Data Analytics Engineering

Summer 2024

USPS Weather Impact Challenge

Anderson Battin
Evans Sarker
Jennifer Carr
Julie Ritenour
Lowell O'Connell
Praveen Kumar

DAEN 690

Project Report

About the Cover

On April 10, 2024, George Mason University proudly inaugurated the Mason Autonomy and Robotics Center (MARC), a cutting-edge, multi-million-dollar facility housed within the College of Engineering and Computing. Strategically situated in Research Hall on the Fairfax campus, MARC boasts a two-tiered aviary dedicated to drone testing, versatile experimental zones, and dynamic spaces designed for both collaboration and individual study.

MARC transcends the traditional boundaries of engineering, serving as a hub for a wide array of colleges to engage in cross-disciplinary initiatives. The center is poised to become a beacon of innovation, offering faculty and students the resources to pursue groundbreaking research and providing unparalleled educational experiences. These efforts are aimed at tackling both local and international challenges in the realms of autonomy, embedded artificial intelligence (AI), and robotics.

Embracing an integrated approach, the center's interdisciplinary activities merge diverse fields such as computer science, electrical and mechanical engineering, systems engineering, psychology, philosophy, and policy. This synergy fosters a robust technology development program, proven and repeatable, that catalyzes collaboration among faculty, students, government entities, and industry leaders.

Highlighting the ceremonial opening, Mason President Dr. Gregory Washington, alongside MARC co-directors Dr. Missy Cummings and Dr. Jesse Kirkpatrick, were assisted by a Boston Dynamics Spot robot in the symbolic ribbon cutting. This momentous event marks the commencement of a new era of technological advancement and educational excellence at George Mason University.

Contents

Table of Contents

Table of Figures

Table of Tables

This page intentionally left blank

Abstract

Abstract

With the increase in severe weather due to climate change, the United States Postal Service (USPS) has seen the on-time percentage scores of mail and packages drop. The USPS is proactively exploring a machine learning solution for predicting weather-related delivery delays so they can better prepare for severe weather events and reduce delays. Previous studies exploring supply chain disruptions caused by extreme weather events used machine learning techniques such as Support Vector Machines and Artificial Neural Networks [1], as well as Bayesian Hierarchical Models [2]. These methods prove that machine learning techniques can yield desirable results when predicting and managing supply chain disruption challenges. We explored multiple machine learning methods to predict the impact of severe weather on mail and package delivery, including classification, regression and clustering. Ultimately, we found that the XGBoost Classifier achieved optimal performance. Using the XGBoost Classifier, we were able to predict if a package was on-time or late with 89.7% accuracy, and if a piece of mail was on time or late with 93.2% accuracy. Feature importance analysis showed that the most influential features for packages included the amount of rain at the destinating facility and the Mail Class code of the packages. The most influential features for mail included the amount of snow at the destinating facility and the max temperature at the originating facility. These results show that various weather factors do have a significant impact in the on-time delivery of USPS mail and packages. The model's accurate forecasting can help USPS allocate resources effectively and update their weather-preparedness plan to ensure more reliable service. This research acts as a framework for future logistics and supply chain management analytics and contributes to the limited research pertaining to supply chain disruptions resulting from weather.

This page intentionally left blank

Report

Section 1: Problem Definition

1.1 Background

According to the National Oceanic and Atmospheric Administration's (NOAA) National Center for Environmental Information (NCEI), 2023 was a record-breaking year for billion-dollar weather and climate disasters. The United States experienced 28 billion-dollar disasters in 2023 with a price tag totaling $92.9 billion [3]. The previous record was 22 billion-dollar disasters in 2020. The map in Figure 1 below is from the NOAA National Center for Environmental Information and provides an overview of the types of billion-dollar weather and climate disasters experienced throughout 2023.

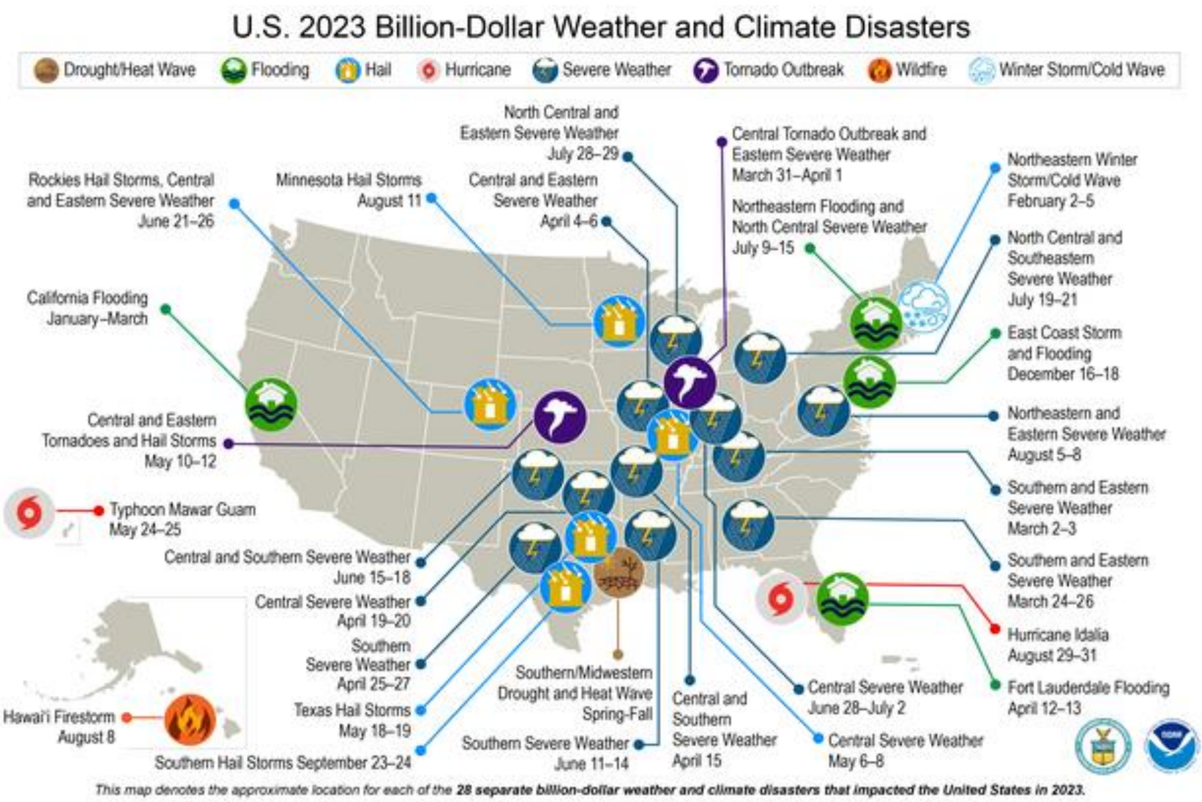

Figure 1: A Map of the 2023 Billion-Dollar Weather and Climate Disasters

Each disaster has a devasting impact on the supply chains within the affected region. Supply Chain Risk Management (SCRM) has become an incredibly important area of focus from both an economic and efficiency perspective across many domains and involves being able to correctly identify, predict, mitigate, and monitor any incoming risks to a supply chain's reliability and performance, two of the most important metrics to consider when analyzing SCRM practices. The risks encountered by a supply chain can fall into three categories organizational, network-related, and environmental.

The United States Postal Service (USPS) has faced issues in all three categories, but there have been significant concerns regarding USPS mail delivery delays due to weather conditions. The United States Postal Service is prioritizing a predictive model for weather-related delivery delays in a step towards adopting proactive risk management strategies.

The increased delays followed USPS's implementation of a new regional distribution system in late 2023. This new distribution system aimed at consolidating mail processing through 60 regional centers to promote efficiency and lower costs to USPS. Instead, it created significant disruptions within the regional centers and the delivery routes, significantly delaying mail and packages delivered by USPS. These operational changes have been looked upon negatively by politicians within both parties. As a result, USPS delays are becoming a growing bipartisan concern for State and National officials since many believe that weather-related environmental challenges need to be properly addressed prior to making any operational changes.

The realignment of facilities has caused mail to travel longer distances before delivery, contributing to delays. Significant delays in the Houston region, particularly involving the USPS processing center in Missouri City, have drawn scrutiny from congressional representatives. Staffing shortages, equipment problems, and poor communication are cited as primary issues for the delay in delivery. In regions like Richmond, Houston, and Atlanta, where the new system was first implemented, delays have been more pronounced. Richmond, for example, saw such severe delays that officials advised residents to use drop boxes instead of mail for voting following the March 2024 audit on the effectiveness of the new regional processing and distribution center in Richmond, VA that was conducted by the USPS Office of Inspector General (OIG) [4]. As demonstrated in Figure 2 below, the year-over-year nationwide average of on-time delivery for first-class mail has been down since September 2023. This trend has remained consistent thru May 2024 [5].

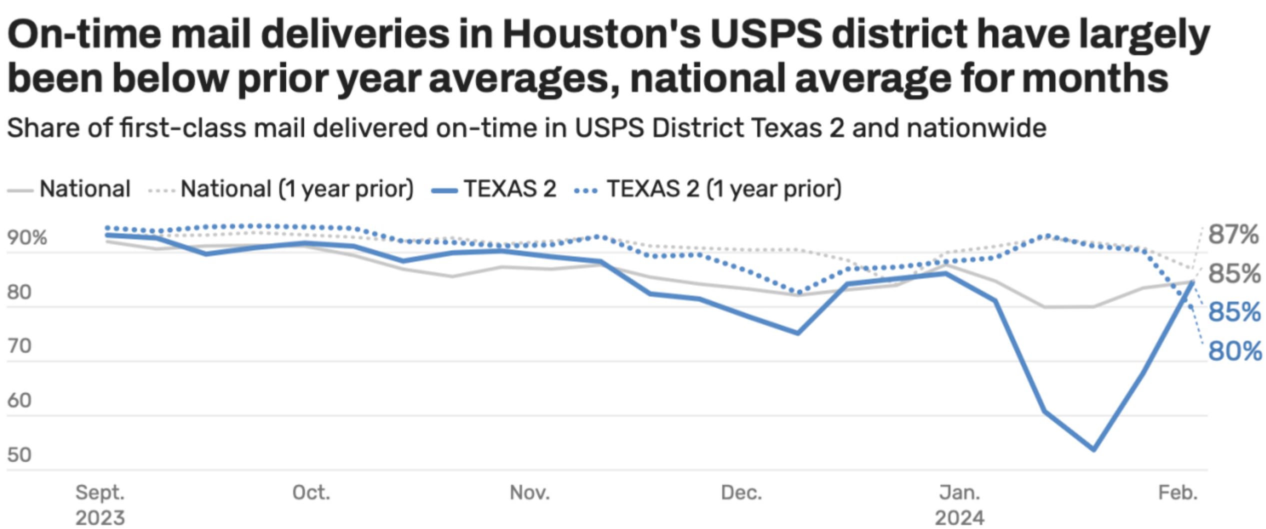

Figure 2: USPS data for Service Performance

January 2024 was the most dramatic decline in the average share of first-class mail delivered on-time when compared to the previous year. District 2 in Texas was particularly affected. On January 20, 2024, the reported share of on-time delivery for first-class mail was 54%, compared to 91% in January 2023. Understanding the reasoning behind these dramatic delays allows USPS to make more informed decisions when planning for changes to their logistics network.

The main contributors behind the substantial delays being faced by the USPS can be broken down into two types: internal USPS operation reforms and severe weather conditions. Adverse weather conditions cannot be prevented. However, the USPS can mitigate some disruptions caused by adverse weather conditions if the conditions are correctly predicted with enough lead-time. Weather-related delivery complications impact all aspects of the USPS operations. As a result, elected officials are pressing the USPS for detailed reports regarding why these delays are being reported at higher rates within their state or district and how the USPS plans to address the increase in delivery delays nationwide.

The USPS has been asked to provide more transparency surrounding their operational changes and provide an explanation for the year-over year trend of rising delivery delays that has plagued the USPS since 2023. Lawmakers from affected regions have demanded accountability and better planning from USPS management. There have been calls for more transparency and adjustments to the new operational strategies to minimize service disruptions. These changes have led to initial setbacks and delays, as seen in areas like Tenessee, Virginia, and Texas [6].

While the USPS has been working with political representatives to address the concerns of their constituents in addition to providing reports regarding weather-related delays, Congress is still convinced that the United States Postal Service is not providing their customers with the quality of service that they promise. This is according to a variety of audits conducted in districts across America that show an increase in complaints to officials regarding USPS delivery delays. The Postal Service has a *Service Delays* website that was created to provide information to the specific region expecting a type of severe weather, but as of May 24, 2024 the specific weather delay notifications in Figure 3 does not provide any detail on what kind of climate or weather event is impacting the areas mentioned nor does USPS provide insight into how long the delays may be [7].

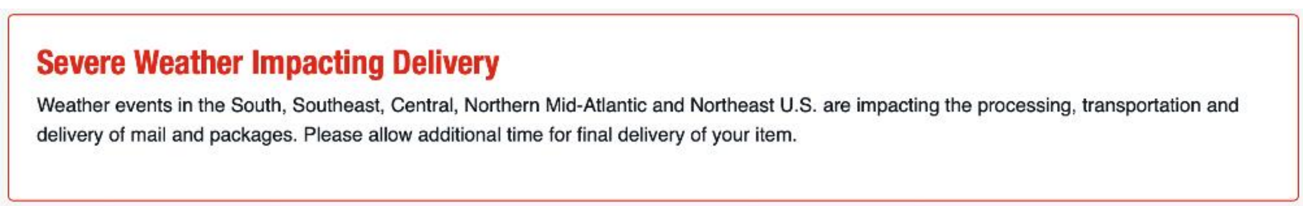

Figure 3: USPS Service Alerts for Weather Delays as of May 31, 2024

The USPS alert from Figure 3 has been active since January, warning residents in states from the South, Southeast, Central, Northern Mid-Atlantic and Northeast United States that there may be weather-related delivery delays. This not only notifies the consumer of a late shipment but allows plants, centers, and distributors to plan accordingly along the entire delivery route from the originating point to the destination point. However, there is no specific date or weather event specified on USPS' *Service Delays* website. The message lacks specific information regarding the type of weather delay and any precise estimate regarding the amount of time mail and packages will be delayed. The message encapsulates the fact that weather-related events are not only impacting the final USPS vehicle route to the destination for mail and packages but also the processing and transportation facilities as well. The current message "please allow for additional time for final delivery of your item" needs to be improved as this does not allow for necessary preparation in cases of retail and other large consumers, as well as those expecting time-sensitive mail or packages [7].

There are four retail and delivery *Postal Areas* defined by USPS: Atlantic, Central, Southern, and Western-Pacific. This is parred down from their previous number of seven *Postal Areas*. Besides these four Postal Areas, the current USPS network also contains four regions of logistical operations and two regions of processing operations, Eastern and Western [8] . These four regional operations areas and the additional logistical and processing regional areas differ in both physical size but also in the number of delivery points, employees, plants and active post offices. In addition, there is more seasonality in certain areas, and some are affected more frequently by severe weather events than others. In Figure 3*,* USPS warns users about weather-related package delays in the South, Southeast, Central, Northern Mid-Atlantic and in the Northeast. While the South and Central regions are USPS defined *Postal Areas,* the Southeast, Northern Mid-Atlantic and Northeast are *divisions* of areas of the USPS network, and these divisions are how USPS further subdivides the United States in a way that optimizes package and mail delivery [9].

Despite USPS not providing the details regarding the weather event(s) that are causing the delays across the operational and transportation board in varying areas of the United States, a reporter sought out the specifics by accessing the weather reports given the timing of Figure 3, concluding the first weather delays were due to cold weather. The USPS states on their website that:

> *Mail delivery service may be delayed or curtailed whenever streets or walkways present hazardous conditions to our carriers and/or vehicles. The Postal Service™ curtails delivery only after careful consideration, and only as a last resort. We appreciate your understanding of our responsibility for the safety of our employees, as well as of our customers* [8].

The second delay is attributed to strong winds and flooding in the Southeast including widespread rain, severe thunderstorms, and flash floods [10]. This impacts not only the ability of the USPS delivery driver to access your mailbox, which is a commonly cited reason for delayed deliveries in icy conditions, it can also lead to a devastating chain event down the line if not properly prepared for and anticipated. For example, if the travel-ending traffic and dangerous roads causes delays for the driver transporting a large volume of mail and packages to a processing center, this will naturally cause the final delivery to be delayed. Furthermore, if the processing center is unprepared to handle the increased volume from the backlogged deliveries, then this excess can generate further delays. As delays in movement through the USPS network continue, a snowballing effect occurs as this generates increasing delays because not only is the initial delay passed on to the distribution and sorting centers, but these locations also face the same potential problem as the other hubs, as they may lack the operational capacity necessary to handle additional volume [11].

The most recent disaster that USPS is citing as a reason for delay is the collapse of the Francis Scott Key Bridge in Maryland, impacting both the Northern Mid-Atlantic and the Northeast with delays in the supply chain faced by all suppliers, and given the United States' increased interest rate the cost of larger inventories and expansion has also risen, making this a major economic problem for exports [12].

USPS delivers to approximately 160 million delivery points across the United States [13]. Weather or climate disasters can delay USPS's ability to deliver a piece of mail or a package on time. The definition of "on-time" can vary depending on the specified service standard for the particular type of mail. Climate and weather disasters overall have been trending upwards, and therefore, USPS will face increasing challenges to their ability to deliver mail or packages within the period determined by their service standards. Delays due to weather or climate disasters may be unavoidable so it is important to manage the customers' expectations when such delays occur. For some, USPS is a delivery service relied upon for necessary items such as medication. One such case was experienced by Irene Ramirez, who told NBC news in an April 2024 interview that, "[H]er 89-year-old father's heart medication from the VA spent more than 18 days stuck in the mail in Houston. She said she even tried calling the White House to get help after spending hours on the phone with the Postal Service, only to be repeatedly disconnected" [14]. To manage the customers' expectations when unavoidable delays occur, USPS needs to be able to predict which delivery points will be impacted by a pending weather or climate disaster and provide advanced notice.

Current material from the USPS website indicates the following mail and package delivery standards in addition to how long a customer should wait to contact USPS customer service with a concern for each mail class. In Figure 4 below, note that the only mail class guaranteed to arrive within the Delivery Standard window is Priority Mail Express. In addition, USPS provides the amount of time to wait to contact customer service following a package delay, and customers with delayed non-Priority Mail Express deliveries must wait a minimum of five days from the date of mailing, as seen in Figure 4. This requires the user to have knowledge of the date of mailing, whereas a delayed non-Priority Mail Express customer should be able to contact customer service the same day that their delivery did not arrive. These USPS service standards guarantee the arrival of a Priority Mail Express delivery within two days, whereas the gap between other mail classes is much more dramatic e.g., for Periodicals, the *Delivery Standard* is between three to nine business days. However, weather events have caused delays, and USPS has not been able to meet their delivery standards as frequently as they promised when they first introduced a change in delivery network operations to cut costs and improve efficiency [15].

MAIL CLASS	DELIVERY STANDARD	CONTACT CUSTOMER SERVICE AFTER
Priority Mail Express®	1-2 calendar days (guaranteed)	6 PM (local time) on the Guaranteed Delivery Date
Priority Mail®	1, 2, or 3 business days (not guaranteed)	5 or more days from the date of mailing
First-Class Mail®	1-5 business days (not guaranteed)	5 or more days from the date of mailing
USPS Ground Advantage™	2-5 business days* (not guaranteed) **If it contains Hazardous Materials / Live Animals* 2-8 business days* (not guaranteed)	5 or more days from the date of mailing **If it contains Hazardous Materials / Live Animals* 14 or more days from the date of mailing
Parcel Select®	2-8 business days* (not guaranteed)	14 or more days from the date of mailing
Package Services: • *Media Mail®* • *Bound Printed Matter* • *Library Mail*	2-8 business days* (not guaranteed)	14 or more days from the date of mailing
Periodicals	3-9 business days* (not guaranteed)	14 or more days from the date of mailing
USPS Marketing Mail®	3-10 business days (not guaranteed)	14 or more days from the date of mailing

Note: *Except Alaska, Hawaii and US Territories - estimate provided by the postage price calculator.

Figure 4: USPS Delivery Standards

The goal of making sure that USPS manages to meet its predefined standards, and in the case of a weather event provide them with the tools and analysis necessary to be able to accurately predict how this will impact the USPS logistics network's ability is the first step towards improving their weather-related *Service Alert* reliability and usefulness. The ability to understand what types of severe weather events impact mail and package delivery throughout the stops along the USPS delivery route, for example at the origin point versus at a processing or transfer center depending on the type of mail or package, which as specified in Figure 4 is relative to the USPS delivery standard, allows for greater insight into the impact of each weather event. This analysis would provide USPS with ideally, a way to improve their current understanding of how weather events impact their ability to provide mail and delivery services and accurately determine by how much time and allow customers to be more proactive.

1.2 Problem Space

Leaders at USPS aim to increase knowledge about how weather events impact mail and package processing, transportation and delivery by building out an AI/ML model. The USPS has partnered with George Mason University students to help build out such a model that will use publicly acquired weather event data and USPS historical data to figure out mail and package delivery impacts by weather events. We can calculate the impact of weather events by comparing the delivery times between deliveries controlled for weather and deliveries impacted by certain weather events. These calculations done by the model will be able to inform USPS and help quantify the impact of weather events.

USPS Leaders are the biggest stakeholders in this problem. They are hoping to increase customer satisfaction by utilizing AI/ML to accurately forecast the delivery date to reduce delays in delivery day projection. USPS Stakeholders and the project sponsor are expecting a created AI/ML that will judge the impact of different weather events on package delivery time. They are specifically asking for these weather events to be from different publicly available data sources: NOAA, NWS, FEMA, USPS websites, and even USPS provided data sets. This model should be able to quantify and show numerically the impact of different weather events. USPS expects these numbers to come from comparing the OTP Scores from package deliveries with weather events and deliveries not impacted by weather events.

Major pain points for USPS stakeholders are storage costs, data gathering, and user dissatisfaction. With an AI/ML model, there will be costs of retaining the training and testing data for such a model, as well as processing and storing the outcomes of said model. That is why stakeholders have restricted the size of the scope of this project, to keep a working model in a concise space. Data gathering will also be a pain point for the stakeholders. This problem revolves around deliveries and supply chain usage which in turn could have many different reasons for why a package could be delayed. Finding datasets and building a model that will accurately use AI/ML to only see weather event impacts and ignore the rest of the possible reasons for package delay is a cause for concern for stakeholders. Customer dissatisfaction is the biggest pain point for USPS stakeholders. Since 2006, USPS has been forced by congress to 'pre-pay' future health benefits to retirees. USPS needs to make sure customer satisfaction is maintained to help zero out the huge cost that the US Government has burdened them with [16].

1.3 Research

1.3.1 Weather

Severe weather is defined as "any dangerous meteorological phenomenon with the potential to cause damage, serious social disruption or loss of human life" [17]. This can include thunderstorms, tornadoes, floods, lightning, hail, wind and winter weather. A narrower definition of severe weather includes only weather phenomenon relating to severe thunderstorms. Severe weather can be categorized into either general or localized. Nor'easters are storms over the East Coast of the U.S. with winds typically from the northeast [18] and are classified as general severe weather as the weather phenomenon associated with the storm covers a large geographic area. Weather such as thunderstorms, tornadoes and downbursts are more geographically localized and classified as severe weather. Severe thunderstorms are defined by hail one inch in diameter or greater, winds from 58 to 75 miles per hour or a tornado [19]. This project is focused on the serious social disruption due to severe weather such as thunderstorms, namely in relation to USPS mail delivery.

Weather can affect USPS mail and package delivery in multiple ways along the route. Entire airports and interstates can be shut down from major storms and this greatly increases the amount of time it will take for the product to reach the destination. Mail carriers can have great difficulty delivering mail to doorsteps when there is hail, ice or flooding. In extreme conditions mail carriers will not go out on their routes to ensure their safety. USPS will even send out notices for residents to clear heavy snow from sidewalks and mailboxes, to allow the mail carrier easier, safer access to homes. These extreme weather events, increasing in frequency, will continue to greatly impact USPS mail and package deliveries.

1.3.2 Climate Change

Since the beginning of modern weather record keeping in 1950, the frequency and intensity of these such weather events has been increasing due to the increase in greenhouse gases in the atmosphere. As carbon dioxide and methane levels increase, the gases become a blanket in Earth's atmosphere, trapping heat close to the planet. This overall increase in heat affects the water cycle, shifts weather patterns and melts land ice [20]. All of these increases can make severe weather worse and more frequent. The increase in atmospheric temperature means there is more moisture in the air, which provides more available water for heavy rainfall, snowfall and hail. Heavy precipitation is highly likely to cause serious social disruption including package and mail delivery from USPS.

1.3.3 USPS Climate Resilience

A recent study conducted by the Government Accountability Office (GAO) in September 2021 found that the U.S. Postal Service could enhance the climate resilience of its facilities by using weather and climate data. The GAO identified around 10,168 USPS facilities were geographically located in areas at risk to flooding, wildfires, storm surge, and sea level rise, it also found that about 3 percent of all USPS facilities sustained damage from weather-related natural disasters from fiscal years 2015 through 2019. The damages, ranging from broken flag poles to collapsed sections of buildings, were caused by floods, hurricanes, winter storms, and other weather-related disasters. These incidents ultimately cost the USPS over $30 million, putting the processing and distribution facilities' ability to process and deliver mail at risk [21]. The report highlights how important it is to account for severe weather in the context of mail and package delivery. See Figure 5 below for an overview of types of facilities that may be affected by the effects of climate change.

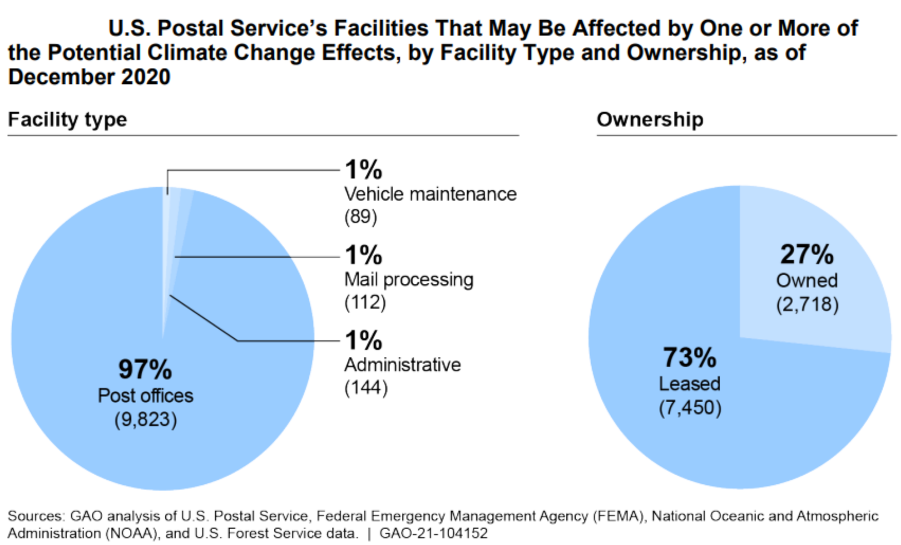

Figure 5: The USPS Facilities affected by climate change

Additionally, the USPS Office of Inspector General (OIG) responded to a congressional inquiry by reviewing the U.S. Postal Service's actions before, during, and after Winter Storm Elliott to address potential preparedness and response weaknesses following a weather emergency and published this audit report on March 12, 2024 [22]. The evaluation of the preparedness of a Postal Service facility involves the OIG assessing how quickly the location can return to operating at its previous level of efficiency following a severe weather event in addition to other factors which involve thorough analysis of management's actions taken. These items include the amount of time management waits before deciding to implement their integrated emergency management plans (IEMPs) that are developed for each facility "which are then used to prepare for and respond to emergencies that occur" by the USPS Office of Continuity and Preparedness, a group of emergency preparedness specialists that work with, prepare for, and respond to emergencies alongside the emergency management teams (EMTs) at area and district levels to implement IEMPs [22].

1.3.4 Related Work on Supply Chain Disruption Analytics

In recent years, several studies have explored the application of advanced analytics techniques to predict and mitigate supply chain disruptions, particularly those caused by extreme weather events. Pandya et al. (2020) [1] demonstrated the effectiveness of machine learning models, specifically Support Vector Machines and Artificial Neural Networks, in predicting freight delivery capacity and delay times on blocked and unblocked streets in Ahmedabad, India. Their approach showcases the potential of these techniques to optimize resource allocation and decision-making in urban freight transportation.

Similarly, Ali et al. (2022) [2] proposed a decision support system based on a Bayesian hierarchical model to assess supplier disruption risks in the face of insufficient data, a common challenge in supply chain risk management. By identifying key risk categories and variables, their framework enables organizations to quantify the impact of disruptions on revenue and inform mitigation strategies.

Brito et al. (2017) [23] provided valuable insights into the far-reaching consequences of extreme weather events on supply chains in Brazil, highlighting the need for comprehensive risk assessment and management practices. Their findings underscore the importance of considering both direct and indirect impacts, as well as the role of media coverage in shaping risk perceptions and preparedness.

Collectively, these studies offer a foundation for the USPS Weather Impact Challenge project, as they demonstrate the viability and value of advanced analytics in predicting and managing supply chain disruptions. By adapting and building upon the methodologies and insights from these works, the team can develop a robust model to forecast the impact of weather events on mail delivery operations, optimize resource allocation, and ultimately enhance the resilience and efficiency of the USPS supply chain. The research also highlights the importance of considering a wide range of risk factors, from transportation disruptions to information gaps, to develop a comprehensive and effective solution.

1.3.5 Enhancing USPS Weather Impact Project through Related Research

Recent studies on supply chain disruptions offer valuable methodologies and frameworks that can inform and enhance the USPS Weather Impact Challenge project. For instance, Garvey et al. (2015) introduced a Bayesian network framework for analyzing risk propagation within supply networks. This approach involves mapping network nodes (e.g., suppliers, distribution centers, delivery points) and links (transportation and communication lines), modeling dependencies and conditional probabilities to predict how initial disruptions, such as weather events, might affect other network parts. Adapting this Bayesian network approach allows the USPS to understand cascading effects of weather disruptions, prioritize risk mitigation efforts, and minimize overall service delivery impacts [24].

Fattahi et al. (2017) developed a model focused on designing supply chain networks that are both responsive and resilient to operational and disruption risks. Their strategies include increasing redundancy, enhancing flexibility, and improving visibility, which align with the USPS's goal to maintain service levels during adverse weather. Implementing these strategies will enable the USPS to optimize operations and enhance customer satisfaction during weather-related disruptions [25].

Saxena et al. (2017) reviewed clustering techniques, useful for identifying patterns and grouping similar data points. Applying clustering to categorize weather events based on their impact on delivery times helps the USPS develop tailored mitigation strategies for different weather scenarios. By understanding specific impacts, USPS can allocate resources more efficiently and improve response planning [26].

Zhang and Goh (2016) introduced Multivariate Adaptive Regression Splines (MARS) for modeling complex, non-linear relationships. For the USPS, MARS can model relationships between weather conditions and OTP scores, capturing intricate dynamics and allowing better preparation and response to weather disruptions. This approach helps identify significant weather variables affecting delivery performance and implement measures to mitigate their impact [27].

Barbosa et al. (2018) reviewed human mobility models, which describe how people move within and between regions. Integrating these models into the USPS's AI/ML framework enhances predictions of delivery delays caused by weather events. By accounting for changes in traffic patterns, road closures, and human behavior during adverse weather, the USPS can optimize delivery routes and schedules to mitigate delays, resulting in more accurate predictions and improved delivery reliability [28].

Hsieh et al. (2005) proposed a method for distinguishing random environmental fluctuations from significant disruptions. Adapting this method allows the USPS to differentiate between minor weather fluctuations and major disruptions, prioritizing resources based on severity. Focusing efforts on the most critical disruptions ensures effective resource allocation and enhances overall network resilience [29].

Altogether, these studies provide a comprehensive array of methods, models, and insights that significantly enhance the USPS Weather Impact Challenge project. By integrating Bayesian network analysis, MARS modeling, clustering techniques, human mobility models, and methods for distinguishing disruption severity, USPS can develop a sophisticated AI/ML model. This prospective model will accurately predict the impact of weather events on delivery times and OTP scores, offering actionable insights for mitigating disruptions.

The focus on resilience, responsiveness, and customer satisfaction in these studies aligns well with USPS stakeholders concerns and objectives. By incorporating these considerations into their model and decision-making processes, the USPS team can address technical challenges of weather-related disruptions, enhance service quality, reduce operational costs, and maintain customer loyalty during adverse weather conditions. This holistic approach ensures the USPS's response to weather impacts is effective and aligned with broader organizational goals, contributing to a more resilient and customer-focused delivery network.

1.4 Solution Space

Our solution delivers value to its users by being able to more accurately predict if a weather event will delay mail and packages. Users will derive value from these predictions by being able to convey better delivery estimates to consumers. Users will also derive value by being able to better prepare for unexpected volume across the delivery route. We expect our solution will provide essential data that will allow the United States Postal service to run more efficiently.

1.5 Project Objectives

The project objective is to quantify the impact of weather events on USPS mail and package delivery times in days using an AI/ML model. The likely outcome of this project will be that major weather events such as Nor'easters or tornadoes will negatively affect the delivery times of mail and packages but smaller weather events such as rainstorms will not change the delivery times. This is simply because USPS is designed to meet service standards despite inclement weather. However more severe weather brought on by climate change will most likely negatively impact delivery times. The solution to this objective lies in either Random Forest or Neural Network machine learning techniques as seen in previous research. Once this project is completed, the team will fully understand severe weather impacts on USPS mail and package delivery times. This understanding will highly benefit USPS and its service partners. The knowledge gained from quantifying the impact of weather events is invaluable to the stakeholders. With this knowledge they can better prepare for such events by assigning more staff or trucks when needed and maybe even investing in the necessary equipment to combat these weather conditions such as plow trucks. USPS may even be able to divert packages through unaffected hubs in case of weather. The value of the knowledge gained from this project is endless.

1.6 Primary User Stories

As a User, I want to predict if mail and packages with be delivered on-time based on weather predictions by the National Weather Service.

1.7 Product Vision

1.7.1 Scenario #1: Proactive Planning for Severe Weather Events

As a mail processing manager at USPS, Stephan is responsible for ensuring that mail and packages are delivered efficiently and effectively. He has been tracking severe weather patterns in his region/hub and knows that heavy rain or Snow can cause significant delays and disruptions to the mail and package delivery process.

Stephan wants to be proactive in planning for any disruptions by using our solution to predict when severe weather events are likely to occur and its potential disruptions, allowing him to plan for the following:

- Adjust mail and package processing schedules to minimize delays
- Prioritize critical packages (e.g., by mail class) for expedited delivery
- Coordinate with other USPS teams and departments (e.g., logistics, customer service) to ensure a smooth response to any potential disruptions due to weather related issues

By using our solution Stephan can reduce the risk of weather-related mail and package delays and ensure that essential services are delivered on time. This scenario highlights the value of proactive planning and preparation in mitigating the impacts of severe weather events.

1.7.2 Scenario #2: Reactive Response to Unforeseen Weather-Related Delays

As a supervisor at a major USPS processing facility, Elizabeth is responsible for managing day-to-day operations. She's been affected by an unexpected snowstorm that has caused significant delays and backlogs in her facility.

Elizabeth wants to quickly respond to the situation using our tool to:

- Identify the affected areas and prioritize mail processing efforts
- Adjust staffing levels and workflow to accommodate increased volumes or delayed deliveries

- Communicate effectively with customers, management, and other stakeholders about the status of mail delivery

By using our solution's predictive capabilities and data analysis, Elizabeth can rapidly respond to unexpected weather-related disruptions and minimize the impact on her facility's operations. This scenario highlights the value and importance of a reactive response strategy in quickly adapting to changing weather conditions.

Section 2: Datasets

2.1 Overview

Data is important to USPS because it allows information to be known about mail and package deliveries. Especially when it comes to the impact that weather has on all steps of the mail and package delivery. To make an AI/ML model that can help accurately determine the numerical impact of each weather event on mail and package deliveries we are using two main datasets: Global Historical Climatology Network Daily and USPS Dataset.

The Global Historical Climatology Network (GHCN) Daily is a database that provides historical climate data such as historical daily temperature, precipitation, and snowfall amounts. It is a composite of multiple sources that have been merged and reviewed [30].

The USPS Dataset was provided by our customer the United State Postal Services. It contains data on both mail and packages that originated from or were delivered to either the Memphis, Nashville, or Music City Annex facilities with an expected delivery data between 2024-01-08 and 2024-01-21. The data was shared with via a compressed file using Microsoft SharePoint.

The package data was partitioned into three text files:

- one file that contains the originated packages,
- one file that contains the barcode scans associated with those packages,
- and one that contains the destinating packages.

We expect a fourth file containing the barcode scans associated with the destinating packages. However, the USPS has run into difficulties pulling that data at this time.

The mail data was partitioned into ten csv files:

- two csv files that contain the originated mail,
- two csv files that contain the destinating mail,
- three csv files that contain the scans associated with the originated mail,
- and three csv files that contain the scans associated with the destinating mail.

2.2 Field Descriptions

The GHCN Daily Database includes roughly 40 weather metrics with the metrics used in this project outlined below, in addition to providing metadata for the stations reporting the weather [31]. For the purpose of this project, we pulled four weather metrics, PRCP, SNOW, TMAX and TMIN, for zip codes related to our mail and package data via the NOAA API.

- PRCP = Precipitation (inches) (Floating Point)
- SNOW = Snowfall (inches) (Floating Point)
- TMAX = Maximum temperature (Fahrenheit) (Integer)
- TMIN = Minimum temperature (Fahrenheit) (Integer)

The USPS mail delivery data contains the following data fields for mail and package piece and scan data.

- Mail
  - Piece Output
    - UNIQUE_IDENTIFIER = Unique ID for the piece (Integer)
    - START_THE_CLOCK_DATE = USPS possession date of the mail (Date)
    - ORIGIN_FACILITY = Facility where mail was received (String)
    - ACTUAL_DLVRY_DATE = Last scan datetime (Date)
    - EXPECTED_DLVRY_DATE = Estimated delivery date (Date)
    - EXPECTED_DESTINATION_FACILITY = Expected destination facility of the piece (String)
    - MAIL_CLASS = First Class Presort, Single Piece First-Class, USPS Marketing Mail, Periodicals (String)
    - MAIL_SHAPE = Letter, Card, or Flat (String)
  - Scan Output
    - UNIQUE_IDENTIFIER = Unique ID for the piece (Integer)
    - SCAN_DATETIME = Date and time of the piece scan (Date)
    - SCAN_FACILITY = Facility name of where the scan occurred (String)
    - OP_CODE = Operation code of the scan (Integer)
- Package
  - Piece Output
    - MailPieceID = Unique identifier for a barcode (Integer)
    - ServiceTypeCode = Service Type Code of mail piece (Integer)
    - StartTheClockZipCode = ZIP Code where the mail piece was received (Integer)
    - StopTheClockDate = Date of the first valid scan (Date)
    - ScheduledDeliveryDate = Date the mail piece was expected to be delivered (Date)
    - StopTheClockZipCode = ZIP Code where the mail piece was delivered (Integer)
    - MailClassCode = USPS mail class code of the mail piece (String)
      - BB = Bound Printed Matter
      - BL = Library Mail
      - BS = Media Mail
      - EX = Priority Mail Express
      - FC = First-Class Mail
      - G0 = Ground Advantage Lightweight
      - G1 = Ground Advantage Heavyweight
      - PM = Priority Mail
      - S2 = USPS Marketing Mail
      - SA = Saturation Mail
  - Scan Output
    - MailPieceID = Unique identifier for a barcode (Integer)
    - TrackingEventID = Unique identifier for a scan (Integer)
    - EventZIPCode = ZIP Code of where the scan occurred (Integer)
    - PTSEventTimestamp = Timestamp of the event (Time)
    - PTSEventCode = USPS Event Code of the scan event (String)
    - PTSEventDesc = USPS Event Code description of the scan event (String)

2.3 Data Context

The first version of the GHCN database was released in 1992 with the primary focus being building a database that would allow for the analysis of climate and weather conditions at both the regional and global level by compiling data from multiple sources. The most recent version update of the GHCN-Daily has kept this focus in mind while improving upon the thoroughness and coverage the weather database, e.g., increase in the number of stations reporting, climate-related metadata, and updating the database more frequently [31].

The historical weather data used needs to be as geographically specific as possible, as USPS operations are done at a digit and zip code-level. Barcode scanning along the USPS network provides enough of the relevant information to act as a basis for building USPS's historical package and mail delivery data database, which will be used to pull a sample dataset for training. To ensure the use of appropriate weather data for the USPS weather delay model the type of weather and climate data collected as well as the frequency of data collection and station updates are important, but it is essential to use data that reports the weather from as many locations as possible. It is the responsibility of USPS to provide mail service in every region in America, and there are areas of the country that are much more remote than others [32].

For our model to accurately predict how much of an impact a weather-related delay will have on the number of days the mail or package is delayed at various USPS locations or along routes around the country we need to integrate weather station data with the historical service performance data provided by USPS. The GHCN-Daily database will allow us to do this and is updated on a scheduled, consistent basis seven-days a week and distributed by the *NOAA National Centers for Environmental Information (NCEI)* [33].

One of the first reported temperature anomalies for the year 2024 is an Arctic Blast that occurred from January 12 until January 16 as reported all over the United States and had a heavy impact on the USPS operations in the Tennessee area. Per the January 2024 Tennessee State Climate Summary:

> *Record-setting snow fell over parts of Tennessee this month, with a total of 279 weather stations reporting measurable snowfall in January, and an additional 11 stations reporting a trace of snow (less than 0.1-inch). For most areas of the state this month brought more snow than the entire year of 2023. Beyond the impressive snow totals, the arctic air that settled into the state after the snowstorm led to prolonged snow coverage, with many areas having snow on the ground for over a week!* [34]

Additionally, January 15, 2024, is included as a notable date in the NCEI's Monthly National Climate Report for January 2024 as Nashville received over six inches of snow, "more than an entire winter's worth of snow for the city" [35]. The Storms Events Database contains records as entered by NOAA's National Weather Service (NWS) from January 1950 to March 2024. This data documents rare and unusual weather events, the frequency of deadly storms and other damaging weather phenomena, and meteorological metadata surrounding the event date (e.g., the maximum temperature). The main purpose of the Storm Events Database is to provide accurate descriptions of weather events.

The service performance data is collected and held by USPS and is not available to the public. The data collected is generated automatically using machines and code-scanning technology and this raw data is processed by USPS based on business rules and practices so that routine diagnostics can be run, and reports can be generated that benefit not only USPS but the greater postal service network. The USPS data sample used in this training model was pulled by USPS and a team generated a sample of their data consisting of a dataset of package and mail data from the area of Tennessee impacted for more than just the period of the *Arctic Blast* in Tennessee, and the sample covers all types of mail and packages beginning January 8, 2024 and ending January 21, 2024.

The datasets host the piece-level data for packages and mail. Each dataset is a combination of the originating pieces and destinating pieces data for the pieces of mail and packages. For mail data, USPS provided the originating and destination data on the scan-level for the mail data but has not yet provided this for packages. The same data is to be expected for the same locations at the package level.

2.4 Data Conditioning

For this project, we are integrating USPS mail delivery data with weather data to predict mail and package delays. This involves several crucial data transformation steps: dealing with missing values by either filling or discarding incomplete records, normalizing the data to maintain consistency, and merging datasets based on shared attributes like date and geospatial location such zip codes. We also focus on detecting and removing outliers to minimize the impact of anomalies. By ensuring our data is of high-quality through these conditioning steps, we can potentially improve the accuracy and reliability of our predictive model, leading to more precise forecasts of mail delays due to severe weather conditions.

To prevent data leakage, we select features that would ensure the integrity and reliability of our mode. For example, when building the model to predict potential package delays, we exclude the expected delivery date since it directly influences the prediction outcome. Instead, we rely on features like origin and destination zip codes, weather conditions at origin and destination, and historical delivery patterns. This approach helps the model learn to predict delays based on data patterns.

2.4.1 Weather Data

The Package weather data is sourced from the NOAA API. The API uses a 5-digit zip code and a date range as the input, and returns the daily minimum and maximum temperature, as well as snow and precipitation levels if available. An example of the API call in python is shown in Figure 6 below.

Input:

datasetid : GHCND (The Global Historical Climatology Network daily (GHCNd) is an integrated database of daily climate summaries from land surface stations across the globe.)

locationid: ZIP (5 digit Zipcode)

startdate, enddate: date (YYY-MM-DD)

Output:

TMIN: Temperature Minimum in Fahrenheit

TMAX: Temperature Maximum in Fahrenheit

PRCP: Precipitation in inches

SNOW: Snowfall in inches

```python
def get_weather_data(api_key, zip_code, date):
    base_url = "https://www.ncdc.noaa.gov/cdo-web/api/v2/data"
    headers = {'token': api_key}
    params = {
        'datasetid': 'GHCND',
        'locationid': f'ZIP:{zip_code}',
        'startdate': date,
        'enddate': date,
        'datatypeid': ['TMIN', 'TMAX', 'PRCP', 'SNOW'],
        'units': 'standard',
        'limit': 1000
    }
```

Figure 6: Example NOAA API Function in Python

2.4.2 USPS: Package Originating Data

In the Package Origin Scans data, all events associated with each package throughout its lifecycle from acceptance to delivery are captured. The packages are identified by its **'MailPieceID'**. 49,432 records of Scans data '**EventZIPCode'** with Zip3 '370', '371' and '372' (Nashville area) along with **'PTSEventCode_Numbers'** of 7 and 10 are selected. To process possible features from the Scans data the records are further consolidated to get the MailPieceID's minimum scan date, maximum scan date, number of distinct event scans **'Distinct_event_scans'** and a time duration **'time_delta_minutes'** between minimum and maximum scan date. The resulting 44,476 records of aggregated Scans data are then combined with the Package Origin Piece data.

The Originating Package Piece data contains 37,290 records of packages with **'StartTheClockZipCode'** and **'StopTheClockZipCode'** with Zip3 of '370', '371', '372' (Nashville area) between 2023-11-08 and 2024-05-11. After inner joining Scans data on **'MailPieceID'** and valid weather data sourced from the NOAA API on Zip code and Date we get 17,503 records to build the predictive model. A package is considered 'late' when the **'StopTheClockDate'** is greater than the **'ScheduledDeliveryDate'.** A late package is represented by 1 and an on-time mail as a 0.

The fields that make up the final data frame are shown in the table below. Fields with **"_O"** represent package Origin **"_D"** represent package Destination. **TMIN** represents minimum temperature, **TMAX** represents maximum temperature in Fahrenheit. **PRCP** represents precipitation in inches. **SNOW** represents snowfall in inches.

Data Conditioning Phase Package Scans	# of Records
Raw File Package Scans	43,524,579
Only Zip3 of '370', '371', '372' And Event Code of 7, 10	49,432
Consolidation of records	44,476

Table 1: Data Conditioning Package Scans

Data Conditioning Phase Package Piece	# of Records
Raw File Package Piece	1,115,279
Only Zip3 of '370', '371', '372	37,290

Combining Weather, Scans	17,503

Table 2: Data Conditioning Package Pieces

ServiceTypeCode	String
MailClassCode	String
Distinct_event_scans	Numeric
time_delta_minutes	Numeric
late	Boolean
Zip_O	Numeric
TMIN_O	Numeric
TMAX_O	Numeric
PRCP_O	Numeric
SNOW_O	Numeric
Zip_D	Numeric
TMIN_D	Numeric
TMAX_D	Numeric
PRCP_D	Numeric
SNOW_D	Numeric

Figure 7: Attributes included in Cleaned Package Dataset

2.4.3 USPS: Mail Data – Originating Data

The Originating Pieces files contains 42,174,946 records of mail originating from either the Nashville [1441275], Memphis NDC [137672], Memphis [1441274] or Music City Annex [1532174] between 2024-01-08 and 2024-01-21.

We took the following steps to condition the Originating Pieces data:

1. Removed rows with null values
2. Removed rows with duplicate unique identifiers
3. Removed rows where the Actual Delivery Date was before the Start Clock Date

Data Conditioning Phase	# of Records
Raw File	42,174,946
Removal of Null Values	41,160,323
Removal of Duplicate IDs	41,160,140
Removal of Date Errors	38,832,134

2.4.4 USPS: Mail Data – Destinating Data

The Destinating Pieces files contain 44,210,232 records of mail delivered to the Memphis, Nashville, or Music City Annex facility between 2024-01-08 and 2024-01-21.

We took the following steps to condition the Destinating Pieces data:

1. Removed rows with null values
2. Removed rows with duplicate unique identifiers
3. Removed rows where the Actual Delivery Date was before the Start Clock Date

The first step of conditioning the Destinating Pieces data is removing the nulls values from the dataset. The Destinating Pieces files contained null values in three columns: START_THE_CLOCK_DATE, OZIP3 and EXPECTED_DELIVERY_DATE. The figure below provides a breakdown of null values in each of these columns. One interesting observation from this is that the same number of null values in the START_THE_CLOCK_DATE column and the EXPECTED_DELIVERY_DATE which indicates that records that are missing the value in one column is also missing the value in the other.

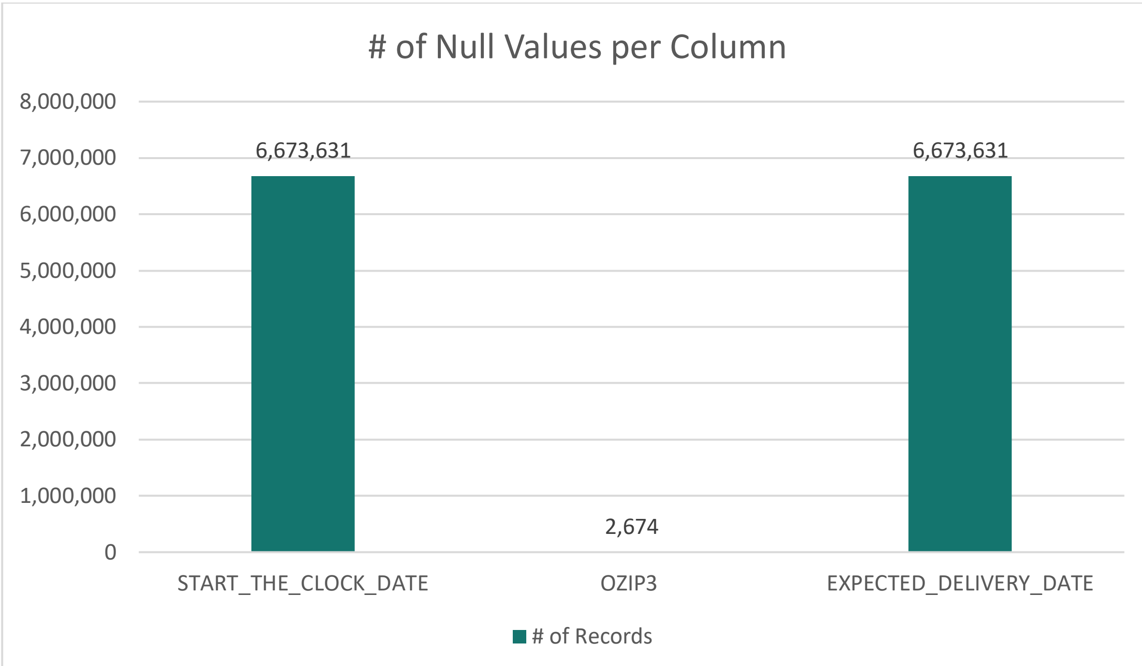

Figure 8: # of Null Values per Column, Destinating Pieces

The next step to condition the Destinating Pieces dataset was removing records that have duplicates in the UNIQUE_IDENTIFIER column. In the Destinating Pieces dataset, there are 614 unique identifiers that were associated with more than one record within the dataset. Upon further investigation, we found that these duplicates typically have varying values in the ACTUAL_DLVRY_DATE, EXPECTED_DESTINATION_FACILITY, and dzip3 columns. Since we are unable to determine which records contains the correct values, we have removed all records that contain a unique id that is associated with more than one records. During this phase of conditioning, we removed 3,693 records with duplicate unique ids.

The last step of conditioning the Destinating Pieces dataset was removing records where the ACTUAL_DLVRY_DATE value was before the START_THE_CLOCK_DATE value. These records were discovered during initial exploratory data analysis, and they need to be removed because a piece of mail cannot have a delivery date that is prior to when the USPS received the mail in the first place. The Destinating Pieces dataset contained 388,444 records where the delivery date was before the start the clock date.

After conditioning the Destinating Pieces dataset, the dataset contained 44,716,693 records. The table below shows the number of records the dataset contained after each phase of conditioning.

Data Conditioning Phase	# of Records
Raw File	53,733,124
Removal of Null Values	47,058,984
Removal of Duplicate IDs	47,055,291

Removal of Date Errors	44,716,693

Table 3: # of Records per Phase of Conditioning

The Destinating Scans files contain 139,005,571 records of scan associated with mail delivered to the Memphis, Nashville, or Music City Annex facilities between 2024-01-08 and 2024-01-21. The only steps required to condition the Destinating Scan data was to remove the rows with null values.

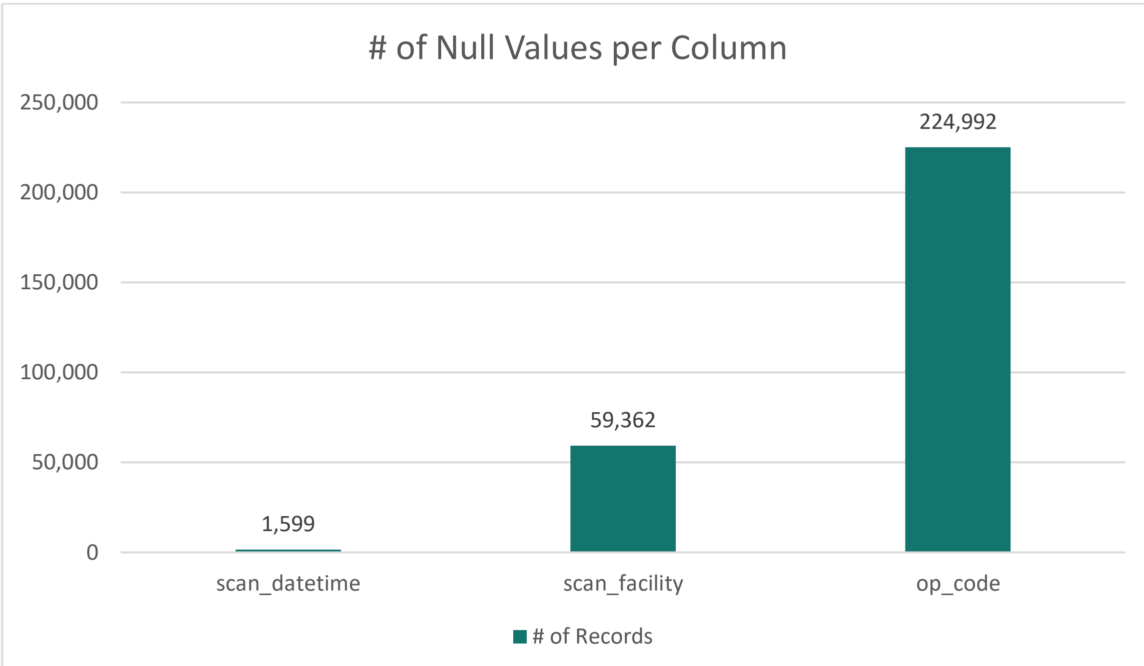

Figure 7: # of Null Values per Column, Destinating Scans

After removing rows that contained null values from the Destinating Scans dataset, the dataset contained 138,722,816 rows.

2.5 Data Quality Assessment

The GHCN-Daily database has built-in, automated *quality assurance* (QA) checks within the core GHCN-Daily system including checks for repetition and duplication, world record exceedance checks, geographic-based outlier checks, climatology outlier checks, internal and temporal consistency checks, spatial consistency checks, spatial corroboration checks, and system performance checks for daily climate observations [36].

Researchers from the National Climatic Data Center note the difficulty in ensuring quality assurance for a database that collects a lot of severe weather event information in such a way that none of the actual extreme meteorological events are lost in this quality check process.

> *Based on an assessment of each individual check and a final evaluation for each element, the system identifies 3.6 million (0.24%) of the more than 1.5 billion maximum/minimum temperature, precipitation, snowfall, and snow depth values in GHCN-Daily as errors, has a false-positive rate of 1% - 2%, and is effective at detecting both the grossest errors as well as more subtle inconsistencies among elements* [33].

The GHCN Dataset:

- Completeness – The dataset does contain null values. For this analysis, we assume that nulls are zero.
- Consistency – Some weather stations do not collect all datapoints, and therefore, can vary in consistency across zip codes.
- Uniqueness – Yes, all attributes within the data set are unique.
- Integrity – All the relationships are complete for this data. All requested dates were provided.
- Conformity – Yes
- Accuracy – Yes, the data is accurate and displayed in the correct format.

The USPS Dataset:

- Completeness – The dataset contains null values that must be removed prior to modeling the data.
- Consistency – The formatting and scale of the values within the USPS data are consistent between features.
- Uniqueness – All attributes within the datasets are unique, with the exception of the unique_identifier which is used to join the files.
- Integrity – The relationships between the features are complete and the data has high integrity.
- Conformity – The data has a high level of conformity.
- Accuracy – Within the mail data, there were scans with invalid operation codes. Also, some data had an actual delivery date that was before the start date. These inaccuracies will be removed prior to modeling.

2.6 Other Data Sources

During the experiment design phase, we considered using USPS survey data. The idea was that we could use survey data to build a model that could predict the impact severe weather had on survey data. We discussed the possibility with the client and were told that it would be possible to get survey data for the mail and packages included in our model. However, we decided not to pursue this model further because it was outside the scope of the client's original request for a model that could predict the number of days mail and packages would be delayed by severe weather.

2.7 Storage Medium

This project will utilize two sets of data that will be merged to create one set of data to train and test our model. The weather dataset that has been pulled from the National Centers for Environmental Information (NCEI) is small (only 11 KB). Therefore, we decided to store the dataset in an excel file on GitHub. The team decided that keeping the data in the team GitHub repository was ideal for ease of access.

The second dataset this project will utilize will be provided by USPS. Based on the discussions we have had with our USPS client; we expect this dataset to be very large. Also, as per our client's request, we will need to ensure that the data set will not be accessible by the public. The client hosted the data in a USPS sharePoint folder. We were able to download the data from the folder on to our personal laptops via two-factor authentication.

2.8 Storage Security

There is currently no required security for the weather dataset in this project, but we have limited access to edit the team's GitHub repository to team members. However, our client from the United State Postal Service has requested that access to the USPS dataset be restricted to team members only.

2.9 Storage Costs

There are no storage costs associated with the weather data required for this project as Git Hub is a free resource. There is also no storage costs for the USPS data since it will be hosted on the USPS SharePoint site as well as on our personal laptops.

Section 3: ML Model Exploration & Selection

3.1 Exploratory Data Analysis

The data from weather sources, coupled with the package and mail data from the USPS has provided insight into the potential relationship between severe weather events and the delivery performance of the Postal Service. Comparing the USPS data provided for the period and region impacted by a specific severe weather event to the average delivery time during a period in which there was no reported weather event provides insight into the potential influence of weather types on the ability for USPS to deliver mail and packages on time. Additionally, knowledge of how different types of weather can alter delivery time and by how much, coupled with an understanding and ability to calculate and quantify this impact will provide USPS with insight into how weather changes affect delivery time compared to how organizational and network related effects impact operations.

The weather event selected is the January 2024 winter storm in Tennessee that impacted operations from January 8 until January 21. The Postal Service reported the following service performance scores for the period from January 1 through January 19 in a January 24 publication:

> ***First-Class Mail:*** *84.0 percent of First-Class Mail delivered on time against the USPS service standard, a decrease of 1.9 percentage points from the fiscal first quarter*
> ***Marketing Mail:*** *92.1 percent of Marketing Mail delivered on time against the USPS service standard, a decrease of 1.6 percentage points from the fiscal first quarter*
> ***Periodicals:*** *80.6 percent of Periodicals delivered on time against the USPS service standard, consistent with performance from the fiscal first quarter* [37]

In addition to providing service performance metrics for specific mail types, USPS reported that in 2024 during the relevant period, the "average time for the Postal Service to deliver a mail piece or package across the nation was 2.8 days" [37].

3.1.1 USPS Package Data

Exploratory data analysis on the USPS Package data we have found some relation with how certain weather events impact the delivery of packages.

In the preprocessed Package data, the overall late percentages for packages were about 19%. A breakdown of the dataset per MailClassCode, along with the late rate, can be seen in Figure 9 below.

MailClassCode	Percentage	LatePercentage
G0	49.63%	21.65%
PM	25.15%	18.74%
G1	18.46%	14.95%
BS	3.46%	8.93%
S2	1.19%	0.48%
FC	0.87%	48.37%
missing	0.58%	0.00%
EX	0.39%	7.25%
BB	0.25%	4.65%
BL	0.02%	0.00%
SA	0.01%	0.00%

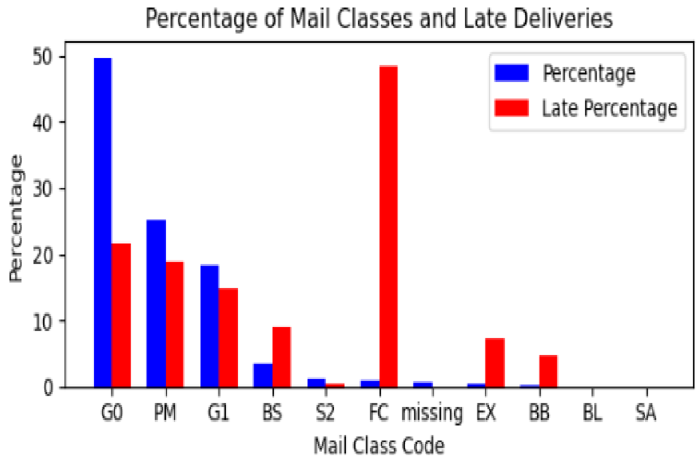

Figure 9: USPS Packages: MailClassCode and Late Percentages

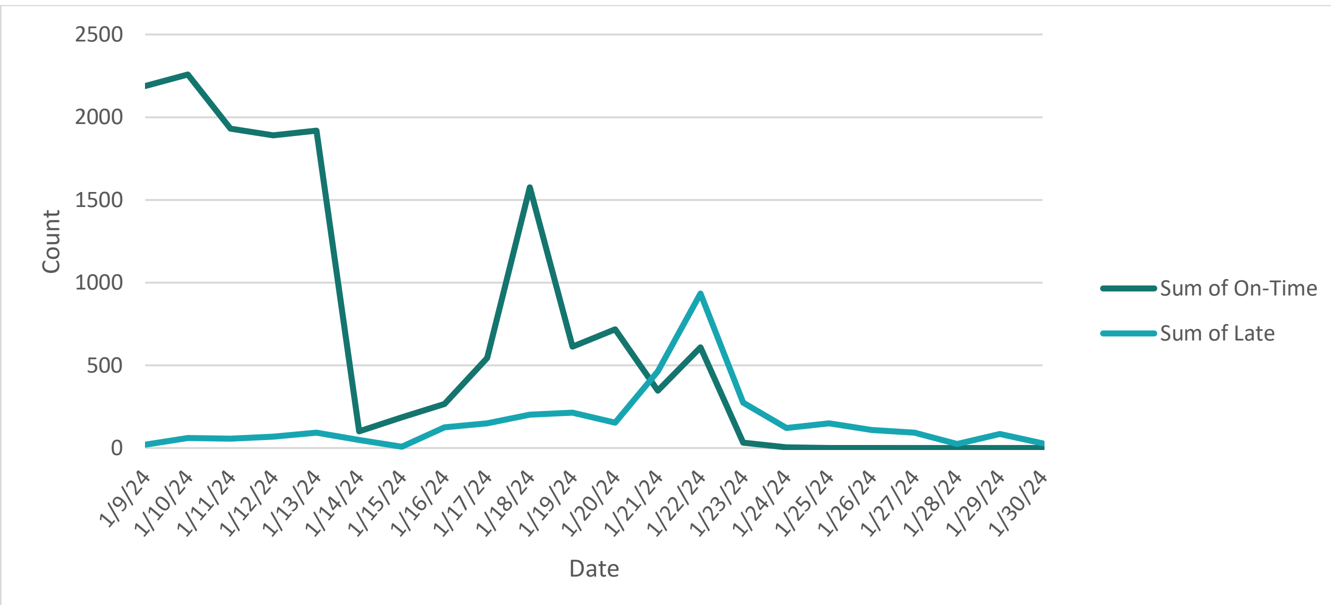

Figure 10: USPS Packages: Late vs On-Time for Zip3 of '370', '371', '372' (Nashville Area)

Figure 10 above shows the volume of on-time and late packages that went through the Music City Annex facility from the start of the timeframe of our experiment, January 8, 2024, thru the end of the month. From the chart we can see that there was a spike in late packages starting on January 20th, and the volume of late packages remains higher than the volume of on-time packages through the end of the month. Figure 11 below shows the average levels of snow and precipitation, in inches, around the Music City Annex during the month of January 2024. We can see that the snow level in the area shot up between January 14th and January 18th. If we compare Figure 10 and Figure 11, we can see that the volume of packages being processed at the Music City Annex dropped significantly during the snowstorm. The package volume then rose up again quickly on January 19th once the snowstorm had ended, with the volume of late packages steadily rising over a series of days. From this observation, we could possibly conclude that package volume dropped due to the snowstorm and operations at the Music City Annex facility were unable to handle the subsequent rise in volume following the snowstorm which resulted in the rise of late packages. We cannot make this assertion, however, since we are not experts in

USPS operations. However, from Figure 10 and Figure 11 it is clear that there was a snowstorm and late packages did follow.

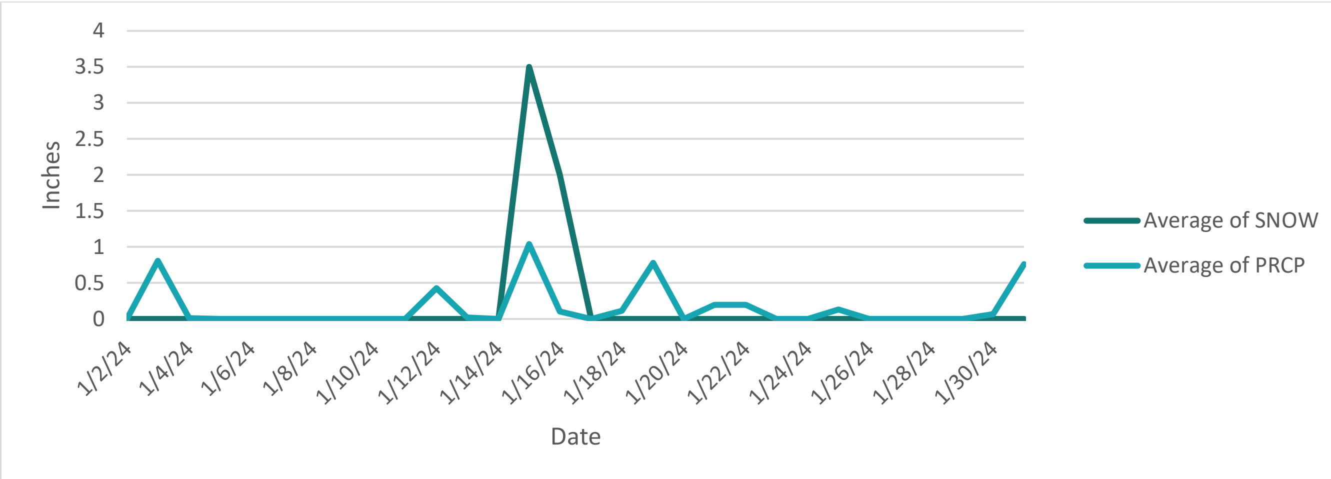

Figure 11: Average Snow and Precipitation Over Time for Zip3 of '370', '371', '372' (Nashville Area)

Two other weather metrics that are included in our analysis are daily minimum and maximum temperature in the area around the Music City Annex Facility. Figure 12 below shows the minimum and maximum temperature during the duration of our experiment. What we can gather from Figure 12 is that the minimum daily temperature dropped during and after the snowstorm that impacted the Music City Annex facility. The minimum temperature dropped below 0°F on January 16th and remains under 0°F until January 23rd. This sustained drop in temperature could also contribute to the increase in late packages we see starting on January 20th.

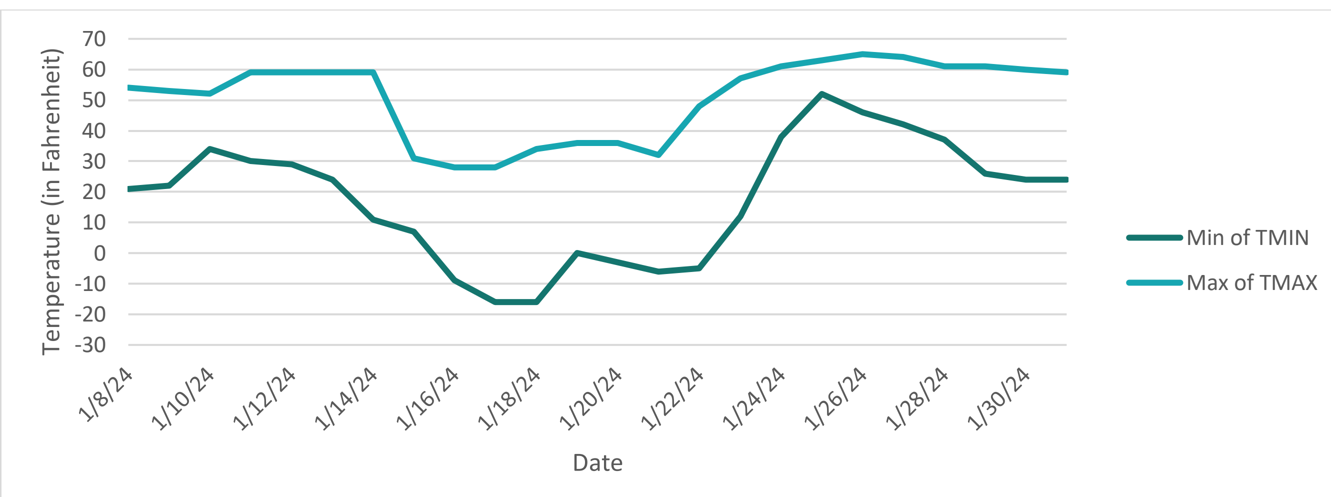

Figure 12: Temperature Over Time for Zip3 of '370', '371', '372' (Nashville Area)

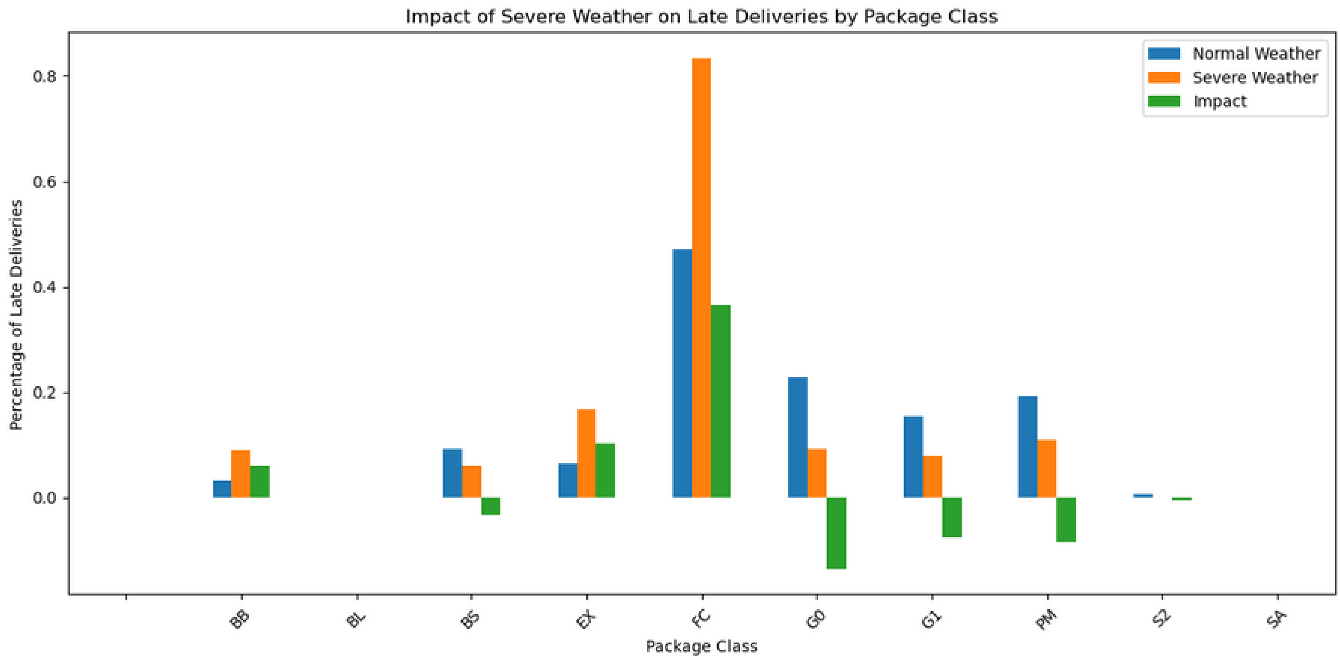

Figure 13: Severe Weather Impact on Late Deliveries by Package Class

Figure 13 above shows the highest percentage of late deliveries, especially during severe weather, broken out by Package Class. Some classes like G0, G1, and PM show a decrease in late deliveries during severe weather (negative green bars). Classes like BB and EX show an increase in late deliveries during severe weather.

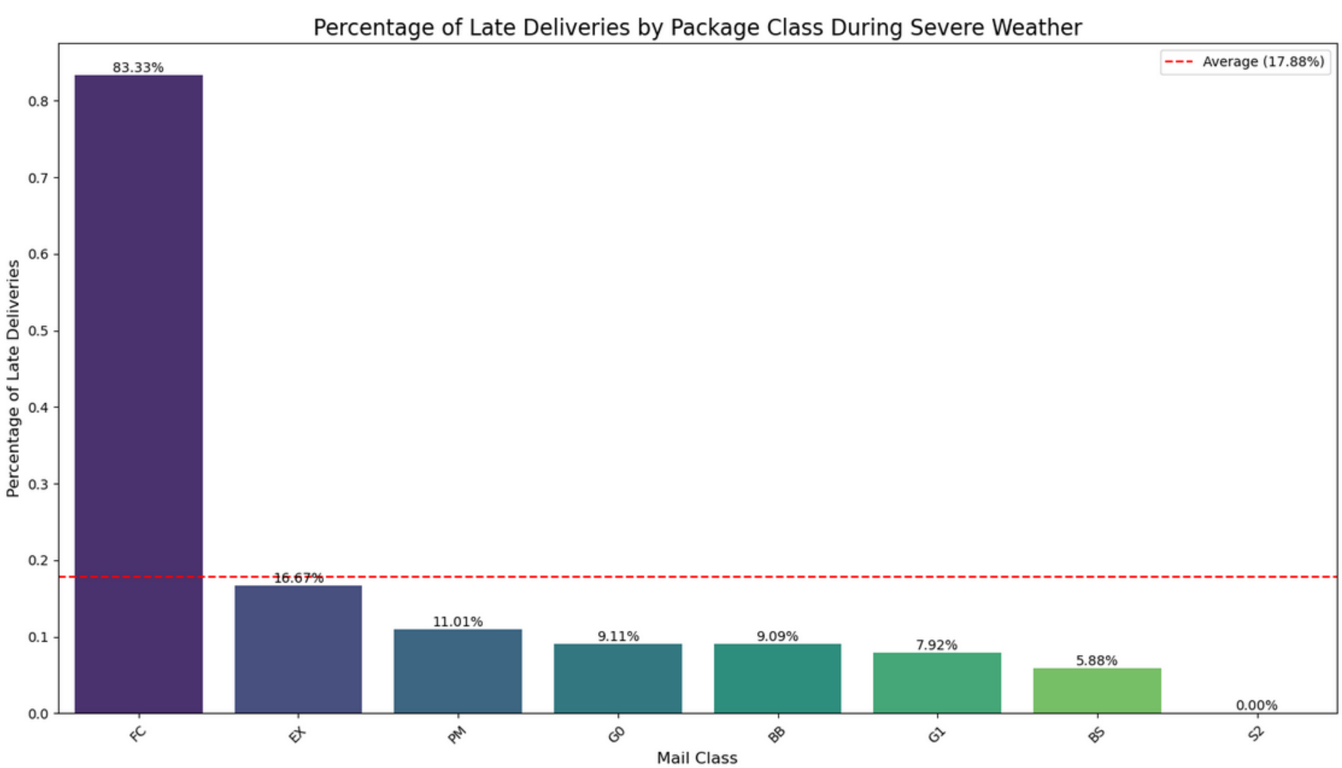

Figure 14: Percentage of Late Deliveries by Package Class During Severe Weather

Figure 14 above shows the late delivery percentages for different mail classes during severe weather. The FC class has the highest rate at over 80%, while other classes like PM, PO, and GX have much lower rates around 10-20%. This suggests FC packages are most affected by severe weather delays.

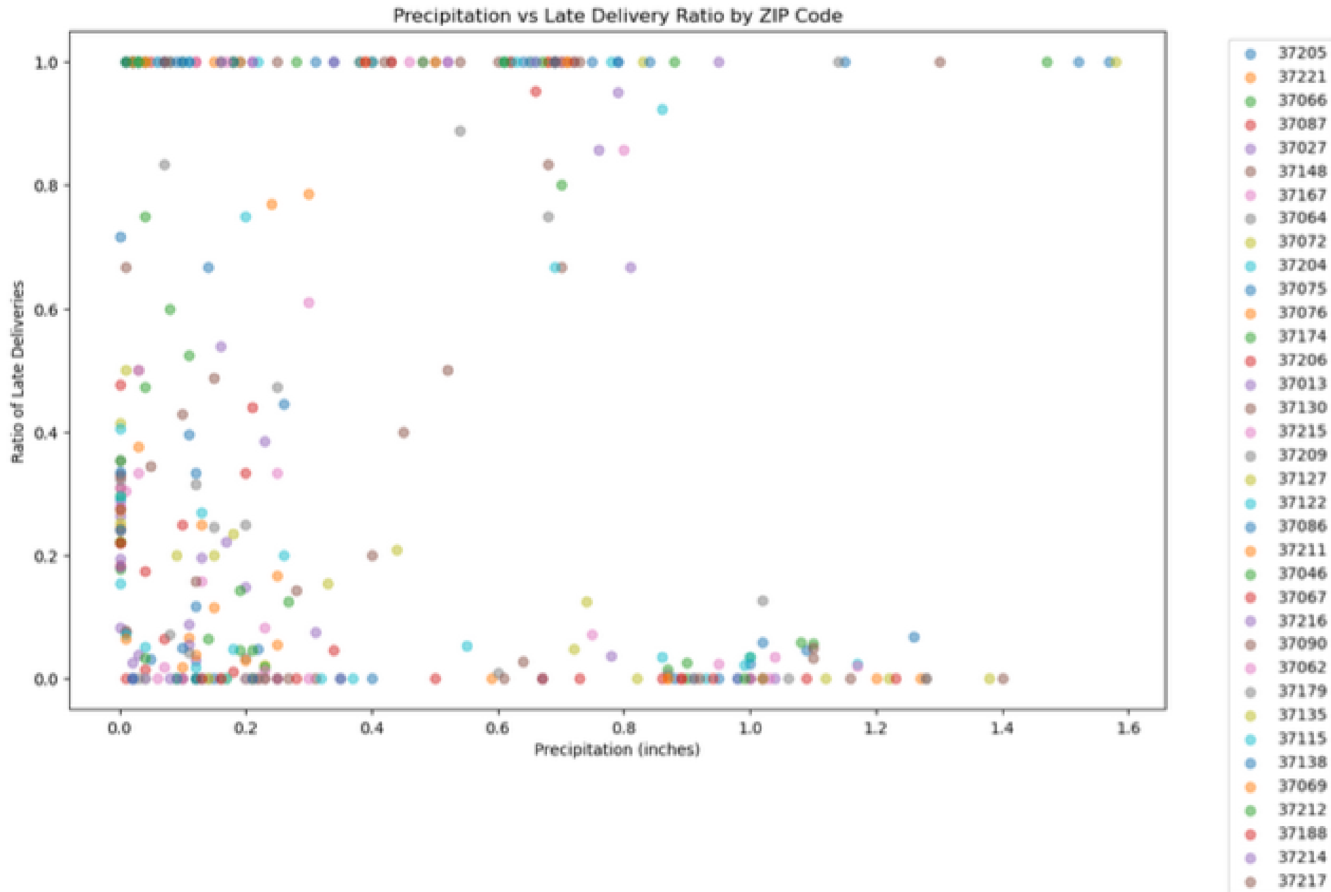

Figure 15: Precipitation vs Late Delivery Ratio by ZIP Code

The scatter plot in Figure 15 displays the relationship between precipitation and late deliveries across ZIP codes. There's significant variation, with some areas showing high late delivery rates even at low precipitation levels. This implies factors beyond just precipitation influence delays. The clustering of points suggests certain ZIP codes may be more susceptible to weather-related delays.

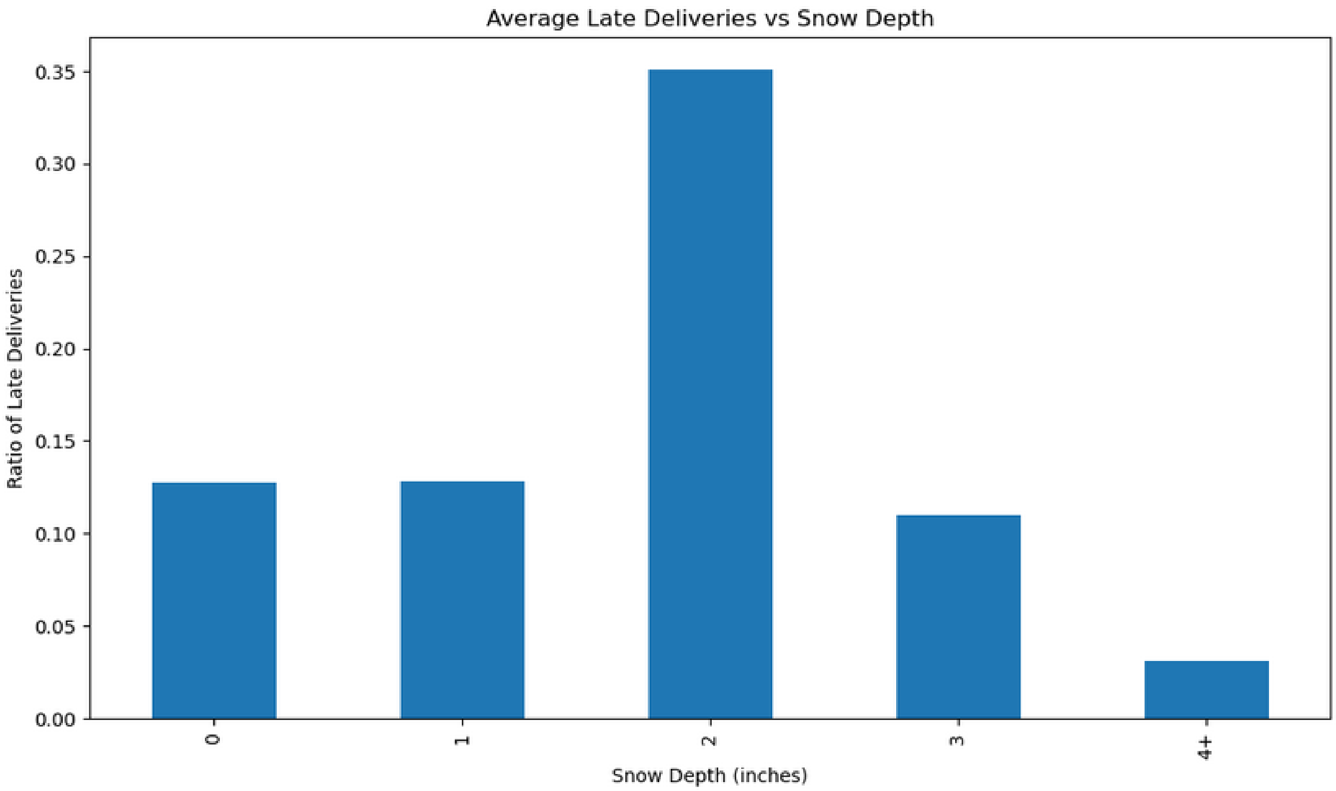

Figure 16: Average Late Deliveries vs Snow Depth

The bar graph in Figure 16 illustrates how the level of snow depth in inches impacts late deliveries. Interestingly, 2 inches of snow corresponds to the highest late delivery rate (about 0.35), while deeper snow does not necessarily mean longer or more delays. This could indicate that moderate snowfall causes more disruption than heavy snowfall, possibly due to differing response and preparation measures. Additionally, the ratio of late deliveries when there is no snow on the ground could indicate an operational issue within the area or USPS facility or that it is weather related and the cold and icy conditions are impacting delivery.

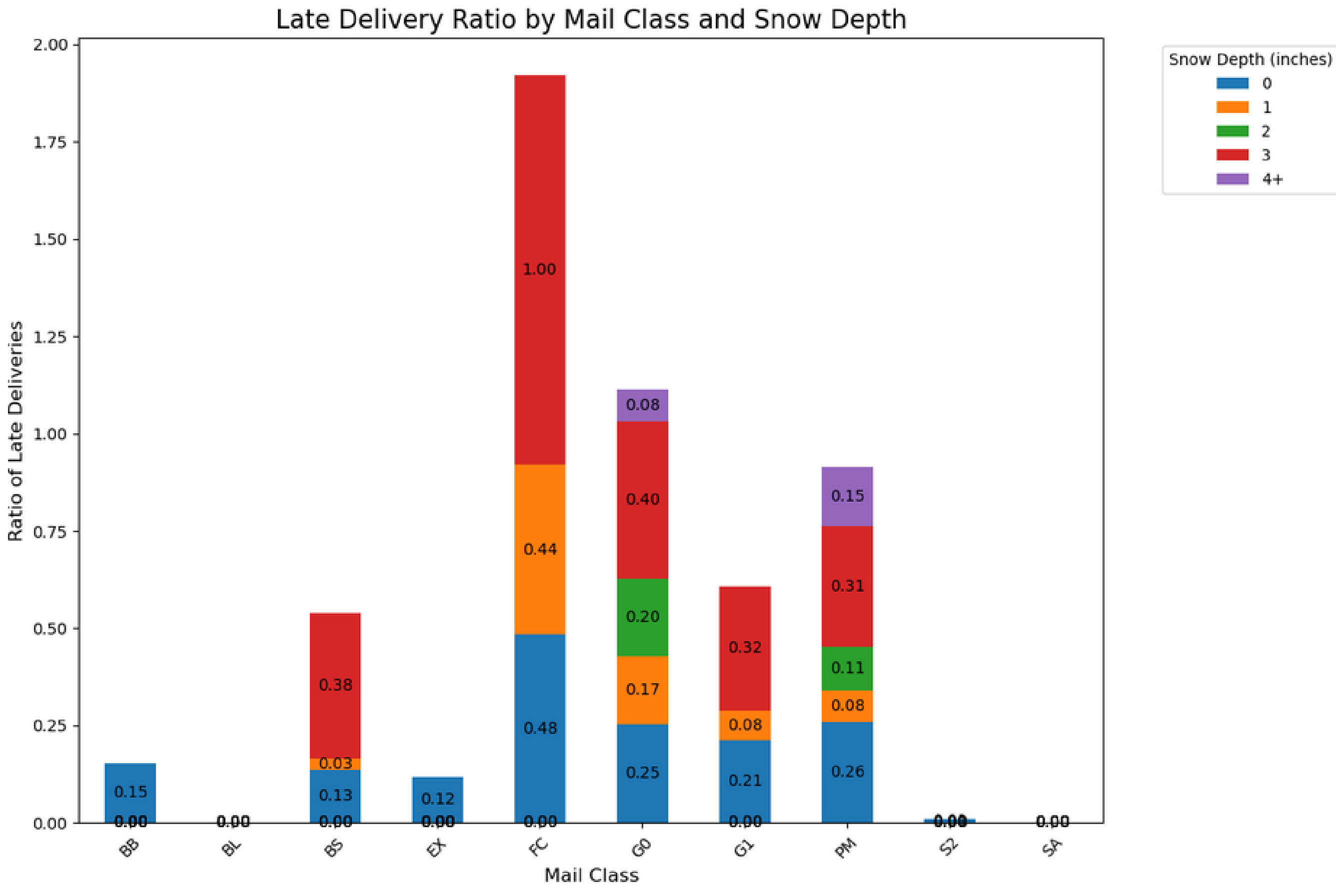

Figure 17: Late Delivery Ratio by Mail Class and Snow Depth

The stacked bar chart in Figure 17 above combines mail class and snow depth to show their joint effect on late deliveries. FC again shows high late delivery rates across snow depths. For most mail classes, deeper snow generally increases delays, but the relationship isn't strictly linear. This visualization potentially helps identify which mail classes are most resilient to snow-related delays.

Figure 18: Precipitation, Snow Depth and Temperature by Zip Code: On Time Deliveries

Figure 19: Precipitation, Snow Depth and Temperature by Zip Code: Late Deliveries

Figure 18 and Figure 19 above show the weather data in the form of precipitation, snow depth, and temperature over the zip codes highlighted throughout this analysis. The most obvious trend is that on time deliveries generally have higher temperatures associated with them. There is not an obvious correlation between precipitation and snow depth when it comes to whether a delivery is late.

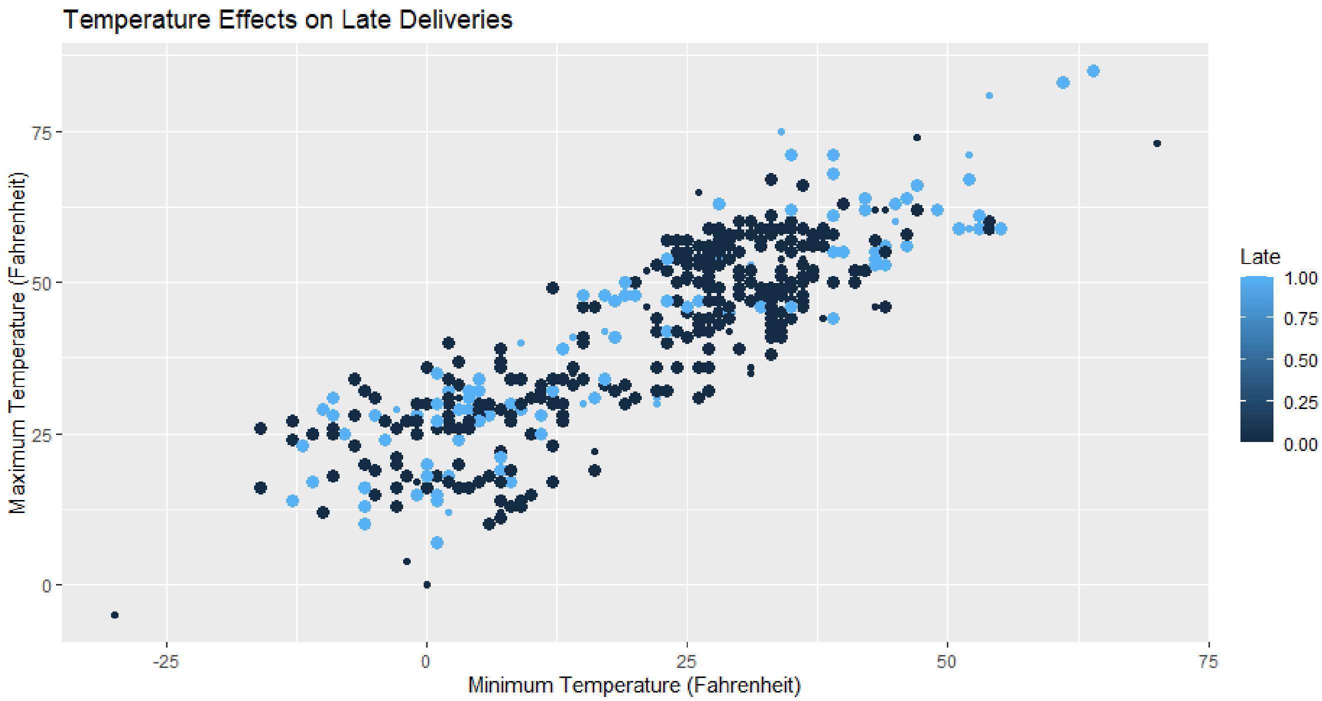

Figure 20: Temperature Extremes and Late Deliveries

Figure 20 above depicts the minimum and maximum temperatures effect on late deliveries. There is an on-time cluster around a minimum temperature of 30°F and a maximum temperature of 50°F. This temperature range likely indicates above freezing conditions and less ice or snow present. There is a small cluster of late deliveries around a minimum temperature of 5°F and a maximum temperature of 30°F. This temperature range likely leads to below freezing temperatures and the presence of ice or snow.

3.1.2 USPS Mail Data

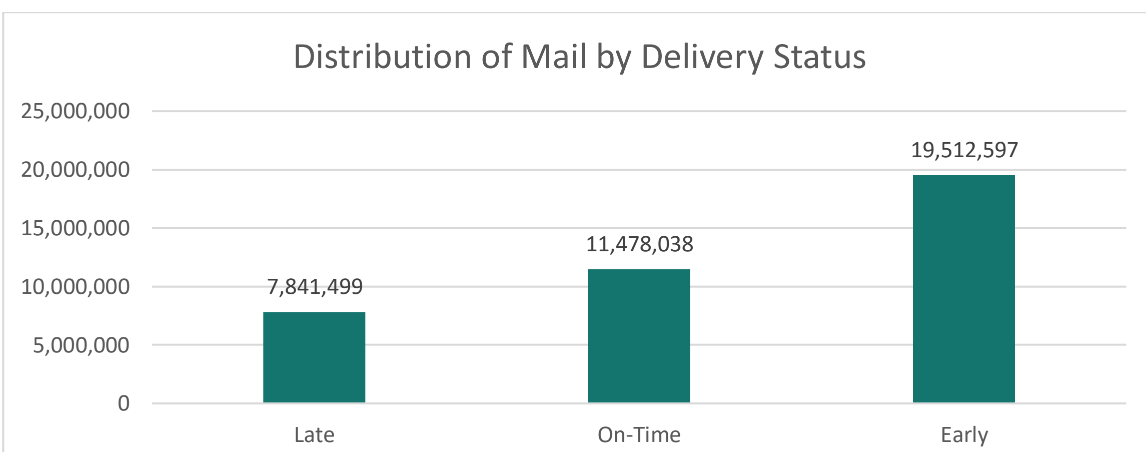

Figure 21: USPS Mail by Delivery Status

In piece-level originating mail data, we were given 38,832,134 records. Of these 19,512,597 had arrival dates that were earlier than the expected, 11,478,038 had on-time arrival dates, and 7,841,499 had arrival dates that were late. We can see this breakdown of Figure 21 above. Based on this we can determine that approximately 20.2% of the mail records are late which is consistent with the package records.

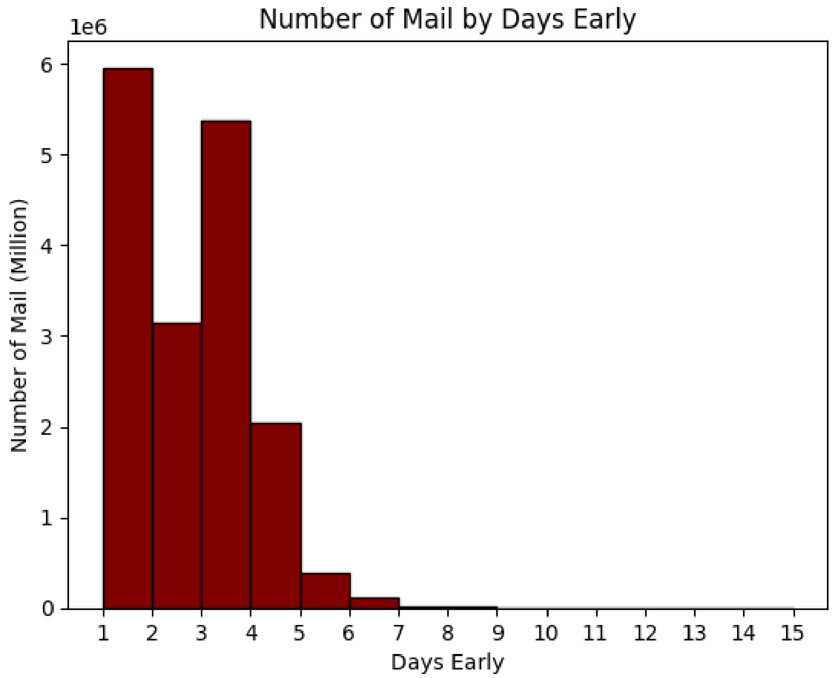

Figure 22: Days Early Distribution

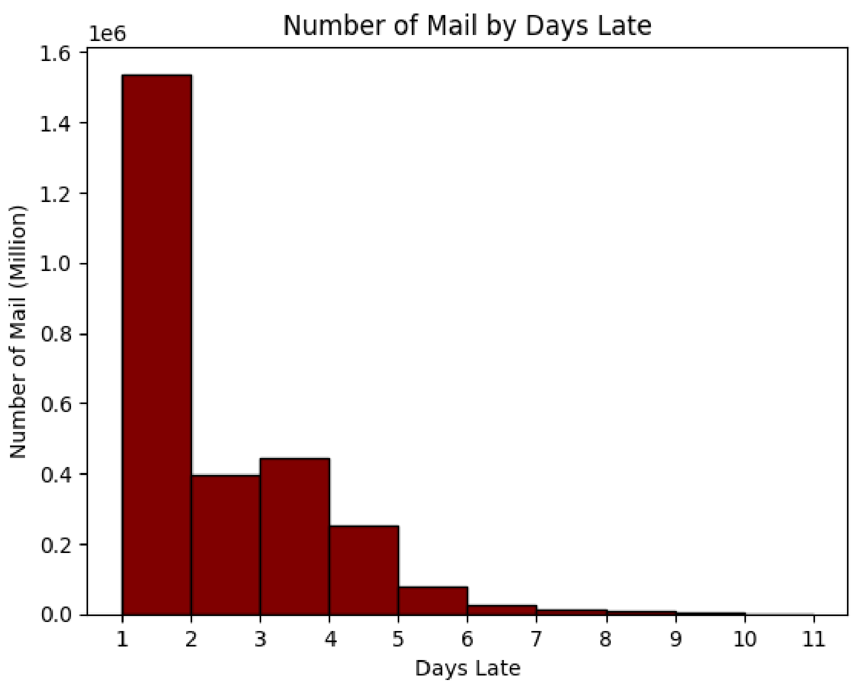

Figure 23: Days Late Distribution

Figure 22 above shows the number of pieces of mail that has arrived early by the number of days it arrived early. We see that most of the mail, about 84.9%, arrives up to 3 days earlier than the expected delivery date. Then

Figure 23 shows the frequency of mail that has arrived late by the number of days it arrived late. Compared to mail delivered early, late mail only has 86.3% of mail being delivered late within 3 days.

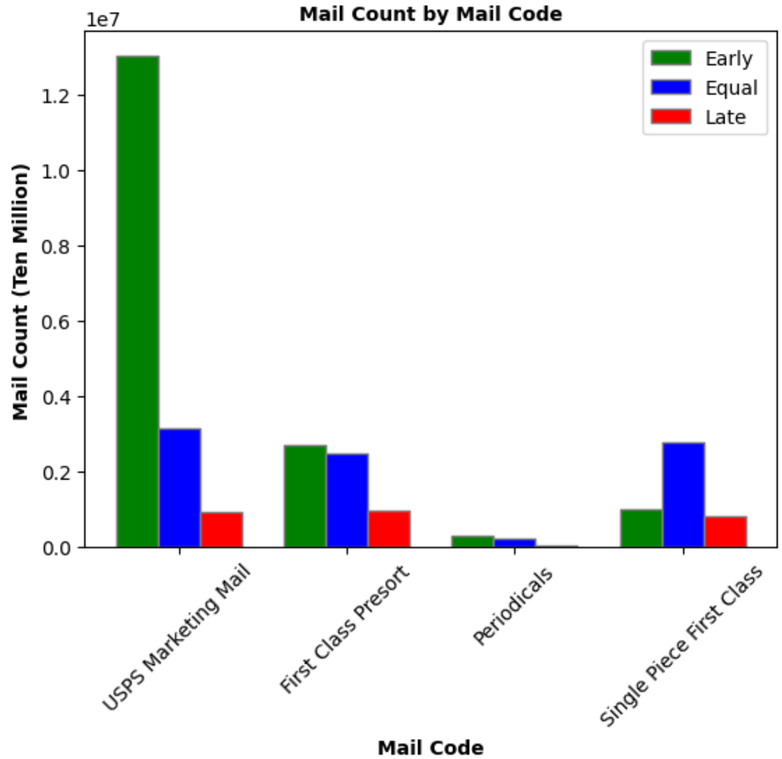

Figure 24: Mail Count by Mail Code

Mail_Codes	Percent On-Time	Percent Late
USPS Marketing Mail	0.944897	0.055103
First Class Presort	0.842167	0.157833
Periodicals	0.950183	0.049817
Single Piece First Class	0.823593	0.176407

Figure 25: On-Time Percentage by Mail Code

Figure 24 demonstrates that most mail types are delivered indeed delivered early; the USPS Marketing Mail is a huge contributor to the 56.96% of mail being delivered early. While the other mail codes still have more mail being delivered early than mail being late.

Compared to the block quote back in Section 3.1 we can compare results. First-Class Mail was said to have an 84.0 percent on-time, and we see that reflected in our analysis. In our analysis we see the First-Class Mail to have an on-time of 84.2 percent.

For Marketing Mail the percentage of on time against the USPS service standard was 92.1 percent. In our analysis, the data shows 94.4 percent. With only a difference of 2.3 percentage points, the difference could be explained by some the differences in the data given to us by USPS.

From our analysis Periodicals had an on-time percentage of 95.0 but with what we researched the percentage of Periodicals was 80.6. This massive gap of 14.4 percent is explained through our cleaning efforts. Periodicals had the biggest number of null values missing in the "EXPECTED_DESTINATION_FACILITY" and thus Periodicals took the biggest hit of entries being removed which did contain late mail deliveries. [37]

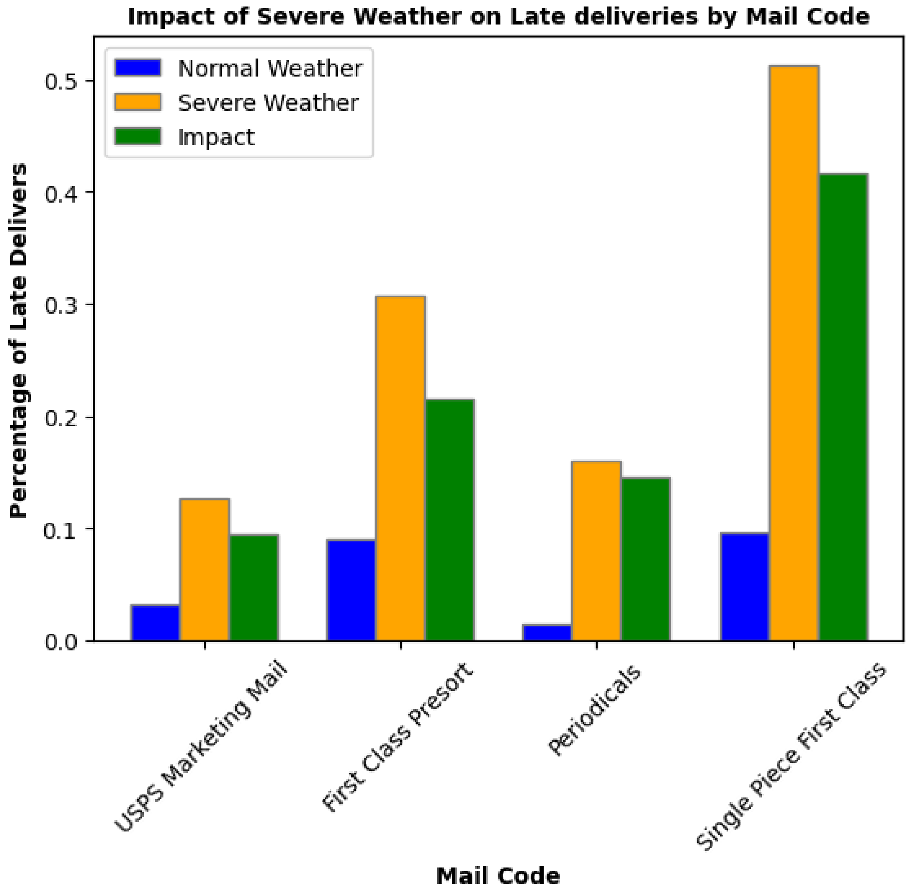

Figure 26: Severe Weather Impact by Mail Code

Mail_Codes	Percent Difference
USPS Marketing Mail	0.598528
First Class Presort	0.542833
Periodicals	0.829020
Single Piece First Class	0.682976

Figure 27: On-Time Percentage Change by Severe Weather

Severe weather events have a significant impact on the volume of late mail deliveries. As can be seen in Figure 26, all mail codes have an increase in their percentage of late mail. Periodicals are impacted the most with 82.9% increase in late deliveries. While First-Class is impacted the least with a 54.3% increase.

3.2 Solution Approach

3.2.1 Systems Architecture

The solution involves the integration of the weather data from the NOAA's GHCN-Daily and the USPS provided mail and package data. GitHub was utilized for weather data storage due to its small size. Private data storage facilities were used to store and process data. Data was exchanged with the customer on a secure web-based server.

Python libraries such as pandas, numpy, matplotlib, seaborn, scikit-learn (including its various modules for preprocessing, model selection, and evaluation), XGBoost, LightGBM, CatBoost, and pgmpy. Spark and Polars will be used for large-scale data pre-processing.

3.2.2 Systems Security

Given the sensitivity of the USPS data, several security measures are implemented in the solution approach. Access to the USPS data is restricted to team members only, using a secure sharing application depending on the size of the data to maintain data access control. The data will be backed up regularly to ensure its integrity and prevent potential data loss.

3.2.3 Systems Data Flows

1. Data Ingestion: Pulled necessary weather data from the NOAA GHCN-Daily, the Storm Events Database and the USPS provided mail and packages data from internal systems.
2. Data Preprocessing: Cleaned and merged the datasets, addressed missing values, addressed outliers and erroneous data entries, removed duplicate entries.
3. Model Training, Exploration and Evaluation
4. Model Selection and Deployment

Figure 28 below demonstrates the above systems data flows described above.

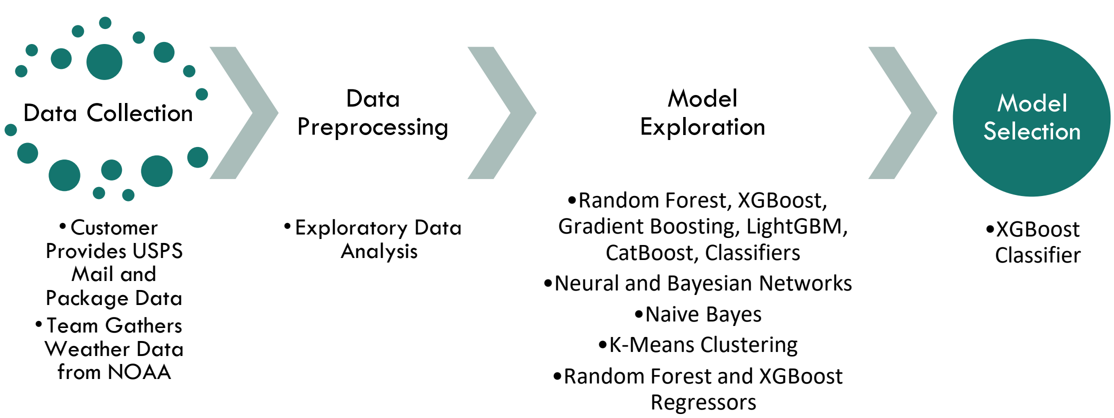

Figure 28: Model Solution Data Flow

3.2.4 Algorithms & Analysis

The solution approach taken for algorithms and analysis is extensive model exploration to determine the best machine learning algorithm for the given data and problem statement. Given the knowledge provided by the exploratory data analysis an optimal algorithm can be selected to fit the system.

The project utilizes the following algorithms and models to predict delivery delays by combining the USPS mail and package data with the NOAA's GHCN-Daily and the Storm Events Database:

- Random Forest Regressor [Framework: Scikit-learn]
  - Use Case: Predicting delivery delays based on weather and package data.
  - Hyperparameters: $n\_estimators = 100$ $random\_state = 42$
  - Deployment: Batch predictions
- K-Means Clustering [Framework: Scikit-learn]
  - Use Case: Grouping similar weather and package data points
  - Hyperparameters: $\mathrm{n\_clusters} = 3, \mathrm{random\_state} = 42$
  - Model Deployment: Batch processing to analyze clusters periodically
- Bayesian Network Analysis [Framework: pgmpy]
  - Use Case: Modeling probabilistic relationships between weather conditions and delivery delays
  - Model Tuning: Cross-validation, alternative model structures tested
  - Model Deployment: Can be used for probabilistic inference on new data
- Random Forest Classifier [Framework: Apache Spark]
  - Use Case: Predicting late shipments using Spark for logistics planning and shipment efficiency
  - Hyperparameters: $\mathrm{numTrees} = 10, 20, 30, \mathrm{maxDepth} = 5, 10, 15.$
  - Model Deployment: Deployed for real-time or batch predictions in a Spark environment or exported to other systems.

3.2.5 Machine Learning

The machine learning process involves two key components, the first is the model exploration and selection and following this is the model training and evaluation. After these two components are completed and the model's performance meets USPS's goals, the trained models can be deployed for batch or real-time processing.

3.3 Machine Learning

3.3.1 Model Exploration

In our goal to forecast package delivery delays, predicting late packages accurately is crucial. We undertook an extensive and detailed examination of a diverse array of machine learning models. In the following sections, we provide a comprehensive analysis of the performance of each model:

Model	Accuracy	AUC	Precision	Recall	F1 Score
Random Forest Classifier	0.8903	0.9221	0.7914	0.6098	0.6888
XGBoost Classifier	**0.8940**	**0.9327**	**0.8900**	**0.8900**	**0.8900**
Gradient Boosting Classifier	0.8715	0.8977	0.8580	0.4247	0.5681
LightGBM Classifier	0.8886	0.9260	0.8315	0.5524	0.6628
CatBoost Classifier	0.8906	0.9271	0.8384	0.5581	0.6701
Neural Network	0.8629	0.8628	0.6887	0.5681	0.6226
Naïve Bayes	0.2482	0.5263	0.2076	0.9857	0.3430

Table 4: Summary of Model Results

Random Forest Classifier: Commencing with the Random Forest Classifier, an ensemble model known for its resilience. It exhibited promising metrics: an accuracy of 0.8903 and an impressive AUC score of 0.9221. Precision for late deliveries stood at 0.7914, indicating a correct prediction rate of approximately 79%, whereas the recall was 0.6098, suggesting some missed late deliveries. The F1 score, providing a balanced view, settled at 0.6888.

XGBoost Classifier [38]: Next, the XGBoost model, another potent ensemble technique, achieved the highest accuracy at 0. 8940 and boasted the highest AUC score of 0. 9327, underscoring its adeptness in distinguishing between on-time and delayed deliveries. Precision for precision, recall, and F1 score all yielded 0.8900 therefore making this the model of choice [38].

Gradient Boosting Classifier [39]: The Gradient Boosting Classifier mirrored similar performance trends with an accuracy of 0.8715 and AUC of 0.8977. Notably, it demonstrated the highest precision for late deliveries at 0.8580 but the lowest recall at 0.4247, resulting in an F1 score of 0.5681, indicative of its cautious approach in predicting delays.

LightGBM Classifier [40]: Employing LightGBM, recognized for its computational efficiency, yielded commendable results with an accuracy of 0.8886 and AUC of 0.9260. It achieved a precision, recall, and F1 score for late deliveries of 0.8315, 0.5524, and 0.6638, respectively, aligning closely with XGBoost's performance metrics.

CatBoost Classifier [41]: CatBoost, tailored for categorical feature handling, exhibited robust performance metrics with the second highest accuracy of 0.8906 and second-highest AUC of 0.9271. Precision, recall, and F1 score for late deliveries were 0.8384, 0.5581, and 0.6701, respectively, establishing its competitiveness.

Neural Network [42]: Introducing a Neural Network yielded moderate results, falling behind tree-based models with an accuracy of 0.8629 and AUC of 0.8628. Its precision, recall, and F1 score for late deliveries were 0.6887, 0.5681, and 0.6226, respectively.

Naive Bayes [43]: Naive Bayes struggled significantly, achieving an accuracy of 0.2482. While exhibiting high recall (0.9857) for late deliveries, precision was markedly low at 0.2076, resulting in an F1 score of 0.3430, suggesting an inclination towards over-predicting late deliveries.

K-Means Clustering [44]: Exploration of K-Means clustering, an unsupervised method, provided insightful but not directly comparable results. The silhouette score of 0.0765 indicated loosely defined clusters, reflecting the complexity inherent in our prediction task.

Bayesian Network [45]: Bayesian Network analysis unveiled probabilistic insights; for instance, during concurrent snow and rain conditions, it predicted a 95.61% chance of on-time delivery and a 4.39% likelihood of severe delay. This counterintuitive finding prompts further scrutiny into data and model assumptions.

Random Forest Regressor [46] and XGBoost Regressor [47]: Transitioning to regression models, the Random Forest Regressor achieved an R-squared score of 0.4795, slightly surpassed by the XGBoost Regressor at 0.4956. These scores underscore the challenge of precise delay prediction, favoring classification approaches for our task.

Moreover, after exhaustive model assessment, the XGBoost Classifier emerged as the preferred choice for our package delivery delay prediction. Its consistent strong performance across multiple metrics, notably the high AUC score indicating robust discriminative ability, and a balanced trade-off between precision and recall for late deliveries, solidify its selection. Moving forward, our focus will be on refining this model and preparing it for deployment in our prediction system.

This decision leverages XGBoost's capability to handle intricate feature interactions effectively, reinforcing its suitability for our operational needs. In the subsequent phase, optimization efforts will aim to enhance its predictive accuracy and operational efficiency.

3.3.2 Model Selection

USPS Packages:

To predict late packages the XGboost classification model produced relatively good results. The XGBoost model is a machine learning algorithm based on the gradient boosting framework. To predict the late USPS packages the parameters identified for the XGBoost model are: colsample_bytree set to 0.9, learning_rate set to 0.2, max_depth set to 5, n_estimators set to 300, and subsample set to 0.9. These parameters were fine-tuned to achieve optimal performance using a cross-validation accuracy of 0.897, indicating that nearly 90% of the predictions were correct during the cross-validation phase. On the test data, the model's accuracy was slightly lower at 0.894, which is still quite high. The Area Under the Curve (AUC) score was 0.933, demonstrating the model's strong ability to distinguish between the classes of late vs non late packages. The classification report provides detailed metrics: for class 0 (on time packages), the precision was 0.91, recall was 0.97, and the F1-score was 0.94; for class 1(late packages), the precision was 0.82, recall was 0.60, and the F1-score was 0.69. The overall accuracy was 0.89, with a macro average F1-score of 0.81 and a weighted average F1-score of 0.89.

Feature importance analysis showed that the most influential features were PRCP_D (precipitation destination), MailClassCode, and time_delta_minutes (duration between first and last scan for a package), among others. Some features, like Distinct_event_scans (number of distinct event scans for a package), had no importance in this model, indicating they did not contribute to the predictive power.

```
Best parameters: {'colsample_bytree': 0.9, 'learning_rate': 0.2, 'max_depth': 5, 'n_estimators': 300, 'subsample': 0.9}
Best cross-validation accuracy: 0.8973724690161676
Accuracy: 0.8940302770636961
AUC: 0.9326720180435
Classification Report:
              precision    recall  f1-score   support

           0       0.91      0.97      0.94      2804
           1       0.82      0.60      0.69       697

    accuracy                           0.89      3501
   macro avg       0.86      0.78      0.81      3501
weighted avg       0.89      0.89      0.89      3501

                 Feature  Importance
12                PRCP_D    0.128770
1          MailClassCode    0.127190
3     time_delta_minutes    0.116278
0        ServiceTypeCode    0.089193
5                 TMIN_O    0.087774
10                TMIN_D    0.085041
7                 PRCP_O    0.067663
9                  Zip_D    0.061041
11                TMAX_D    0.059322
4                  Zip_O    0.050921
8                 SNOW_O    0.046543
13                SNOW_D    0.040666
6                 TMAX_O    0.039598
2   Distinct_event_scans    0.000000
```

Figure 29: XGboost Model results to predict Late (1) vs On-time (0) USPS Packages

Section 4: ML Model Training, Evaluation, and Validation

4.1 Overview

This section details the machine learning model training, evaluation, and validation methodologies utilized to develop a solution to help the Postal Service to predict mail and package delivery delays caused by weather events. These are important steps in determining the most appropriate model for predicting the impact of severe weather on package delivery and mail delivery performance and providing a way to quantify performance by analyzing the volume of late deliveries. As determined through prior literary review and extensive rounds of testing and exploration as detailed in Section 3 of the Project Report, the most effective, efficient, and accurate models for predicting delays in the USPS package data and mail data as a result of weather conditions are generated using the Extreme Gradient Boosting (XGBoost) library. The following section details the development of robust models using XGBoost and highlights all tools, methodologies and processes used in not only the generation of these machine learning models but also those used to ensure the reliability of the model in addition to the rationale for model selection and the hyperparameter tuning processes undertaken.

The foundational step of training the model involves feeding the model historical data from which relationships are learned and evaluated in this case this involves historical weather data and USPS package and mail data. The *Model Training* section will explore the intricate steps taken to preprocess the data and detail handling of missing values and feature engineering. To ensure the model in question is not only accurate but also reliable, the *Model Evaluation* section will detail the validation strategies employed. Lastly, in the *Model Validation* section details regarding the testing of verifying that the model performs well on other data samples and confirming its able to be applied by USPS and that it meets the standards described by the client. Using XGBoost, solutions are achieved by leveraging historical weather data and USPS provided mail and package tracking data, and an analysis of these results is provided to improve the Postal Service's understanding of the impact of weather on their delivery network. Using these models and results, USPS can improve their ability to predict, mitigate, and manage logistical and operational processes in the face of weather-related.

4.2 Visualizations

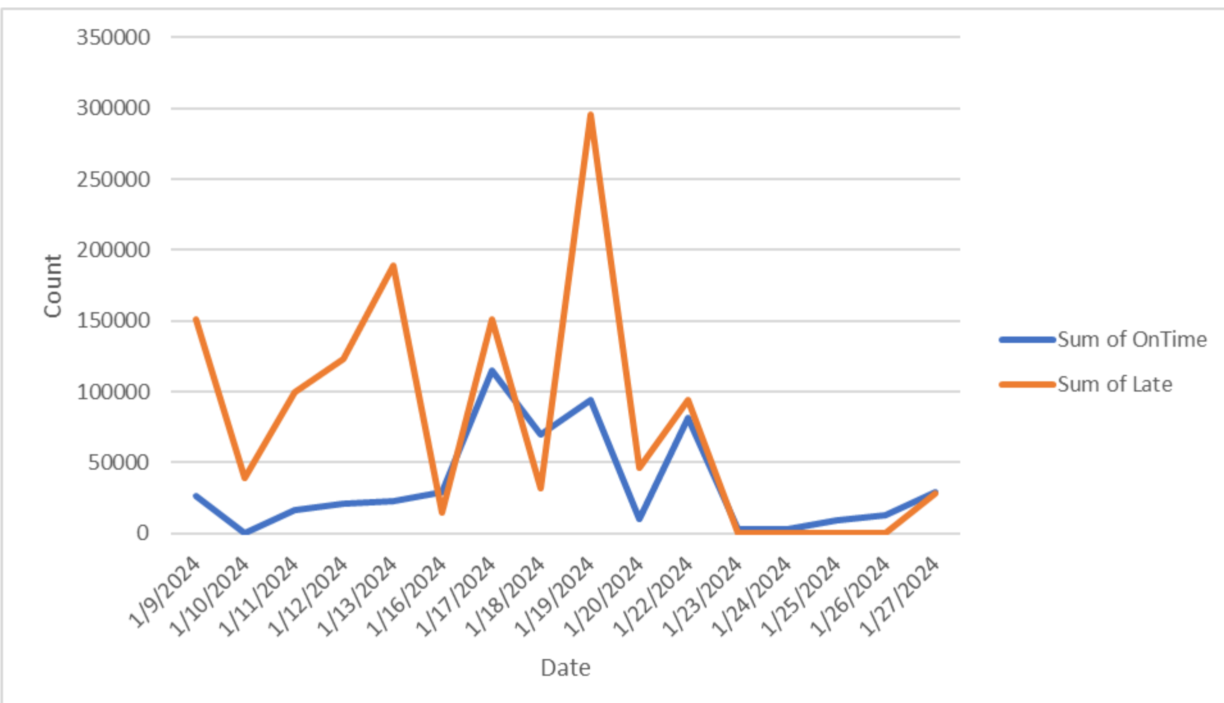

Figure 30: USPS Mail - Late vs On-Time deliveries

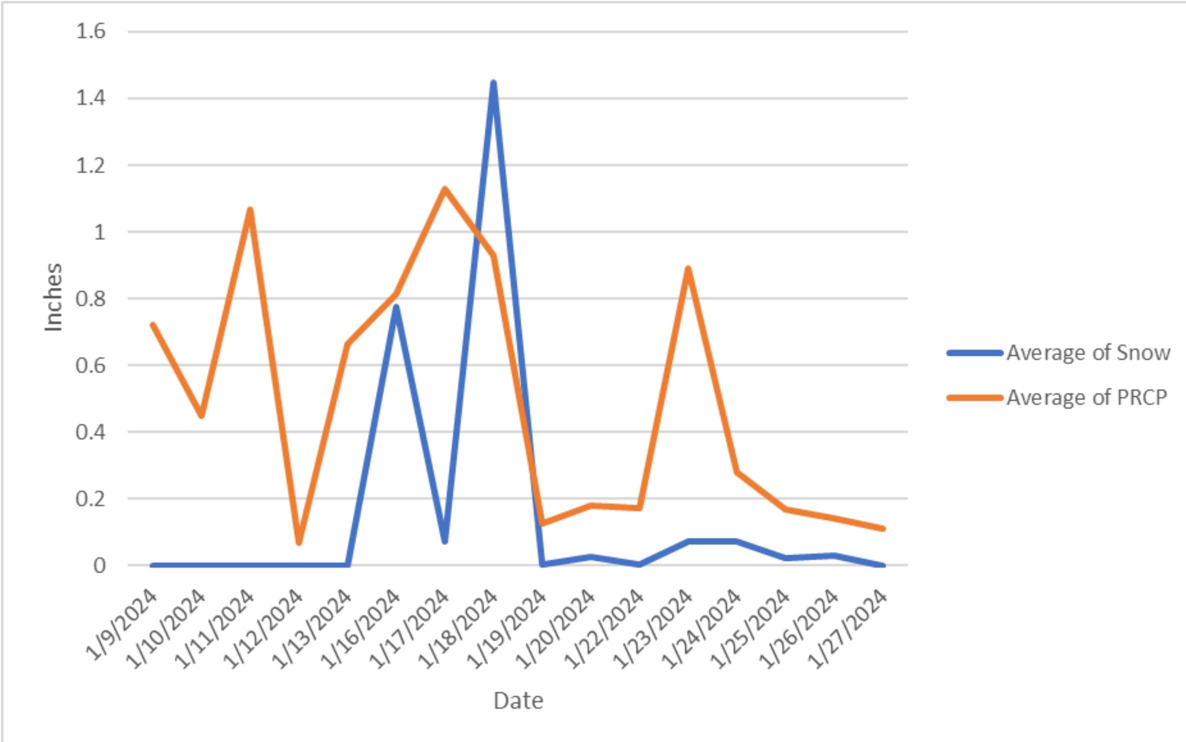

Figure 31: Average Snow and Precipitation Over Time

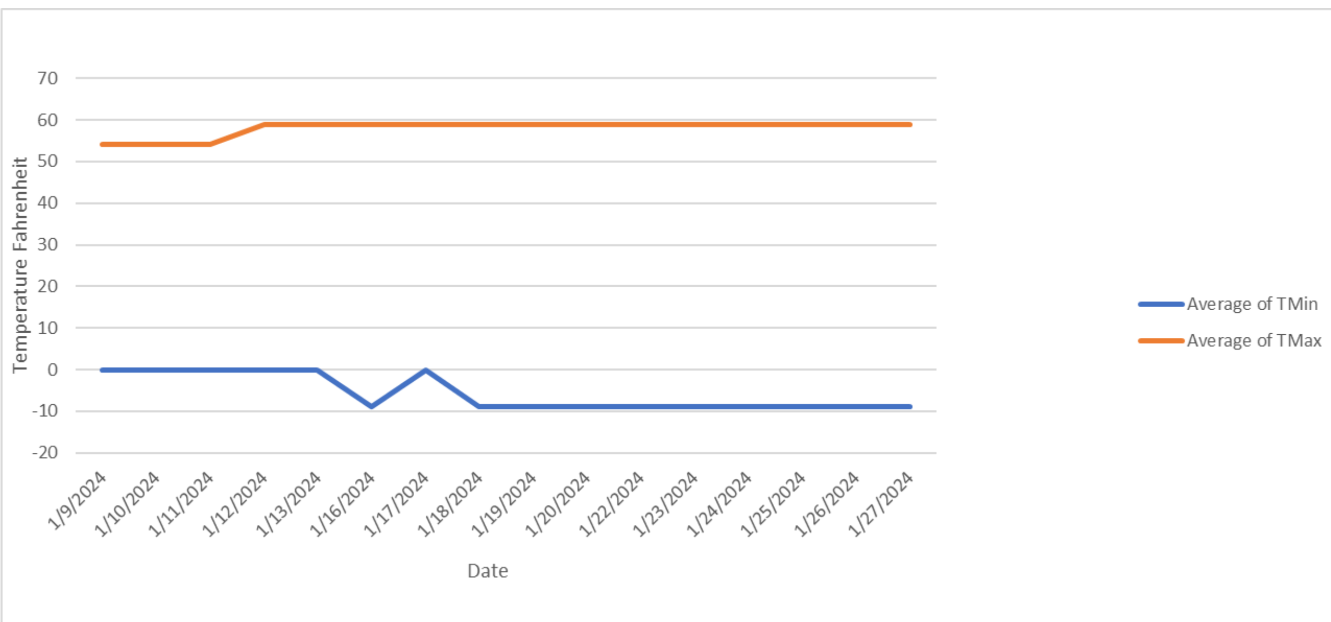

Figure 32: Average Temperature Over Time

We wanted to compare the data that we have received from USPS between the different types of deliveries: Mail and Packages. When comparing Figure 12 to Figure 30 we see a similar spike in late mail deliveries in late January. We also see for mail deliveries, unlike for packages, a high percentage of late deliveries in early January.

The Average snow and precipitation for mail and packages have similar results. With a spike of snow inches around middle of January, and precipitation fluctuating through all January. With these results we can see that the spike in late deliveries in January correlates to the spike in snowfall due to the nature of deliveries being late while snow was on the ground.

Comparing the average temperature for mail and packages we can see that mail deliveries typically have a lower temperature. With TMin for mail deliveries hovering under 0-degree Fahrenheit while for packages you have the lowest TMin around 30-degree Fahrenheit and the highest TMin being 60-degree Fahrenheit. This lower temperature for Mail deliveries can be used to reason why there is a bigger differential between on-time and late deliveries for mail vs packages.

4.3 Machine Learning

4.3.1 Model Training

Once we had selected the XGBoost model based on the results of experiments on package data, our next step was to see how the XGBoost model performed on mail data. We performed our model experiments on the package data because it contained significantly less records than the mail data. As can be seen in Table 5 below, the package dataset contained less than 1% of the data in the mail dataset once the datasets were filtered down to only mail and packages that both originated and destinated to the Music City Annex facility. Due to the large volume of mail data, both Pandas and Pyspark were both inefficient packages for handling the data. Therefore, we utilized the Polars python package to prepare the mail data for model training.

Mail/Packages	# of Records
Mail	1,821,507
Packages	17,503

Table 5: Mail vs Packages Comparison

We decided to test the model on the mail data because if the model performed well on mail data, then it indicated that a single solution could be built for both mail and packages. If the model performed poorly on mail data, however, then further research would be required to determine a solution for mail data.

The first step in preparing the mail data for training was to clean the mail piece data based on the steps outlined in Section 2.4.3 above. Once we had a clean dataset, we filtered the dataset down to only mail pieces where both the origination and destination facility was the Music City Annex. This narrowed down the number of zip codes associated with the mail data, and thereby minimized the computational expense of pulling the weather data.

```
In [78]: zip3

Out[78]: [385, 422, 374, 307, 371, 372, 384, 373, 421, 370]
```

Figure 33: 3-digit Zip Codes - Music City Annex

Unlike the package data, where every record contained an associated 5-digit zip code, the mail data only provided a 3-digit zip code for each record. To pull weather data for the mail data, we needed to pull data for every 5-digit zip code within a 3-digit zip code area. By narrowing the data down to only pieces that originated and destinated from the Music City Annex facility, we reduced the number of 5-digit zip codes from over 43,000 to 482. We pulled 4 weather metrics via the National Oceanic and Atmospheric Administration (NOAA) API: amount of precipitation in inches (PRCP), amount of snow in inches (SNOW), the maximum temperature for the day in Fahrenheit (TMAX), and the minimum temperature for the day (TMIN). Because the NOAA API does not capture all weather metrics for every zip code, the resulting dataset contains many zeros and nulls. To prevent the weather data from getting drowned out by zip codes that did not have a significant weather event, we did not use the mean to aggregate the data for each 3-digit zip. Instead, we aggregated the data using the maximum value for precipitation, snow and maximum temperature, and we used the minimum value to aggregate the data for minimum temperature.

```
weatherdf = weatherdf.group_by(['date', 'zip3']).agg([pl.max('PRCP'),
                                                      pl.max('SNOW'),
                                                      pl.max('TMAX'),
                                                      pl.min('TMIN')])
[98]
```

Figure 34: Aggregation of Mail Weather Data

The origination weather data was joined to the dataset based on the START_THE_CLOCK_DATE feature. The destination weather data was joined to the dataset based on the ACTUAL_DLVRY_DATE. Originally, we joined the destinating data based on EXPECTED_DELIVERY_DATE, however, that model only performed with an accuracy of 75%. Therefore, for this experiment we switched the join column to the ACTUAL_DLVRY_DATE. However, in an implementation scenario the EXPECTED_DELIVERY_DATE attribute would be used to find the weather at the destinating facility since the ACTUAL_DLVRY_DATE attribute would not be available.

The final dataset, seen in Figure 35 below, contained 9 predictive attributes, and 1 predictor label.

	DZIP3	o_PRCP	o_SNOW	o_TMAX	o_TMIN	d_PRCP	d_SNOW	d_TMAX	d_TMIN	isLate
0	384	0.17	0.0	34.0	5.0	2.52	0.0	62.0	48.0	1
1	370	1.27	0.0	57.0	27.0	0.15	0.0	52.0	-13.0	1
2	384	0.17	0.0	34.0	5.0	1.09	0.0	66.0	40.0	1
3	385	1.27	0.0	57.0	27.0	2.11	0.0	67.0	34.0	1
4	371	0.45	0.0	54.0	23.0	1.08	0.0	59.0	24.0	1

Figure 35: Training Dataset - Mail Data

Before we could move on to training the model, we needed to fix the significant class imbalance between late and on-time pieces of mail. Only 6% of the finalized mail data was labeled as late. This means that the model could label every single piece of mail as on-time and still perform with 93% accuracy. Therefore, the dataset would need to be resampled to balance out the predictor classes. Due to the large size of the mail dataset, we used an undersampling technique to reduce the total number of records in the dataset. Figure 36 below shows the number of records in each class before and after the implementation of undersampling.

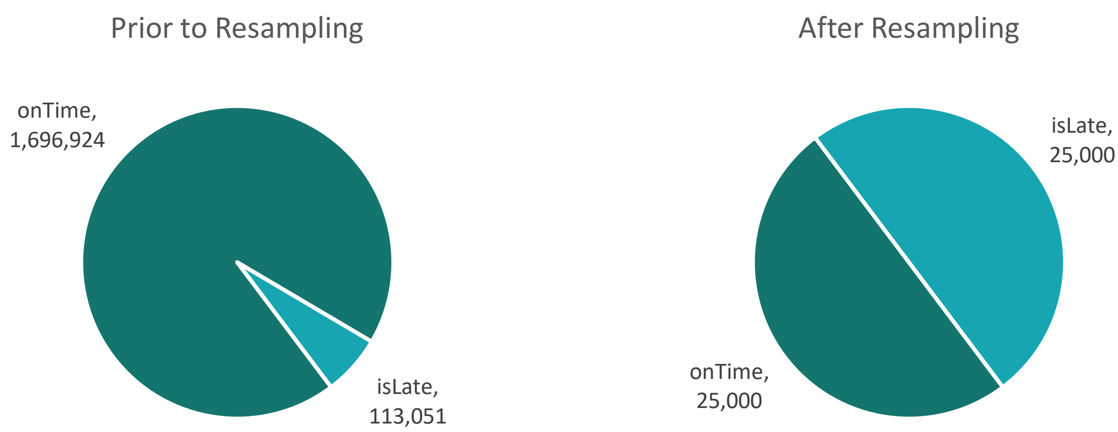

Figure 36: # of Mail Records Pre and Post Resampling

After the data was resampled, we were ready to train the model. We split the dataset into a training dataset and a testing dataset. The training dataset contained 80% of the resampled dataset, and the testing dataset contained the remaining 20% of the training dataset. When splitting the dataset, we ensured that the class balance of the predictive label remained the same in the training and testing dataset as seen in Figure 37 below.

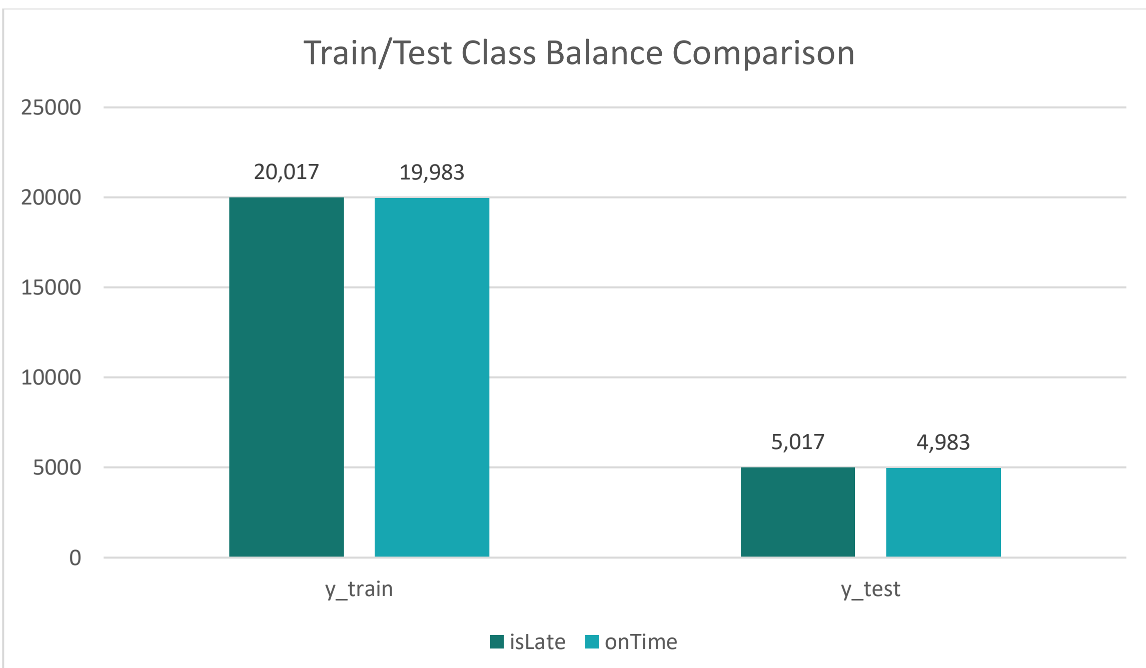

Figure 37: Train/Test Class Balance Comparison

Like we did for the package data, we used a grid search to perform hyperparameter tuning and 5-fold cross validation to train the model.

4.3.2 Model Evaluation

Once the model is trained, we can evaluate the optimal hyper tuning parameters and accuracy determined by the 5-fold cross validation training process. As can be seen in Table 6 below, the turning process for both models mostly selected the same hyperparameters. This is significant because it indicates that the same model and the same hyperparameters can be used to predict the on-time delivery status of mail and packages. Possibly eliminating the need to separate mail and package data in future research.

Parameter	Package Model	Mail Model
colsample_bytree	0.9	0.9
learning_rate	0.2	0.2
max_depth	5	5
n_estimators	300	300
subsample	0.9	0.7
accuracy	0.897	0.937

Table 6: Model Evaluation Comparison

The mail model performed slightly better than the package model. This could be accounted for by the significant differences between the Mail and Package data provided by the USPS. Further research should be conducted to align the datasets and retest the results of the model.

	Feature	Importance
6	d_SNOW	0.205985
3	o_TMAX	0.185548
8	d_TMIN	0.155902
7	d_TMAX	0.128020
1	o_PRCP	0.117358
5	d_PRCP	0.105329
4	o_TMIN	0.047817
2	o_SNOW	0.035249
0	DZIP3	0.018792

Figure 38: Mail Model Important Features

Figure 38 above shows the importance of each feature after training the model. In this model, the amount of snow in the area on the day that a piece of mail is delivered has a 20.6% impact on the model's prediction. The max temperature on the day that the USPS receives the piece of mail has an 18.6% impact on the model's prediction. And the minimum temperature on the day the piece of mail is delivered has a 15.6% impact on the model's prediction.

4.3.3 Model Validation

Next, we took the trained model and used it to predict the delivery status of pieces of mail in the testing dataset. The model performed with similar accuracy on the test data, with an accuracy of 93.2%. The area under the curve score for the model is 98.6%. This indicates that the model's ability to predict between classes is almost perfect. As can been seen in the classification matrix in Figure 39 below, the model is correct over 90% of the time when it is predicting if a piece of mail is onTime (0) or isLate (1). Overall, the results show that the XGBoost model that was selected using Package data can also be used on mail data.

```
Classification Report:
              precision    recall  f1-score   support

           0       0.95      0.92      0.93      5017
           1       0.92      0.95      0.93      4983

    accuracy                           0.93     10000
   macro avg       0.93      0.93      0.93     10000
weighted avg       0.93      0.93      0.93     10000
```

Figure 39: Mail Classification Report

Section 5: Findings

5.1 Overview

Based on the work the group has completed thus far, the following sections detail the major findings regarding the development and exploration of the machine learning model's capacity to predict delayed delivery of USPS mail and packages based on severe weather conditions in additon to insights found during data analysis. Key factors of weather that impact USPS delivery scheduling were identified in the exploratory data anaylsis, conducted in *Section 3.1 Exploratory Data Analysis*, and further explored in the feature imporatance analysis of the machine learning model. Our findings indicate a tremendous amount of practical applications that the XGBoost model and analysis provide to USPS.

5.2 Optimal Model Selection and Model Performance

While multiple machine learning methodologies were explored, as detailed in Table 4 of the report, the XGBoost Classification Model proved itself as the optimal model for predicting delivery delays for both packages and mail for the Postal Service.

The XGBoost Classification Model has shown promising results in predicting delivery delays based on weather data and other relevant features. Our XGBoost model demonstrated high accuracy in predicting delivery delays. For package data, the model achieved an accuracy of 89.4% and an impressive AUC score of 0.933. The mail data model performed even better, with an accuracy of 93.7% during cross-validation in predicting whether the scheduled delivery would arrive on time or late.

The XGBoost Classifier Model generated results that support the use of this machine learning model for the desired purpose and application of USPS, predicting delivery delays due to severe weather events and has demonstrated scalability such that integration and or implementation on the side of USPS will not impede the predictive accuracy.

5.3 Severe Weather Impact Analysis

Our model and analysis confirm that weather events play an important role in model predictions of mail and package being late or on time. Our XGBoost Classifier model indicates that these severe weather events such as temperature extremes and heavy precipitation in fact cause delays in the USPS logistics network.

Specifically, the reported snowfall and precipitation at the destination facility, as well as maximum temperature reported at the originating facility had significant impact on the status for delivery of mail and packages in the USPS network.

Not only do our model predictions highlight the impact of weather on delivery delay patterns for mail and packages, but the results indicate that severe weather influences the USPS logistics network deeply and at both the transportation and processing phases of the mail distribution process. Further insight into this may provide USPS with the foundation for what is neccessary to develop a model to demonstrate the difference between the impact of the weather on the delivery of packages at each step of the USPS delivery process.

Depending on where in the delivery process the mail or package modeled by the data is, the results of the model will provide further insight into how the intended destinating or originating location is able to handle the impact of severe weather and can re-allocate resources and assess employee count in addition to providing USPS delivery service workers with the equipment and necessary manpower to process and or transport the mail and packages.

5.4 Feature Importance Analysis

The feature importance analysis indicates that the difference in the type of severe weather and the variation of the weather variables results in differing levels of impact on the mail and package delivery delay times.

Feature importance analysis indicates that for mail deliveries, the reported snowfall at the destination facility and the reported maximum temperature at the origin facility are the most influential factors in delivery delay of mail. For packages however, we found that the reported level of percipitation at the destination facility and the USPS mail class code were critical factors in predicting delays in package deliveries. In addition, feature importance analysis revealed that the different mail classes are affected differently by weather events, as some classes show more resilience than others.

5.5 Data Quality and Conditioning

The model's accuracy is contingent on the quality of the data. The rigorous data cleaning, normalizing and merging of data sources underwent in order to prepare and process the data frame were essential to the notably successful performance of the XGBoost model. The integration of USPS dataset with weather data has provided a more comprehensive view of delivery performance.

Data quality and consistency are crucial for accuracy in predictions, therefore we have implemented robust data cleaning and preprocessing steps to understand the impacts of weather on mail and package delivery in one specific area of origination and destination.

Due to the quality of the data, consistent data conditioning steps are crucial to ensuring the predictive accuracy and reliability of the XGBoost model. USPS should improve their data collection and cleaning proccesses in order to reduce erroneous predictions by the model due to poor data quality and instead ensure high quality data for model input.

5.6 Practical Implications

The model's capabilities and predictive accuracy are directly aligned with the project's *User Stories* defined in our initial objectives solidified with the client. The ability to accurately model the impact of weather events on the USPS mail and package delivery delays can help the USPS proactively manage schedules and prioritize critically affected packages and mail to mitigate the contributing factors which coincide with severe weather. The insights the model provides are major contributing factors to mitigate delivery risks and delays.

The use of the XGBoost machine learning model to predict USPS delivery delays based on severe weather conditions would allow the Postal Service access to improved accuracy in delay predictions and can be used as a foundation for an enhanced and more advanced system to inform consumers of anticipated weather-related delays by integrating the model with the USPS notification systems. This will vastly improve a current and persistent problem faced by USPS regarding their percieved reliability by bolstering trust in USPS in addition to minimizing the quantity and consistency of consumer complaints and late-delivery inquiries. Additionally, with the findings and precision of the model, it can be used to improve consumer communication with more accurate *Service Alerts* by specifying not only the weather event but the predicted delays for USPS packages and mail due to the type of severe weather event.

Additionally, the Postal Service can implement the XGBoost model to allow them to improve their understanding of the level of potential impact an upcoming or past weather event had on the USPS mail and package delivery delay as well as improve their current *Severe Weather Preparedness* plans in order to allow for more proactive and flexible weather-response plans for the entire USPS network, noting that these are often based on the specific USPS facility. Additionally it provides USPS with the capacity to reroute mail and package deliveries; in addition, the model’s insights provide USPS with a tool to alter the scheduling of deliveries such that when severe weather occurs, USPS can allocate resources and adjust routes and scheduling such that the mail and packages that are most impacted by the predicted weather can be prioritized and delays of critical mail can be mitigated.

The XGBoost model provides USPS with a tool to improve both their current proactive and reactive operational planning. The proven predictive capabilities and accuracy of the model demonstrate that the model can help USPS proactively manage resources, prioritize critical mail classes, and improve communication during severe weather events.

The model can be used by USPS to predict not only the areas and USPS facilities to be impacted by a severe weather event but also can be adopted to identify areas in which severe weather events have a larger scale impact on delivery delay and help USPS determine if this is operational or if this is due to an issue within the delivery networks efficiency in the area overall.

Section 6: Summary

Severe weather events like hurricanes, snowstorms, floods and extreme temperatures significantly impact USPS and partners' operations. These severe weather events cause delays, disrupt delivery routes and pose safety risks to postal workers. Understanding the impact of severe weather events is crucial for improving USPS service resilience and planning effective response strategies.

6.1 Project Achievements

The team successfully developed a machine learning model to predict and quantify the impact of weather events on USPS mail and package delivery times using NOAA weather data and USPS provided mail and package data. The following significant milestones have been achieved:

Data Integration: We successfully cleaned and normalized the USPS and weather datasets based on geospatial location and zip code. Then we were able to effectively merge USPS mail and package data with weather data from the GHCN Daily database, creating a comprehensive dataset for analysis.

Model Development: After exploring numerous algorithms, we selected XGBoost Classifier as our primary modeling approach. We have successfully trained and optimized the model for both mail and package data.

Model Performance: Our XGBoost model has demonstrated high accuracy in predicting delivery delays. For package data, the model achieved an accuracy of 89.4% and an impressive AUC score of 0.933. The mail data model performed even better, with an accuracy of 93.7% during cross-validation.

6.2 Key Takeaways

BY addressing the following key areas, this project effectively gave USPS guidance on how to model the challenges posed by severe weather events, ensuring more reliable and resilient strategies for mail and package delivery services.

Data Quality: Clean and quality data is essential to model accuracy. The accuracy and reliability of the predictions depend on the quality of the data analyzed. Ensuring that data from weather sources like NOAA and USPS is clean, accurate and regularly updated is essential. The seamless integration of weather data with USPS operational data systems will enhance the efficacy of the model.

Weather: Severe weather impacts the delivery of mail and packages as proven through exploratory data analysis. Our analysis revealed crucial insights into how severe weather events affect mail and package delivery performance across different mail classes and weather conditions. Severe weather significantly impacts USPS delivery performance during the processing and transportation phases. The amount of snow at the destination facility and the maximum temperature at the originating facility were significant predictors of delays in mail delivery. The amount of rain at the destination facility and the Mail Class code were critical for predicting package delays.

Mail Classes: Some classes of mail were more weather resistant than others. Certain priority mail types were less affected by servere weather than standard mail types.

Model Performance: The XGBoost Model performed well on both mail and package data. We were able to successfully develop a robust predictive model using XGBoost Classifier, which demonstrates the high accuracy in forecasting delays caused by adverse weather conditions for both mail and package data. XGBoost Classifier outperformed the other machine learning techniques explored including Random Forst Classifier, LightGBM and Neural Networks.

6.3 Recommendations for USPS

The analysis indicates that specific mail classes are more affected by severe weather than others and that select USPS mail classes have demonstrated more resilience to severe weather conditions. This knowledge will allow USPS to manage and allocate their resources and machinery in preparation for bad weather forecasts.

Based on the performance of the XGBoost model, the next steps in *Section 7: Future Work* are laid out to best enable USPS to utilize the machine learning model in real-time to provide USPS with data-driven insights that have demonstrated in testing, the ability to help circumvent the impact of weather on delivery delay for mail and packages.

Section 7: Future Work

7.1 Overview

To ensure that the model can easily be adopted by USPS and eventually be implemented across the entire USPS network to enable real-time weather delay predictions and allow for better response planning, the following steps are recommended to be taken to ensure that the model can be ideally adopted into the USPS network and employed throughout the enterprise:

- Testing Model Robustness
- Pilot Testing
- USPS Adoption
- Continued Monitoring

7.2 Test Model Robustness

The next step of this project is to expand the model application by applying the XGBoost model to data from different originating and destination locations to validate its overall robustness across different geographic regions with diverse weather conditions.

Additionally, testing of a data sample comprised of mail and packages with distinctly different origination and destination points is essential to confirm that the XGBoost model can maintain its high predictive accuracy and classification precision of delayed mail and packages with USPS data comprised of USPS mail and packages with distinctly different origination and destination points.

Lastly, we recommend testing the model's performance across different seasons to ensure it maintains accuracy throughout the year. Weather patterns can vary significantly between seasons, and this testing will confirm the model's reliability across all weather conditions USPS may encounter.

If the model performs well for mail and packages as observed in previous experiments detailed in Section 4 of the report, we recommend that USPS conduct a pilot test of the model using a real-time severe weather event.

7.3 Pilot Testing

The phase of pilot testing during a real-time severe weather event is important as it will allow the model’s real-world applicability to be thoroughly evaluated and allow the model to be fine-tuned if necessary.

This is especially important as the model deals with dynamic weather conditions and needs to be able to adapt. Following these pilot tests, the model will be refined at this point such that the model’s predictions are worked into the USPS operational workflow.

7.4 USPS Adoption

Enterprise-wide adoption is our recommendation to USPS assuming successful pilot testing, as the model should be integrated into existing USPS systems and across the network, successful pilot testing and validation support the scalability of the XGBoost machine learning model.

7.5 Continued Monitoring

We recommend the Postal Service invest in continuous data quality improvement and regular model retraining for the XGBoost model to maintain high predictive accuracy and ensure the predictive capability of the model is not eroded as new data becomes available as well as institute checks. The importance of continuous monitoring and model retraining is prevalent in current supply chain disruption literature.

We also recommend establishing a feedback loop with USPS operations teams. This will enable continuous validation and improvement of the model based on real-world outcomes. Regular input from those on the front lines of mail and package delivery will ensure the model remains practical and aligned with operational realities.

Appendix

Appendix A: Glossary

Term	Definition
annex	A subsidiary building separate from a parent network mail processing facility or delivery unit (called carrier annex) that supports the need for additional operational floor space [48].
average days to deliver	The average number of days it took the mail from collection to delivery in your geographic area. Actual delivery days include all days in which market-dominant products are eligible for delivery, not including Sundays, holidays, and other exclusions [49].
canceled trips	It occurs when the Postal Service cancels a trip for assorted reasons, or when the contractor fails to perform the scheduled trip [50].
destinating mail	Incoming mail arriving for its point of final delivery (destination) through a processing facility [48].
destination	(1) The intended or actual final delivery point for mail. (2) A qualifier that identifies where mail is to be delivered such as destination ZIP Code or where mail is to be entered such as destination delivery unit [48].
extra trips	An unplanned additional truck transportation trip for an existing route, and this leads to increased transportation costs [50].
late trips	Transportation truck trips arriving or departing after scheduled time [50].
Network Distribution Centers (NDCs)	A highly mechanized and automated mail processing facility, NDCs consolidate the processing of mail to increase operational efficiency, decrease costs and maintain service while expanding the surface transportation reach. Formerly Bulk Mail Centers (BMCs) [48].
Nor’easters	Are storms over the East Coast of the U.S. with winds typically from the northeast [18] and are classified as general severe weather as the weather phenomenon associated with the storm covers a large geographic area.
originating mail	Outgoing mail and local mail that enter the mailstream (that is, the point of origin) for mail processing and delivery [48]
origin facility	The point of entry used by a mailer presenting a mailing [48].
origin/optional entry three-digit	A presort level in which the ZIP Code in the delivery address on all pieces begins with one of the three-digit ZIP Code prefixes processed at the sectional center facility (SCF) in whose service area the mail is verified/entered. Subject to standard, a separation (i.e., separate bundle or container) is required for each three-digit ZIP Code prefix area regardless of mail volume. Can also be written as origin/optional 3-digit(s) [48].
origin ZIP Code™	The ZIP Code in which mail is prepared or the mail is entered into the mailstream [48].
Processing and Distribution Centers (P&DC)	A central mail facility that distributes and dispatches part or all of both incoming mail and outgoing mail for a designated service area. It also provides instructions on the preparation of collection mail, dispatch schedules, and sorting plan requirements to mailers. The facility is usually a sectional center facility or a general mail facility, but it can also be a dedicated mail processing facility without a station or branch [48].

service standards	Stated delivery performance goals for each mail class and product that are usually measured by days for the period of time taken by USPS to handle the mail from end-to-end (that is, from the point of entry into the mailstream to delivery to the final destination). Established service standards also include destination entry standards for mail entered by the mailer at or near a postal destination facility [48].
severe thunderstorms	Storms with hail one inch in diameter or greater, winds from 58 to 75 miles per hour or a tornado [19].
severe weather	Any dangerous meteorological phenomenon with the potential to cause damage, serious social disruption or loss of human life [17].

Table 7: Glossary Table

Appendix B: GitHub Repository

Overview

Below is the link to the GitHub repository for this project, and the contents of the Read.md file associated with the repository.

GitHub Repository Link

https://github.com/JDRitenour/teamkangaroo

Installation

The following packages are required to run the files in this repository: Pandas, Numpy, Matplotlib, Seaborn, Polars, Sklearn, XGBoost, LightGBM, CATBoost, Pgmpy, shap.

Project Goals

The United States Postal Service (USPS) aims to increase knowledge about how weather events impact mail and package processing, transportation and delivery. The USPS has partnered with George Mason University students to build a machine learning model that will use publicly acquired weather data to predict if a piece of mail or a package will be late.

Can the weather predict if a piece of mail or a package will be late?

File Descriptions

Package_Notebooks

- Package_EDA.ipynb - Exploratory Data Analysis on USPS Packages
- package_scans_cleaned.ipynb - Data Conditioning on the Package Scans Data
- package_data_final.ipynb - Data Conditioning to combine package and weather data
- MODEL_EXPLORATION.ipynb - Assesses a Random Forest, XGBoost, Gradient Boosting, LightGBM, CatBoost, Neural Network, and Naive Bayes model on the USPS Package Data
- RF&XG_REGRESSOR_MODEL.ipynb - Assesses a Random Forest Regressor and XGB Regressor Model on the USPS Package Data
- XGBOOST_USPS_Packages.ipynb - Assesses a XGBoost model on the USPS Package Data

Mail_Notebooks

- mail_data_cleaned.ipynb - Clean and pull the weather data for the USPS mail data
- xgboost_mail.ipynb - Assess the XGBoost Model on the USPS mail data

Results

We found that weather does play a key role in determining is a piece of mail or a package will be on time or late. The XGBoost model performed the best. The package data model had an accuracy of 0.894 and an AUC of 0.933. The mail data model had an accuracy of 0.93 and an AUC of 98.6%.

Licensing, Authors, Acknowledgements

The data used in this project is the property of the United State Postal Service, and the National Oceanic & Atmospheric Administration. This project was completed by students at George Mason University.

Appendix C: Risks

Sprint 1 Risks

Risk	Description	Probability	Impact	Mitigation
Scope Management	The project scope may expand beyond the initial problem statement, leading to delays or incomplete deliverables.	Medium	Medium	Clearly define and document project scope, regularly review and validate scope, prioritize tasks based on impact and alignment, establish a change management process, and communicate the impact of scope changes to stakeholders.
Gap in Domain Knowledge	Team members may lack specific domain knowledge or technical skills required for certain aspects of the project	Medium	Low	Identify gaps early, allocate resources for training, encourage knowledge sharing, and leverage documentation and external resources.
Availability and Quality of Data	Required data from NOAA, NWS, FEMA, or USPS may be incomplete, inconsistent, or difficult to obtain.	Medium	High	Prioritize critical data sources, establish clear communication with USPS, and develop a data validation process.
Biased Sample Dataset	The USPS-provided sample data for mail tracking and service performance may contain biases, leading to inaccurate conclusions about the impact of weather events on mail delivery.	Medium	High	Examine sample data for biases, collaborate with USPS to understand data collection, employ statistical techniques to address biases, and communicate limitations or concerns.

Table 8: Sprint 1 Risks

We identified 4 major risks that we have encountered throughout this sprint or expect to encounter throughout this project. The first risk we encountered is scope management. Scope creep is a concern in any project. Therefore, we made sure to clearly identify the client's expectations for the outcome of this project. One example of where lack of communication can cause scope creep did occur during this sprint. The initial project request explained that increased delays caused by inclement weather impact the USPS on-time percentage scores. Therefore, our initial project objectives included an updated calculation of on-time percentage scores. After presenting our initial objectives to the client, it was clarified that they did not expect our solution to provide an update on-time percentage score, but instead they wanted to be able to predict the number of days mail or packages would be delayed. This shows that clear communication is essential for defining a projects objectives and solution space.

The second risk we identified is our gap in domain knowledge. None of the team members have any experience with the logistical and operational systems within the United States Postal Service. Therefore, we could not identify what kind or severity of weather would cause delays with the USPS deliveries of mail and packages. As a result, we are relying on the expertise of our client to identify weather events that we should focus on for this project.

Our third risk is the availability and/or the quality of data. Our project relies on being able to merge data provided by the USPS with data we can gather regarding weather events. Since we must rely on the USPS to pull the data we require for this project, it could be difficult to ensure that we have the proper fields to merge the data. We will have to continue to work closely with our client to ensure that the data provided meets the needs of our project.

The last risk we identified in Sprint 1 is the risk of a biased dataset. Weather is not the only impact on the operational and logistical systems within the United States Postal Service. During our research, we discovered that the USPS has been undergoing system and process changes that are designed to increase the efficiency of the USPS. Change management is complex, and some changes have resulted in increased transit time for mail and/or packages. We will have to work closely with the client as data is gathered to ensure that the facility and timeframe we select for our project is not also being impacted by the changes within the USPS system.

Sprint 2 Risks

Risk	Description	Probability	Impact	Mitigation
External Data Quality	Potential quality issues with gathered weather data.	Low	Medium	Vet the dataset thoroughly, consider quality checks on ingested data, lastly have contingency plan for data quality issues.
Data Scalability	Challenges in processing and storing large volumes of weather and USPS data as the project scales.	Low	Medium	Assess data storage and processing requirements, leverage big data technologies (i.e., Cluster), and consider cloud-based solutions.
Data Consistency	Inconsistencies across weather stations in data points collected.	Medium	Low	Focus on most consistently available data points and document limitations as they are encountered.
Completeness of Dataset	Dataset may contain null values and higher number of outliers, which can impact analysis if not handled appropriately.	Medium	Low	Develop clear strategy for handling null values (i.e., imputation) also testing model sensitivity.

Table 9: Sprint 2 Risks

We identified four major risks that we have encountered throughout this sprint or expect to encounter throughout this project. The first risk we encountered is external data quality. Potential quality issues with the gathered weather data could affect our project's outcomes. Therefore, we have implemented thorough vetting of datasets, established quality checks on ingested data, and created a contingency plan to address any data quality issues that arise.

The second risk we identified is data scalability. As the project scales, we may face challenges in processing and storing large volumes of weather and USPS data. To mitigate this risk, we assessed our data storage and processing requirements early on. We have decided to leverage big data technologies, such as clusters, and are considering cloud-based solutions to ensure our infrastructure can handle the growing data demands.

Our third risk is the consistency of data. There are inconsistencies across weather stations in the data points collected, which can impact our analysis. To address this, we are focusing on the most consistently available data points and are documenting any limitations we encounter. This approach will help us maintain the integrity of our analysis despite these inconsistencies.

The last risk we identified in Sprint 2 is the completeness of the dataset. Our dataset may contain null values and a higher number of outliers, which can impact our analysis if not handled appropriately. We have developed a clear strategy for handling null values, such as imputation, and are testing model sensitivity to ensure our analysis remains robust.

In summary, during Sprint 2, we focused on identifying and mitigating risks related to external data quality, data scalability, consistency of data, and the completeness of the dataset. By addressing these risks proactively, we aim to maintain high data quality and scalability, ensuring reliable project outcomes.

Sprint 3 Risks

Risk	Description	Probability	Impact	Mitigation
Model Overfitting	Random Forest Classifier may overfit training data which can lead to weak genera lization on unseen data.	Medium	High	Focus on proper cross validation techniques, carefully tune hyperparameter, and do regular testing on a held-out validation set.
Spark Infrastructure Issues	Spark instance for model training and deployment may experi ence technical issues.	Low	High	Implement regular backups and have an alternative plan that uses a different computational resource
Feature Engineering Complexity	Issues in combining weather, geographical, and USPS data into useful features for the model.	Medium	High	Allot additional time for feature engineering given the large dataset size with merge happening, con sult with client, and refine feature based on model performance and feedback
Data Inconsistency	Variations and noise in USPS provided data can lead to poor model predictions.	Medium	High	Perform thorough data cleaning while looking at data summaries along the way and use techniques to handle missing or inconsistent data.
Interpretability of Model	Complexity of models may make it harder to interpret and explain to client.	Low	Medium	Apply feature importance metrics, clearly document along the way, and consider alternative models that may deliver simpler interpretations.

Table 10: Sprint 3 Risks

We identified five major risks that we have encountered or anticipate encountering throughout this sprint and the project as a whole. The first risk we addressed is model overfitting. The Random Forest Classifier may overfit training data, which can lead to weak generalization on unseen data. This risk has a medium probability but a high impact. To mitigate this, we focused on proper cross-validation techniques, carefully tuned hyperparameters, and conducted regular testing on a held-out validation set. We're pleased to report that this risk has been successfully mitigated.

The second risk we identified is Spark infrastructure issues. While the probability is low, the potential impact is high as Spark instances for model training and deployment may experience technical issues. To address this, we implemented regular backups and developed an alternative plan that uses a different computational resource. This risk has also been successfully mitigated.

Our third risk is feature engineering complexity. We faced medium probability and high impact challenges in combining weather, geographical, and USPS data into useful features for the model. To mitigate this, we allotted additional time for feature engineering given the large dataset size. We're also continuously consulting with the client and refining features based on model performance and feedback. This risk has been successfully addressed.

The fourth risk we identified is data inconsistency. There are variations and noise in USPS provided data that can lead to poor model predictions, presenting a medium probability and high impact risk. To combat this, we performed thorough data cleaning while looking at data summaries along the way and employed techniques to handle missing or inconsistent data. We're pleased to report that this risk has also been mitigated.

Lastly, we recognized a risk in the interpretability of our model. The complexity of our models may make it harder to interpret and explain to the client. While this risk has a low probability, it has a medium impact. To address this, we applied feature importance metrics, clearly documented our process, and considered alternative models that may deliver simpler interpretations. This risk has been successfully mitigated as well.

In summary, during Sprint 3, we focused on identifying and mitigating risks related to model overfitting, Spark infrastructure issues, feature engineering complexity, data inconsistency, and model interpretability. By addressing these risks proactively and implementing effective mitigation strategies, we have successfully navigated all identified challenges. This approach has allowed us to maintain high-quality data processing, robust model development, and clear communication with the client, ensuring reliable project outcomes.

Sprint 4 Risks

Risk	Description	Probability	Impact	Mitigation
Model Performance Degradation	The model's performance may deteriorate when applied to new, unseen data	Medium	High	Implement regular model retraining, monitor performance metrics, and use cross-validation methods.

Integration Challenges	Difficulties in integrating the trained model into existing USPS systems	Medium	High	Collaborate closely with USPS data team, develop a detailed integration plan, and conduct thorough testing
Data Drift	Changes in data distribution over time may affect model accuracy	Medium	Medium	Implement data drift detection mechanisms, regularly update the model with new data
Interpretability Issues	Difficulty in explaining model predictions to stakeholders	Low	Medium	Use interpretable ML techniques, develop clear visualization tools, and provide thorough documentation
Model Deploying Latency	Potential delays in real-time predictions due to model complexity	Low	High	Optimize model for inference speed, consider model compression techniques, and implement efficient serving infrastructure

Table 11: Sprint 4 Risks

In Sprint 4, we identified five key risks associated with the ML model training, evaluation, and validation phase of our project.

The most critical risk is the potential for model performance degradation when applied to new, unseen data. This risk has a medium probability but high impact, as it could significantly affect the accuracy of our weather-related delay predictions. To mitigate this, we are implementing regular model retraining and monitoring performance metrics closely.

Integration challenges pose another significant risk. As we prepare to deploy our model into existing USPS systems, we anticipate potential compatibility issues. We are addressing this by collaborating closely with the USPS IT team and developing a detailed integration plan.

Data drift is a medium probability, medium impact risk that could affect our model's accuracy over time. We are implementing data drift detection mechanisms and planning for regular model updates to mitigate this risk.

Moreover, we recognize the importance of model interpretability for our stakeholders. We're addressing this low probability, medium impact risk by incorporating interpretable ML techniques and developing clear visualization tools.

Lastly, regarding model deployment latency, the probability is low, the impact could be high if our model can't provide predictions quickly enough for real-time decision making. We are mitigating this by optimizing our model for inference speed and considering model compression techniques.

Sprint 5 Risks

Risk	Description	Probability	Impact	Mitigation
Presentation Time Management	Difficulty in condensing project information into the allotted presentation time	Medium	High	Conduct multiple rehearsals, prioritize key findings, and use visual aids effectively
Inconsistencies in Final Report	Discrepancies or contradictions in different sections of the report due to multiple contributors	Low	Medium	Implement a thorough review process with cross-checking by team members
Technical Difficulties During Presentation	Potential issues with presentation software or hardware during the showcase	Low	High	Prepare backup options (for instance, offline version of slides), test equipment in advance
Complexity in Communicating Technical Results	Difficulty in explaining complex machine learning concepts and results to a non-technical audience in the final presentation	Medium	Medium	Develop clear, simplified explanations and visual aids; practice presenting to non-technical team members for feedback

Table 12: Sprint 5 Risks

For Sprint 5, we identified four key risks associated with the final report preparation and showcase presentation phase of our project.

The most critical risk is presentation time management. With a medium probability and high impact, condensing our extensive project information into the allotted presentation time poses a significant challenge. To mitigate this, we are conducting multiple rehearsals, prioritizing key findings, and developing effective visual aids. This approach ensures we deliver a concise yet comprehensive overview of our project within the time constraints.

Inconsistencies in the final report present another risk, albeit with low probability and medium impact. Given multiple contributors to the report, there's a possibility of discrepancies or contradictions across different sections. We're addressing this by implementing a thorough review process with cross-checking by team members, ensuring a cohesive and accurate final document.

Technical difficulties during the presentation, while having a low probability, could have a high impact on our showcase. To mitigate this risk, we're preparing backup options such as offline versions of our slides and thoroughly testing all equipment in advance. This preparation ensures we can smoothly deliver our presentation even if unexpected technical issues arise.

Lastly, we recognize the complexity in communicating technical results as a medium probability, medium impact risk. Explaining intricate machine learning concepts and results to a potentially non-technical audience during the final presentation could be challenging. To address this, we're developing clear, simplified explanations and visual aids, and practicing our presentation with non-technical team members to refine our communication approach.

By proactively addressing these risks, we aim to ensure smooth, effective final report delivery and showcase presentation, effectively concluding our project on a high note.

Appendix D: Agile Development

Scrum Methodology

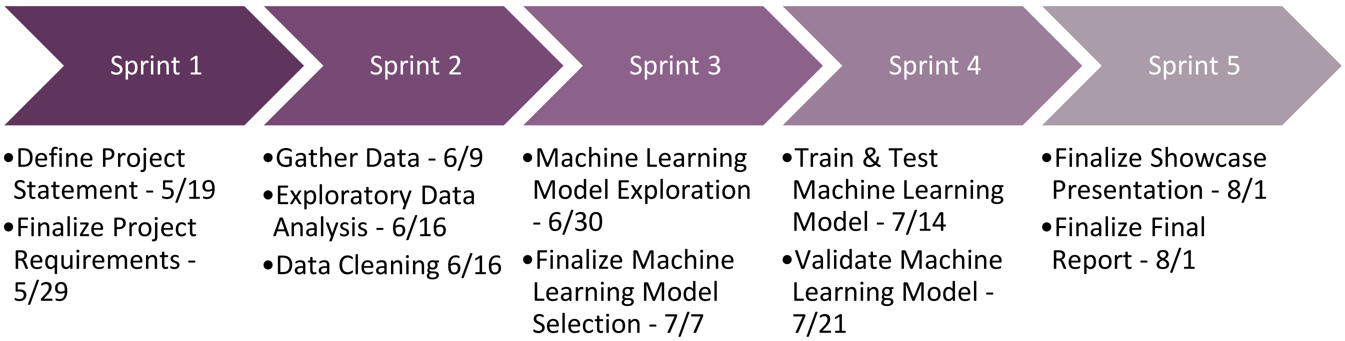

Figure 40: Sprint project dates

Team Kangaroo conducted this project in 5 sprints over a 12-week period. We conducted our scrum stand-ups virtually via a discord chat. Overall, this approach worked well. However, we suggest the team put a recurring weekly meeting on the calendar so that all team members are available if the team needs to meet. We used the YouTrack dashboard to track UserStories throughout the sprints. This was a good tool and allowed the team to see what needed to be accomplished to each sprint, what was already being worked on, and what was already completed. Below we provided a breakdown of our methodologies and lessons learned from each sprint.

Sprint 1 Analysis

During Sprint 1 we focused our user stories on tasks that were required to complete the deliverables for Sprint 1. This included defining and refining the Problem Statement, conducting research into the project background and the solution space, and building out our solution space, primary user story and product vision. We performed weekly scrum meeting weekly where we assigned each team member a primary task based on the deliverables for that week. The team was in constant communication via the team's discord where we shared ideas and research to ensure that all team members were familiar with the growing body of knowledge associated with this project. One challenge with this part of the project is that we only met with the client once a week and one week we did not meet at all. During this initial sprint, our understanding of the of the problem evolved quickly and meeting with the client more often would have allowed us to be better prepared for Sprint 2.

Sprint 2 Analyis

Our User Stories in Sprint 2 were focused on the tasks that we required to gather the required information to complete Section 2 of the Final Report. We pulled the weather data that will be required to completed our project from the Nation Centers of Environmental Research. The dataset required very minimal conditioning. Ultimately we were not able to complete everything we wanted to in this sprint because we did not receive the mail and packages dataset from our client, the United States Postal Service. Therefore, we will need to continue to follow-up with our client during the initial part of Sprint 3. Fortunately, Sprint 3 is three weeks long so we should be able to accomplish any lingering tasks from Sprint 2 as well as complete the necessary tasks in Sprint 3.

Sprint 3 Analysis

For Sprint 3, we decided to keep our user stories broad and focus them on the tasks that needed to be completed for Section 3 of the Final Report. We decided on this broad approach in order to allow team members the opportunity to be creative in their experiments for a solution to our problem statement. For our solution we explored a Linear Regression model, a Random Forest Regression model, a Random Forest Classifier, a Gradient Boosting Classifier, a LightGBM Classifier, a XGBoost Classifier, and Naive Bayes Model, and a Nueral Network Model. Ultimately we found that the XGBoost Classifier Model performed the best across a wide range of metrics.

Sprint 4 Analysis

In Sprint 4, our primary focus was on applying the machine learning model training, evaluation, and validation methodologies to develop a solution for predicting mail and package delivery delays caused by weather events. These steps are crucial for determining the most suitable model for predicting the impact of severe weather on delivery performance and quantifying the volume of late deliveries. These steps are crucial in determining the most suitable ML model for our usecase to predict USPS mail and package delivery delays caused by weather events. Using the XGBoost library, we preprocessed and integrated historical NOAA weather and USPS data, addressed class imbalances, applied feature engineering, and performed hyperparameter tuning. The models were evaluated and validated, showing high accuracy and reliability in predicting delays, with key weather factors like precipitation, snow and temperature significantly impacting predictions. The results indicate that a unified model for both mail and package deliveries is feasible, providing USPS with a robust tool to manage weather-related delivery challenges.

Sprint 5 Analysis

Durring Spring 5, we focused on building the PowerPoint presentation for our Final Presentation, finished the final sections of the Final Report, and commited the finalized versions of our code to GitHub. Our biggest challenge during this sprint was getting our presentation within the 20-25 minutes timeframe. We found that using a script during the presentation helped ensure that each team member’s presentation was kept within the required timeframe.

This page intentionally left blank

Reference

[1] P. Pandya, R. Gujar and V. Vakharia, "Modeling and Prediction of Freight Delivery for Blocked and Unblocked Street Using Machine Learning Techniques," *Transporation Research Procedia,* pp. 555-561, 15 September 2020.

[2] S. M. Ali, A. B. M. Mainul Bari, A. A. M. Rifat, M. Alharbi, S. Choudhary and S. Luthra, "Modelling supply chain disruption analytics under insufficient data: A decision support system based on Bayesian hierarchical approach," *International Journal of Information Management Data Insights,* vol. 2, no. 2, 2022.

[3] A. B. Smith, "2023: A historic year of U.S. billion-dollar weather and climate disasters," climate.gov, 8 January 2024. [Online]. Available: https://www.climate.gov/news-features/blogs/beyond-data/2023-historic-year-us-billion-dollar-weather-and-climate-disasters. [Accessed 26 May 2024].

[4] E. Katz, "As USPS institutes network reforms, mail delivery hits a 3-year low," Government Executive, 22 February 2024. [Online]. Available: https://www.govexec.com/management/2024/02/usps-institutes-network-reforms-mail-delivery-hits-three-year-low/394388/. [Accessed 29 May 2024].

[5] P. Warren, "USPS data reveals Houston-area performance has dropped below previous standards since November," Houston Chronicle, 24 February 2024. [Online]. Available: https://www.houstonchronicle.com/news/houston-texas/trending/article/usps-mail-delays-houston-data-18679802.php. [Accessed 31 May 2024].

[6] E. Katz, "As USPS institutes network reforms, mail delivery hits a 3-year low," Government Executive, 22 February 2024. [Online]. Available: https://www.govexec.com/management/2024/02/usps-institutes-network-reforms-mail-delivery-hits-three-year-low/394388/. [Accessed 26 May 2024].

[7] United States Postal Service, "Service Alerts," United States Postal Service, 24 May 2024. [Online]. Available: https://about.usps.com/newsroom/service-alerts/. [Accessed 28 May 2024].

[8] 21st Century Postal Worker, "Learn about UPSP areas, regions," 21st Centry Postal Worker, 20 January 2021. [Online]. Available: https://www.21cpw.com/learn-about-usps-areas-regions/. [Accessed 29 May 2024].

[9] Postal Work, "USPS Customer Service District Offices," Postal Work, [Online]. Available: https://postalwork.net/district_offices/#EASTERN_AREA. [Accessed 29 May 2024].

[10] United States Postal Service, "USPS.com FAQs," United States Postal Service, 20 March 2024. [Online]. Available: https://faq.usps.com/s/article/No-Mail-Delivery . [Accessed 30 May 2024].

[11] K. Coleman, "USPS Warns Customers to Allow Extra Time for Delivers - Here's Where," Best Life, 27 March 2024. [Online]. Available: https://bestlifeonline.com/usps-deliveries-extra-time-warning-news/. [Accessed 30 May 2024].

[12] D. J. Lynch, "Bridge collapse disrupts key trade lane, promting new supply chain woes," Washington Post, 27 March 2024. [Online]. Available: https://www.washingtonpost.com/business/2024/03/26/baltimore-bridge-supply-chain/. [Accessed 30 May 2024].

[13] United States Postal Service, "A decade of facts and figures," United States Postal Service, 2023. [Online]. Available: https://facts.usps.com/table-facts/. [Accessed 26 May 2024].

[14] S. Pettypiece, "Major mail delivery delays raise concerns about voting in the 2024 elections," yahoo!news, 6 April 2024. [Online]. Available: https://www.yahoo.com/news/major-mail-delivery-delays-raise-130000630.html. [Accessed 30 May 2024].

[15] United States Postal Service, "Delayed mail and packages?," United States Postal Service, 21 January 204. [Online]. Available: https://faq.usps.com/s/article/Delayed-Mail-and-Packages. [Accessed 28 May 2024].

[16] USPS, "Retiree Health Benefits Prefunding," 2023. [Online]. Available: https://about.usps.com/who-we-are/financials/annual-reports/fy2010/ar2010_4_002.htm.. [Accessed 26 May 2024].

[17] Wikipedia, "Severe weather," Wikimedia Foundation, Inc, 23 May 2024. [Online]. Available: https://en.wikipedia.org/wiki/Severe_weather. [Accessed 26 May 2024].

[18] National Weather Service, "What is a Nor'easter?," National Oceanic and Atmospheric Administration, [Online]. Available: https://www.weather.gov/safety/winter-noreaster. [Accessed 26 May 2024].

[19] National Weather Service, "Severe Weather Definitions," National Oceanic and Atmospheric Administration, [Online]. Available: https://www.weather.gov/bgm/severedefinitions. [Accessed 1 June 2021].

[20] SMD Content Editors, "Extreme Weather and Climate Change," NASA, March 2024. [Online]. Available: https://science.nasa.gov/climate-change/extreme-weather/. [Accessed 26 May 2024].

[21] United States Government Accountability Office, "Better Use of Climate Data Could Enhance the Climate Resilience of Postal Facilities," September 2021. [Online]. Available: https://www.gao.gov/assets/gao-21-104152.pdf. [Accessed 26 May 2024].

[22] United State Postal Service Office of Inspector General, "U.S. Postal Service Emergency Preparedness Winter Storm Elliott," United States Postal Service, 2024.

[23] R. Peregrino de Brito, P. Laczynski de Souza Miguel and S. Pereira, "Connecting Extreme Weather Events and Supply Chains," *Academy of Management Proceedings,* August 2017.

[24] M. D. Garvey, S. Carnovale and Y. Sengun, "Ananalytical framework for supply network risk propagation : A Bayesian network approach," *European Journal of Operational Research,* pp. 618-627, 2015.

[25] M. Fattahi, K. Govindan and E. Keyvanshokooh, "Responsive and resilient supply chain network design under operational and disruption risks with delivery lead-time sensitive customers," *Transportation Research Part E,* pp. 176-200, 2017.

[26] A. Saxena and M. Prasad, "A Review of Clustering Techniques and Developments," *Neurocomputing,* pp. 664-681, 2017.

[27] Z. Wengang and A. T. Chee, "Multivariate adaptive regression splines for analysis of geotechnical engineering systems," *Computers and Geotechnics,* pp. 82-95, 2013.

[28] H. Barbosa, M. Barthelemy and G. Gourab, "Human Mobility: Models and Applications," *Physics Reports,* pp. 1-74, 2018.

[29] C.-h. Hsieh, S. M. Glaser, A. J. Lucas and G. Sugihara, "Distinguishing random environmental fluctuations from ecological catastrophes for the North Pacific Ocean," *Nature,* pp. 336-340, 2005.

[30] National Centers of Environmental Information, "GHCN (Global Historical Climatology Network) – Daily Documentation," National Centers of Environmental Information, [Online]. Available: https://www.ncei.noaa.gov/pub/data/cdo/documentation/GHCND_documentation.pdf. [Accessed 16 June 2024].

[31] M. J. Menne, I. Durre, R. S. Vose, B. E. Gleason and T. G. Houston, "An Overview of the Global Historical Climatology Network-Daily Database," *Journal of Atmospheric and Oceanic Technology,* vol. 29, no. 7, pp. 897-910, 1 July 2012.

[32] W. Weissert and Z. Miller, *Postal Service Reform Act signed into law. Here's what it does.,* ABC, 2022.

[33] National Climatic Data Center, "Global Historical Climatology Network - Daily (GHCN-Daily), Version 3," NOAA National Centers for Environmental Information, 30 June 2023. [Online]. Available: https://www.ncei.noaa.gov/access/metadata/landing-page/bin/iso?id=gov.noaa.ncdc:C00861. [Accessed 7 June 2024].

[34] W. Tollefson and A. Joyner, "January 2024 Tennessee State Climate Summary," East Tennessee State University, 2024.

[35] National Centers for Environmental information, "Monthly National Climate Report for January 2024," National Oceanic and Atmospheric Association, 2024.

[36] I. Durre, M. J. Menne, B. E. Gleason, T. G. Houston and R. S. Vose, "Comprehensive Automated Quality Assurance of Daily Surface Observations," *Journal of Applied Meteorology and Climatology,* vol. 49, no. 8, pp. 1615-1633, 2010.

[37] United States Postal Service, "Latest USPS Report Shows Sustained Service Performance," 29 January 2024. [Online]. Available: https://about.usps.com/newsroom/national-releases/2024/0129-latest-usps-report-shows-sustained-service-performance.htm. [Accessed 27 June 2024].

[38] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining,* pp. 785-794, 2016.

[39] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," *Annals of Statistics,* vol. 29, no. 5, pp. 1189-1232, 2001.

[40] G. Ke, "LightGBM: A highly Efficient Gradient Boosting Decision Tree," *Advanced Neural Information Processing System,* vol. 30, pp. 3146-3154, 2017.

[41] A. Verikas, "CatBoost: Gradient Boosting with Categorical Features Support," *Adv. Intelligent Syst. Comput.,* vol. 873, pp. 309-318, 2019.

[42] Y. LeCun, Y. Bengio and G. Hinton, "Deep Learning," *Nature,* vol. 521, no. 7553, pp. 436-444, 2015.

[43] C. M. Bishop, in *Pattern Recognition and Machine Learning*, New York, NY, Springer, 2006, pp. 173-177.

[44] J. MacQueen, "Some Methods for Classification and Analysis of Multivariate," *Fifth Berkeley Symp. Math. Stat. Probability,* vol. 1, pp. 281-297, 1967.

[45] J. Pearl, Probabilitc Reasoning in Intelligent Systems: Networks of Plausible Inference, San Mateo, CA: Morgan Kaufmann, 1988.

[46] L. Breiman, "Random forests," *Machine Learning,* vol. 45, no. 1, pp. 5-31, 2001.

[47] T. C. a. C. Guestrin, "XGBoost: A Scalable Tree Boosting System," *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining,* pp. 785-794, 2016.

[48] United States Postal Service, "Publication 32 - Glossary of Postal Terms - Contents," United States Postal Service, Washington, DC, 2013.

[49] United State Postal Service, "How to use this website - Service Performance," United State Postal Service, 2024. [Online]. Available: https://spm.usps.com/#/main. [Accessed 31 May 2024].

[50] United State Postal Service Office of Inspector General, "Effectiveness of the New Regional Processing and Distribution Centers in Richmond, VA," United States Postal Service, 2024.

[51] National Weather Service Institute, "NATIONAL WEATHER SERVICE INSTRUCTION 10-1605," National Oceanic & Atmospheric Administration, 2021.

[52] G. K. e. al., "LightGBM: A Highly Efficient Gradient Boosting Decision Tree," *Advances in Neural Information Processing Systems, vol. 30,* pp. 3146-3154, 2017.

This page intentionally left blank