparation pathways is based on a **structured qualitative coding of programs offered at 11 teacher-training institutions (TTIs)**. The purpose of the analysis was to generate **comparative diagnostic evidence** on

the coherence, depth, and coverage of core pedagogical domains across institutions, rather than to construct a course-by-course inventory or a machine-replicable curriculum database.

Data inputs

Program documentation was collected and compiled from official institutional sources. Materials included:

- program structures and course lists
- credit allocations
- practicum sequencing and duration
- syllabi and course descriptions (where available)

These materials were consolidated into structured spreadsheets alongside original source documents and provided to the World Bank analysis team. The detailed location of input information for table B7 in the Appendix is below:

B7	1. Hoa Lu University: Page 38 and 39 of Course outline 2. Vietnam National University – UED: Page 43-49 3. Hanoi National University of Education: Across the course outline 4. Ho Chi Minh City University of Education: Page 10-13 5. Quang Nam University: Page 55 6. Quy Nhon University: Page 14-15 7. University of Danang-University of Education (DNE): Page 17-19 of Course outline (Minimum credits is on page 1) 8. Tan Trao University – Diploma program: Page 19 – 24 9. Hanoi National College of Education: Page 6 – 11 10. Ho Chi Minh City College of Education: Page 4 -10 11. Kontum Community College: Page 14
-----------	--

Methodology for classifying ITE course credits.

Program documents were reviewed course by course and credits were reclassified into four standardized categories to enable comparison across institutions. Institutional labels such as “general,” “foundation,” “specialized,” “professional,” or “major” were not used automatically because TTIs apply these terms inconsistently. Instead, each course was coded according to its primary purpose.

Category	Coding rule
General knowledge	General education and broad foundational courses, including political education, law, foreign language, ICT, research methods, educational psychology, general education, arts foundations, and other non-ECEC-specific foundation courses.
Professional knowledge	ECEC-specific and teacher-preparation content, including child development, care and education, curriculum, assessment, subject-method courses, classroom management, inclusive education, family/community engagement, and specialist pedagogy.
Pedagogical training/practice	Early professional practice courses focused mainly on observation, school familiarization, planning, guided practice, or limited teaching responsibility.
Internship	Extended practicum or final placement courses in which students are expected to implement activities and act more fully as classroom teachers.

Where courses combined multiple purposes, they were coded according to the dominant stated purpose in the course title and description. Graduation theses or equivalent capstone projects were excluded from professional-experience counts unless explicitly described as school-based teaching practice. Physical education and national defense credits were excluded where institutions stated they were not included in the total program credit load.

The outputs of this coding constitute the data used in the paper. These outputs are presented in Tables B.7–B.8

Software Requirements

Microsoft Excel, QGIS (version 3.38)

Folder Structure

- Folder 1: Source Data
 - + 1. ECEC_National_level_data
 - + 2. ECEC_province_level_data
 - + 3. ECEC_Regional_level_data
 - + 4. TLEC_Respondent_level_data
 - + 5.1. LEY_Teacher responses
 - + 5.2. LEY_Manager responses
 - + 6. ECEC_district_level_data
 - + 2022 GDP per capita
 - + 2022 Urbanization rate
 - + 2023 TTI training target
 - + Ethnic ECE teachers
 - + Population – Ethnicity by provinces – 2019
 - + B6
 - + B5b
 - + Folder: TTI Course outlines
- Folder 2. QGIS
- Folder 3. Outputs
 - + Figure 7
 - + Figure 8
 - + LEY and TLEC figures
 - + MOET Administrative data_Tables and figures
 - + B3
 - + Appendix

Annex 1: Data received from the Ministry Of Education And Training

I. File Annex – Universalization scheme (file: Annex-PreK Universalization scheme 9.2023.xlsx)

1. *Data on the total and disaggregated numbers (by socio-economic regions) 2021-2025 of the following:*
 - i. Annex 2: Number of ECE facilities 2015-2021, the total and disaggregated numbers (by socio-economic regions)
 - ii. Annex 3: Number of non-public ECE facilities 2015-2021, the total and disaggregated numbers (by socio-economic regions)
 - iii. Annex 4: Total enrollment rate (nation-wide) 2015-2021, disaggregated by daycare and kindergarten
 - iv. Annex 5: Percentage of children attending 2 sessions/day, the total and disaggregated numbers (by socio-economic regions)
 - v. Annex 6: Percentage of malnutrition (stunting, underweight) 2015 -2020, the total and disaggregated numbers (by socio-economic regions)
 - vi. Annex 7: Percentage of ECE teachers meeting qualification standards 2016-2021, the nation's total
 - vii. Annex 8: Teacher/class ratio 2015-2021, the total and disaggregated numbers (by socio-economic regions)
 - viii. Annex 9: Percentage of preschools meeting national standards 2015-2021, the nation's total
 - ix. Annex 10: Stocking of classrooms by types (shared, permanent, semi-structure, temporary) 2016-2021, the nation's total
 - x. Annex 11: Number of supporting facilities and equipment, playsets 2016-2021, the nation's total
2. *Data on the total and disaggregated numbers (by province) in the school year 2020-2021 of the following:*
 - i. Annex 12: ECE institutions - academic year 2020-2021, by (i) provinces, (ii) nurseries and kindergartens, (iii) public schools, private schools, school sites, people-founded, independent institutions, (iv) meeting national standards, (v) getting independent assessments.
 - ii. Annex 13: Number of groups/classes AY 2020-2021, by (i) private vs. public, (ii) nurseries vs. kindergarten and (iii) provinces, socio-economic regions.
 - iii. Annex 14: Number of preschool children AY 2020-2021, by private vs. public, (ii) nurseries vs. kindergarten and (iii) provinces, socio-economic regions.
 - iv. Annex 15: Number of ECE teachers, administrators, and support staff AY 2020-2021, by provinces, types, shortage, meeting qualification standards or not, targets.
 - v. Annex 16: Stocking of ECE institutions' facilities AY 2020-2021, by types, shortage.
3. *Data on the projected/estimated numbers by 2030 of the following:*
 - i. Annex 17: Kindergarten children population (ages 3-5), by provinces.
 - ii. Annex 18: Total preschools, public/private preschools, by provinces.
 - iii. Annex 19: Children at preschools and the mobilization rate, by private vs. public schools & provinces.
 - iv. Annex 20: Classes in public vs. private schools, by provinces.
 - v. Annex 21: Teacher shortage, teachers and teachers and management staff requiring upskilling to meet professional standards, by provinces.
 - vi. Annex 22: Additional budget needed for investment in classrooms and facilities, by provinces.
 - vii. Annex 23-24: The ratio of commune/ward/town/district achieving kindergarten education universalization, by provinces.

II. Data on disadvantaged areas (file: Annex – ECE for disadvantaged areas Scheme)

1. ECE institutions - academic year 2020-2021, by (i) provinces, (ii) nurseries and kindergartens, (iii) public schools, private schools, school sites
2. Number of groups/classes AY 2020-2021, by (i) private vs. public, (ii) nurseries vs. kindergarten and (iii) provinces, socio-economic regions.
3. Number of children AY 2020-2021, by (i) private vs. public, (ii) nurseries vs. kindergarten and (iii) provinces, socio-economic regions.

4. Number of ECE teachers, administrators, and support staff AY 2020-2021, by provinces, types, shortage, meeting qualification standards or not, định mức, ethnic minority teacher (including number of teachers who can teach both Vietnamese and ethnic dialects).
5. Stocking of ECE institutions' facilities AY 2020-2021, by types, shortage, provinces.
6. Forecast of public and private ECE schools in 2030 by provinces.
7. Forecast of children in nurseries and kindergartens in 2030 in disadvantaged and extremely disadvantaged areas by provinces.
8. Forecast of classes in nurseries and kindergartens in 2030 in disadvantaged and extremely disadvantaged areas by provinces.
9. Forecast of teacher shortage in 2030 by provinces.
10. Forecast of investment need for facilities till 2030 by provinces.
11. Forecast of number of children needed Vietnamese strengthening till 2030 by provinces and by ethnic groups.
12. Forecast of the need for teachers' training 2023 - 2030.

III. Data on 15 provinces piloting preK universalization (Folder: 15 pilot provinces)

In each province's worksheet:

- i. population ages 3-5
- ii. groups of children or classes, children at preschools and the mobilization rate
- iii. Children attending 2 sessions
- iv. Children receiving school meals
- v. Stunting, malnutrition, obesity rates
- vi. classrooms, outdoor playsets, teaching aids, and instructional materials
- vii. Teachers by types of contracts, qualification and meeting standards or not.
- viii. Teachers/class ratio
- ix. Projection – outreach
- x. Projection – additional investments in classrooms and headcounts

Annex 2: ECEC Data Stocktaking

Data	Unit	Time period
1. National-level data		
Number of schools, by level (all/nursery/kindergarten); number of independent settings; number of classes/groups; number of children, by age group, gender, ethnic minority, disability; number of managers, teachers, support staff; number of teachers, by gender, ethnic minority, tenure, meeting standards; students/class ratio; students/teacher ratio; teacher/class ratio, by level (nursery/kindergarten) -- all by ownership (public/non-public)	school; class; child; staff; classroom	2014-2022, every school year
Admission targets of ECE teacher training institutions in 2023 (by institution)	students	2023
2. Province-level data (Coverage: nationwide)		
Number of ECE schools, by type (public/non-public); school sites; independent settings; schools meeting national standards; schools got independent assessment	school	school year 2020-2021
Number of ECE groups/classes, by type (public/non-public) and age group (nursery/kindergarten)	class	school year 2020-2021
Number of ECE children and coverage, by age group (nursery/kindergarten); children/class ratio, by type of institution (public/private); % of children in private institutions	child	school year 2020-2021
Number of kindergarten children, by age (3/4/5); kindergarten enrollment rate, by age (3/4/5)	child; percent	school year 2020-2021
Number of ECE staff, by type (administrators/teachers/support staff); teacher quota; teacher shortage; teacher/class ratio; number of teachers meeting standards [in public institutions only]	Staff	school year 2020-2021
Number of classrooms and supporting facilities, by type; classrooms and supporting facilities shortage [in public institutions only]	facility	school year 2020-2021
Percentage of districts and communes meeting kindergarten universalization standards	percent	2022-2030, every year
Forecast of ECE schools, by type (public/non-public) by 2030	school	school year 2020-2021; school year 2029-2030
Forecast of preschool children population, by age (3/4/5) by 2030	child	school year 2020-2021; 2030
Forecast of enrollment, by type of institutions (public/non-public); coverage (kindergarten); coverage (non-public institutions)	child; percent	2030
Forecast of ECE classes, by type of institutions (public/non-public) by 2030	class	2030
Forecast of teacher shortage and training demand by 2030	staff	2030
Forecast of investment demand by 2030 [in public institutions only]	facility; expenditure (VND)	2030
Forecast of percentage of communes meeting kindergarten universalization standards	percent	2022-2030, every year
3. Province-level data (Coverage: 15 provinces¹)		
Forecast of teacher shortage by 2030; teacher quota	teacher	2030
4. District-level data (Coverage: 14 provinces²)		

¹ 15 provinces include: Vinh Phuc, Bac Ninh, Hai Phong, Ninh Binh, Lao Cai, Dien Bien, Son La, Hoa Binh, Da Nang, Kon Tum, Ba Ria – Vung Tau, Long An, Soc Trang, Bac Lieu, Ho Chi Minh City

² 14 provinces include: Vinh Phuc, Hai Phong, Ninh Binh, Lao Cai, Dien Bien, Son La, Hoa Binh, Da Nang, Kon Tum, Ba Ria – Vung Tau, Long An, Soc Trang, Bac Lieu, Ho Chi Minh City

Data	Unit	Time period
Preschool children population (total, handicapped, female, ethnic minority, eligible for preschool) by age (nursery/3/4/5), number of classes by age, number of students (enrolled, attending full-day, semi-boarding, malnutrition) by age, facilities, teachers (by type of contracts, whether meeting standards), teacher/class ratio	child; teacher; facility	school year 2020-2021
5. Institution-level data (Coverage: 6 provinces³)		
List of ECE institutions, with information on number of classes by age (nursery/3/4/5), number of students (enrolled, attending full-day, semi-boarding, malnutrition) by age, facilities, teachers (by type of contracts, whether meeting standards), teacher/class ratio	child; teacher; facility	school year 2020-2021

³ 6 provinces include: Vinh Phuc, Dien Bien, Da Nang, Kon Tum, Ba Ria – Vung Tau, Ho Chi Minh City