Final Report

for

Maine Trust for Local News (METLN) Project

By Ben Cox, Sarah Crane, Andrew Fumarola

DS5110 – Essentials of Data Science, Professor Bockmon

Northeastern University, Roux Institute

December 9, 2025

Table of Contents

1. Introduction

The goal of our project for Maine Trust for Local News (METLN) was to conduct a preliminary examination of the data we were provided by the client. Initially, we were provided sample data from February to October of 2024. However, additional 2025 data was recently added to the repository for the project. Our group felt strongly that using the most recent data would give us the best snapshot as an overview of the kinds of analysis that we could explore using the client's data.

During the planning stages for this project, our group (consisting of completely remote students) decided to divide up our areas of analysis and chose to examine the following topics: Geographic Analysis, Digital vs. Print Subscriptions, and Temporal Trends. By dividing our analysis tasks, each team member was able to take ownership of one area of the project, and each member was able to become a subject matter expert on their area of analysis. The 'Results' and 'Discussion' sections of this report have been divided into three analysis areas to provide further details into each portion of this project.

In addition to this report, which summarizes each subtopic we explored, this project consists of a Github repository to house our individual code files (file types include Jupiter notebooks, Python files, markdown files, etc.), and the presentation that we completed on December 2. A copy of the presentation slides will be saved in the Github repository, along with any other pertinent files.

Link to the Github Repository: https://github.com/benrcox/ds5110_final_project_METLN.git

2. Background

This project was brought to the Roux Institute, of Northeastern University, on behalf of the client, Maine Trust for Local News (METLN), with the goal of gaining better insight into potential areas for growth for their current subscription offering and geographic reach. Presently, METLN oversees several newspapers and three websites that focus on covering state and regional news for the urban communities surrounding Portland, Lewiston, Augusta, and Waterville. Each current publication is distributed through various means, including a mix of print and digital subscription offerings. Out of Portland, the Portland Press Herald and Maine Sunday Telegram serve as statewide news, as well as focusing on community news for Yok and Cumberland counties and the Midcoast region. The Lewiston Sun Journal focuses on community news for Androscoggin County and Western Maine. Finally, the Morning Sentinel and Kennebec Journal focus on news pertaining to the communities surrounding Waterville and Augusta,

respectively.

3. Methods & Analysis

To get a sense of the type of data we would be working with, one of the first things we did was look at the shape of the data. The “METLN_data.csv” file is a combined file of each of the data files from the data repository for this project. We can clearly see that the basic demographic information is organized into the following columns: “Publication”, “AccoutID”, “Status”, “Bill Method", "Dist ID”, “Route ID”, “Day Pattern", "City”, “State”, “Zip”, “Rate Code”, “LastStartDate”, “OriginalStartDate”, and “Date” (see Figure 3.1 for a snapshot of the first 5 lines of data).

```
# Read in file & view first 5 lines
metln = pd.read_csv('/content/METLN_data.csv')
print(metln.head())
```

```
  Publication  AccoutID  Status    Bill Method Dist ID Route ID Day pattern  \
0      MTM_MS    505360  Active  Auto Pay - BD    DST2   WTV058     SXmmmmX
1      MTM_PT    528751  Active     Office Pay       2   144077     SoTWTFS
2      MTM_PT   2446127  Active     Office Pay       2   144077     SoTWTFS
3      MTM_PT    522993  Active     Office Pay       2   144077     SoTWTFS
4      MTM_PT   2487779  Active  Auto Pay - CC       2   144077     SoTWTFS

          City State    Zip           Rate Code LastStartDate  \
0   Waterville    ME  04901     MATHER_7DAY_830      04-01-18
1  Scarborough    ME  04074     MATHER_7DAY_740      04-01-18
2  Scarborough    ME  04074     MATHER_7DAY_910      05-24-22
3  Scarborough    ME  04074    MATHER_7DAY_1170      04-01-18
4  Scarborough    ME  04074  PT_7Day_R_FullRate      11-29-24

  OriginalStartDate      Date
0          04-01-18  102025.0
1          04-01-18  102025.0
2          05-24-22  102025.0
3          04-01-18  102025.0
4          11-29-24  102025.0
/tmp/ipython-input-771356103.py:2: DtypeWarning: Columns (4,5,9) have mixed types. Specify dtype option on import or set low_memory=False.
  metln = pd.read_csv('/content/METLN_data.csv')
```

Figure 3.1: Overview of Combined CSV Data for METLN

For identifying cancellations and new accounts over time, it was important to look at how the daily snapshot logs changed. The file LogCompiler.py contains methods for cleaning and compiling this data. This includes formatting the dates to match and concatenating data sets together. The format followed the columns as defined by the original data sets.

The most crucial parts of cleaning the data were to adjust date formats and to clean the city labels so as to properly aggregate the data. The start dates were converted to a DateTime format and the city names were all converted to entirely lower case so that there was no difference between a “portland” or a “Portland” or a “PORTLAND”. The state labels were converted to upper case for the same reason. It was also vital to use a 'zfill’ method on zip codes so that the leading 0 in many Maine zip codes were not omitted (e.g. 04017 might be converted mistakenly to 4017).

4. Results

4.1 Geographic Analysis

In order to evaluate the number of subscriptions across different locations, a general geographic analysis of the data was performed. This portion of the project mainly focused on analyzing the number of subscribers ("AccoutID" subscriptions) across cities and states. This began with an analysis of subscriptions inside and outside of Maine (Table 4.1.1), subscriptions by Maine city (Table 4.1.2). The full versions of these tables are too long to include here but can be found in the GitHub code. Significant data processing was required to create visualizations and insights from the data, with missing values, erroneous entries, and capitalization issues. Line Shapefile data was combined with the original data to create visualizations utilizing the Geopandas library to display the number of subscriptions by city area (Figure 4.1.1) and the rate of accounts per population for each city area (Figure 4.1.2). The latter of these included U.S. Census Bureau population information as well. An additional chart was created to display the percentage of digital versus print subscriptions by area (Figure 4.1.3) as well, though this is furthered in the next section concerning Digital vs. Print subscriptions.

Figures 4.1.1, 4.1.2, and 4.1.3 include Line Shapefiles provided by the U.S. Census Bureau. These Shapefiles are defined as 5-Digit ZIP Code Tabulation Areas, or ZCTAs. These ZCTAs are approximate representations of the similar ZIP codes, as used by the U.S. Postal Service, but they are more standardized and match up with respective population, demographic, and other Census Bureau datasets. For this data, only areas that contained a respective geographic ZCTA value matching a ZIP code were considered.

State	ME	MA	FL	NY	NH
# Subs	55190	1108	640	462	413

Table 4.1.1: Top 5 states by number of subscribers

City	Portland	South Portland	Brunswick	Lewiston	Scarborough
# Subs	6731	2421	2153	2117	2081

Table 4.1.2: Top 5 cities in Maine by number of subscribers

Unsurprisingly, the most populous state is Maine with approximately 10% of all subscriptions based in Maine (55190 in Maine, 5329 outside of Maine). Also as expected, towns and cities contain the highest number of subscriptions ("# Subs").

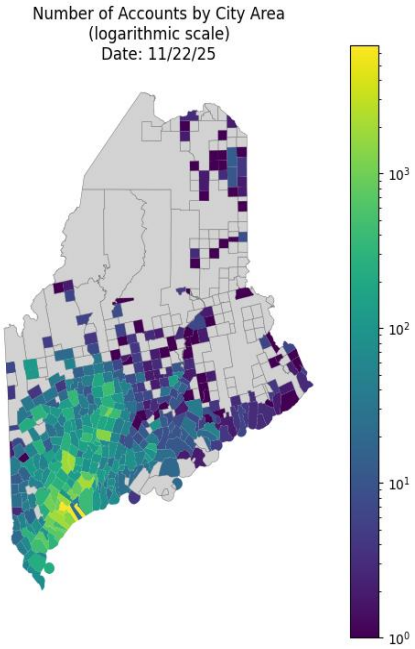

Figure 4.1.1: Number of subscriptions by city area.

Because there is a divergence between the ZIP code used by the USPS and Maine Trust for Local News and the ZCTA values, this geographic image and the following figures only take into account cities that have an associated ZCTA. This does mean some areas may not be accounted for, although many of these are in Northern Maine where there is a small population, and few news subscribers anyways.

The coloring of this graphic makes it difficult to read, with the greater Portland area and other larger cities resizing the scale for the other areas. This is even with the logarithmic scale for subscriptions to color in ZCTA areas. We can help solve this coloring issue in the following figure by accounting for population.

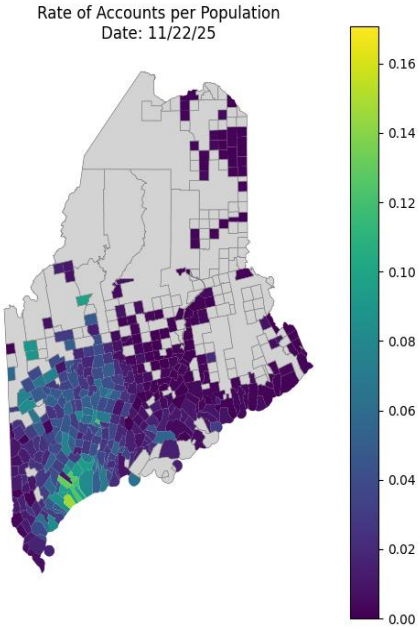

Figure 4.1.2: Number of Accounts by Population (Rate of Accounts)

After the inclusion of population data (also from U.S. Census Bureau), the rate of accounts per population unit (1 person) can be illustrated by dividing each ZCTA area by its respective population. For example, a rate value of 0.12 for a moderately well-populated ZCTA area (such as areas in the Southern Maine region) would have approximately 12 subscribers per 100 people. This helps with the coloring issue in the previous Figure 4.1.1 and begins to illustrate different regions of interest within Maine: Southern Maine, the New Hampshire Border, Downeast, and the Mid-Coast.

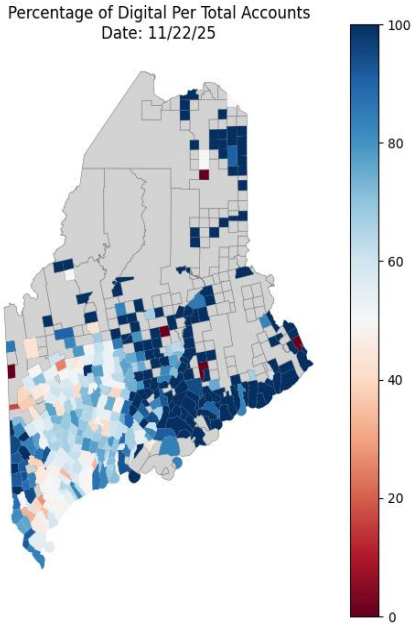

Figure 4.1.3: Percentage of Digital vs Print Per Total Subscriptions (11/22/2025)

The above Figure 4.1.3 shows the percentage of digital versus Print per total number of subscriptions. This was created by dividing the number of print subscriptions ("07Day" pattern in the "Day Pattern" column) by the total number of subscriptions per a city's designated area (ZCTA). This is a preliminary geographic look at print versus digital subscriptions, which is defined in-depth in the following section.

4.2 Digital vs. Print

To better examine the subscriptions trends of digital versus print subscribers, this portion of the project was divided into two areas of analysis, namely: Data Exploration & Static Visualizations, and then creating Map Visualizations. The Data Exploration mainly consisted of exploring the range of the data: getting an idea for the publication counts (see Figure 4.2.1), count of day patterns (see Figure 4.2.2), percent of digital subscriptions (see Figure 4.2.3), trends in digital subscriptions (see Figures 4.2.4 and 4.2.5), state subscribership (see Figures 4.2.6 and 4.2.7), and a comparison between the highest and lowest cities regarding digital subscriptions (see Figures 4.2.8 and 4.2.9).

The Map Visualizations were created using the resources of the 'Geopy' and 'Folium' libraries to create interactive maps that would reflect the METLN data present within the CSV file. After comparing static graphs of the highest and lowest digital subscriptions by city, the first

map was created to look for geographic trends in the areas with the highest and lowest digital subscribership (see Figure 4.2.10). Finally, an interactive map was created to examine the overlap between the highest and lowest cities with respect to their Digital vs. Print subscribership (see Figure 4.2.11).

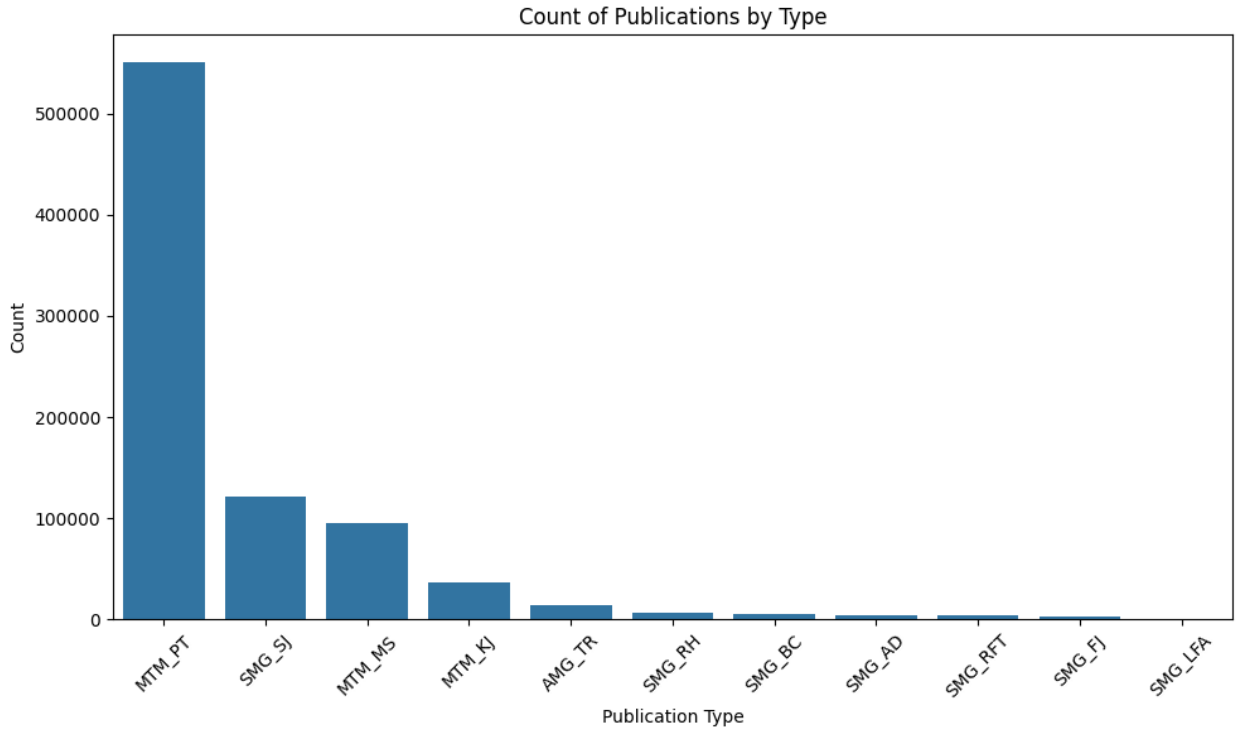

Figure 4.2.1: Count of Publications by Type

Looking at the Publication type (Figure 4.2.1), we can see that "MTM_PT" had the most subscribers, regardless of whether those subscriptions were Digital vs. Print. When we look closer at 'day pattern', we see that digital subscriptions, listed as "07Day" pattern (Figure 4.2.2), make up the bulk of subscriptions. The pie chart for day pattern indicates that digital subscriptions make up 61.3% of all subscription types (Figure 4.2.3).

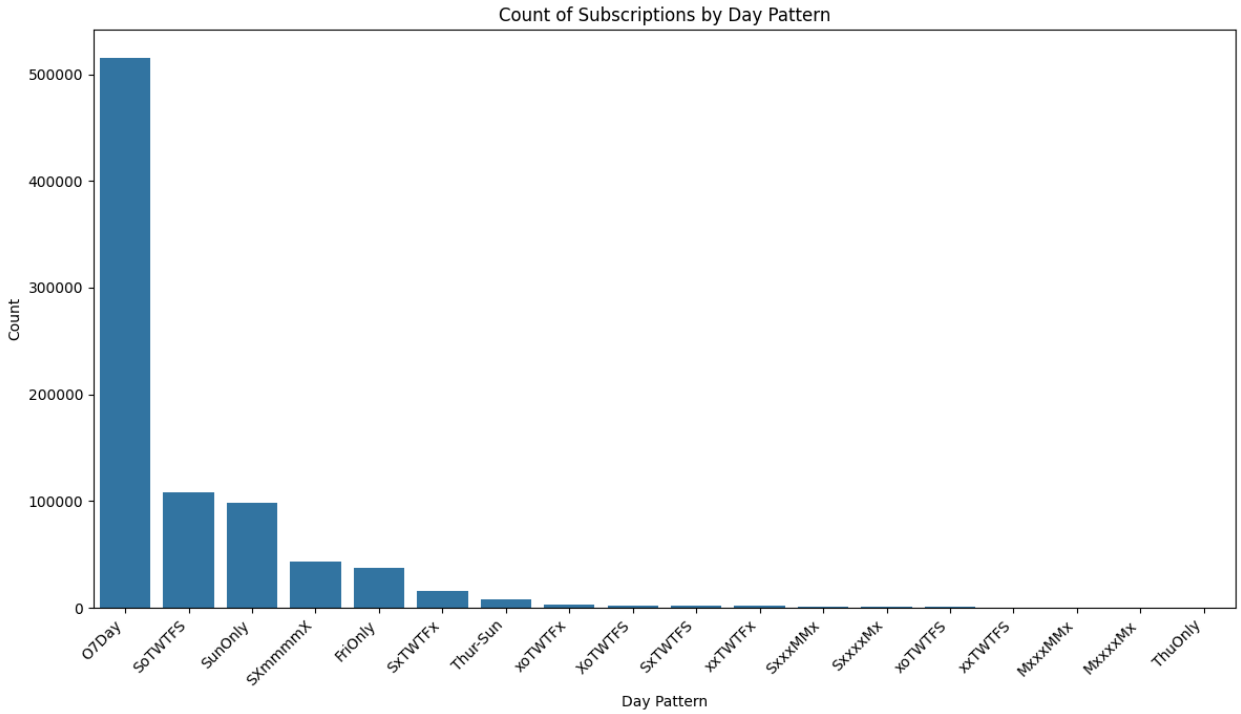

Figure 4.2.2: Count of Subscriptions by Day Pattern

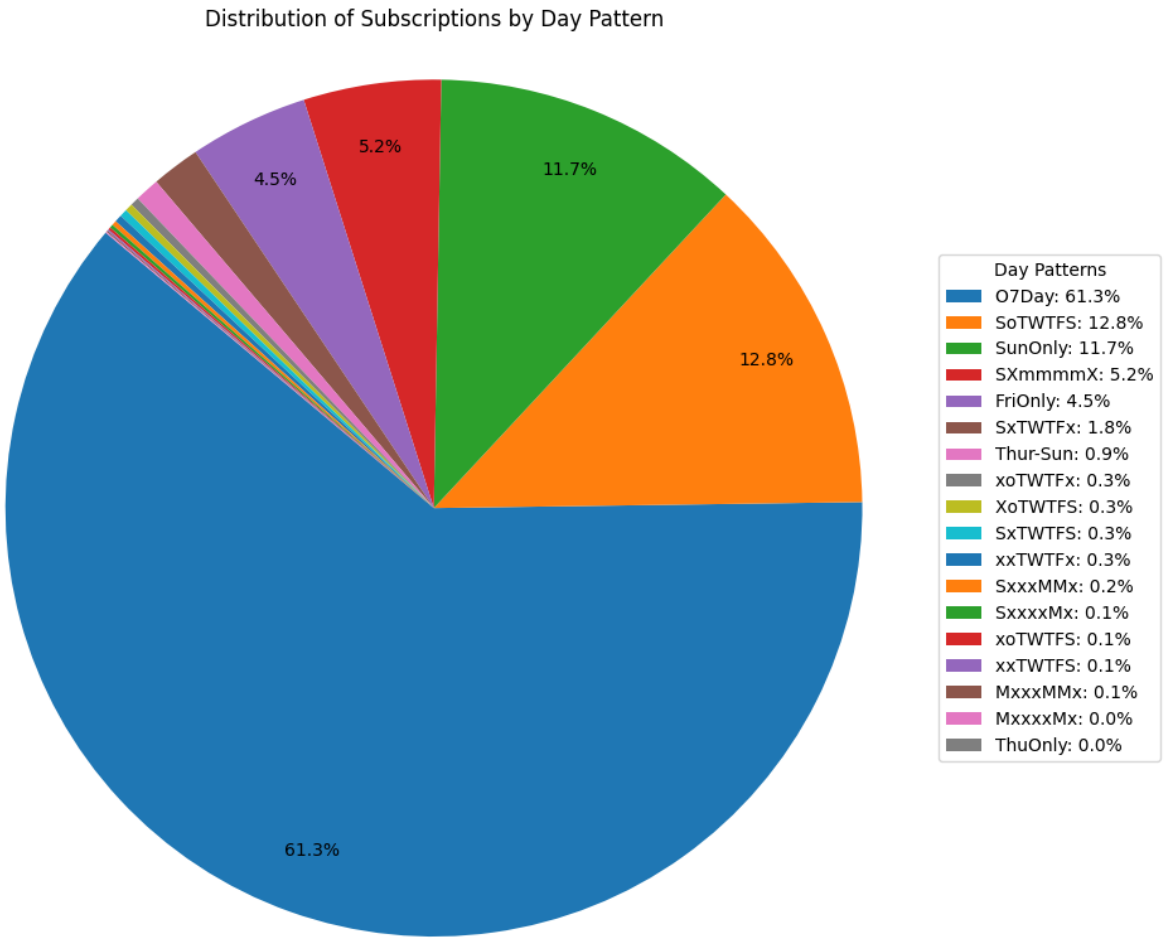

Figure 4.2.3: Pie Chart for the Distribution of Subscriptions by Day Pattern

Next, it was interesting to look at trends over time regarding digital subscriptions. Figures 4.2.4 and 4.2.5 attempted to see if there were any major increases or decreases in digital subscriptions over time. It is notable that there appears to be a ‘spike’ in digital subscriptions at the start of 2018; it is unclear whether this is due to when METLN started offering digital subscriptions or if that simply represented the start date of the data provided. There appears to be a smaller second ‘spike’ in 2020, with a bit of a drop off, and then steadily increasing digital subscriptions over time.

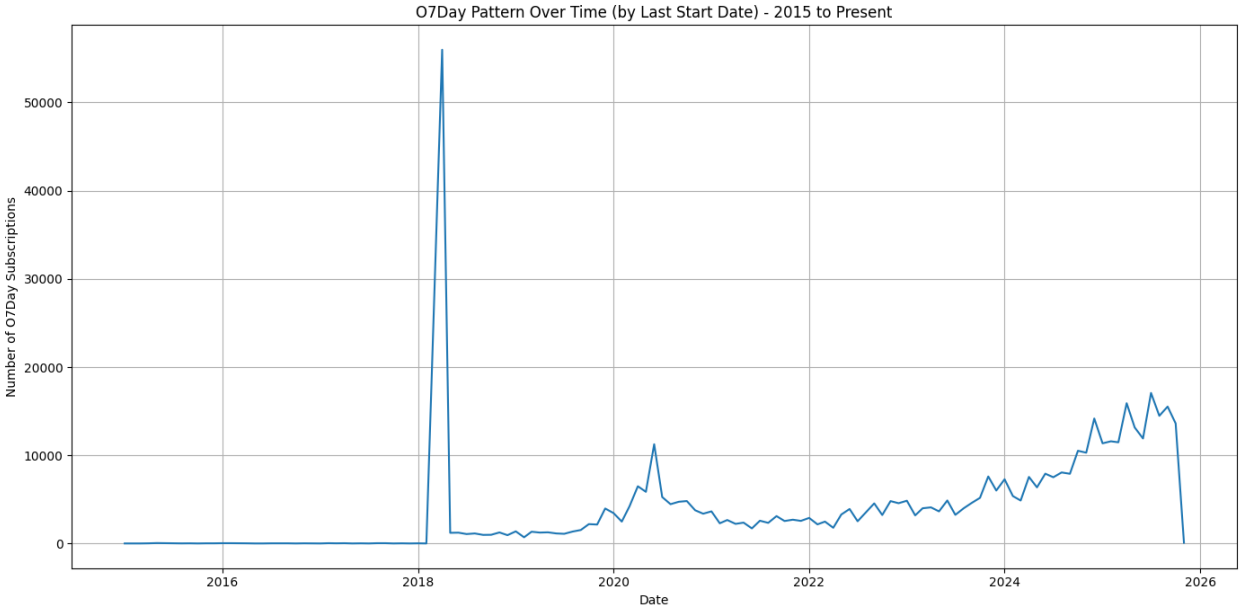

Figure 4.2.4: 07Day (Digital) Pattern Over Time (by Last Start Date) - 2015 to Present

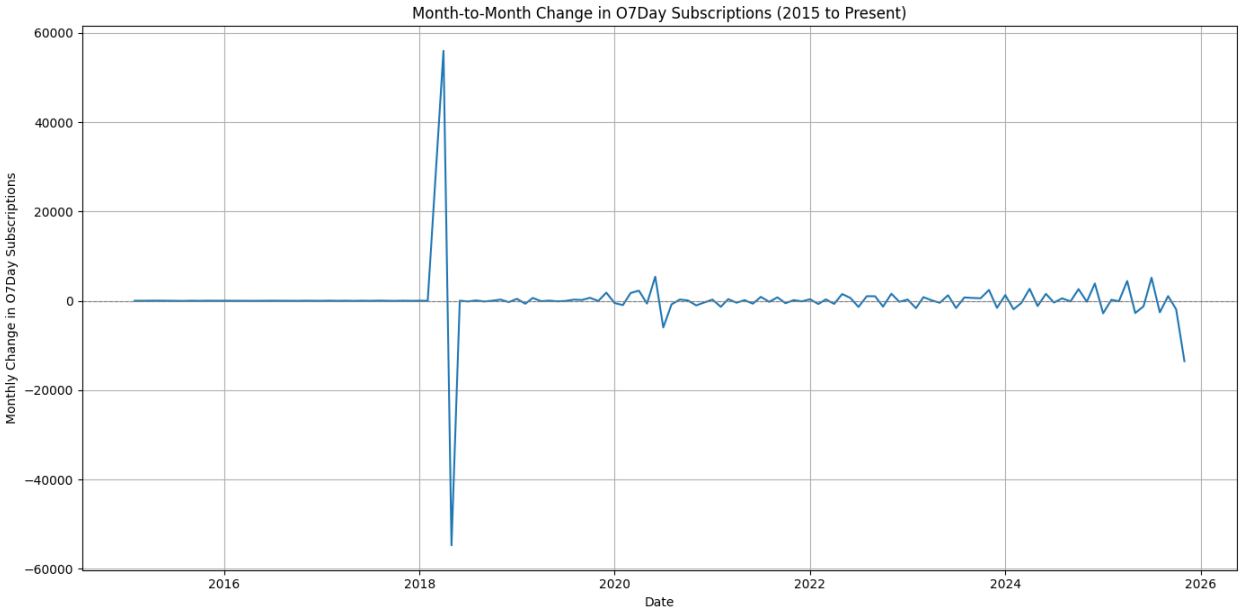

Figure 4.2.5: Month-to-month Change in 07Day (Digital) Subscriptions (2015 to Present)

One area that was surprising was to see the variety of subscriptions that exist outside the state of Maine. Figures 4.2.6 and 4.2.7 show that 91.2% of all subscriptions are attributed to Maine, and the remaining subscriptions are spread throughout the United States across 24 additional states.

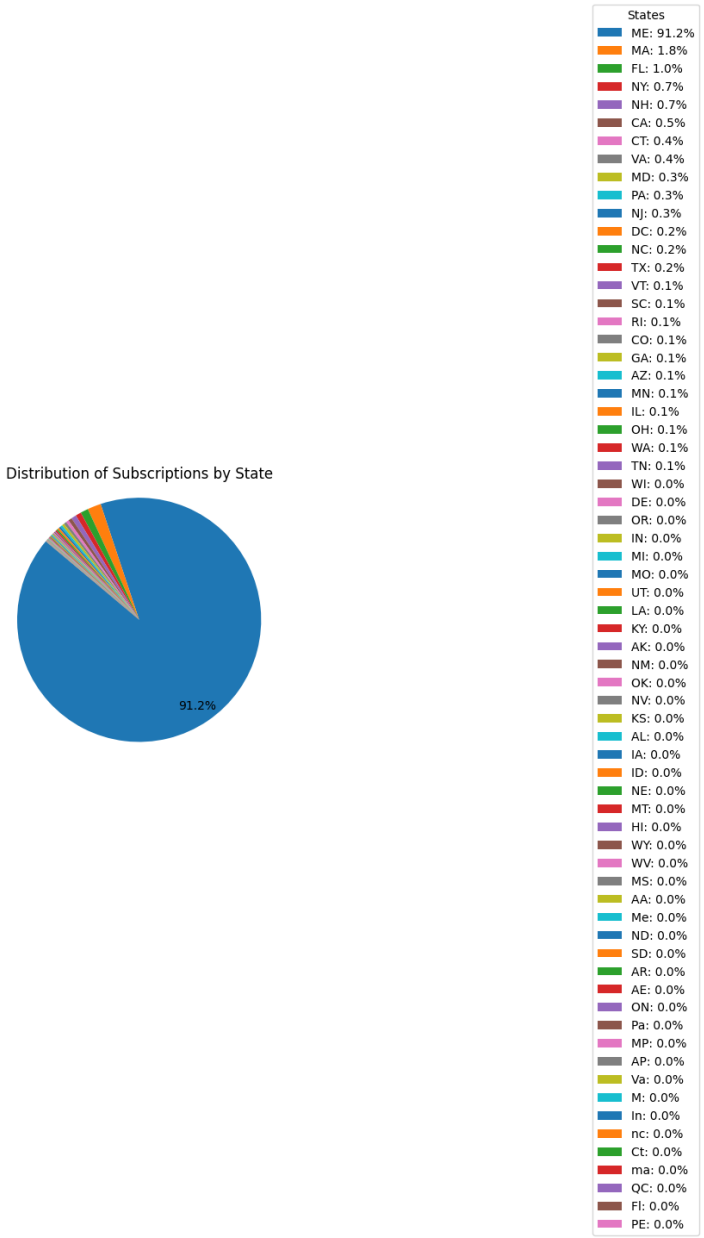

Figure 4.2.6: Pie Chart for Distribution of Subscriptions by State (Digital & Print)

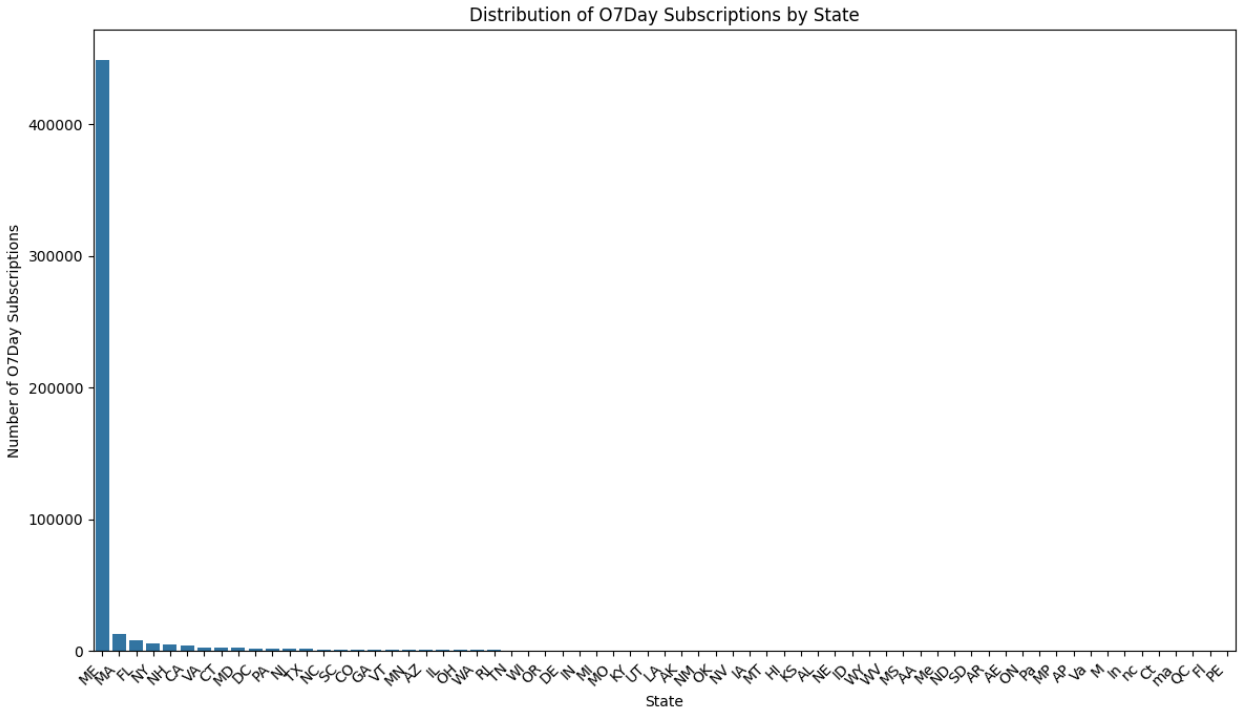

Figure 4.2.7: Distribution of 07Day (Digital) Subscriptions by State

The last static visualizations completed (Figures 4.2.8 and 4.2.9) pertained to shifting from a state view to a ‘city’ view of the highest and lowest 25 ‘cities’ within the state of Maine. For the purposes of these visualizations, ‘city’ is used as a general term referring to the way ‘city address’ information was included within the columns of the CSV file. No distinction between ‘town’ size appears to apply here; rather ‘city’ is used for location purposes. It is important to note that some inconstancies within the raw data were found when exploring city information. If you look at the left side of Figure 4.2.8 you will notice that ‘Portland’ and ‘PORTLAND’ appear as the first two ‘cities’ within the top 25 list, however they should be counted as the same. To address this issue, the data was re-run using normalized city data that put every column entry into lowercase lettering, so that the counts would more accurately reflect the totals of the highest and lowest cities for digital subscriptions.

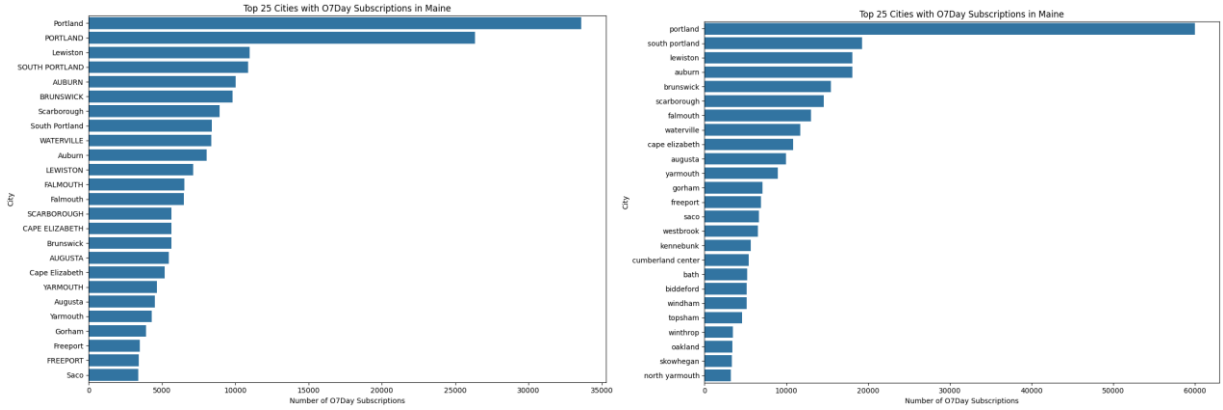

Figure 4.2.8: Top 25 Cities with 07Day (Digital) Subscriptions in Maine (Raw Data – Left, Normalized Data – Right)

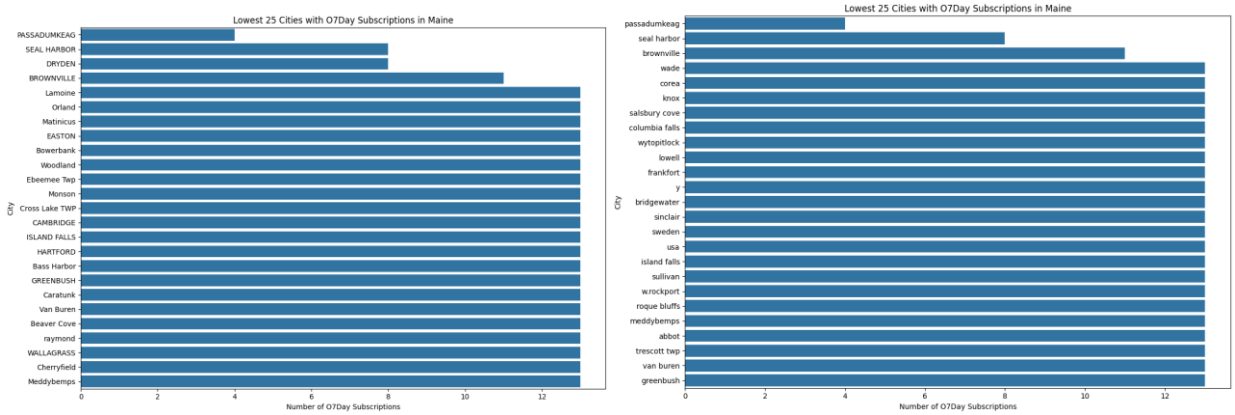

Figure 4.2.9: Lowest 25 Cities with 07Day (Digital) Subscriptions in Maine (Raw Data – Left, Normalized Data – Right)

When you look at the interactive map visualizations, you start to get a clearer picture of how widespread digital subscriptions appear. Within Figure 4.2.10 you can see that the 'top' cities (with blue map pins) are congregated around the southern and central portions of the state of Maine, surrounding the major 'cities' such as Portland, Lewiston/Auburn, Augusta, and Waterville (following the major I-95 highway corridor). This is in sharp contrast to the 'lowest' cities (with red map pins) in the state of Maine, which are spread throughout the state, particularly along the mountains towards the west (beyond Lewiston/Auburn), the 'down east' area (think Machias and Eastport), and more northern parts (such as surrounding Van Buren and Caribou).

When you overlap the statistics surrounding Digital vs. Print subscriptions (see Figure 4.2.11), you can see that the areas with the 'top' cities have a healthy combination of digital

(purple) and print (orange) subscriptions. Once you get beyond Waterville, there is a bit of a shift where digital subscriptions begin to outperform print subscriptions. There are, of course, some exceptions, with points on the map that indicate that there are some rural communities that are still receiving print subscriptions, however the majority of rural communities (those outside of the more metropolitan areas of Waterville, Augusts, Lewiston/Auburn, and Portland) appear to be served more via digital subscriptions.

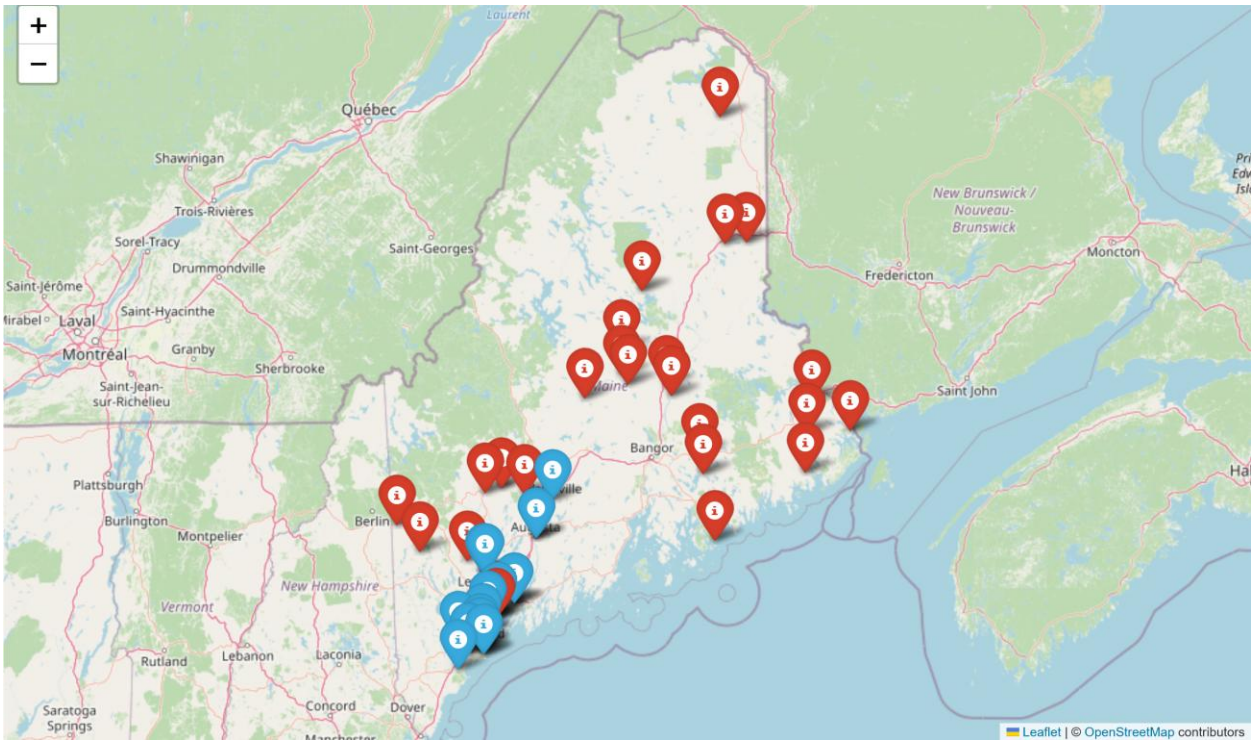

Figure 4.2.10: Map of Highest (blue) and Lowest (red) 25 Cities for Digital Subscriptions in Maine (11/30/25)

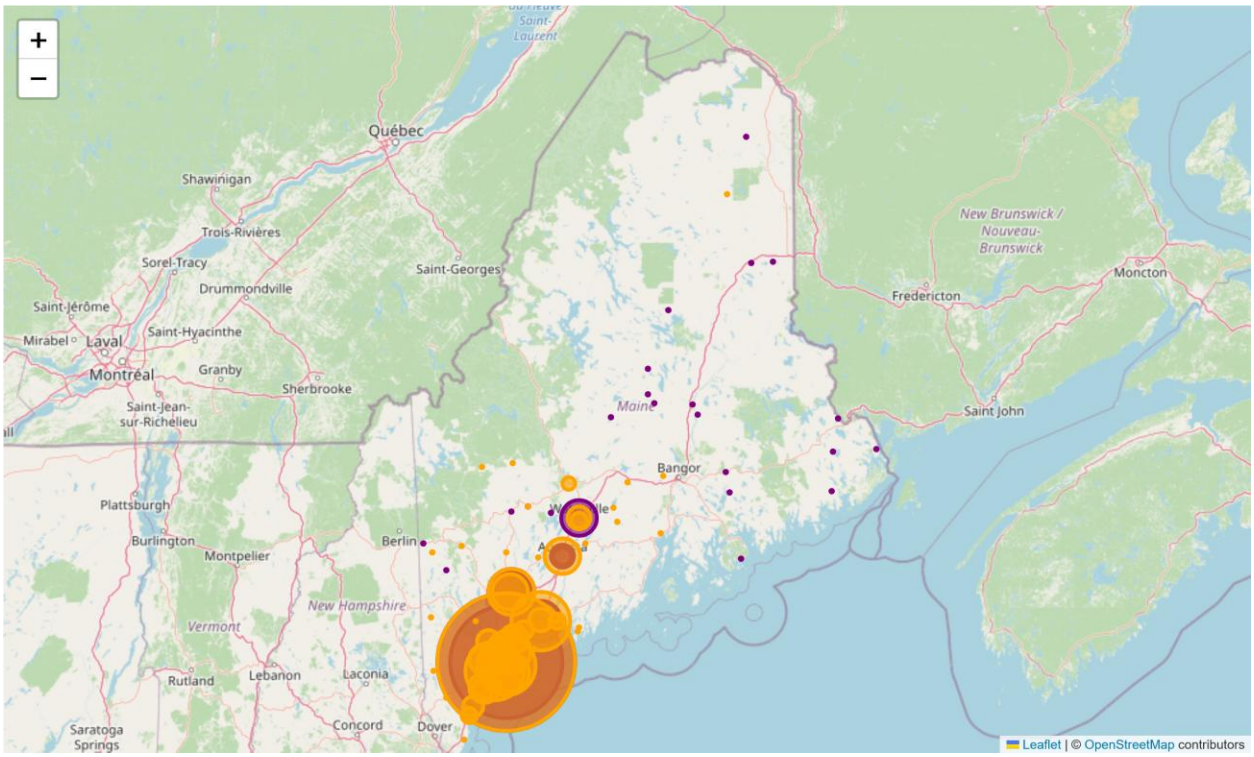

Figure 4.2.11: Map of the Highest and Lowest 25 Cities for Digital (purple) vs. Print (orange) Subscriptions in Maine (11/30/25)

4.3 Temporal Trends

The analysis on temporal trends was conducted on the date range of October 25, 2025 to November 21, 2025. While this is a small snapshot in a larger business cycle, the methods are designed to handle whatever date range is inputted into the LogCompiler script.

Analysis on temporal trends is contained to the LogAnalysis.py and ChurnRetentionRates.py files. For viewing daily active users, I aggregated the data by log date (which was added in the LogCompiler methods) and created a data frame containing daily counts. This method would work over any collection of logs that have been compiled in the LogCompiler script. This data frame contains counts of users on each day available in the log and is simple to graph from there using MatPlotLib. Figure 4.3.1 shows the results of this analysis. The chart indicates that daily active users decreased by about 600 accounts over this period, with only several days reporting higher daily active users than the day before. While the graph may look grim, it is important to note that 600 is about 1% of the total users at the end of this period.

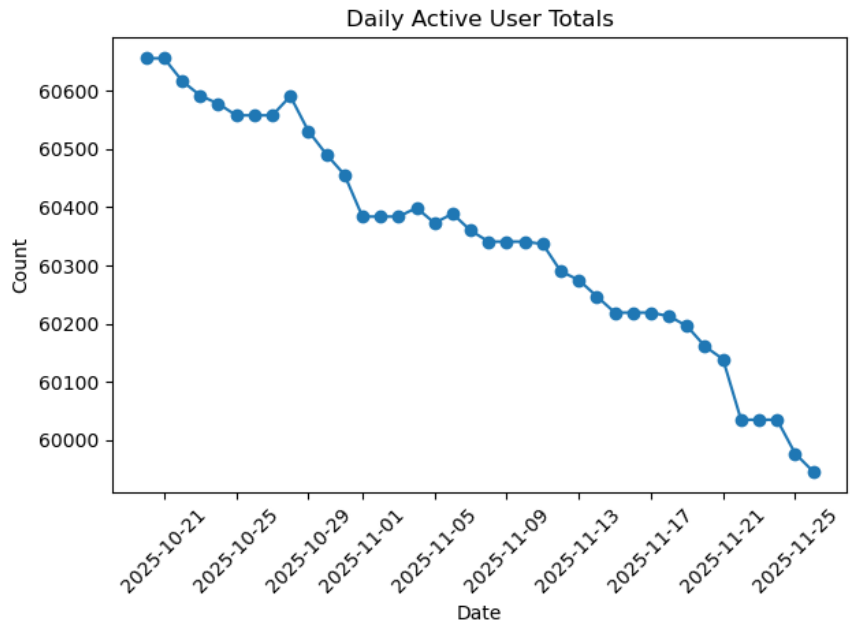

Figure 4.3.1: Line chart of daily active users over the course of October 21, 2025 to November 25, 2025.

In order to identify trends in new accounts and cancellations over time, I had to add labels to each entry in the greater log. The 'last_active_day' method was used to determine if this entry would be the final time this account id is seen in the log. This worked by sorting the log by Account ID and then Log Date. A new column called 'next_log' was created to point to the next instance of that AccountID based on the sorted list. I added another column called 'last_active_day" which is True when the 'next_log' variable is NA, meaning there is no next active day, meaning that the user cancelled their subscription at this time. When compiling cancelled accounts later, it was important to make sure to not include the last day in the log, since the 'next_log' would be NA for all these rows, but that is only because it is the latest data.

Similarly, another method 'first_active_day_label' identified the first instance of any account. This was done using a cumcount() method on the sorted log. When the cumulative count is equal to 0, we know this is the first instance of this account id. When analyzing this later, it was again important to disqualify entries on the first day of data, since this method would categorize every log on the earliest date as the first instance of the Account ID.

Now that the data contains labels indicating a first or last active day, I was able to compile the number of new accounts and cancelled accounts by date. By isolating the new and cancelled accounts, I could plot these impacts over time. Viewing the results visually in Figure 4.3.2, it is clear that there is a weekly cyclical activity occurring, which may make sense depending on when subscriptions can start and end based on the specific payment system. It would be interesting to look at a larger time period and see week-to-week changes. Seeing the daily changes shows that there may be anywhere from 0 to 250 cancellations on a given day, and

0 to 150 new accounts on a given day. Seeing these two lines side-by-side contributes greater context to the Daily Active Users chart (figure 4.3.1)

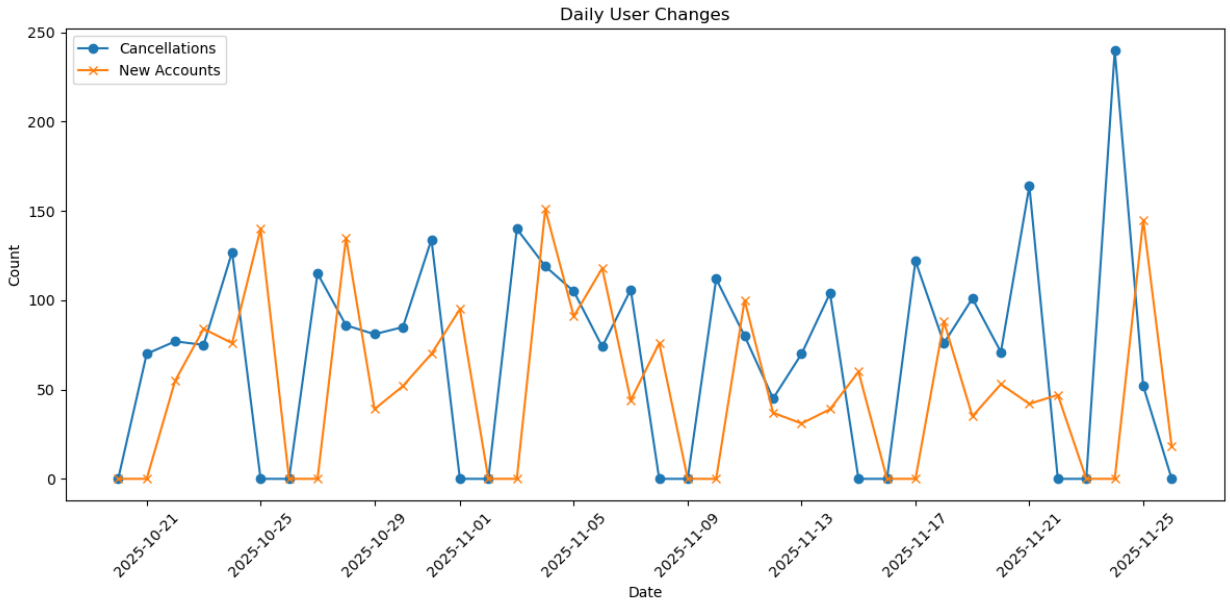

Figure 4.3.2: Line chart of account cancellations and account creations from October 21, 2025 to November 25, 2025.

This data could further be used to isolate and calculate churn and retention rates over this date range. The script ChurnRetentionRates.py contains methods to calculate these rates by zip code and then sort them. Churn rate is calculated as the number of customers lost divided by the number of customers at the beginning. Retention rate is the percent of customers at the beginning of the period that remain through the end of the period. This script specifically calculates these rates for each zip that appears more than 500 times in the log. This number can be manipulated depending on the client's desired scope. For this time period, the zip codes with the greatest churn rates were:

Zip	City	Churn Rate (%)
04941	Freedom	18.75
04087	Waterboro	15
04943	Hartland	14.29
04544	East Boothbay	13.33
04583	Boothbay Harbor	12.66

Figure 4.3.3: Top five zip codes by churn rate from October 21, 2025 to November 25, 2025.

Breaking the churn rates down by zip code was a more granular choice than city, since larger cities like Portland and South Portland contain multiple zip codes. For instance, the 04101 zip code (part of Portland) had a 3.47% churn rate. In total, there were 85 zip codes that showed a churn rate of 5% or larger.

5. Discussion

5.1 Geographic Analysis

The results from the Geographic Analysis figures are an illustration of Maine's regions: Southern Maine, the New Hampshire border, the Downeast region, and the Mid-Coast all appear as relatively distinct areas. Maine by far contains the highest number of subscribers by state, and because only about 10% of subscribers live out-of-state, this analysis will predominantly consider Maine and leave out other regions. It is worth noting that the anecdotal "snowbirds" residing in Florida are apparent here, as Florida has the third-highest number of subscriptions by state, only surpassed by Massachusetts. The high number of MA subscriptions is likely due to the larger nearby population, especially in and around the greater Boston area. These subscriptions could be Mainers or others working in Massachusetts while still taking interest in their northern neighbor.

The Portland, Lewiston, South Portland, and greater surrounding areas contain many of Maine's population centers. These areas in Southern Maine have high rates of subscription per population. Broader efforts to increase subscriber rates in these areas might be easier due to the density of urban centers, but could contain a limit on how much of a city's population would purchase a subscription. Outside of these high-rate areas are the surrounding suburban zones with a middling population. This group sees a gradual reduction in subscriber rates but could prove good areas for expanding current readership with focused targeting. The final areas along the borders of New Hampshire, Downeast, and Mid-Coast regions contain very low subscriber rates. General outreach here may prove useful, but it may be more optimal to consider specific, targeted efforts to establish a base of readership.

These trends show the lower the population, the lower the rate of subscriptions per population as well. Societal trends and pressure could be the explanation for this: if everyone in town has a subscription, you may feel more obligated to purchase one. Further examinations into classifying each city or ZCTA locale as one of these three types of regions could help build specific types of customers for growth opportunities across Maine. These analyses in combination with the following Digital vs. Print and Temporal Trends sections can help to define areas of support and opportunities to provide the three different types of customers throughout the regions of Maine.

5.2 Digital vs. Print

When you look at the results of the Digital vs. Print subscription analysis, you get the sense that rural communities tend to favor digital subscriptions. Areas outside of the more metropolitan communities (i.e. Portland, Lewiston/Auburn, and Augusta; essentially those cities surrounding the I-95 corridor), such as those communities in the mid-coast region, Down East Maine, along the New Hampshire border, and Northern Maine communities tend to be more rural and favor digital subscriptions.

These results could be used to better target or promote campaigns toward potential subscribers in these areas. More rural communities (which may have more limited access to mail service, hence hindering the relevance of print subscriptions), may be better served through digital subscriptions. Communities in southern Maine, particularly the greater Portland area, may still find use for both print and digital subscriptions.

It might be useful to further examine Maine population data pertaining to households that have access to high-speed internet among rural communities. This could be helpful before targeting subscription campaigns towards rural communities. Additionally, it would be helpful to get a better understanding of the reasons behind the ongoing trend towards having print and digital subscriptions in southern Maine. Surveying current subscribers to get a better understanding of their subscription preferences might be helpful and provide further insights into the reasons behind the trends in those areas.

5.3 Temporal Trends

The temporal trends for this time frame indicate a churn and retention rate that nearly balance out, but there is an overall downward trend in daily active users. Over this period, there 85 zip codes that had a churn rate over 5% and 13 zip codes with a churn rate over 10%. Based on these results, we recommend that METLN investigate these towns more deeply to see if there are any common reasons for cancelling. Further research could also attempt to cluster and characterize these townships with especially high churn rates.

Overall, this pipeline could be considered for any length of time for which there is data. As a tool, identifying churn and retention rates over time for any demographic, should prove to be a useful tool for METLN to identify strong and weak business areas and populations.

6. Limitations

One issue within the data we had to work with was the inconsistency of the names of the cities and use of abbreviations. Extra data pre-processing was needed to ensure that cities, like ‘Portland’, were accurately represented within the data and not divided into differently labeled cities of the same name. For instance, ‘Portland’ and ‘PORTLAND’ would be considered two different labels unless normalized and changed to all lowercase lettering.

This also proved to be an issue when it came to geocoding the latitude and longitude of certain cities using the Geopy and Folium libraries to create interactive maps. Cities that had been labeled with unnecessary abbreviations like “Belgrade lks” (which should read “Belgrade Lakes”) were not able to be properly geocoded and thus were dropped from being included within the interactive mapping.

Certain demographic information would be helpful for running more precise and more meaningful insights. For instance, looking at churn and retention rates, we can only cluster this activity based on geographic location and publication. Furthermore, the range in size of zip codes might make comparisons difficult. For that reason, it would have been interesting to see if METLN would consider customer surveys to gather information on age, political affiliation, gender, race, or other metrics to gain more insights into what factors affect their daily active users. This could also help build the different general types of customers to determine the best approach to grow readership across regions or demographics.

Finding more information about customers' relationships with Maine and reasoning for their subscription would help to analyze subscribership outside of Maine. These may be cases such as if they have familial connections in Maine, if it is their home state, or if they are part of the Floridian "snowbirds" category. This opens opportunities for furhter analysis outside of Maine, evaluating not just those living in the state currently but Mainers nationwide.

7. Future Work

There are several different research areas that could be viable and worth exploring to expand the reach of Maine Trust for Local News but were beyond the scope of the preliminary data analysis accomplished for this project. With the addition of more historical subscriber data, it might be interesting to explore various Machine Learning (ML) applications. One idea that could be explored would be: could ML be used to predict which accounts are likely to cancel?

With more data (both to-date and historical) it would be prudent to examine long-term trends. One question to explore includes: what cycles contribute to churn and retention rates? Time-based trends that could provide relevant insights include delving into analysis of various events, such as elections, annual cycles, and major local/national milestones.

The bulk of the data examined for this project revolved around the state of Maine and did not really delve into any external geographic locations at a national (or global) level. It might be worth considering the following questions: how do out-of-state subscribers contribute; is there a correlation with state migration; and are there summer increases in subscribership, i.e. due to tourism or people with second or vacation homes?

Looking at possible areas to expand reach within Maine could include trying to pinpoint areas of future subscribership. This might warrant an investigation into areas surrounding institutions of higher education, as well as local primary and secondary schools. Could there be a pool of untapped subscribers from these demographics?

Finally, here are some suggestions to consider that would make future work with this data easier. It would be helpful to be more selective with your use of abbreviations (town names should probably be recorded in full or abbreviated with some sort of standard convention and key to easily identify and interpret abbreviated data points). Better collection and organization of

subscriber demographic information would be helpful. This might include surveying current subscribers or collecting additional data at the point of registering for new subscriptions that could be used to get a better sense of the reach of your subscriptions (for example, household size, etc.). Lastly, there were some inconsistencies that we found within the data we worked with, it might be useful to implement steps within your data collection that ensures standardizing the data through consistent, valid entries for each variable (such as use of capitalization, spacing, etc.), as well as placeholder variables for 'Null' items.

8. References

City and Town Population Totals: 2020-2024, https://www.census.gov/data/tables/time-series/demo/popest/2020s-total-cities-and-towns.html

Folium, https://folium.readthedocs.io/en/latest/

Geopandas: https://geopandas.org/en/stable/

Geopy, https://geopy.readthedocs.io/en/stable/

Maine Trust for Local News, https://www.metln.org/

U.S. 2023 ME Shapefiles, https://www2.census.gov/geo/tiger/TIGER_RD18/STATE/23_MAINE/