Data Guide: Aggregated Parking Transactions by Pittsburgh Parking Authority

Ryan Negron

LIS 2188: Open Government Data & Information

Dr. Nora Mattern

December 9, 2024

Pittsburgh Parking Authority Aggregate Parking Transactions Data User Guide

I. Original purpose and application

According to the Pittsburgh Parking Authority (PPA) via the Western Pittsburgh Regional Data Center (WPRDC), this dataset records city of Pittsburgh parking transactions for street parking and parking lots that are owned and operated by the PPA, aggregated in ten-minute bins by parking zone. Each row includes the start and end times of the ten-minute bin, the parking zone or lot, the number of transactions, whether the transaction is a mobile or meter payment, and the dollar amount of the payments made during that time span. The data is updated numerous times a day and has temporal coverage from January 2, 2014, until the present day. The data is sourced from multiple mobile app partners and physical meters. This leads to the creation of an enormous dataset in need of warehousing.

Originally, this dataset was created in partnership with the WPRDC to house real-time data about parking transactions to maintain the duty of the Pittsburgh Parking Authority to keep record of meaningful data to impact decision making, urban planning, and budgeting. Using the WPRDC data repository that is ran on the open-source data management software CKAN, data stewards have stored this large dataset and created ways to download subsets of the data to avoid long wait times to be able to use parts of the data.

II. Organizational history and context

The Pittsburgh Parking Authority was established in 1947 to satisfy needs for the City of Pittsburgh outlined in the Parking Authority Law of Pennsylvania, enacted on June 5, 1947. According to the PPA, “The Authority was created for the purpose of conducting the necessary research activity to maintain current data leading to efficient operation of on-street parking facilities, and is authorized by law to plan, design, locate, acquire, hold, construct, improve, maintain and operate, own and lease, either in the capacity of lessor or lessee, land and facilities to be devoted to the parking of vehicles of any kind…” (Pittsburgh Parking Authority).

The law establishing the Parking Authority gives it the power to collect fees, sell leases, rent out commercial spaces within their facilities, and gives it the power

of eminent domain. The Authority can acquire properties through this power of eminent domain and any property it obtains is exempt from real estate taxation. The body is governed by a board consisting of five members all appointed by the Mayor of the City of Pittsburgh. This board is responsible for utilizing the data collected to provide updates to facilities, inform decisions, and balance the Authority's budget.

In more recent history, trends captured within this data show that parking payment is evermore increasing towards mobile only payments (Gobep, 2019) and representatives from the Pittsburgh Parking Authority expect that trend to continue until eventually parking payments are exclusively accepted via mobile app. This context positions the parking authority to invest more of the ascertained budget and resources to support mobile payments, mobile data collection, and infrastructure.

III. **Workflow**

Visitors, residents, and any in need of parking within the city limits of Pittsburgh may seek out longer term parking lots, short term street parking, or a mix of the two. They will need to pay a parking fee for the duration of the parking in that zone or lot. This is where the data originates. A customer paying for their parking fee will either submit payment at a kiosk/meter or online via one of the mobile payment apps that are partnered with the Parking Authority. The PayMobile and GoMobilePGH apps are the two main sources of online payments for the PPA.

Unfortunately, due to licensing issues, the addition of more mobile payment apps has caused for disturbances in data collection for this dataset, which will be revisited later. The data that can be extracted, is done every ten minutes. The data stewards credited with assembling this data are IT Staff within the Pittsburgh Parking Authority. Once IT Staff have collected the data it is compiled into an updated version of the running CSV file that houses over eight million data points. Without explicit knowledge, the assumption is that this data is raw and unedited. This data is then fed to the repository at the Western Pennsylvania Regional Data Center, where it can be accessed openly.

The data is also accessible via Data.gov, the United States national government's website for open government data. This provides users with a direct link to the datasets. Having the data discoverable via Data.gov, allows for more exposure nation wide for the dataset and can provide research data for interested parties to experiment with, compare to, or analyze Pittsburgh's parking transaction data.

IV. **Things to know about the data, including limitations**

As the partnerships between the Pittsburgh Parking Authority, the City of Pittsburgh government, and the WPRDC have evolved, there have been increasing concerns and questions about how the PPA is maintaining their responsibilities to this collection of datasets. There has been staff overturn, budget adjustments, and decision making that have led to the weakening of the partnership between the PPA and WPRDC. Staff members at the WPRDC attempting to ensure the quality of data published for these datasets, are concerned about the PPA's continued decision making about changing payment methods for the parking transactions (Gradeck, 2024). This change or rather addition of vendors who support mobile transactions for the PPA has occurred without attainting the proper licensing from the apps to collect data. For this main reason there are several limitations apparent in this dataset:

- The WPRDC provides a data advisory on this dataset that reads, "As of April/May 2021, the Parking Authority changed its mobile payments system, and since then we have not been able to extract all mobile transactions. We hope to fix this in the future." (WPRDC) There is no additional information about the correction of this problem. One must assume that data collection as it pertains to mobile payments is skewed for the past 4.5 years.
- This dataset grows exponentially, with data being collected and published in increments of 10 minutes. This has led to the culmination of 8 million+ data points. If one is seeking to use this data in whole, they must be prepared for long waited download times and must ensure that they have proper storage quotas to retain this data. The WPRDC has provided abbreviated sets of the data to speed up download times for parts of the data.
- The WPRDC also provides a warning that the link they currently have available to download and access the full dataset will assuredly timeout. They have not provided an update but simply have stated that they are looking into changing this in the future.

- On the webpage of the dataset, there is advertised a parking dashboard, which allows for the data to be viewable in a tangible format for users trying to get a glimpse of the data or filter parts of it. Unfortunately, the link to the HTML dashboard is non-functioning and it can be assumed that this dashboard no longer exists.
- It is unclear without exploring the data whether there is missing data. It is evident in this dataset, however, that there is underrepresented data due to the lapse in coverage for some of the mobile apps. In the dataset, there exists values of zero (0). Where a zero exists, one can assume that it means data was collected but the result was zero. In contrast, if one were to discover null, n/a, or "none" values existing in the dataset, one must assume the data was not collected or available for that period and must be treated as missing, null, or unknown, not as a zero.

V. Additional applications and resources

It is difficult to find any evidence of this exact dataset being used for a particular research study or use case. The nature of the dataset is that it is large and requires truncation to perform any type of analysis. With that said, combining what we know about the responsibility of the Pittsburgh Parking Authority to maintain up to date parking data to inform decision making, one can speculate that this data is being used in reports to the Board of Directors for the PPA. Additional insight can be taken from analyzing reports published by the PPA on their website, however, specific mentions of the data does not exist.

In recent news, the City of Pittsburgh government has voted to extend "purple curb" parking indefinitely. This program was a pilot initiative to monitor high traffic parking areas and seek to limit the duration an individual can park in these areas (KDKA, 2024). The hope with the program was to prohibit drivers from occupying busy areas for extended periods and increase the availability of parking. Using data from this dataset, policy makers can gain insight into where these parking spaces will be best served in an effort to expand the program. Additionally, this data can be used to analyze the effectiveness of the program and whether the incentive to pay the short-term pricing, outweighs the benefits to the user of simply occupying the space and risking the consequences.

Real time parking data can inform traffic patterns as well. A thoughtful article written by AI, Robotics, and smart cities expert Naveen Joshi, details how real time parking data can support sustainable transportation efforts. They outline that this data can impact travel planning, urban planning, support carpool initiatives, public transit efforts, EV infrastructure, and more. The aggregated parking data in this dataset is updated every ten minutes and is quite close to real time. An early effort to research Joshi's claims in the City of Pittsburgh can leverage this dataset, though an expansion of the study may require data refreshed at a higher rate.

VI. Appendix

Below are direct links to where the datasets are available. In these portals, one can find additional information on the coverage, fields included, and data types of the data.

- https://data.wprdc.org/dataset/parking-transactions/resource/1ad5394f-d158-46c1-9af7-90a9ef4e0ce1
- https://catalog.data.gov/dataset?publisher=Pittsburgh%20Parking%20Authority
- https://tools.wprdc.org/downstream/1ad5394f-d158-46c1-9af7-90a9ef4e0ce1

VII. Sources/Acknowledgements

Darnay, M. (2024, November 26). Pittsburgh's purple curb parking zone program has been extended indefinitely. *KDKA News*. Retrieved December 1, 2024, from https://www.cbsnews.com/pittsburgh/news/pittsburghs-purple-curb-parking-smart-loading-zone-program-extended/.

Gobep, R. (2019, July 5). Half Of All Parking Transactions Are Made Through the City's Mobile App. *Development & Transportation*. broadcast, Pittsburgh, PENNSYLVANIA; WESA.

Gradeck, B. (2024). *Open Government Data*. *LIS 2188 Week 12*. Pittsburgh; University of Pittsburgh.

History and Purpose. Public Parking Authority of Pittsburgh. (n.d.). https://www.pittsburghparking.com/About

Joshi, N. (2023, September 23). *10 Ways Real-Time Parking Data Can Support Sustainable Transportation*. December 7, 2024, https://www.linkedin.com/pulse/10-ways-real-time-parking-data-can-support-naveen-joshi/

WPRDC. (2019, January 19). *Aggregated parking transactions - transactions by zone and time of day*. CKAN. https://data.wprdc.org/dataset/parking-transactions/resource/1ad5394f-d158-46c1-9af7-90a9ef4e0ce1

VIII. Additional research plan

An additional research plan to expand on this data guide would require input from informed individuals within the City of Pittsburgh government, the Pittsburgh Parking Authority, and the Western Pennsylvania Regional Data Center. The goal would be to gain insight into decision making and the unique relationship between all three entities. Beginning with the City of Pittsburgh government, information about laws, policies, and amendments that impact open data initiatives, parking infrastructure, or budgeting would be helpful to the expansion of the data guide.

Subsequently, interviewing members of the Board of the Authority in the PPA would provide additional context and use cases for the data as it pertains to the entity that stewards the data. Is this information being leverage to impact decision making like outlined in the About page of the Authority's website? This insight could support expansion of the data guide and provide more concrete examples of uses for the data.

Finally, speaking with the individuals responsible for the technical work on the dataset would be paramount to the expansion of the data guide. IT Staff within the PPA as well as staff from the WPRDC could supply information about the collection, governance, and short comings of the data. There is a lot of mystery behind the trustworthiness of the data provided in this dataset. Due to the multiple data advisories warning users about trouble with data collection, a user could have doubts about the integrity of the data provided. Gaining an understanding through reaching out to the groups/individuals mentioned previously would be a crucial aspect of expanding and implementing the data guide in practice.