Northeast Big Data Innovation Hub & National Student Data Corps
Welcome to the Transportation Data Science Project (TDSP) and Community!
Help Make Roads Safer for Vulnerable Road Users in Disadvantaged Communities

The Problem: [According to the U.S. Department of Transportation Federal Highway Administration](#), vulnerable road users, also known as pedestrians, bicyclists, or other cyclists, account for about 19% of all U.S. roadway fatalities. Over 76,000 pedestrians and 47,000 bicyclists are injured in roadway crashes annually. We invite you to join the Transportation Data Science Community and tackle the challenge to better understand where and why these fatalities happen, and how we can help prevent them.

Source: US Department of Transportation, Federal Highway Administration, FHWA Highway Safety Programs, Pedestrian & Bicycle Safety: https://highways.dot.gov/safety/pedestrian-bicyclist. This DOT website was last updated June 16, 2023

How You Can Help: Focusing on vulnerable road users, this data science project will teach you how to analyze transportation data and create real-world recommendations so that *you* can help improve road safety for all. Jump into the project below to join this global community of transportation data science explorers!

Learners from around the world, from all educational backgrounds and skill levels, are invited to participate.

Project Overview: Leveraging a [New York City OpenData transportation dataset](#), transportation data science learners will use data science tools and techniques to develop data-driven insights on how roads can be made safer for all. Participants will learn to assess for potential bias in the data by comparing neighborhoods or zip codes, an important element of data science ethics. They will also create pertinent data visualizations and graphs to communicate what the data shows and develop analytical models that can be deployed in the real world.

This project was designed to inspire everyone interested in the use of data in transportation, from beginner to advanced learners. This is an opportunity for you to learn data science techniques in a supportive environment!

🚗**Choose Your Route:**🚗

This project is offered in two different formats depending on your experience level. Each format can be accomplished in 6 to 10 weeks (or less) depending on your experience. We invite you to participate in one or both of the following TDSP routes!

➤ **The Explorer TDSP** has six Milestones, including guided transportation research into your own community!
  ○ **Prerequisites:** There are no prerequisites for the Explorer TDSP, although some basic knowledge of data analysis and Python can be helpful! We will guide and support you throughout this project.
➤ *Coming Soon!* **The Navigator TDSP** has eight Milestones, including self-guided research and discovery on transportation engineering and data science.
  ○ **Prerequisites:** Intermediate data science and Python programming experience will be helpful while completing this project. We recommend you complete the Explorer TDSP first, though not required.

Monthly mentoring calls with the Transportation Data Science Community will support both the Explorer and the Navigator TDSP participants.

TDSP Key Highlights:
- You'll gain practical experience in data cleaning, exploration, and preparation.
- You'll learn about time series analysis, geospatial analysis, and data storytelling techniques.
- You will develop a transportation research question and a written summary of your insights.
- Virtual mentoring sessions will be available monthly to answer your questions, provide guidance, and address any challenges you encounter.

Final Submission & Incentives:

At the end of the project(s) you choose, you will submit a [one-page virtual poster board](#) to display your final visualizations and insights to the Northeast Big Data Innovation Hub and National Student Data Corps (NSDC) HQ team at Columbia University, and the U.S. Department of Transportation/Federal Highway Administration (DOT/FHWA). This final submission will showcase your data analysis skills and your ability to communicate findings effectively.

Once your project is submitted, the NSDC:
- Will provide you with a virtual certificate for your participation.
- May post select visualizations on the NSDC website as part of a project showcase.
- May invite you to present your findings to DOT/FHWA during a live event, hosted by the NSDC!

Collaboration and Mentored Learning:

We encourage you to share ideas and seek feedback from your peers and the NSDC HQ team throughout the project. You may complete the project(s) individually or with a group of 2-4 people. The NSDC will host monthly virtual mentoring sessions over Zoom, and asynchronously over [Slack](#).

Tools You Will Need:
- A Gmail Account to access the [Google Colab Notebook](#) (required)

Important Links:
- [Slack Channel for Discussion and Guidance](#) (**#tdsp-community** channel)
- nsdc@nebigdatahub.org (email us with any questions!)

🚦 **[Register to Participate!](#)** 🚦