CSCI 370 Final Report

bpx: PLX Behavior Benchmarking - Optimizing Routing and Well Prioritization with Agentic AI

White Shirt, Blue Tie

Karl Eisenbarth
Alex Oh
Shane Ritter
Ethan Tran

Delivered June 14, 2025

CSCI 370 Summer 2025

Mr. Scott Jensen

Table 1: Revision history

Revision	Date	Comments
New	5/14/25	Created Document
Rev – 2	5/17/25	Added Introduction, Functional Requirements, Non-Functional Requirements, Risks, Definition of Done, Team Profile, and Appendix
Rev – 3	5/24/25	Added System Architecture
Rev – 4	5/31/25	Added Software Test and Quality, and Ethical Considerations
Rev – 5	6/7/25	Added Project Completion Status, Future Work, and Lessons Learned, Updated Software Test and Quality
Rev – 6	6/8/25	Added References, fixed grammar, tense, and improved readability
Rev – 7	6/13/25	Incorporating feedback and polishing for final submission

Table of Contents

I. Introduction

About bpx energy

bpx energy is the U.S. onshore oil and gas subsidiary of bp, focused on transforming the traditional energy model by using automation, AI, and digital tools to make operations safer, more efficient, and more sustainable. With hundreds of field operators managing thousands of sites across Texas and Louisiana, routing optimization is a major operational challenge. By improving how operators are guided to their daily tasks, bpx aims to reduce fuel use, cut down on administrative burden, and increase productive time on-site. This project directly supports bpx's goal of modernizing field operations through data-driven tools like PLX, making it a key step in their broader digital transformation strategy [1].

Purpose

This document outlines the objectives for the PLX (Perfect Lap Execute) Routing Optimization and Behavioral Benchmarking project, conducted in partnership with bpx energy, a company in the oil and gas industry. The project aims to analyze the historical driving behavior data from bpx energy field operators, employees who regularly drive to oil wells to perform inspections and maintenance, and use it to build a reliable benchmark that evaluates the effectiveness of a newly deployed routing tool.

The tool, PLX, is an AI-powered voice assistant that guides operators through optimized routes, with the goal of reducing travel inefficiencies, increasing time spent on maintenance tasks, and minimizing administrative overhead. By comparing historical routing patterns with PLX-driven behavior, we aim to quantify improvements in route efficiency, decision-making, and regulatory compliance. Our goal is to create a benchmark for Pre-PLX data, creating concrete metrics to base future PLX behavior on and visualize Post-PLX improvements in an intuitive and user-friendly manner.

Project Overview

To support this goal, we utilize GPS telemetry data (breadcrumb data) collected from NAUTO devices installed in bpx energy vehicles. This data includes detailed information on vehicle movement, stop duration, and location coordinates, enabling us to reconstruct daily operator behavior.

The project focuses on:

- Identifying route inefficiencies
- Suggesting improvements to the GIS-based routing network
- Quantifying operational metrics like wrench time, administrative time, and commute time
    - **Wrench time**: Time spent parked/stopped on a pad that has oil/gas well heads
    - **Administrative time**: Time spent parked outside of a pad or at a bpx energy office
    - **Commute time**: Time spent driving from home to bpx energy office, and office to designated PLX route areas.

Deliverable includes key performance indicators (KPIs), interactive dashboards, and documentation to support evaluation of the PLX system's impact. The intended stakeholders are bpx energy operations leadership, routing developers, and the field team using PLX in daily work.

This document defines the project's vision, system requirements, risks, and success criteria for completion.

II. Functional Requirements

1. **Benchmark operator driving behavior using NAUTO** breadcrumb data by analyzing time spent driving, stopping, and visiting well sites.

2. **Evaluate routing behavior changes over time** by comparing key performance indicators (KPIs) such as route length, stop frequency, and on-site durations before and after PLX deployment.
3. **Assess GIS network contributions** to routing outcomes to routing outcomes by identifying how map structure, well location, and road access influence route selection.
4. **Develop route efficiency metrics** and visualization tools to measure route efficiency, commute time, wrench time, and administrative time using Python (pandas, matplotlib).
5. **Deliver a spatial dashboard** in ArcGIS Pro that visualizes pre- and post-PLX breadcrumb routes. The dashboard should include layer toggles, time filters, and annotation tools to highlight operator behavior shifts such as new detours or reduced idle time.
6. **Create a final presentation and executive summary** that uses charts, KPIs, and case examples to communicate the impact of PLX to both technical teams (developers, analysts) and non-technical stakeholders (managers, field leads).
7. **Visualize key operator metrics** including:
    a. Wrench time (time at well heads)
    b. Administrative time (engine-on idle time at pads or gates)
    c. Commute time (travel from home to office to first field location)
8. **Detect, explain, and resolve outliers** in GPS data (e.g., impossible routes, abnormal durations) using scripted logic and documented thresholds.

III. Non-Functional Requirements

1. **Dashboards and visuals must be user-friendly**, with clear labels, legends, and filter options accessible to both technical and non-technical users.
2. **All tools must be compatible** with bpx energy's software environment, including ArcGIS Pro and standard CSV/Python workflows.
3. **Code and analysis must be modular and documented**, with inline comments, a README file, and an onboarding guide for future developers or interns.
4. **The system must support scalable data integration**, allowing easy updates with new NAUTO datasets or changes in route structures without major refactoring.

IV. Risks

1. NAUTO breadcrumb data may be large or contain gaps, inconsistencies, or missing timestamps, which could affect the accuracy of reconstructed driving behavior.
2. If current team members lack familiarity with GIS tools, ArcGIS Pro, or routing algorithms, it could hinder progress or require additional onboarding time.
3. Communication gaps or delayed feedback from bpx energy stakeholders or team members could lead to misalignment of expectations or delayed course corrections.
4. Differences between Mines' ITS systems, ArcGIS Pro licensing, or secure file sharing methods may delay access to necessary data or software environments.
5. Mishandling of sensitive data (e.g., vehicle IDs or operator names) could violate NDA terms or result in deliverables being unusable by the client.

V. Definition of Done

1. A spatial dashboard is provided that shows analytics on NAUTO breadcrumb data Pre-PLX and Post-PLX.
    a. This will be provided as a Jupyter Notebook file.
2. Numerical results are well documented, contrasting the performance of the routing in terms of detours, route length reduction, stop efficiency, and unnecessary backtracking.
3. A presentation or executive summary is written and supplied to bpx energy, detailing how PLX has affected operator routing behaviour, and suggesting where further improvement could be made.
4. Benchmarks are visualized with the help of the available GPS data, presenting operator behavior on different routes and at different points in time.

5. All tools and scripts for analysis and visualization are well-documented, modular, and transferred in a usable format (e.g., Jupyter Notebooks, ArcGIS projects).

VI. System Architecture

Our project is focused primarily on data analysis, which allows for a relatively straightforward system architecture (Figure 1). The simplicity of the architecture reflects the linear nature of our process: data is collected, analyzed, and then transformed into actionable insights through visualizations.

The initial data is gathered by NAUTO systems installed in bpx energy fleet vehicles. These systems transmit breadcrumb data at regular intervals (approximately every 10–30 minutes), capturing key fields such as the operator identity, vehicle identification, GPS coordinates (latitude and longitude), speed, and operational state (moving, parked, or stopped). While the system is designed to handle both pre- and post-PLX datasets, our current analysis focuses solely on Pre-PLX data collected between 2023 and early 2025, as the PLX system has not yet been fully deployed. These systems act as the entry point of our pipeline, providing high-resolution driving data that forms the foundation for all further analysis.

Once collected, the data is centralized within bpx energy's OneMapGIS platform. This platform facilitates spatial organization and enables seamless export to tools like ArcGIS Pro. From ArcGIS, we export the relevant information into Excel spreadsheets, making the data accessible for further processing. Excel acts as a lightweight staging layer, allowing for preliminary inspection and validation of the data structure before it is processed in the notebook environment.

Next, the data is imported into Python-based Jupyter Notebooks, which serve as our main environment for data wrangling. Here, we clean, filter, and manipulate the datasets—removing anomalies, correcting formatting, and engineering features necessary for deeper analysis. Our Jupyter scripts are modular and reusable, allowing us to compute key metrics like stop time, distance traveled, and wrench time, while also generating figures and tables for the final report.

Finally, our insights are communicated through tailored visualizations. These visual outputs are critical to interpreting trends and evaluating route efficiency, operator behavior, and potential impacts of PLX. The streamlined architecture—from data acquisition to visualization—not only matches the structure of our project goals but also allows us to stay flexible and modular as the scope of analysis evolves. Each stage in the architecture builds directly on the previous one: data ingestion ensures completeness, Excel enables inspection, notebooks enable metric computation, and the visualizations synthesize findings into accessible outputs.

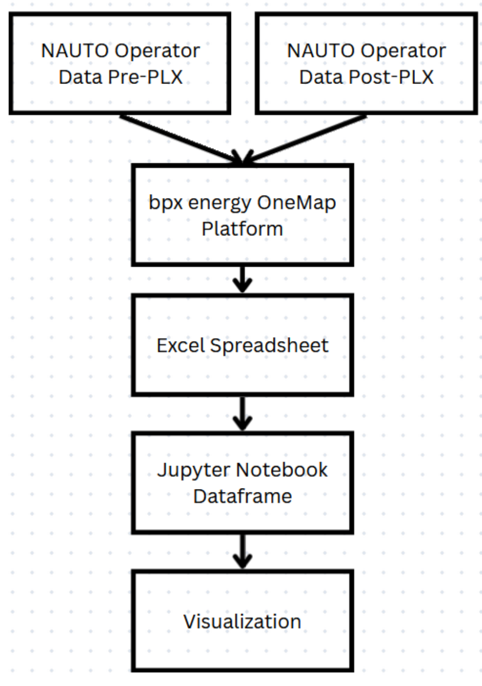


Figure 1 - Overview of Data Inputs

While Figure 1 outlines the overall flow of data through our system, Figure 2 zooms in on the data schema supporting this flow. It represents how the historical breadcrumb data collected from NAUTO systems in bpx energy vehicles is structured and utilized throughout our benchmarking process.

Each table in the schema corresponds to a specific role in the data processing workflow. The raw telemetry tables ('Tilden/LA2_NAUTO_TestRoutes' and 'PLX_TestRoutes_Operator1/Operator2') include fields such as latitude, longitude, timestamp, state (moving, parked, or stopped), and identifiers like plate, operator_name, and route_id. This detailed telemetry allows us to reconstruct operator behavior on a near-continuous timeline.

We also integrate supporting geospatial data from 'NAUTO_HIST_Tildn/LA_PADLocations', which was provided to us by the client via a spatial join. This join mapped each breadcrumb point to its corresponding well pad, allowing us to determine when a vehicle was located on a pad during its route. This information is essential for computing wrench time, which, as defined before, is the time an operator is parked or stopped on a pad while performing work.

After ingesting and cleaning the raw GPS data, we aggregate it into a derived summary table ('Tilden2/LA3_distances'). This summary contains daily-level metrics such as total time moving, total time stopped, time parked, and total distance traveled per vehicle per day. Start and end GPS coordinates are also retained for evaluating route paths.

The simplicity and normalization of this schema support modular and efficient querying in our Jupyter Notebook environment. By breaking the data into distinct, interrelated tables, we can isolate and analyze specific behaviors (e.g., time spent commuting vs. time on-site) and benchmark operator performance over time. Additionally, this schema facilitates correlation analysis between vehicle behavior and well locations, enabling insights into route efficiency, adherence to proposed PLX paths, and time utilization across operational zones.

This layered, structured approach makes our system both flexible and scalable, allowing for easy updates when new Post-PLX data becomes available or when new fields (e.g., fuel usage, ticket counts) are added in the future.

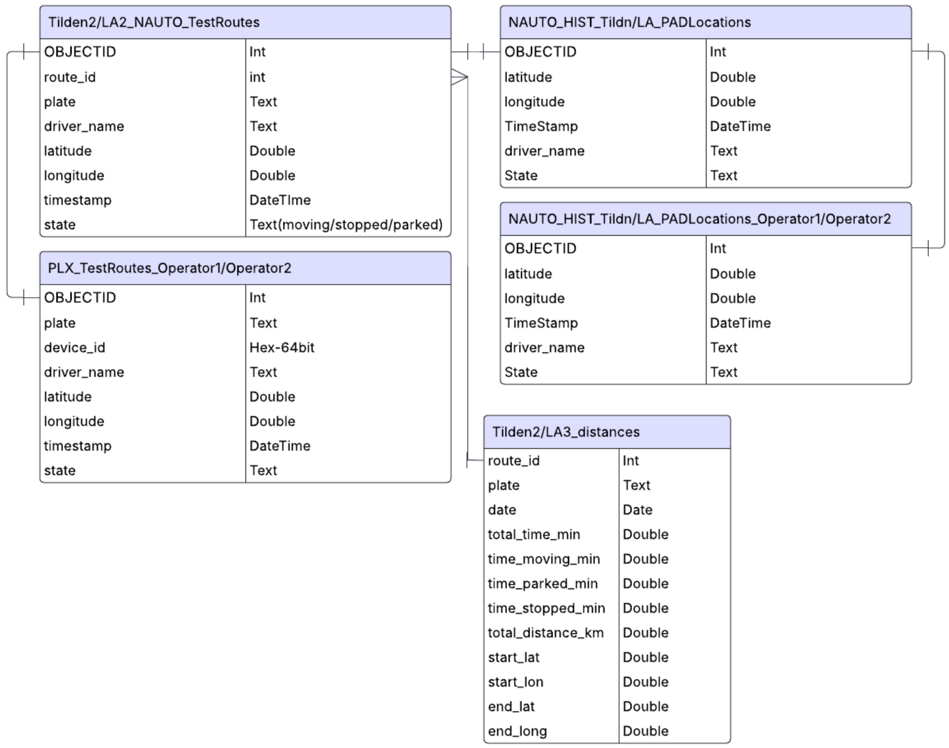


Figure 2 - Data Schema

VII. Software Test and Quality

Due to the project being mainly about data analysis, our software quality plan is based on data integrity, modular design, and usability, in hopes of making the PLX routing analysis tool interpretable and maintainable. The following operations drive our strategy:

1. **Data Validation & Cleaning**
   a. Description: Search systematically through NAUTO breadcrumb data for missing fields, repeated rows, and inconsistent timestamps. Removing columns based on lack of usefulness (empty columns, unrelated to visualizations).
   b. Purpose: Avoiding the pollution of routing metrics and KPIs (drive time, stopping time, parked time, detour count) by bad points of interest data.
   c. Tool: Jupyter Notebook, Python, pandas
   d. Threshold for Acceptability: No duplicate data, data can be parsed correctly for benchmarks, and no large errors in visualizations.
   e. Edge Cases: Data with entirely new rows not in the current operator dataset.
   f. Result of testing: Duplicate and null data removed.
2. **Nonspecific Data Checks**
   a. Description: The code is able to produce correct graphs regardless of operator data, showing that the Jupyter Notebook is made with potentially any operator in mind, rather than being tailored to one operator.
   b. Purpose: To allow easy send off of code, as new Post-PLX data will be acquired and must be put through the same checks to accurately measure the difference in times and values.
   c. Tool: Jupyter Notebook, Python, pandas
   d. Threshold for Acceptability: Graphs do not error out, and the data is in the correct format.
   e. Edge Cases: Post-PLX data may have unforeseen issues as we don’t have it, and thus can’t test it.
   f. Result of testing (for the future): Graphs and data utilizing different Pre-PLX and Post-PLX operator data.

3. **Manual Trace Review**
    a. Description: Visual inspection of GPS traces is conducted to verify the logic behind flagged events such as excessive idling or backtracking.
    b. Purpose: Provides human oversight to avoid false classifications and validate algorithm output.
    c. Tool: ArcGIS Pro, Jupyter Notebook (matplotlib plots)
    d. Threshold for Acceptability: Routes must visually follow known roads or expected paths.
    e. Edge Cases: GPS points slightly off known roads due to signal drift, operators deviating off course due to other work-related matters.
    f. Result of testing: Found redundant and mismatched pad data that needed to be resolved in the code.
4. **Code Documentation**
    a. Description: The Jupyter notebook includes clearly marked sections with markdown explanations and comments throughout Python code.
    b. Purpose: Enhances clarity for stakeholders, future users, and ensures reproducibility.
    c. Tool: Jupyter Notebook, GitHub
    d. Threshold for Acceptability: Code is still clearly commented and easily readable
    e. Edge Cases: Missing imports, cells depending on previous outputs without comments, and unrecognized errors.
    f. Result of testing: Code is readable and easy to understand.
5. **Output Logging**
    a. Description: Intermediate overviews are printed (e.g., stopping times, matched stops) for each level of processing for inspection.
    b. Purpose: Focusing on how you can validate logic throughout the pipeline and ensure data soundness.
    c. Tool: Jupyter Cells, Python (print)
    d. Threshold for Acceptability: Logs show correct and interpretable summaries (e.g, stop counts, durations)
    e. Edge Cases: Long runs with no outputs, mislabeled fields.
    f. Result of testing (for the future): Print caught logic errors; effective for quick validation.

VIII. Project Ethical Considerations

The ethical considerations of this work are especially pertinent because of the application of real-world GPS traces in our project.

1. Privacy & Data Protection - **ACM 1.6: Respect Privacy**
    a. No direct personal identifiers (e.g., names or contact information) are used in our analysis.
    b. Sensitive timestamps and locations are generalized in the case of being rendered in visualizations.
    c. Does not connect with any outside sources, so no risk of data leakage
        i. Only local data is used in generating visualizations
    d. Since data is collected from NAUTO software installed on operator vehicles, informed consent on information on operators is implied.
        i. The dataset is proprietary and private, meaning outside sources can not gain information on bpx energy operators
    e. **Impact**: Complies with ACM/IEEE conditions for user privacy.
2. Open and Transparent Methods - **ACM 1.3: Honesty and Trustworthiness**
    a. All transformation steps are visible in the Jupyter notebook with clear markdown blocks.
        i. Clear what NAUTO data is inputted.
    b. **Impact**: It allows replication and confidence in our results.
3. Human-Centered Framing - **ACM 3.3: Manage Personnel and Resources to Enhance Quality of Work**
    a. This is not surveillance so much as optimization, making field operations more efficient and cutting down on administrative burdens.

b. Our benchmarks are produced with the goal of not only speeding up times but also providing easy-to-read information for whoever reads them; the graphs produced are understandable to people just joining the project.
c. **Impact**: Contributing to the responsible use of AI consistent with the ethical values of beneficence and fairness.

IX. Project Completion Status

1. Unimplemented features
    a. Grouping pads based on location
        i. Due to some pads having exceedingly large wrench times, our team was planning on normalizing times based on the number of oil wells for each designated pad (ex. STS C 16 PAD). This way, pads with 4+ wellheads would be normalized with pads with only 1 or 2 wellheads, allowing us to gain insight into possible outliers.
    b. Post-PLX data in visualizations
        i. PLX is still in the early stages of release, so the dataset is not large enough to be meaningful for comparison against benchmarks.
    c. Opportunities for Enhanced Correlation and Analysis
        i. Due to the limited timeframe of the 5-week session, we were constrained by both time and the availability of clean, structured, and relational data. This limited our ability to perform a deeper correlation analysis. For example, we were unable to explore potentially meaningful relationships such as the correlation between wrench time and the number of tickets or issues reported per pad. Incorporating such data in future iterations could provide more insightful and actionable benchmarks.
2. Performance test results
    a. Generated graphs to identify anomalous variables such as total time stopped, time moving, time parked, wrench time, and compute time.
        i. This helped detect inaccurate data and guided follow-up requests for improved data from the client.
    b. Variables like wrench time and administrative time were leveraged in multiple ways to develop a more robust and meaningful benchmark from the dataset.
        i. This approach helped uncover actionable insights for the client, highlighting how the data could be utilized to inform decision-making once PLX is launched.
    c. Used outlier analysis to identify areas for data quality improvement and potential enhancements in variable accuracy.
3. Summary of testing
    a. Used different operator data to ensure the Jupyter notebook was modular.
    b. Redacted operator information within graphs to preserve privacy.
    c. Made sure that the code was reusable for different datasets that contained the same schema.
    d. Added descriptive summary of graphs and expectations of what would happen Post-PLX.
4. Delivered as a Jupyter Notebook file to the client.
5. Results of usability tests

Figure 3 visualizes how long every operator within the pad areas spent moving, parking, stopping, and the time overall spent. With left and right averages, we are able to pick a date and see whether the average time in each of those driving states increases or decreases. This is meant to be used in tandem with the PLX launch date to see the overall impact of PLX on how long operators spend moving/parked/stopped.

With PLX, time parked should increase, as that implies the operators are spending more time working on a pad and increasing wrench/administrative time. Time moving should decrease, implying operators are getting to their associated pads quicker.

Figure 3 also allows us to see which days have more workers, visualizing peak and trough times for the amount of work throughout the year.

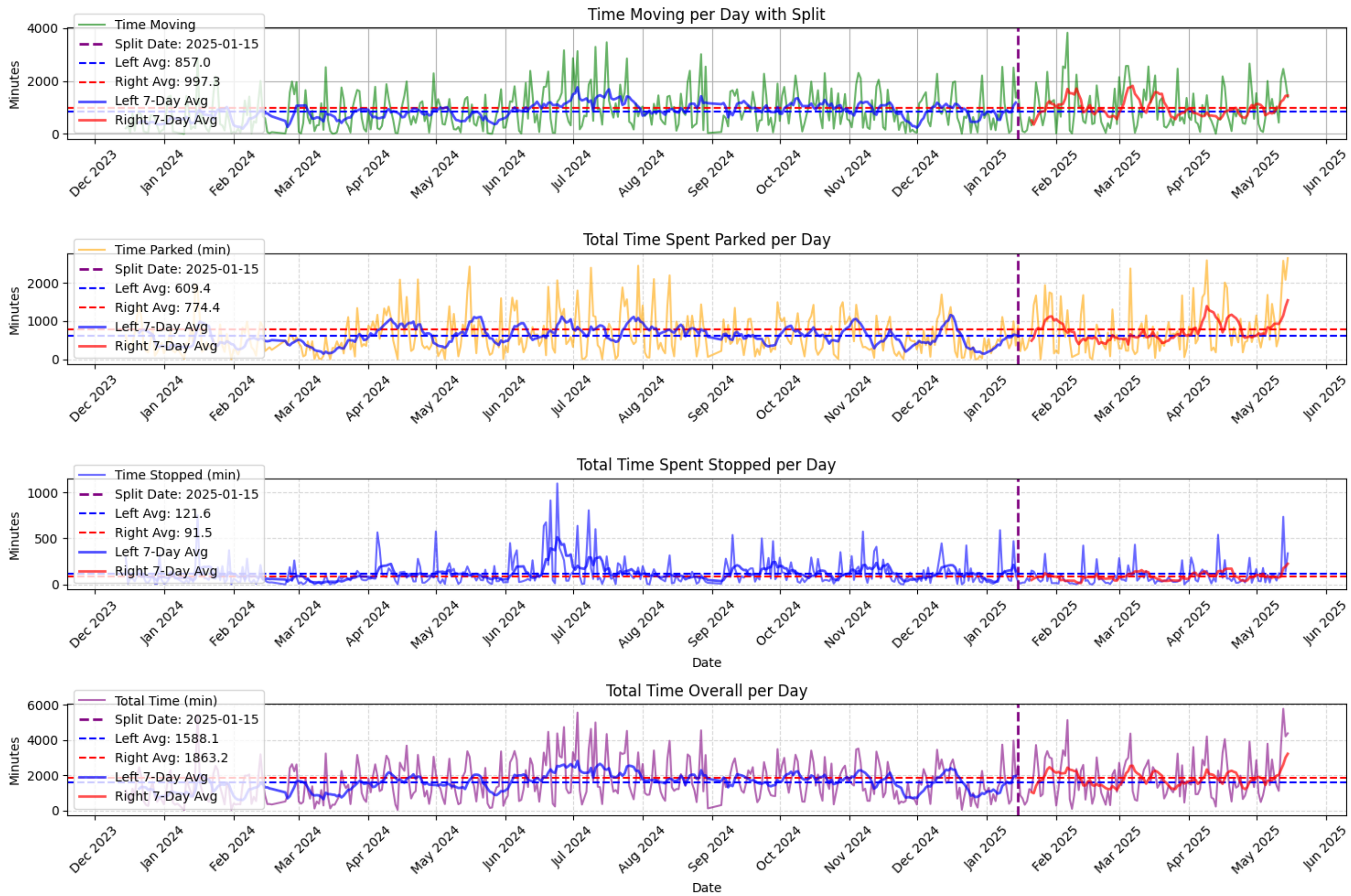


Figure 3 - Time Metrics Graphs From Data

Figure 4 shows the difference in commute times between operators 1 and 2. Our analysis suggests that the commute to the office is far less than the commute to the pad, except for December 2023. This outlier in commute time is most likely due to operations starting and pads being without a need for repair. If PLX is added to these routes, overall commute time would decrease or stay the same, as operators would be driving along the most optimal paths to get to work. There is also insight on what operators did in each of the months, and whether they were heavily focused on administrative time, driving to pad, or a mixture of both.

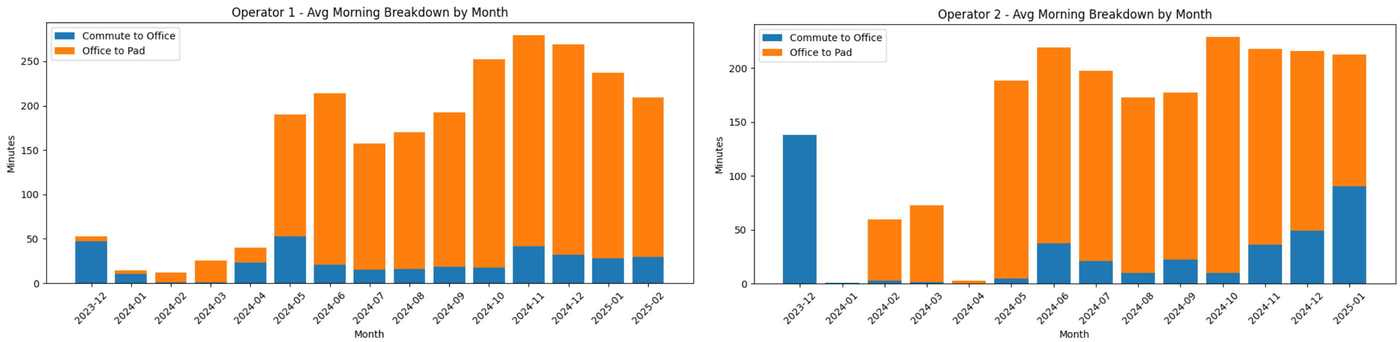


Figure 4 - Commute Time Graphs

Figure 5 allows us insight into how long each specific operator spends at their designated pads parked, equating to wrench time. With direct communication from the operator, it was found that the pad operator 1 spent the most time at STS C 16 PAD, was their highest priority pad. With months spent on each pad listed, a timeline of pad priority can be visualized, allowing us insight into priority changes over time. Once PLX is launched, the goal is to have the overall time spent at each of these pads increase, as that means operators are spending less time in transit and more time working on pad. Priority is also useful here, as this can be put into PLX as priority for which pad the operator needs to hit on their route, influencing the possible routes taken in order to shortcut to the highest priority pads.

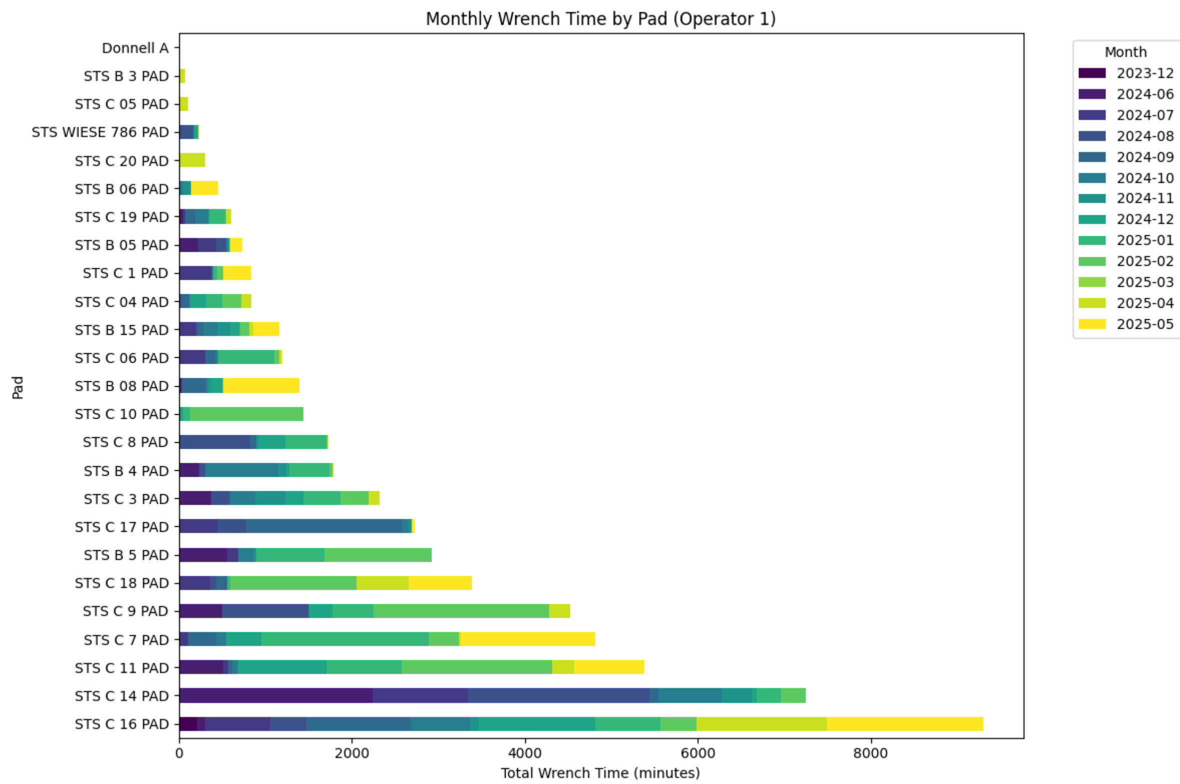


Figure 5 - Wrench Time Operator 1 Graph

Figure 6 depicts data from all operators in LA 3, we spot that the STS C 16 Pad has an abnormally high amount of wrench time, totalling around 167 work days. This could have two reasons, one being NAUTO having incorrect data, or STS C 16 PAD being highest on priority and necessitating year-round work shifts.

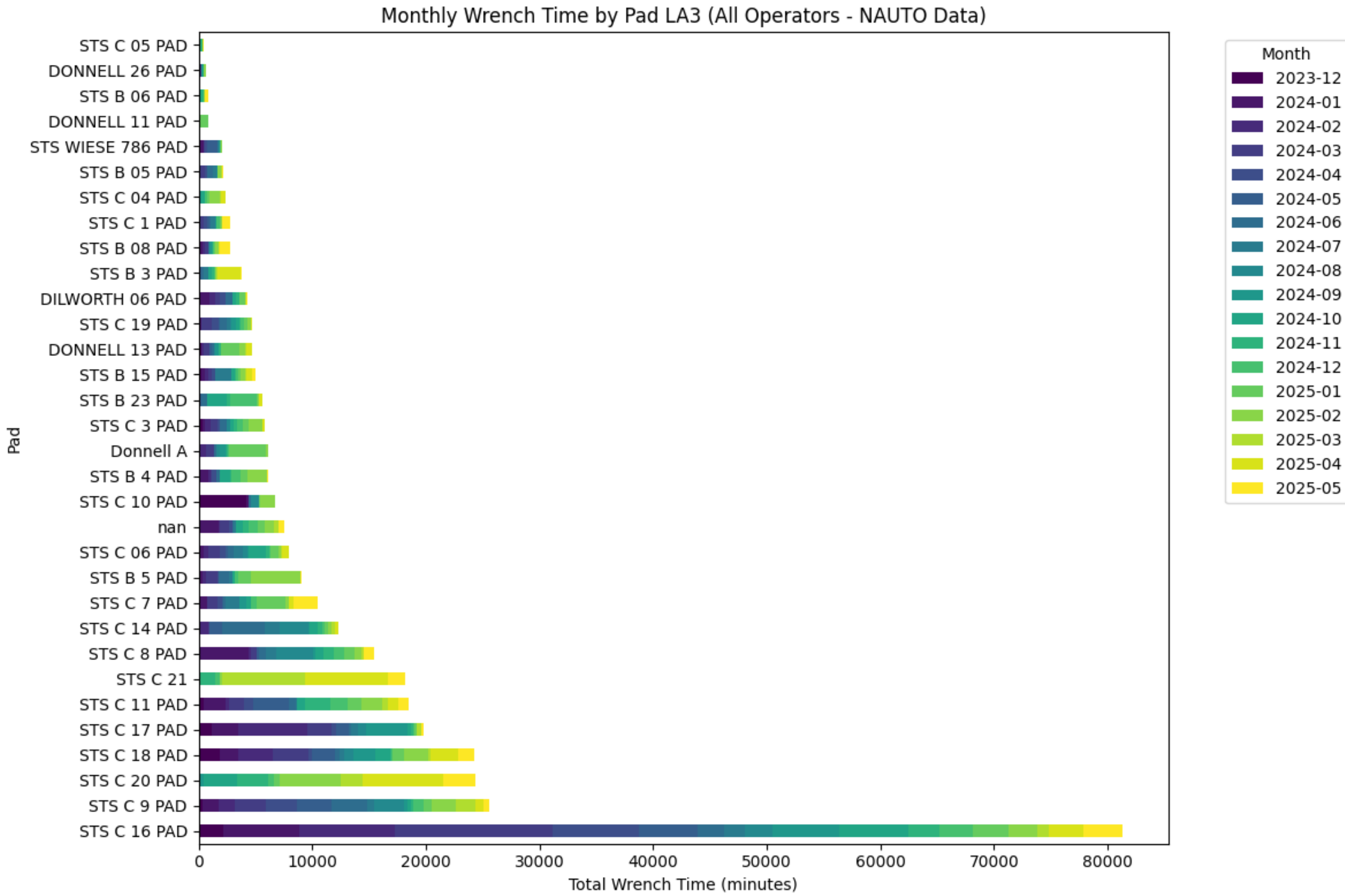


Figure 6 - Wrench Time LA3 Graph

X. Future Work

1. Compare Pre-PLX Benchmarks to Post-PLX Performance
   a. Analyze driving behavior metrics before and after PLX implementation to evaluate improvements in route efficiency, reduced travel time, and operational consistency. This includes metrics such as time moving, total distance, and frequency of stops.
2. Expand PLX Implementation Across Additional Areas
   a. Identify regions where PLX has not yet been deployed and recommend expanding its coverage. Assess whether similar performance improvements can be achieved in these new areas based on observed trends.
3. Collect Additional Pre-PLX Data for Broader Benchmarking
   a. Gather more historical (Pre-PLX) driving data from different regions and operators to build a more representative and robust baseline for benchmarking. This ensures fair comparisons and highlights PLX's overall effectiveness
4. Explore Deeper Correlations Within the Dataset
   a. Conduct advanced analysis such as normalizing wrench time by the number of well heads per pad, or analyzing travel time relative to ticket volume, to uncover hidden patterns and refine benchmarks. There is a long list of possible correlations for each graph in our Jupyter Notebook. Some possible correlations for our main three metrics (wrench time, commute time and admin time) include:
      i. Wrench Time vs. Pad Complexity or Size
         1. Pads like STS C 16 and STS C 14 consistently show high wrench times, which may correlate with larger or more complex pad setups.
      ii. Commute Time vs. Total Wrench Time

            1. High commute or office-to-pad durations may limit time available for productive wrench work.
        iii. Admin Time vs. Wrench Time
            1. Increased admin time may detract from time available for productive pad work.

5. Highlight Operational Improvements
    a. Quantify and showcase reductions in route detours, total route length, unnecessary backtracking, and idle time. These metrics can demonstrate tangible improvements in logistical planning and driving behavior post-PLX.
6. Recommend Spatial Data Enhancements for PLX Optimization
    a. Identify and correct inaccuracies in the geospatial datasets (e.g., pad center points or access road geometry) that affect PLX's routing performance. Improved spatial accuracy can help refine the routing algorithm and user experience.
7. Monitor Adherence to Suggested Routes and Stop Sequences
    a. Measure how closely operators follow the PLX-recommended routes and stop sequences. Analyze deviation frequency, potential justifications (e.g., blocked roads), and areas where compliance could be improved.
8. Track and Analyze Unplanned Detours
    a. Determine how often operators deviate from NAUTO/PLX-suggested routes, the extent of such deviations, and their impact on time and efficiency. This can help improve route trustworthiness and identify routing gaps.

XI. Lessons Learned

We learned about the need for structured validation of data, modular design of code, and clear communication with stakeholders. Real GPS data is noisy and needs to be carefully preprocessed. We also learned about the importance of logging intermediate steps and visual inspection to identify logic errors early. Working as a team prepared us that versioning and documentation are key whenever working with large datasets and GIS tools.

One lesson learned is to push for all data upfront, because the lack of access had a major effect on timing and process. We did not have full GPS data for the first two weeks, which hampered our progress until the meeting with our client, leaving some days where we did not have any idea of how to continue, and wasting time that could have been spent on the project.

We also learned the ethical importance of protecting privacy. By anonymizing outputs and avoiding exposure of sensitive operator or location information, we ensured that our work upheld responsible data handling practices.

XII. Acknowledgments

We would like to acknowledge Scott Jensen, our project advisor, for providing us with great direction and guidance throughout the span of the project and for connecting us with resources at Mines. Though the SQL server ultimately did not pan out, his willingness to intervene if the other side was late/unresponsive demonstrated his commitment to our success and helped keep us moving forward during uncertain stages of development.

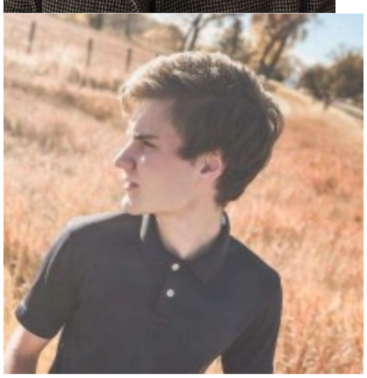
We would also like to acknowledge the bpx energy team: Ayush Rastogi, Sarah Fakult, Alina Shemetova, et al., for being in constant communication and fully willing to support the project's needs. Their support was especially helpful during the initial stages of the project, where we did not have full access to the requisite resources and were unfamiliar with ArcGIS and OneMap.

XIII. Team Profile

Members:

Karl Eisenbarth
Junior
Computer Scientist + Data Science
Hometown: Denver, Colorado
Work Experience: Web Developer at NVOK
Hobbies: Taekwondo, Art, Game Development

Alex Oh
Sophomore
Computer Scientist + Research Honors
Hometown: Castle Rock, Colorado
Work Experience: Web developer at Colorado School of Mines HASS
Hobbies: Art, Writing, Game



Shane Ritter
Senior
Computer Scientist + Data Science
Hometown: Golden, Colorado
Work Experience: lifeguarding, Warehouse Associate
Hobbies: Surfing, Skiing, Video Games

Ethan Tran
Senior
Computer Science + Robotics and Intelligent Systems
Hometown: Denver, Colorado
Work Experience: Data Annotator at Outlier AI
Hobbies: Writing

References

[1] “bpx energy complete modernization of U.S. operations,” BP Newsroom, May 4, 2021. [Online]. Available: https://www.bp.com/en_us/united-states/home/news/press-releases/bpx-energy-completes-modernization.html

[2] “bpx energy - OneMap,” OneMap. (accessed from May 16th to June 8th, 2025)

Appendix A – Key Terms

Includes descriptions of technical terms, abbreviations, and acronyms

Term	Definition
PLX (Perfect Lap Execute)	*An AI-powered voice assistant developed by bpx energy to optimize field operations, automate data entry, and provide routing insights for operators.*
NAUTO Breadcrumb Data	*Historical GPS tracking data collected from field vehicles, including timestamped locations, speeds, and states (e.g., Parked, Driving), used to analyze past operator routes.*
Benchmarking	*Establishing a standard or baseline (in this case, using historical NAUTO data) to compare future routing behavior after the implementation of PLX.*
KPI (Key Performance Indicator)	*A quantifiable metric used to measure the effectiveness of routing behavior, such as drive time, stop efficiency, and route adherence.*
Esri StreetMap Premium	*A commercial routing dataset used by PLX for generating optimized driving routes. It includes enhanced data like speed limits, road types, and access restrictions.*
GIS (Geographic Information System)	*A system for capturing, storing, analyzing, and visualizing spatial or geographic data. Used in this project to interpret and optimize operator movement patterns.*
Wrench Time	*Time spent parked/stopped on an oil well pad*
Administrative Time	*Time spent parked/stopped at a bpx energy office/outside of a well pad*
Tilden2 / LA3	*CSV files containing NAUTO route data used in benchmarks*
Designated Pad (i.e STS C 16)	*Specific well pad taken from Tilden/LA3 data that operators work at*

Appendix B – Jupyter Notebook Use

Instructions on using the Jupyter Notebook

Overview

- Jupyter Notebook
- Python ver 3
- Uses CSVs and XLS for accessing data
- Pandas
- matplotlib for graphs

Pip Installs

These have the necessary Python packages that need to be installed to run the notebook.

Data Locations

'area_location' is used to insert total NAUTO breadcrumb data for an entire pad area (i.e., Tilden2 or LA3).

'wrench_time_LA3_PADS' and 'wrench_time_Tilden_PADS' is based on a spatial join between test routes and pad locations off ArcGIS.

This data is found in the Original_PLXData folder. This can be changed in the future to be the same operator's Pre-PLX and Post-PLX data.

For the graph concerning operator moving/stopping/parking over time, there is a field to change the date of left and right averages; change this to the date PLX launches for correct benchmarking.

If data needs to be changed, it is easiest to replace the location strings in this block. Or copy and paste the code with the CSV location changed. Then compare the graphs from there.