

VNUHCM - UNIVERSITY OF SCIENCE

FACULTY OF INFORMATION TECHNOLOGY

CSC14003 – PROGRAMMING FOR DATA SCIENCE



Project Report

Team Plan and Work Distribution

CLASS 23KHDL - GROUP 4

Students:

23127168 – Cao Trần Bá Đạt
23127238 – Trần Hoài Thiện Nhân
23127516 – Bùi Nam Việt

Lecturers:

Mr. Phạm Trọng Nghĩa
Mr. Lê Nhựt Nam
Mr. Nguyễn Thanh Tình

December 24th, 2025

Contents

1 Project Information	2
2 Team Member Information	2
3 Work Breakdown by Member	2
3.1 Student 1 - Cao Trần Bá Đạt (23127168)	2
3.2 Student 2 - Trần Hoài Thiện Nhân (23127238)	3
3.3 Student 3 - Bùi Nam Việt (23127516)	4
4 Collaboration Process	4
4.1 Communication & Coordination	4
4.2 Scrum Meetings Schedule	4
4.3 Workflow	5
5 Project Timeline	5
6 Project Structure	6

1 Project Information

Item	Details
Project Name	Life Expectancy (WHO) - Exploratory Data Analysis & Modeling
Dataset	Life Expectancy (WHO) Fixed
Description	The subject of this dataset is health data of people in countries around the world based on information provided from The Global Health Observatory (GHO) data repository under World Health Organization (WHO).
Class	23KHDL

2 Team Member Information

#	Full Name	StudentID	Role
1	Cao Trần Bá Đạt	23127168	Dataset Documentation & Questions 2, 5
2	Trần Hoài Thiện Nhân	23127238	Project Structure & Questions 4, 6, 7
3	Bùi Nam Việt	23127516	Data Exploration & Questions 1, 3

3 Work Breakdown by Member

3.1 Student 1 - Cao Trần Bá Đạt (23127168)

Tasks	Description	Deadline	Status
Dataset Subject	Describe topic, domain, phenomenon and real-world context	22h 26/11	Done
Data Source	Platform, URL, original authors, publication date	22h 26/11	Done
Data License	License verification and usage restrictions or attribution requirements	22h 26/11	Done
Data Collection Method	Collection method, target population, limitations	22h 26/11	Done
Dataset Selection Rationale	Interests and potential questions	22h 26/11	Done
Question 2	Key factors to improve longevity growth	15h 7/12	Done

Question 5	Factors that help poor countries have high life expectancy	15h 7/12	Done
Key findings	List 3-5 most important insights from your analysis	22h 12/12	Done
Future Directions	Additional exploration and improvement suggestions	22h 12/12	Done
Individual Reflection	Personal challenges, learning, and growth	22h 12/12	Done

Contribution Percentage: ~33%

3.2 Student 2 - Trần Hoài Thiện Nhân (23127238)

Tasks	Description	Deadline	Status
Project Skeleton	Create Jupyter Notebook structure and Table of Contents	22h 26/11	Done
Team Plan	Work distribution and project timeline	22h 26/11	Done
Question 4	The Battle of Development: Schooling vs. GDP in the Preston Curve Context	15h 7/12	Done
Question 6	What factors are needed to reduce infant mortality?	15h 7/12	Done
Question 7	The "Paradox of Affluence" - How Lifestyle (Alcohol, BMI) Impacts Life Expectancy Across Economic Groups	15h 7/12	Done
Merge Code	Integrate all team members' contributions	22h 10/12	Done
Individual Reflection	Personal challenges, learning, and growth	22h 12/12	Done

Contribution Percentage: ~33%

3.3 Student 3 - Bùi Nam Việt (23127516)

Tasks	Description	Deadline	Status
Data Exploration	Comprehensive EDA (preprocessing interleaved)	22h 26/11	Done
Question 1	Why did life expectancy trends 2000–2015 vary across countries and regions, and what factors explain the improvements or declines?	15h 7/12	Done
Question 3	How to assess the importance of groups of factors and build accurate life expectancy prediction models?	15h 7/12	Done
Limitations	Dataset, analysis, and scope limitations	22h 12/12	Done
References	Materials and resources consulted	22h 12/12	Done
Individual Reflection	Personal challenges, learning, and growth	22h 12/12	Done
README File	Project overview, dataset info, findings summary	22h 15/12	Done

Contribution Percentage: ~34%

4 Collaboration Process

4.1 Communication & Coordination

- Primary Communication: Team meetings and online discussions
- Code Management: Git and Github for version control and code sharing
- Task Tracking: Shared task board with deadlines and status updates

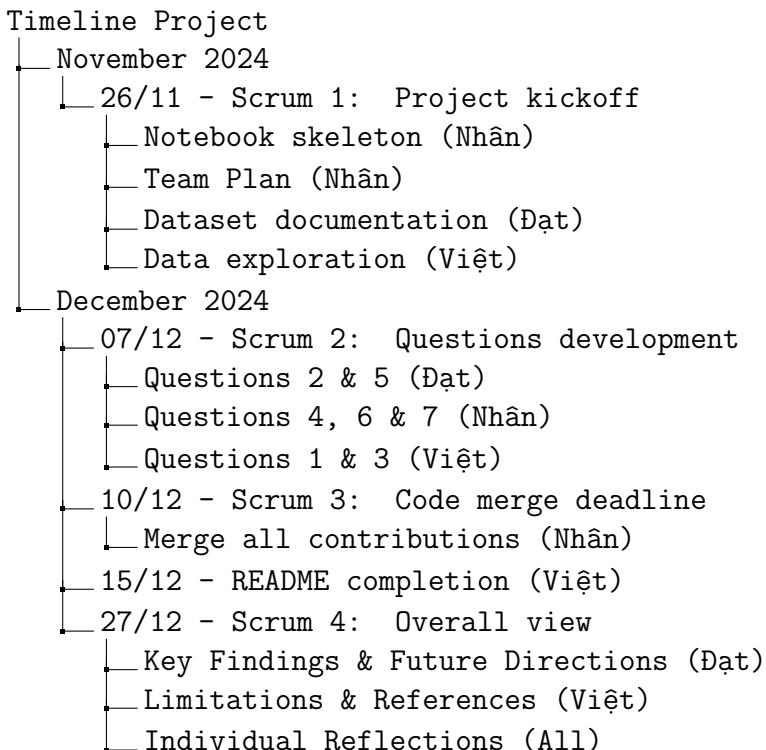
4.2 Scrum Meetings Schedule

Meeting	Date/Time	Purpose
Scrum 1	22h 26/11	Project kickoff, initial task assignments
Scrum 2	15h 7/12	Review questions, discuss next steps
Scrum 3	22h 10/12	Review code and content, merge code
Scrum 4	10h 27/12	Overall review, prepare for submit

4.3 Workflow

1. Individual Work: Each member completes assigned tasks independently
2. Review Sessions: Regular reviews after completing assigned tasks
3. Code Merge: Final integration by designated team member (Trần Hoài Thiện Nhân)
4. Quality Check: All members review the final notebook

5 Project Timeline



#	Question	Assigned to	ML required
Q1	Why did life expectancy trends 2000–2015 vary across countries and regions, and what factors explain the improvements or declines?	Việt	No
Q2	Key factors to improve longevity growth	Đạt	No
Q3	How to assess the importance of groups of factors and build accurate life expectancy prediction models?	Việt	Yes
Q4	The Battle of Development: Schooling vs. GDP in the Preston Curve Context	Nhân	Yes
Q5	Factors that help poor countries have high life expectancy	Đạt	No
Q6	What factors are needed to reduce infant mortality?	Đạt	No
Q7	The "Paradox of Affluence" - How Lifestyle (Alcohol, BMI) Impacts Life Expectancy Across Economic Groups	Nhân	No

6 Project Structure

```
Life-Expectancy-EDA-Modeling/
├── data/
│   └── Life-Expectancy-Data-Updated.csv
├── documentation/
│   └── Team Plan and Work Distribution.pdf
└── src/
    ├── 01_data_exploration.ipynb
    ├── 02_meaningful_questions.ipynb
    └── 03_project_summary.ipynb
├── README.md
└── requirements.txt
```