Thematic Analysis of Causal Factors of U.S. Derailments

2025.11.10 / Philip Situmorang

I. Introduction

Risk analyses commonly aim to describe how factors influence the frequency or severity of a risk outcome. In the context of traffic safety, they are normally conducted to understand how various risk factors influence safety outcomes such as crashes, which informs countermeasures in the form of road/infrastructure design, vehicle design, policy, or driver/operator education. In the United States, the National Transit Database (NTD) captures variables associated with transit agencies that may indicate how factors associated with each agency may indicate transit safety outcomes. This analysis offers a discussion on methodologies suitable for its analysis, an exploration into the dataset, and a thematic analysis into causal factors of train derailments among U.S. transit agencies as a potential springboard for future studies aimed at improving transit safety in the United States.

II. Methodologies

At the core of statistical methodologies into evaluating causal relationships is the measurement of correlations between factor variables (X's) and an outcome variable (Y). Some notable examples in traffic safety are the associations between icy road conditions with higher crash frequencies, and the lack of seatbelt use with higher injury severities. While the strength of correlations between factors and outcomes does not *prove* a causal relationship, the scientific community today regularly evaluates them as its potential signal.

How much this correlation compels as indicator for causality hinges on aspects relating to data quality and analysis methodology. Optimizing data quality is generally achieved through measures directed at minimizing bias (i.e. randomization), capturing and controlling possible confounding factors, achieving a sample size of sufficient statistical power, and protecting data integrity. Analyses are then conducted following standards established either through regulation or convention. For example, randomized control trials (RCT) and A/B testing are standard in pharmaceutical and technology spaces, respectively.

When data collection constraints prohibit experimental methodologies such as RCT and A/B testing, the difference-in-differences (DID) method is commonly employed to mimic the methods using observational readily available data[1]. This analysis began with an exploration into employment of this using transit agency-submitted data available in NTD. Specifically, it began by exhaustively considering each column available in the tables of NTD Datamart schema for causal factors to evaluate (treatment variable, in DID terms) and frequency/severity indicators which may be used as outcome variables (Y) against which the treatment variables (X) may be measured.

III. Data

This analysis considered 139,056 submissions of unique major incidents since 2008 (incidents submitted using S&S 40 instead of S&S 50) found in the 2025 NTD Datamart snapshot. The dataset contains variables which indicate frequency or severity of safety event outcomes among transit agencies (the target or dependent variable), as measured by the number of occurrences of a particular event type or the property damage/fatality resulting from the incidents. Collisions make up most of incidents reported (105,069 records of rail and non-rail collisions), followed by assaults and robberies (20,881 cases). Fires make up the next largest category (1,705 cases), followed by derailments (1,656 cases).

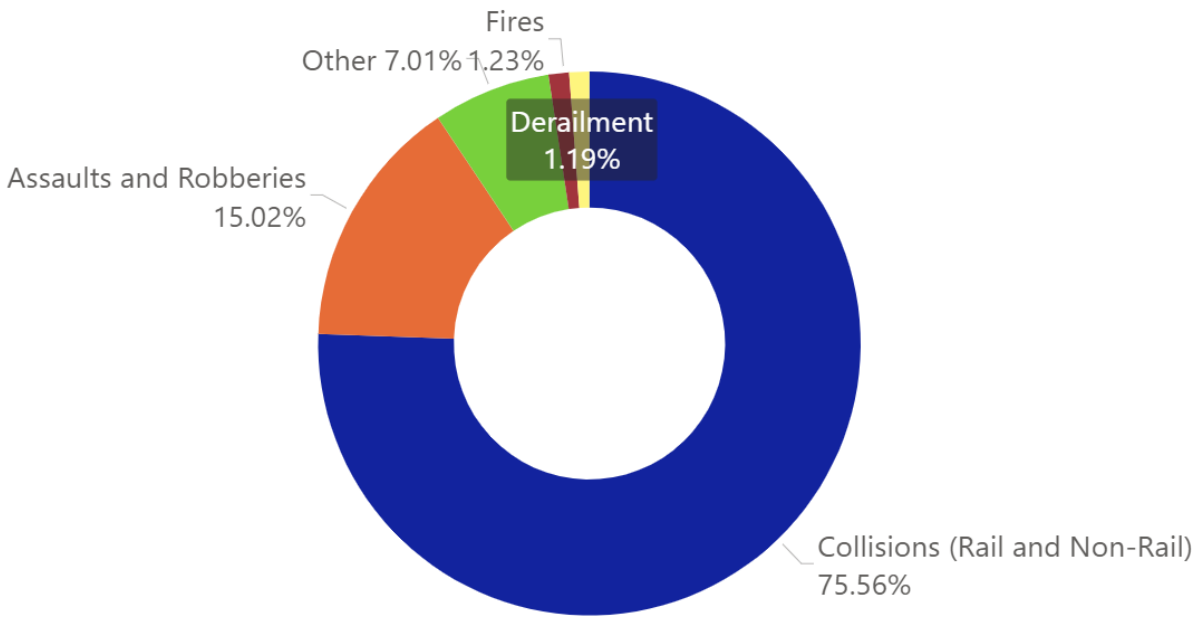

Figure 1. Breakdown of nearly 140,000 major incidents by type in NTD.

The frequency and severity of these incidents serve as potential safety outcome variable by which causal factors may be measured against. This analysis identified in NTD such potential factors in variables relating to an agency's capital allocation, employees, and vehicles. For example, the dataset contains agency attributes such as agencies' maintenance hours and the number of full-time/part-time employees each year, which may serve as indicators for a potential safety outcome such as the number of mechanical breakdowns agencies experience. However, several issues exist which render the difference-in-differences an unsuitable method for causal inference in this analysis.

#1 Lack of plausible link between available factor and outcome variables in the NTD dataset

This analysis finds little to no plausible links between factor and outcome variables available in the NTD dataset. The incidents against which factor variables may be measured against consist primarily of collisions, which are inextricably linked to incident-specific human factor variables not captured in the NTD. To quantify the contribution effect of indirect variables such as an agency's number of maintenance hours or employees towards the number of collisions an agency experiences in a particular year provides little potential to inform countermeasures that can reduce the number of collisions (this would have been different if the outcome variable by which maintenance hours or number of employees are measured against is more suitable, such as the frequency or severity of vehicle breakdowns instead of collisions).

#2 Data Underreporting

The lack of data verification procedures in the NTD induces data integrity issues such as incidents underreporting or missing values, which leads to misleading conclusions or bias against certain agencies based on their reporting practices.

IV. Thematic Analysis

The underreporting of incidents (either through lack of incident submission or underestimation of the severity of incidents) distorts the outcome variable in treatment-control methods such as difference-in-differences, where causal effect is evaluated by measuring the difference in outcomes between the treatment and control groups. In the presence of such limitations, the employment of thematic analyses[2] may be applied to extract insight out of available descriptive data and inform future research questions and data collection decisions which enable treatment-control methodologies. Thematic analysis is a flexible, qualitative method which allows inference of causal effects by discovering themes among incidents, specifically by looking at each description of incidents found in the NTD.

This analysis offers a thematic analysis into derailments among U.S. transit agencies as a sub-study of the entire NTD incidents profile, to illustrate one way of extracting insight in the presence of data issues which render methods such as difference-in-differences inapplicable. Additionally, it offers an insight into the nature of derailments in the United States, which may be used to inform future research questions and the capturing of variables needed to apply statistical methods in answering those questions.

V. Analysis

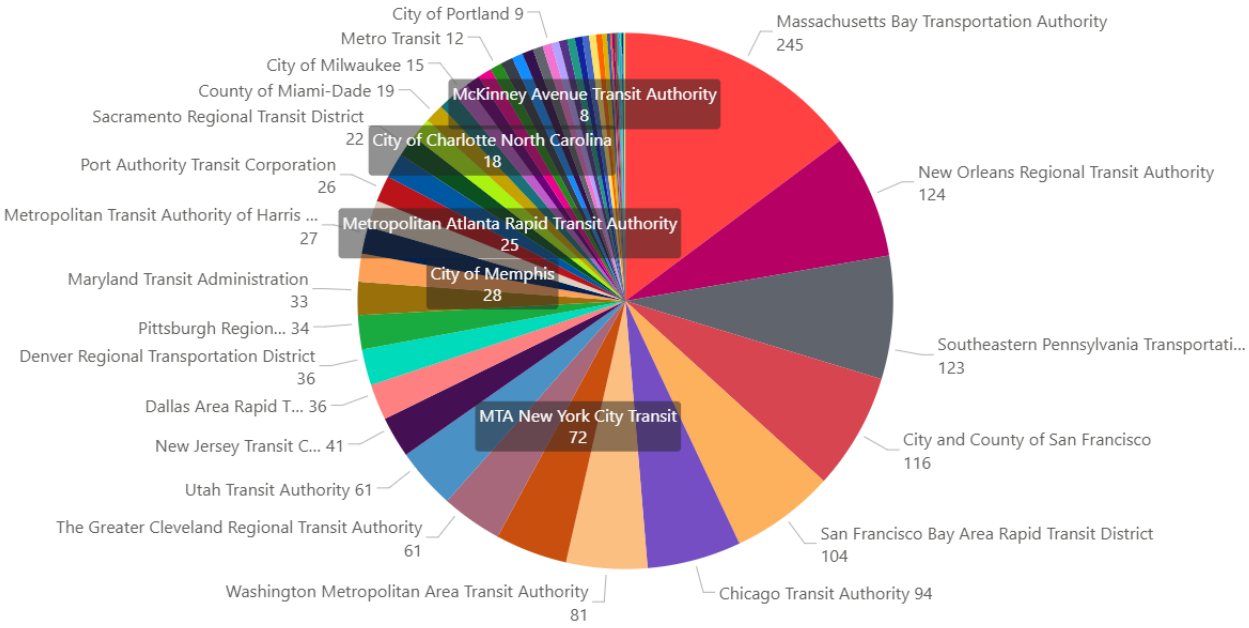

Figure 2. Number of derailments reported since 2008, per agency.

The 2025 NTD Datamart snapshot captures 1,656 cases of derailments since 2008. The application of thematic analysis qualitatively investigates each case description and searches for a *root cause* judgment, as guided by the Train Accident Cause Codes implemented by the Federal Railway Administration. This categorization is based on the submitting agency's explicit attribution of the cause of the derailment and makes no assumptions on cases in which no root cause attribution is made.

Theme 1 – (119 cases) Failure to Inspect Switch/Observe Signal, FRA Code 3

The analysis identified 119 cases where derailment is directly attributed to operator error of failing to inspect switch prior to proceeding over a switch or observing signal prohibiting such procession. The error leads to train car(s) derailing after proceeding over a "gapped" switch or a switch not properly aligned for the operator's intended route. In some cases, derailment occurs during the train operator's attempt to reverse over an improperly aligned switch.

Example Cases

Incident SK	Date	Description
54027	9/26/2014	Operator failed to observe switch points and train signal following a streetcar vehicle entering the yard minutes ahead of him. Both wheel trucks of the streetcar derailed.
98	11/16/2009	The operator of train 14, comprised of a 2 LRV consist, lead LRV 111 and 118 was pulling out of the Light Rail yard for northbound revenue service. The controller noted that the LRV was traveling in the wrong direction and the operator was requested to do a reverse run back into the yard, provided the switches were ok. While performing a reverse run, the LRV derailed from the tracks and the right side of LRV 111 and 118 made contact with a cement pillar and guard railing. The probable cause of the accident and related damage was operating the train through a signal indicating Stop. The error was then compounded by not continuing all the way through the switch, and checking switch alignment, prior to performing a reverse move over the switch.
64798	8/16/2015	At 11:51pm on August 16, 2015, a SEPTA Overhead Power Maintenance Employee was piloting Hyrail Truck #8361 east on the eastbound track in the trolley tunnel (aka green line). The employee came to a stop at a red signal at the LA4 switch. The employee waited for the signal to change and when it didn't change to green the employee proceeded to ignore the signal and proceed forward. The switch was not lined for the route, causing the vehicle to derail. There was no damage to the equipment nor to the switch and no injuries were reported.
109532	1/9/2019	The streetcar operator failed to perform his inspection stop upon his approach to the switch at the above location. Instead, he simply proceeded forward, and as he did so, the "front trucks" came off the tracks, and the streetcar derailed.
44315	8/22/2023	ON AUGUST 22, 2013, AT THE ASHLAND/63RD GREEN LINE RAIL STATION (THE STATION IS NOW TEMPORARILY A RED LINE STATION DUE TO THE REHABILITATION OF THE DAN RYAN RAIL STATIONS) AN EIGHT-CAR TRAIN WITH LEAD CAR #3185 THAT HAD BEEN ROLLING SB DERAILED. A TRUCK ON THE FOURTH CAR, 2878, DERAILED WHEN THE SUPERVISOR TRIED TO PULL THE TRAIN BACK AFTER THE RAIL OPERATOR MOVED THE TRAIN THROUGH A DOUBLE RED SIGNAL OVER THE TRIP AND THE SWITCH. A PRELIMINARY INVESTIGATION INDICATED THE SWITCH WAS SPLIT WHEN THE RAIL OPERATOR INITIALLY MOVED THE TRAIN THROUGH THE SWITCH.

Theme 2 - (58 cases) Ice/Snow or Foreign Objects Impeding Train Movement, FRA Code 5

The analysis observed 58 cases where objects such as ice/snow or foreign objects not originally part of the derailed train obstructed regular movement over track, which led to a wheel climb or prevented a switch mechanism from fully closing gaps. Thirty cases of derailment were attributable to ice/snow accumulation at a track site, while the remaining twenty-eight involved debris or foreign objects.

Example Cases

Incident SK	Date	Description
1262	2/13/2010	Pilot car derailed due to ice in track
1326	11/2/2010	An LRV struck a phone book on the track, causing the left wheel to climb and derail the front truck. As evidenced by debris driven between rail joints and onto the rail bonds, the book appears to have been caught between the rail and lead axle. The derai
133124	12/17/2020	Ice build up from a snowstorm caused the train that was moving through inside loop to climb over the the rail and derail. No injuries and train re-railed.
58094	2/23/2015	At approximately 06:26 hours, the Continuous Welded Rail (CWR) Train (consist of 11 cars) was being transferred from Jamaica Yard to Linden Yard. As the train was traversing Track P7 from the Linden Flyover into Linden Yard all four wheels of the # 1 and #2 trucks of car BCR08 derailed to the east approximately 120 feet north of a hand throw switch. The train traveled approximately 48 feet in a derailed state before coming to rest. There were no reported injuries as a result of this incident. Post incident inspection revealed that the lead car (BCR08) rode up on an accumulation of snow and ice and derailed to the east. Division of Car Equipment (DCE) personnel responded and found no visible damage to car equipment components. The train was re-railed at approximately 09:02 hours. Office of System Safety (OSS) has attributed this derailment to inclement weather conditions that existed in the yard at the time of the incident; however, OSS is reviewing winter operations procedures for snow clearing activities. Property damage is estimated to be $2,528.
113921	7/23/2019	As the operator was pulling away from the stand, proceeding through the switch, the streetcar jumped the track and derailed due to a rock lodged in the trackway.
164526	8/5/2023	At approximately 2230, the operator was proceed northbound through the Union Station switch on streetcar 806 in the exclusive right-of-way leading to the tail track. While proceeding through the switch, which received the signal to advance, the streetcar derailed after running over a shirt that had been dragged into the switch. The truck ended approximately 10’ from the switch in exclusive right of way. There were no injuries and ten people were escorted off the vehicle onto the sidewalk. Maintenance visually inspected and verified switch box, signals, truck, and wheels were undamaged.

Theme 3 – (46 Cases) Defective track or switch mechanism, FRA Code 1

The analysis observed 46 cases where derailments occurred due to broken/damaged/deteriorated track (15 cases), defective switch mechanism (17 cases) or other track-related issues (14 cases).

Example Cases

Incident SK	Date	Description
151146	9/21/2022	On Wednesday, September 21, 2022, at approximately 1211hrs, CTJ-500 (High Rail Bucket Truck) was traveling northbound on North Yard Track 11 when it derailed at North Yard Switch 13. CTJ-500 was operated by a maintenance of way (MOW) employee, who had properly set Switch 13 for a straight move remaining on Track 11 prior to the derailment. The probable cause of the derailment of CTJ-500 was Switch 13 being damaged from being trailed previously by an unknown vehicle. The investigation continues.
129044	8/19/2017	1st truck of 4th car of a 4 car consist derailed in yard due to worn switch rails and lack of lubrication.
29441	9/13/2011	The lead car of a two-car LRV train consist travelling south on the Metro Green Revenue Loop while under the control of a Metro Shopman, derailed the rear axle wheels of its leading A-end truck, while moving over a faulty track switch at approximately 3mph. There were no injuries incurred as a result of the derailment.
44942	5/29/2013	On Wednesday, May 29, 2013, at approximately 5:48 p.m., a southbound #1 train (interval 1726) derailed on track# 2, south of the 125th Street Station. The train operator reported to the Rail Control Center that the #1 truck on the head car (#2460) was derailed due to a broken rail, and traveled approximately 255 feet further down the track until it came to a complete stop, near survey marker 71+48. The train had approximately 425 passengers remained onboard until a reach train was brought in end to end for the evacuation. The passengers were then walked back to 125th Street Station. The crew was sent to MAC for incident testing. There were no reported injuries to the passengers or crew.
35662	6/30/2012	Streetcar derailed at the above intersection due to deteriorating crossties which allowed the track gauge to move out of position. The streetcar was placed back on track and gauge was repaired and re-tested before track was re-opened for revenue service.

Theme 4 – (18 Cases) Excessive speed over switch or curve, FRA Code 3

The analysis identified eighteen cases where train operators proceeded over a curve or switch with power which exceeded the allowable speed, leading to derailment.

Example Cases

Incident SK	Date	Description
91497	10/27/2017	As the operator approached the switch at the above location, he proceeded too quickly, and "split the switch", derailing the streetcar.
137156	3/14/2021	At approximately 16:33hrs. on March 14, 2021 Communications System Technician #523 notified OCC that Train 7 had derailed at the 11th St. Loop traveling Eastbound on Washington St. Train 7 was single tracking from Culver Crossover to the 11th St. loop and was set to diverge from Track 2 to Track 1 utilizing switch 7. Video evidence shows switch 7 was properly set for the diverge move, the signal aspect was functioning properly. Video evidence shows Operator #331 passing the flashing yellow signal for the move from Track 2 to Track 1 and entering the switch at more than 30mph, the speed limit for all diverge moves is 10mph (Operator rule #3.12.3).
187998	5/12/2024	On 14:06 on 5/12/2024, SB Utah Transit Authority TRAX Green Line Train 42 with a four car consist (LRVs 1164, 1120, 1114, 1168) heading to West Valley, over accelerated into the turn at 3100 S 2200 W, West Valley City, UT causing the rear car (LRV 1168) to derail. The rear car derailment struck an overhead catenary pole causing significant disabling damage to the rear rail car. No injuries were reported due to the incident. The 16-passengers in the front three train cars were evacuated on to the right of way to a bus bridge.
109005	1/28/2019	Train #25/LRV #316 derailed at grade crossing/street running intersection. Posted speed for curve is 10 MPH; EDR information indicates approximate speed of 30 MPH. One train passenger was ejected from the LRV; the LRV door frame crushed/amputated the passenger's left foot.
89085	7/9/2017	Operator was moving a three-car consist from the north yard to the south yard; the operator trailed a switch successfully with two of the three LRV's while operating at a speed in excess of allowed yard speed. One LRV was derailed.

Theme 5 – (13 Cases) Parts detached from train and impeding movement, FRA Code 4

The analysis observed thirteen cases of mechanical issues which lead to parts detaching from a train causing derailment. In four cases, derailment occurred after wheel/axle completely detached from a train. In nine cases, train part(s) other than wheel/axle detached and obstructed wheel-track contact leading to wheel climb and derailment.

Example Cases

Incident SK	Date	Description
52148	6/17/2014	ON JUNE 17, 2014, THE THIRD CAR OF NB ORANGE LINE TRAIN 725 DERAILED BETWEEN THE ENTRANCE OF SWITCH 7 AND EXIT SWITCH 15 OF THE TOWER 12 INTERLOCKING, AFTER HITTING A BRAKE CALIPER THAT HAD DETACHED FROM THE THIRD CAR OF THE EIGHT-CAR TRAIN. PASSENGERS WERE UNLOADED FROM THE TRAIN TO THE STATION. THERE WERE NO REPORTED INJURIES.
16372	1/22/2010	A broken slack adjuster dropped down from underneath the trolley and caught between trolley wheel and track causing trolley to wheel to raise up and come slightly off of track.
17956	4/6/2010	On April 6, 2010, at approximately 7:45PM, the Operations Control Center received a report from the Train Operator (T/O), on board eastbound Train #58 (2043-2049), of a derailment approximately seventy yards west of Spring Street Station, La Mesa California. After performing a ground inspection the T/O confirmed that the rear wheels of the front bogie of LRV 2049 had derailed. Prior to the incident, the T/O reported hearing an object dragging from the front of the train. The operator was instructed to perform a ground inspection. He reported an unidentifiable object hanging from the undercarriage. The operator was authorized to continue at a reduced speed to Spring Street Station where the train would be taken out of service. While en route to Spring Street Station the operator reported derailing. A subsequent investigation discovered the track brake from the a-end bogie of LRV 2049 had become dislodged and caused the rear wheels of the bogie to lift from the rail. The train was re-railed at approximately 8:57PM. No injuries were reported at the time of the incident.
88918	7/17/2017	Maintenance vehicle Spot Tamper's left wheel broke off as well as all the bolts and derailed just north of the North Quincy crossover. No reported injuries or damages to rail. The sport tamper was re-railed around 4:46 am. No impact to service.
88302	4/20/2017	A ten car consist traversing a switch in the yard derailed due to a damaged current collector being dislodged and falling into the heel of the switch. There were no reported injuries. Weather was not a factor. Damages are pending.

V. Summary

The observational nature of the NTD Datamart snapshot used this thematic analysis limits conclusions on aspects such as the representativeness of the above themes towards the nature of train derailments in the United States. Despite its limitations, analyses on the dataset's agency-submitted incident narratives offer guidance towards the prioritization of future research questions as well as decisions on data collection and analysis methodologies to answer those questions. Specifically, such analyses can be used to prioritize evaluation on the effectiveness of agency-implemented policies to improve safety and drive the design of an analysis framework for that purpose. Future analysis may be extended to events other than derailment to guide future inquiry into safety/security incidents such as collisions.

VI. Citations

[1] Baker, Andrew et.al. "Difference-in-Differences Designs: A Practitioner's Guide." arXiv, 2025, arXiv:2503.13323

[2] Naeem, M., Ozuem, W., Howell, K., & Ranfagni, S. (2023). A Step-by-Step Process of Thematic Analysis to Develop a Conceptual Model in Qualitative Research. *International Journal of Qualitative Methods*, *22*. https://doi.org/10.1177/16094069231205789