**Project Proposal: An Examination of the Efficacy of Recent Motorcycle Helmet Campaigns in Thailand**

Jeffrey Orcutt

University of Wisconsin – Green Bay

Master in Data Science Capstone Project

December 1, 2025

**Initial Working Title of Project:** Evaluating the Impact of Helmet Promotion and Enforcement on Motorcycle Accident Hospitalizations in Thailand.

**Project Description (Summary/Overview):** Thailand’s road safety is consistently ranked among the most dangerous in the world. Using data provided by the Thai Government on all hospitalizations resulting from vehicle incidents, this project seeks to determine whether increased enforcement and education projects have significantly impacted motorcycle accident outcomes.

**Project Rationale:** In the past, I’ve visited Thailand over a dozen times and made Thailand my home for a year. I was astonished by the lack of enforcement of various traffic laws despite social media and traditional media consistently covering adverse outcomes of traffic accidents. This project can add an updated view of the situation, allowing policymakers to examine whether their current education efforts are working.

**Proposed Project Purpose:** This project is a case study examining whether policy and enforcement efforts on motorcycle helmet requirements in Thailand have had an effect. This project could benefit the Thai population, government, and non-governmental organizations investing in efforts to curb risky driver behavior. Additionally, it may have broader implications for similarly situated countries in the region with similar culturally acceptable risky behaviors.

**Project Objectives (List 3-5 project objectives):**

1. Provide a comprehensive analysis of motorcycle accident trends and their severity in Thailand using government hospitalization data.

2. Using data-driven techniques, identify and evaluate the effectiveness of helmet promotion and enforcement campaigns in changing motorcyclist behavior.

3. Effectively analyze and present insights on the impact of helmet promotion and enforcement efforts, using data visualizations and statistical findings to support conclusions.

**Project Data**: This project's data source is the Thai government’s Department of Disease Control’s Injury Surveillance Data. (*ข้อมูลระบบเฝ้าระวังการบาดเจ็บ (Injury Surveillance)*, 2024) This data contains more than 2 million hospitalizations over the past six years resulting solely from vehicle accidents. The data in its original form is in Thai. I'll combine my language knowledge with Google Translate API to create an English version of the dataset. Furthermore, various entries within the data are significantly missing or incorrect, according to the data definitions and entry guide, so there will need to be a significant amount of time spent in the ETL phase of the project.

**Source of Data:**

*ข้อมูลระบบเฝ้าระวังการบาดเจ็บ (Injury surveillance: IS) - Open Government Data of Thailand*. (2024, December 11). https://data.go.th/dataset/injury-surveillance

**Computational Tools:** During the ETL phase of the project, I’ll use Python, publicly accessible packages, and Internet access for the Google Translate API. I don’t anticipate the need to use distributed computing.

**Data Science concepts/methods/techniques that will be applied:** Following the ETL phase, my analysis plan includes examining regression models concerning helmet usage and employing classification models for expected outcomes of accidents. Regression and Multiple Regression models come from DS 705, various classification models such as Random Forests and XGBoost, and cross-validation of models were taught in DS 740, and visualization methods were taught in DS 730 and DS 745. Then, I will interpret those results and glean insights that will be relayed in a written report with appropriate visualizations.

**Description of Final Report (tentative outline of your capstone paper):** The paper will have an overview, a literature review of 18+ sources I’ve found so far, an introduction to the data, an analysis and discussion of the employment of each model, and conclusions.

**Project Timeline with Activities: (activities you will be engaged in to achieve objectives listed above and timeline of completion of these activities):** Literature review and search – 21 found so far, 13 completed. ETL phase – 80% complete at this point; I think it’ll be done by mid-February; introduction to data/analysis/conclusions to begin and be complete by mid-March.

**Resources and/or Contacts that will be utilized for the project:** I’ve extensively used UWGB Libraries for sources, Google Translate API for on-demand translation needs, and my Thai partner for translation verification and support. Also, Grammarly and other tools will help me with grammar and clarity.

**Project Timeline with Activities:**

|  |  |  |  |
| --- | --- | --- | --- |
| Week | Task # | Assignment Name | Due Date |
| 1 | 3.1 | Make Discussion post | 1/28/2025 |
| 1 | 1.1 | Finish incorporating details from Canvas | 1/30/2025 |
| 1 | 1 | Capstone Project Idea Submission - Required Assignment | 1/31/2025 |
| 1 | 2 | DS735 Project 1 (i.e. mini literature review) | 1/31/2025 |
| 1 | 9.1 | Chapter 2 - Create Literature Review for sources 1-5 | 1/31/2025 |
| 2 | 3 | Introduction Discussion - Required Assignment | 2/2/2025 |
| 2 | 3.21 | Respond to peer 1 | 2/2/2025 |
| 2 | 3.22 | Respond to peer 2 | 2/2/2025 |
| 2 | 6.1 | Chapter 1 - Write Introduction | 2/5/2025 |
| 2 | 6.2 | Chapter 1 - Write Background | 2/5/2025 |
| 2 | 6.3 | Chapter 1 - Statement of Business Problem | 2/5/2025 |
| 2 | 9.2 | Chapter 2 - Create Literature Review for sources 6-10 | 2/8/2025 |
| 3 | 4 | Capstone Project Proposal & Timeline | 2/9/2025 |
| 3 | 4.1 | Incorporate Feedback from the Instructor | 2/9/2025 |
| 3 | 4.2 | Ensure Meets Guidelines | 2/9/2025 |
| 3 | 6.4 | Chapter 1 - Project Objectives | 2/15/2025 |
| 3 | 6.5 | Chapter 1 - Significance of Project | 2/15/2025 |
| 3 | 6.6 | Chapter 1 - Definition of Terms | 2/15/2025 |
| 3 | 6.7 | Chapter 1 - Assumptions, Limitations, Delimitations | 2/15/2025 |
| 3 | 6.8 | Chapter 1 - Conclusion | 2/15/2025 |
| 3 | 9.3 | Chapter 2 - Create Literature Review for sources 11-15 | 2/15/2025 |
| 4 | 5 | Project Discussion | 2/16/2025 |
| 4 | 9.4 | Chapter 2 - Create Literature Review for sources 16-XX | 2/22/2025 |
| 5 | 6 | Project Update 1 | 3/1/2025 |
| 5 | 9.5 | Chapter 2 - Write Introduction | 3/1/2025 |
| 6 | 7 | Chapter 1 submission for faculty comments | 3/2/2025 |
| 6 | 9.6 | Chapter 2 - Conclusion to Chapter 2 | 3/3/2025 |
| 6 | 11.1 | Chapter 3 - Introduction | 3/4/2025 |
| 6 | 11.2 | Chapter 3 - Data Overview | 3/4/2025 |
| 7 | 8 | Project Update 2 | 3/15/2025 |
| 7 | 11.3 | Chapter 3 - Data Cleaning and Processing | 3/11/2025 |
| 8 | 9 | Chapter 2 submission for faculty comments | 3/16/2025 |
| 8 | 11.4 | Chapter 3 - Exploratory Data Analysis | 3/18/2025 |
| 9 | 11.5 | Chapter 3 - Model Development, Deployment, and Evaluation | 3/25/2025 |
| 10 | 11.6 | Chapter 3 - Conclusion | 4/1/2025 |
| 10 | 13.1 | Chapter 4 - Introductions | 4/1/2025 |
| 11 | 10 | Project Update 3 | 4/5/2025 |
| 11 | 11 | Chapter 3 submission for faculty comments | 4/6/2025 |
| 11 | 13.2 | Chapter 4 -Results | 4/8/2025 |
| 12 | 13.3 | Chapter 4 -Conclusion | 4/15/2025 |
| 13 | 12 | Project Update 4 | 4/19/2025 |
| 14 | 13 | Chapter 4 submission for faculty comments | 4/20/2025 |
| 14 | 14 | Turnitin Submission for Chapters 1-4 | 4/20/2025 |
| 14 | 15.1 | Chapter 5 - Introduction | 4/22/2025 |
| 14 | 15.2 | Chapter 5 - Summary of Project | 4/22/2025 |
| 15 | 15.3 | Chapter 5 - Implications of Results | 4/28/2025 |
| 15 | 15.4 | Chapter 5 - Limitations and Next Steps | 4/28/2025 |
| 15 | 15.5 | Chapter 5 - Conclusion | 4/30/2025 |
| 16 | 15 | Final Paper - Client Based Project or Case Study | 5/4/2025 |
| 16 | 17 | Final Documentation - Capstone Website Form | 5/5/2025 |
| 16 | 18 | Final Project Self Reflection Paper | 5/6/2025 |

Summary: I don’t anticipate any significant hurdles to completing this project this semester. My literature review has begun, with 10 of the 21 sources I’ve found being summarized already and in a framework of organization. My next steps are to finish the literature review and incorporate them into my project and with each other. I have a good start on processing the data in the ETL phase, should be able to begin the EDA phase in the next week or two, have concepts of a plan for modeling choices, and don’t anticipate any real issues with analysis or interpretation.