PHILIPPINE JUNIOR DATA SCIENCE CHALLENGE
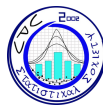
2024

DATASETS

INTRODUCTION

The Philippine Junior Data Science Challenge (PJDSC) 2024, hosted by the UP Data Science Society, invites senior high school and undergraduate students to tackle real-world issues related to Sustainable Urban Mobility through data science. Participants in the Elimination Round will receive datasets to develop data-driven project proposals addressing challenges in public transit, cycling infrastructure, and many others. Insights from quality data can be used to create sustainable solutions that enhance accessibility and optimize transportation systems.

Provided here are datasets that may be of use to the teams as they draft their Data Science Project Proposals.

Note that teams are free to use additional datasets beyond those provided here.

DATASETS

NAME	DESCRIPTION	LINKS
SafeTravel PH Urban Mobility Datasets	A variety of data gathered from vehicle feed data that include geospatial, time, velocity, and boarding/alighting time data from buses deployed and monitored in EDSA, Bacolod, Baguio, and Iloilo from 2022-2023 **IMPORTANT:** Please read the **Open Data license** before using datasets.	**SafeTravel Open Data Drive** **SafeTravel Data Inventory Website**
DOTr Bike lanes Datasets	Geospatial data on the bike lanes on roads across the country. Source: Department of Transportation	**AT_BikeLanes_08282024**

NAME	DESCRIPTION	LINKS
Road Traffic Accident Data of Epifanio delos Santos Avenue, Metro Manila (2007-2016)	Contextual, geospatial, and date and time information on traffic Accidents that occurred in EDSA from 2007-2016. Source: Luz, Leonard; Blanco, Ariel (2019), "Road Traffic Accident Data of Epifanio delos Santos Avenue, Metro Manila (2007-2016)", Mendeley Data, V1, doi: 10.17632/hwbf6n4krw.1	**Road Traffic Accident Data of Epifanio delos Santos Avenue, Metro Manila (2007-2016)**
Project CCHAIN Datasets	The Project Climate Change, Health, and Artificial Intelligence (Project CCHAIN) dataset is a validated, open-sourced linked dataset containing 20 years (2003-2022) of climate, environmental, socioeconomic, and health dimensions at the barangay (village) level across twelve Philippine cities (Dagupan, Palayan, Navotas, Mandaluyong, Muntinlupa, Legazpi, Iloilo, Mandaue, Tacloban, Zamboanga, Cagayan de Oro, Davao). Source: *Project CCHAIN*. (2024, June 25). Kaggle. https://www.kaggle.com/datasets/thinkdatasci/project-cchain/	**Project CCHAIN**