Final Report

GROUP B HEAD Competition Project -
Municipal Parking in the City of Hamilton

Nataly Aranda, Sebastian Bravo, Jolly James, Sophia Valdizon

BUSN - 10197 - 01 - Project Management for Business Analytics

April 4, 2022

Table of Contents

Executive Summary

The city of Hamilton is experiencing changes with its public parking. Some parking locations are being repurposed for development while the city believes the introduction of new technologies will seize unrealized opportunities. In order to find insights into customer behavior and parking lot usage, the city of Hamilton provided us with parking data from 2020- 2021 across 57 lots and 76 Cale terminals contained in 13 datasets.

The project involved cleaning and merging the provided datasets while also gathering other relevant datasets to support our analysis. We collected crime data across the Hamilton area from the Hamilton Police Department. Using analytical tools like MS Excel, Tableau and Python scripting we focused on finding lots with high occupancy density and making recommendations for them. We used ArcGIS to create a map of parking lots and their levels of occupancy density. Lot 9DU in the Dundas area was found to have a high occupancy density. This was true for the Dundas parking lots in general. Lot 82, besides the Hamilton General Hospital also had a high occupancy density. We recommended that some existing lots be expanded and new lots be created.

We also looked into customer behavior, to find notable patterns. We found that the customers who use the parking lot more are weekly users. We defined weekly users as users who use the parking lot at least 4 days in a 7 day period. We found the proportion of weekly users to be a significant 20%. These users spend $165.00 a year on average compared to all other users spending an average of $5.28. We concluded that these were the city's most important customers and should be targeted. We recommended a weekly pass, which will increase overall profit and attract new customers.

The crime data was used with the parking data to find if crime rates affected parking lot usage. We found that lots with high usage had high crime rates, especially in the downtown area. This disproved our hypothesis that high crime rates affected parking lots usage. However, we recommended installation of surveillance cameras to give the customers an added sense of security when using the parking lots. Our findings pleased our stakeholders; the city of Hamilton and the HEAD competition judges.

Project Charter

Higher Education Analytics Data Competition	HEAD Competition – Group B	
Project Sponsor: Professor Saifur Rahman	Target Completion Date- Mar. 15,2022.	
Project Manager: Nataly Aranda	Version No.1.0	Feb. 27, 2022.

Project Purpose and Benefits

As the city grows, commercial space for new buildings and redevelopment of the city is driving parking lots to be repurposed. As a result, parking in Hamilton has become scarce. The main objective is to find out why some parking lots are being used and others are not. Our idea for suggestions is to possibly create a new monthly parking pass and yearly parking pass based on data to optimize profit and benefit the user. As well, we had an idea to use ArcGIS to predict where new lots can be placed.

As a sub-thought, after asking a few citizens of Hamilton their opinion on parking issues, the general consensus is that they avoid the downtown core because of crime. We will investigate the truth of this statement by analyzing crime data within the City of Hamilton that has been provided by Hamilton Police. We will determine which wards and neighborhoods have higher crime rates and see if parking lots within them compare to those that have lower crime rates. What is the difference between these neighborhoods? What services are or are not available? Is there a chance for personal risk as a user to park in these lots? Are there facilities that are drawing people to use some parking lots versus others? Our suggestion would be to implement security in these parking lots to increase usage.

The last thought we had was to see if the events at the First Ontario Center or Tim Hortons Field contribute to higher usage rates on particular days. This might not be feasible. It seems no one has tried to look into this option as of yet.

Goal	Strategy / Deliverables	Performance Measure
Clean and prepare datasets	Use Excel to clean the data. Use python to prepare data.	Merge and clean the 6 Passport Datasets as well as the 2 CALE Datasets. Result should be one clean “Masterfile”
Create and analyze graphs	Use Tableau or Cognos to make graphs. Use MS Word to document graph analysis.	Generate 3-4 Visualizations to demonstrate our analysis of the dataset and tell a story
Phase 1: Prepare a poster and presentation of parking analysis	PDF / PowerPoint Poster and a 10-minute video	48″ x 36″ Poster that includes a summary, data modeling explanation, analysis, conclusion, recommendations, and references.

Phase 2: Prepare a 20 minutes presentation of parking analysis answering the second phase questions.	20 minutes PowerPoint presentation showing our findings and the process to obtain them.	Presentation that not exceed the limit time that includes summary, data modeling explanation, analysis, conclusion, recommendations, and references.

"IN" Scope	"OUT" of Scope
Analyze the parking data provided by Hamilton to describe the problem Prescribe data-backed solutions to improve parking in Hamilton Identify new trends or opportunities for parking in the City of Hamilton	Perform a demographic focused analysis of parking lot users, data on users is not available Assess level and usage of "street parking", data is not available on this. Potentially also not feasible to collect such data

High-Level Milestones	Target Completion Dates (Month/Day)
Data Cleaned Up and Ready for Analysis	February 23rd, 2022.
Initial Analysis and Visualizations Generated	February 28, 2022.
Final Analysis and integration of additional data	March 10th, 2022.
Presentation and Poster completed to summarize findings, analysis, and present visualizations	March 15th, 2022.
Phase 2: Final analysis of data, new findings and final oral presentation.	March 24-25, 2022

Project Costs	Associated dates
Obtaining Crime Data - $5	Obtained February 25th, 2022.
TOTALS	**$5**

Project Team and Stakeholders

Group B- Sophia Valdizon, Sebastian Bravo, Nataly Aranda, Jolly James

Team Member Name and Title	Role on the Project	Required Involvement	
		Duration	Level of Effort
Sophia Valdizon Sebastian Bravo Nataly Aranda Jolly James	Team Member Team Member Team Member / PM Team Member	12 weeks	HIGH

Stakeholder Names	Interests & Needs	Strategies to Manage Expectations
City of Hamilton Open Data Program	The project will provide analytic insight for improving parking in the City of Hamilton	.
HEAD Competition	Project will be a result and presented as part of the HEAD Competition along with the projects of our peers in providing analytics insight into this issue.	Datasets provided come from the City o Hamilton Open Data Program. Data outside that to be gathered is encourage but won't be the same as our peers.
Professor Rahman and Way	Both manage the Project Management Course as part of the Analytics Program at Mohawk and oversee students progress in the project for the HEAD Competition	Open dialogue if we have any questions regarding expectations for the HEAD Competition. Prof Way has demonstrated he can also communicate our questions to the City of Hamilton

Other Related Projects & Initiatives	Interdependency & Impact
Mastery in the Analytics Programming class	The programming for analytics class provides additional skills in python needed to successfully prepare the data.
Mastery in the Tableau Prep class	The Tableau class provides skills needed to analyze data and create effective presentations.

Project Risks	Likelihood (Low/Moderate /High)	Impact (Low/Moderate/High)	Risk Response
Performing analysis with bad data Performing analysis beyond the scope of the project	HIGH HIGH	HIGH HIGH	Thorough data cleaning and preparation Proper planning and understanding of the business problem and the datasets.

Assumptions	Constraints
To obtain the parking revenue we assume: ● The daily rate is paid when the hourly charge exceeds the value of the daily rate. ● The hours before the opening or after the closing hours of the parking lot are not charged.	● Difficulty obtaining crime data in parking areas in time to include them in the analysis. ● Difficulty obtaining First Ontario Centre event days data in time to include them in the analysis.

Business Understanding

Business Objectives

1.1- Background

Public parking is rapidly changing in the City of Hamilton, Ontario. Parking lots are being repurposed for development for public activities such as cycling and pedestrian facilities, trees, restaurant patios, etc., while new advances in technologies are creating unrealized opportunities. The City of Hamilton has incorporated connected technologies to improve operational efficiency but not the user experience.

Live data is being collected for every transaction in public surface parking spaces operated through pay-by-plate technology. Soon, this will be expanded to every parking space operated by the City using a variety of systems. It should be noted that parking payment transactions are not an exact indicator of parking occupancy, with many payments exceeding the time the user is actually parking in the spot and others are insufficient payments, however there is a correlation between payments and utilization.

1.2- Business Objectives

The city would like to use the data that is being collected to benefit the user, particularly as public parking becomes less visible and harder to find.

1.3- Business Success Criteria

Giving useful insights into the data and a method that will benefit the user so that the two issues presented are solved: the scarcity of parking being created by repurposing lots and finding a pattern amongst the data that will generate a recommendation that will ultimately save the user money.

Assess Situation

2.1- Inventory of Resources

There are many great resources available to the team. All team members have access to tools such as: Tableau, Python, Excel, ArcGIS Pro, and SPSS.

2.2- Requirements, Assumptions, Constraints

The project requires the data to be harnessed and used in a way that will benefit the user.

The assumptions made are:

- The duration is an accurate measure of the length of stay
- The rate per day can be divided into rate per hour
- If a user stays in the lot for less time than the minimum duration the chargeable minutes is automatically rounded up to the nearest multiple of minimum duration until the maximum daily rate is reached.
- Once the daily rate is reached, the fee cannot increase until the next operating opening hour of operation is reached
- After the next day is reached the count begins again and the maximum daily rate is met once again
- Monthly permits are available but none of the users in this particular dataset are monthly users

The constraints of the project are that each team member has a lot of different responsibilities this semester making time to dedicate to this project a constraint. Another constraint is the possibility of building a predictive modeling using the dataset or combining it with a new one. The optimal prediction would be to determine which factors contribute to choosing a parking lot using demographic data on the user. However, demographic data on the user is not available. Demographic data for the parking lot is available but will not be used by this team. We decided early on that demographic data of the parking lot will not tell us how the parking lots are being used but instead build a picture of what kind of people surround the lots.

2.3-Risks and Contingencies

Each of the team members has a lot of work to do when it comes to other courses, the capstone project and life. This project is not a requirement for graduation and thus might end up last on the list of priority. This would lead to a time constraint and less motivation to devote time to the

work it would take to complete this project. The best solution is to have weekly meetings scheduled to try and hold each other accountable to the progress of the project.

Another risk is attempting to obtain a dataset that is not so readily available for download. Such as the data that was obtained from Hamilton Police. The solution to this problem is to have alternative ideas flowing in case the dataset does not come in time.

A third risk is that the project is being led in an incorrect direction. The solution to this is to remain open to pivot the idea and find a new angle on the dataset.

2.5- Costs and Benefits

The costs to the business in this case is potentially zero and the benefits are high. The cost is zero because the college is taking on students that are working at no cost to the business. If the students are able to find a great solution to implement, the business could potentially benefit by optimizing the user experience and increasing usage in their parking lots.

Determine Data Mining Goals

3.1- Data Mining Goals

The primary objectives of the project are the following:

- Identify underused parking lots based on population density and see if there are any patterns contributing to some lots being used more than others
- Identify a subpopulation of the user that could be catered to with a new permit, such as a weekly pass, weekend pass, or three day pass at a discounted rate that will optimize the user experience.

3.2- Data Mining Success Criteria

Identifying patterns in the data that will lead to a great recommendation will be the criteria to determine success. Even making it into phase 2 of the competition would show that the analysis of the data was effective.

Produce Project Plan

4.1-Project Plan

Stage 1- Clean the data
Duration: 3 weeks
Resources Required: Excel, Python, Tableau
Input: Dirty dataset that needs to be cleaned
Output: Clean accurate dataset that will give us good predictions

Stage 2- Use the rules of each parking lot to determine the fee per transaction
Duration: 3weeks
Resources Required: Python
Input: Clean dataset with all required field inputs to accurately determine the fee per transaction
Output: Accurate transaction fees

Stage 3- Use and test different criteria to determine the top 10 used parking lots and the bottom 10 underused parking lots
Duration: 2 weeks
Resources Required: Tableau, ArcGIS Pro and Excel
Input: Complete, clean dataset
Output: Graphs and maps identifying top ten used parking lots and bottom 10 parking lots along with terminals that were not on the spatial data map that was downloaded on the Hamilton website.

Stage 4- Determine rules to define a subset of the population of users to see if there are any patterns in the users behaviours.
Duration: 2 weeks
Resources Required: Tableau, Python and Excel
Input: Complete, clean dataset
Output: Graphs and a business case with some statistics that will defend our recommendation for the subset of population

Stage 5- Analyze the results and summarize the findings on a poster
Duration: 1 week
Required Resources: Powerpoint and Microsoft Word
Input: Finalized graphs, maps and analytical text
Output: A beautiful poster.

4.2- Initial Assessment of Tools and Techniques

Tools that will be used for this project are: ArcGIS Pro, Python, Tableau, Excel, Powerpoint, Microsoft Word. The techniques that will be used are analyzing data. It was decided early on that a prediction was not possible. The only possibility would be to try and predict where another parking lot could be built using ArcGIS. This was not possible due to the lack of data.

Data Understanding

Data Collection Report

1. **Existing Data**

Dataset ID	Dataset	Source Location	Owner	Description	Intended Use
1	2021 Passport Parking 40340-40354	https://datacompetition.mohawkcollege.ca/2022-data-links/	City of Hamilton	Records of use of the municipal parking lots 40340 to 40354 for 2021	Calculation of payed fee and identification of Top 10 and

					Bottom 10 used Parking lots.
2	2021 Passport Parking 40320-40339	https://datacompetition.mohawkcollege.ca/2022-data-links/	City of Hamilton	Records of use of the municipal parking lots 40320 to 40339 for 2021	Calculation of payed fee and identification of Top 10 and Bottom 10 used Parking lots.
3	2021 Passport Parking 40300-40319	https://datacompetition.mohawkcollege.ca/2022-data-links/	City of Hamilton	Records of use of the municipal parking lots 40300 to40319 for 2021	Calculation of payed fee and identification of Top 10 and Bottom 10 used Parking lots.
4	2020 Passport Parking 40340-40354	https://datacompetition.mohawkcollege.ca/2022-data-links/	City of Hamilton	Records of use of the municipal parking lots 40340 to 40354 for 2021	Calculation of payed fee and identification of Top 10 and Bottom 10 used Parking lots.
5	2020 Passport Parking 40320-40339	https://datacompetition.mohawkcollege.ca/2022-data-links/	City of Hamilton	Records of use of the municipal parking lots 40320 to 40339 for 2020	Calculation of payed fee and identification of Top 10 and Bottom 10 used Parking lots.
6	2020 Passport Parking 40300-40319	https://datacompetition.mohawkcollege.ca/2022-data-links/	City of Hamilton	Records of use of the municipal parking lots 40300 to 40319 for 2020	Calculation of payed fee and identification of Top 10 and Bottom 10 used Parking lots.
7	CALE Terminals 2021	https://datacompetition.mohawkcollege.ca/2022-data-links/	City of Hamilton	Records of use of Cale Web Terminal for 2021	Calculation of payed fee and identification of Top 10 and Bottom 10 used Parking lots.
8	CALE Terminals 2020	https://datacompetition.mohawkcollege.ca/2022-data-links/	City of Hamilton	Records of use of Cale Web Terminal for 2020	Calculation of payed fee and identification of Top 10 and Bottom 10 used Parking lots.
9	CALE Terminals List	https://datacompetition.mohawkcollege.ca/2022-data-links/	City of Hamilton	Cale Web Terminal list with ID, Location, Node	Calculation of payed fee and identification of Top 10 and Bottom 10 used Parking lots.

10	Space Count by Lot	https://datacompetition.mohawkcollege.ca/2022-data-links/	City of Hamilton	Number of spaces by Lot for the Municipal Parking Lots	Calculation of occupation
11	Passport Parking Zone List	https://datacompetition.mohawkcollege.ca/2022-data-links/	City of Hamilton	Information of Zone, Paking Lot ID and Lot Address for the Parking Lots	Identification of Passport Parking zones
12	Hamilton Municipal Parking Lots	https://open.hamilton.ca/datasets/d56d996d4725499da2a5555aa5e5b651_5/about	Open data source from Open Hamilton	Location of Municipal Car Parks in Hamilton.	Identification of Passport Parking zones
13	Map of Hamilton Municipal Parking Lots	https://spatialsolutions.maps.arcgis.com/apps/webappviewer/index.html?id=18eda2d0ea0b4920b84abc5b1bb93d6c	Open data source from Open Hamilton	Map of Municipal Car Parks	Localization of the municipal Parking lots.

2. Purchased Data

Dataset ID	Dataset	Source Location	Owner	Description	Intended Use
14	Crimes Data	Hamilton Police Department database	Hamilton Police Department	Records of crimes occurences and location in Hamilton for 2020 and 2021	Correlation of hot crime zones with underused parking lots.

3. Additional Data

Dataset ID	Dataset	Acquisition Method	Owner	Description	Intended Use
15	Map of Hamilton Parking Meters in Hamilton	Open data: https://open.hamilton.ca/datasets/3538fee1b2eb413581659aed3d1c8ef6_4/explore	Open data source from Open Hamilton	Map of Parking Meters in Hamilton	Insights of use of municipal parking lots and correlations.
16	CALEMASTERFILE.xslx	Compilation of all the data	Group B	Compilation of Terminals and Cale clean data redy to analysis	This is the final dataset to be used in the analysis.

Data Description Report

1. Descriptive Characteristics

Dataset ID	Source Format	Acquisition Method	Dataset Type	Granularity
1	Excel Document	Downloaded	Multivariate dataset	Minutely parking duration by vehicles in every parking lotper transaction
2	Excel Document	Downloaded	Multivariate dataset	Minutely parking duration by vehicles in every parking lot per transaction
3	Excel Document	Downloaded	Multivariate dataset	Minutely parking duration by vehicles in every parking lot per transaction
4	Excel Document	Downloaded	Multivariate dataset	Minutely parking duration by vehicles in every parking lot per transaction
5	Excel Document	Downloaded	Multivariate dataset	Minutely parking duration by vehicles in every parking lot per transaction
6	Excel Document	Downloaded	Multivariate dataset	Minutely parking duration by vehicles in every parking lot per transaction
7	Excel Document	Downloaded	Multivariate dataset	Minutely parking duration by vehicles in every terminal per transaction in 2021
8	Excel Document	Downloaded	Multivariate dataset	Minutely parking duration by vehicles in every terminal per transaction in 2020
9	Excel Document	Downloaded	Multivariate dataset	Location, terminal ID, and installation status of every terminal
10	Excel Document	Downloaded	Multivariate dataset	Street address, location, lot number and space of every parking lot
11	Excel Document	Downloaded	Multivariate	Street address, zone number, and lot number of parking lot
12	CSV File	Downloaded	Multivariate	Street address, location and lot number of parking lots
13	ArcGis map server	Web Map Service	Multivariate	Geographic representation of parking lots in the city of Hamilton
14	Excel document	Downloaded	Multivariate	Location, specification and date of crimes in the city of Hamilton
15	Excel document	Downloaded	Multivariate	Geographic representation of parking meters in the city of Hamilton

16	Excel document	Compilation of all data	Multivariate	Street address, iD of parking lots and terminals. Minutely parking duration of vehicles in parking lots and terminals

2. Quantitative Characteristics

Dataset ID	Size	Number of Records	Number of Columns	Period Covered
1	89KB	17,564	9	2021
2	582KB	10,783	9	2021
3	488KB	9,945	9	2021
4	243KB	4,360	9	2020
5	158KB	2,736	9	2020
6	256KB	4,477	9	2020
7	7,128KB	156,003	6	2021
8	13,374KB	290,319	6	2020
9	12KB	76	4	N/A
10	12KB	57	4	N/A
11	13KB	55	3	N/A
12	4KB	59	4	N/A
14	1,280KB	15,531	6	2020 - 2021
16	97,226KB	496,187	11	2020-2021

3. Data Dictionary

Dataset ID: 1										
Field Name	**PK**	**FK**	**Data Type**	**Value Type**	**Length**	**Description**	**Format or Coding**	**Data Range**	**Analytically Relevant**	**NULL Ok**
Rate Name	N	N	String	NominaL	9	Description of the lot fee rate	text|text |	N/A	Y	N
Zone	N	N	Integer	Categorical	4	Zone numbers for the lots	N/A	40340 - 40354	Y	N
Zone Name	N	N	String	Categorical	9	The name of the zone	N/A	N/A	N	N
Vehicle	y	N	String	Categorical	7	ID of different vehicles using parking lots	N/A	N/A	Y	N
Entry Time	Y	N	Date	Continuous	10	Start time of a parking reservation	MM/DD/YYYY HH:MM	N/A	Y	N
Exit Time	Y	N	Date	Continuous	9	End time of a parking reservation	MM/DD/YYYY HH:MM	N/A	Y	N
Duration	N	N	integer	Continuous	8	Duration in minutes of time spent parked	N/A	N/A	Y	N
Method	N	N	String	Categorical	6	Platform of payment for lot fee	N/A	N/A	Y	N
Total Extensions	N	N	Number	Continuous	16	Number of extensions of parking time per transaction	N/A	N/A	Y	N

Dataset ID: 2										
Field Name	**PK**	**FK**	**Data Type**	**Value Type**	**Length**	**Description**	**Format or Coding**	**Data Range**	**Analytically Relevant**	**NULL Ok**
Rate Name	N	N	String	NominaL	9	Description of the lot fee rate	text|text |	N/A	Y	N
Zone	N	N	Integer	Categorical	4	Zone numbers for the lots	N/A	40320 - 40339	Y	N
Zone Name	N	N	String	Categorical	9	The name of the zone	N/A	N/A	N	N
Vehicle	y	N	String	Categorical	7	ID of different vehicles using parking lots	N/A	N/A	Y	N
Entry Time	Y	N	Date	Continuous	10	Start time of a parking reservation	MM/DD/YYYY HH:MM	N/A	Y	N
Exit Time	Y	N	Date	Continuous	9	End time of a parking reservation	MM/DD/YYYY HH:MM	N/A	Y	N
Duration	N	N	integer	Continuous	8	Duration in minutes of time spent parked	N/A	N/A	Y	N
Method	N	N	String	Categorical	6	Platform of payment for lot fee	N/A	N/A	Y	N
Total Extensions	N	N	Number	Continuous	16	Number of extensions of parking time per transaction	N/A	N/A	Y	N

Dataset ID: 3										
Field Name	**PK**	**FK**	**Data Type**	**Value Type**	**Length**	**Description**	**Format or Coding**	**Data Range**	**Analytically Relevant**	**NULL Ok**
Rate Name	N	N	String	NominaL	9	Description of the lot fee rate	text|text |	N/A	Y	N
Zone	N	N	Integer	Categorical	4	Zone numbers for the lots	N/A	40300 - 40319	Y	N
Zone Name	N	N	String	Categorical	9	The name of the zone	N/A	N/A	N	N

Vehicle	y	N	String	Categorical	7	ID of different vehicles using parking lots	N/A	N/A	Y	N
Entry Time	Y	N	Date	Continuous	10	Start time of a parking reservation	MM/DD/YYYY HH:MM	N/A	Y	N
Exit Time	Y	N	Date	Continuous	9	End time of a parking reservation	MM/DD/YYYY HH:MM	N/A	Y	N
Duration	N	N	integer	Continuous	8	Duration in minutes of time spent parked	N/A	N/A	Y	N
Method	N	N	String	Categorical	6	Platform of payment for lot fee	N/A	N/A	Y	N
Total Extensions	N	N	Number	Continuous	16	Number of extensions of parking time per transaction	N/A	N/A	Y	N

Dataset ID: 4										
Field Name	**PK**	**FK**	**Data Type**	**Value Type**	**Length**	**Description**	**Format or Coding**	**Data Range**	**Analytically Relevant**	**NULL Ok**
Rate Name	N	N	String	NominaL	9	Description of the lot fee rate	text|text |	N/A	Y	N
Zone	N	N	Integer	Categorical	4	Zone numbers for the lots	N/A	40340 - 40354	Y	N
Zone Name	N	N	String	Categorical	9	The name of the zone	N/A	N/A	N	N
Vehicle	y	N	String	Categorical	7	ID of different vehicles using parking lots	N/A	N/A	Y	N
Entry Time	Y	N	Date	Continuous	10	Start time of a parking reservation	MM/DD/YYYY HH:MM	N/A	Y	N
Exit Time	Y	N	Date	Continuous	9	End time of a parking reservation	MM/DD/YYYY HH:MM	N/A	Y	N
Duration	N	N	integer	Continuous	8	Duration in minutes of time spent parked	N/A	N/A	Y	N
Method	N	N	String	Categorical	6	Platform of payment for lot fee	N/A	N/A	Y	N

Total Extensions	N	N	Number	Continuous	16	Number of extensions of parking time per transaction	N/A	N/A	Y	N

Dataset ID: 5										
Field Name	**PK**	**FK**	**Data Type**	**Value Type**	**Length**	**Description**	**Format or Coding**	**Data Range**	**Analytically Relevant**	**NULL Ok**
Rate Name	N	N	String	NominaL	9	Description of the lot fee rate	text|text |	N/A	Y	N
Zone	N	N	Integer	Categorical	4	Zone numbers for the lots	N/A	40320 - 40339	Y	N
Zone Name	N	N	String	Categorical	9	The name of the zone	N/A	N/A	N	N
Vehicle	y	N	String	Categorical	7	ID of different vehicles using parking lots	N/A	N/A	Y	N
Entry Time	Y	N	Date	Continuous	10	Start time of a parking reservation	MM/DD/YYYY HH:MM	N/A	Y	N
Exit Time	Y	N	Date	Continuous	9	End time of a parking reservation	MM/DD/YYYY HH:MM	N/A	Y	N
Duration	N	N	integer	Continuous	8	Duration in minutes of time spent parked	N/A	N/A	Y	N
Method	N	N	String	Categorical	6	Platform of payment for lot fee	N/A	N/A	Y	N
Total Extensions	N	N	Number	Continuous	16	Number of extensions of parking time per transaction	N/A	N/A	Y	N

Dataset ID: 6										
Field Name	**PK**	**FK**	**Data Type**	**Value Type**	**Length**	**Description**	**Format or Coding**	**Data Range**	**Analytically Relevant**	**NULL Ok**
Rate Name	N	N	String	NominaL	9	Description of the lot fee rate	text|text |	N/A	Y	N
Zone	N	N	Integer	Categorical	4	Zone numbers for the lots	N/A	40300 - 40319	Y	N

Zone Name	N	N	String	Categorical	9	The name of the zone	N/A	N/A	N	N
Vehicle	y	N	String	Categorical	7	ID of different vehicles using parking lots	N/A	N/A	Y	N
Entry Time	Y	N	Date	Continuous	10	Start time of a parking reservation	MM/DD/YYYY HH:MM	N/A	Y	N
Exit Time	Y	N	Date	Continuous	9	End time of a parking reservation	MM/DD/YYYY HH:MM	N/A	Y	N
Duration	N	N	integer	Continuous	8	Duration in minutes of time spent parked	N/A	N/A	Y	N
Method	N	N	String	Categorical	6	Platform of payment for lot fee	N/A	N/A	Y	N
Total Extensions	N	N	Number	Continuous	16	Number of extensions of parking time per transaction	N/A	N/A	Y	N

Dataset ID: 7										
Field Name	**PK**	**FK**	**Data Type**	**Value Type**	**Length**	**Description**	**Format or Coding**	**Data Range**	**Analytically Relevant**	**NULL Ok**
Vehicle_Plate	Y	N	String	Categorical	13	ID of different vehicles using parking lots	N/A	N/A	Y	N
Start_Date	Y	N	Date	Continuous	10	start day of parking reservation	YYYY-MM-DD	N/A	Y	N
Start_Time	Y	N	Date	Continuous	10	start day of parking reservation	HH:MM:SS	N/A	Y	N
End_Date	Y	N	Date	Continuous	8	start day of parking reservation	YYYY-MM-DD	N/A	Y	N
Endt_Date	Y	N	Date	Continuous	10	start day of parking reservation	HH:MM:SS	N/A	Y	N
Terminal_	Y	N	String	Categorical	9	Terminal ID	N/A	N/A	Y	N

Dataset ID: 8										
Field Name	**PK**	**FK**	**Data Type**	**Value Type**	**Length**	**Description**	**Format or Coding**	**Data Range**	**Analytically Relevant**	**NULL Ok**
Vehicle_Plate	Y	N	String	Categorical	13	ID of different vehicles using parking lots	N/A	N/A	Y	N
Start_Date	Y	N	Date	Continuous	10	start day of parking reservation	YYYY-MM-DD	N/A	Y	N
Start_Time	Y	N	Date	Continuous	10	start day of parking reservation	HH:MM:SS	N/A	Y	N
End_Date	Y	N	Date	Continuous	8	start day of parking reservation	YYYY-MM-DD	N/A	Y	N
Endt_Date	Y	N	Date	Continuous	10	start day of parking reservation	HH:MM:SS	N/A	Y	N
Terminal_	Y	N	String	Categorical	9	Terminal ID	N/A	N/A	Y	N

Dataset ID: 9										
Field Name	**PK**	**FK**	**Data Type**	**Value Type**	**Length**	**Description**	**Format or Coding**	**Data Range**	**Analytically Relevant**	**NULL Ok**
Terminal ID	Y	N	string	Nominal	11	Terminal ID of terminals	None	N/A	N	N
Location	N	N	String	Nominal	8	Address of terminals	None	N/A	Y	N
Node	N	N	String	Nominal	4	Terminal node	None	N/A	Y	N
Terminal Installation Status - Name	N	N	String	Ordinal	35	Status of the terminal	None	N/A	Y	N

Dataset ID: 10										
Field Name	**PK**	**FK**	**Data Type**	**Value Type**	**Length**	**Description**	**Format or Coding**	**Data Range**	**Analytically Relevant**	**NULL Ok**
Lot #	Y	N	String	Nominal	5	Lot Number	N/A	N/A	Y	N
Municipal Address	N	N	String	Categorical	17	Parking lot address	N/A	N/A	Y	N

Dataset ID: 11										
Field Name	**PK**	**FK**	**Data Type**	**Value Type**	**Length**	**Description**	**Format or Coding**	**Data Range**	**Analytically Relevant**	**NULL Ok**
Zone #	Y	N	Integer	Continuous	6	Zone number	N/A	N/A	Y	N
Lot Number	N	N	Integer	Continuous	10	Lot Number	N/A	N/A	Y	N
Lot Address	N	N	String	Categorical	11	Address of lots	N/A	N/A	Y	N

Dataset ID: 12										
Field Name	PK	FK	Data Type	Value Type	Length	Description	Format or Coding	Data Range	Analytically Relevant	NULL Ok
OBJECTID	Y	N	Integer	Continuous	8	ID of the dataset	N/A	321 - 379	N	N
LOT_NUMBER	N	N	Integer	Continuous	10	Lot number	N/A	N/A	Y	N
LOCATION	N	N	String	Nominal	8	Location of parking lot	N/A	N/A	N	N
ADDRESS	N	N	String	Nominal	7	Address of parking lot	N/A	N/A	Y	N

Dataset ID: 14										
Field Name	**PK**	**FK**	**Data Type**	**Value Type**	**Length**	**Description**	**Format or Coding**	**Data Range**	**Analytically Relevant**	**NULL Ok**
Class	N	N	String	Categorical	5	ID of different vehicles using parking lots	N/A	N/A	Y	N
First Date	N	N	Date	Continuous	11	when the crime occured	YYYY-MM-DD	N/A	Y	N
RD	N	N	String	Categorical	2	Area in the city of Hamilton where the crime occured	HH:MM:SS	N/A	Y	N

STREET	Y	N	String	Categorical	6	Street where the crime occured	YYYY-MM-DD	N/A	Y	N
LOCATION TYPE	N	N	String	Categorical	14	What kind of building did the crime occur in	HH:MM:SS	N/A	Y	N

Dataset ID: 16										
Field Name	**PK**	**FK**	**Data Type**	**Value Type**	**Length**	**Description**	**Format or Coding**	**Data Range**	**Analytically Relevant**	**NULL Ok**
Rate Name	N	N	String	NominaL	9	Description of the lot fee rate	text|text |	N/A	Y	N
Zone	N	N	Integer	Categorical	4	Zone numbers for the lots	N/A	40340 - 40354	Y	N
Zone Name	N	N	String	Categorical	9	The name of the zone	N/A	N/A	N	N
Vehicle	y	N	String	Categorical	7	ID of different vehicles using parking lots	N/A	N/A	Y	N
Entry Time	Y	N	Date	Continuous	10	Start time of a parking reservation	MM/DD/YYYY HH:MM	N/A	Y	N
Exit Time	Y	N	Date	Continuous	9	End time of a parking reservation	MM/DD/YYYY HH:MM	N/A	Y	N
Duration	N	N	integer	Continuous	8	Duration in minutes of time spent parked	N/A	N/A	Y	N
Method	N	N	String	Categorical	6	Platform of payment for lot fee	N/A	N/A	Y	N
Total Extensions	N	N	Number	Continuous	16	Number of extensions of	N/A	N/A	Y	N

						parking time per transaction				
Zone	N	N	Integer	Categorical	4	Zone numbers for the lots	N/A	40300 - 40319	Y	N
Zone Name	N	N	String	Categorical	9	The name of the zone	N/A	N/A	N	N

Data Exploration Report

Datasets Used	Exploration Technique	Exploration Results	Promising Attributes	Hypothesis	Further Investigation
Crime data	Geospatial mapping	Crime data around parking lots in city of Hamilton	Possible correlation between crime rates and parking lot low usage	Crime parking lot usage	What areas of the city are most affected and what areas aren't
CALEMASTER data	Graphs and Programming	Discovered top 10 most frequent users of parking lots. Discovered top 10 used parking lots.	Patterns between top users to show what time and what day they use parking lots most.	Users who use pacific parking lots significantly work around the area.	Look for patterns in their behavior to find data to support a possible weekly pass recommendation.

Data Quality

Problem ID	Dataset ID	Problem Type	Field Affected	Description
1PASS	Passport Files (2020 and 2021) - All of them	Coding Inconsistencies	"Rate Name"	Too much information in this field - Rate hr - Operating Days - Operating TImes - Minimum Minutes - All Day Rate. Can be broken down.
2PASS	Merged Dataset (CALE + PASSPORT Files)	Wrong Datatype	Lot #, Min Minutes, Mins for Daily Rate, etc.	Dataset needs a review when being loaded into Tableau as it's reading data types to INT when they should be string values as these are labels, and calculations won't be done on them.
1CRIME	Accurint_Crime_Records	Merged Data	PUBLIC_ADDRESS	Field contains both district name and street name. In order for analysis to be done need Street Name Separated

2CRIME	Accurint_Crime_Records	Merged Data	Street	Street Still merged as there are “Intersections” as denoted by an &. Also it seems like the cardinal direction of the first street in an intersection is after the 2nd street as well. Needs investigating
1CALE	CALE_Terminal	Merged Data	Terminal_ID	Terminal ID needs to be split to find the lot number and form there we can derive rate/hr, operating days and hours and additional lot information for analysis

Data Preparation

Data Cleaning

Problem ID	Resolution	Outcome	Completed
1PASS	Split the “Rate Name” field into multiple fields - Rate/hr, Operating Days, Lot Opening Hour, Lot Closing Hour, Minimum duration, All Day Rate	The solution resulted in the creation of 5 additional fields at minimum that allows for easier to manipulate and categorize data for analysis	Completed!
2PASS	Identified all values being read when loaded into Tableau for visualization purposes and ensured they were being read correctly. Changed INT values to string where appropriate (Lot# for example) as calculations would not be done on this data necessarily.	Datatype is now changed to it’s appropriate and intended type and proper analysis can be done to the data	Completed!
1CALE	Split the “Terminal_ID” field to obtain the Lot Number. With the lot number add the information of Rate/hr, Operating Days, Lot Opening Hour, Lot Closing Hour, Minimum duration, and All Day Rate of the Dataset ID 12	The solution resulted in the creation of 8 additional fields at minimum that allows for easier to manipulate and categorize data for analysis	Completed!
1CRIME	Separated the Street and District Names by use of a function in Excel as the city’s were limited to Hamilton, Stoney Creek, Flamborough, etc.	Resulted in two separate fields for street names and we only wanted data that concerned Hamilton. Can then manipulate Street Data when necessary	Completed!

2CRIME	In Crime Data Street Name was still combined with intersections (two streets) as denoted by an & and the cardinal direction of the first street was after the second street. Separated the field by & and then used a formula to bring cardinal direction (N, S, W, E) to the first street	Resulted in a new separate field if the crime occurred on an intersection by way of “Intersecting Street” and Street Names made more accurate by being combined with correct direction (Main St to Main St E for example)	Completed

Modeling

Test design

Test Number: 1-Data preparation				
Test Goal: Prepare datasets from terminals and passports for analysis.				
Step	Dataset ID	Activity	Description	Result
1	1-6	Merge all Passport data in one dataset and clean it.	After merging all 6 datasets we did clean the data as described in 1Pass Problem ID. Additionally, the duration was calculated for each record using the entry and exit hours.	First dataset: All the passport data ready.
2	7-9	Merge all Cale Terminals data in one dataset and clean it.	After merging the datasets we did clean the data as described in 1CALE Problem ID. Additionally, the duration was calculated for each record using the entry and exit hours.	Second dataset: All the Cale Terminals data are ready.
Test Result Summary: Datasets were prepared for analysis.				

Test Number: 2-Fees calculation				
Test Goal: Obtain the fee paid by each user. A python script was written to obtain the fees				
Step	Dataset ID	Activity	Description	Result
1	Passports	Calculate the fee paid by users (all records) in the Passport dataset.	A python script was written to obtain the payment fee. It was taking into account the weekends and holidays using the	The fee paid in each record was calculated for all the Passports datasets.

			Business-duration package from Python.	
2	Cale terminals	Calculate the fee paid by users (all records) in the Cale Terminals dataset.	A python script was written to obtain the payment fee. It was taking into account the weekends and holidays using the Business-duration package from Python.	The fee paid in each record was calculated for all the Cale Terminals datasets.
3	16	Cale Terminals and Passports data sets were merged to obtain one dataset. Some fields were deleted and fields as method and type were included to identify the source of data (Cale terminals or passports).	Some fields were deleted and fields as method and type were included to identify the source of data (Cale terminals or passports).	The final dataset was called CALEMASTERFILE.XSLS, and it is ready for analysis.
Test Result Summary: Datasets include the fee field and the two datasets were merged into one.				

Test Goal: Crime locations data in Hamilton				
Step	Dataset ID	Activity	Description	Result
1	14	Calculate the location for crime data in Hamilton.	The obtained crime data was cleaned and prepared following the 1CRIME and Street Cardinal Directions and Intersections Problems Id.	The crime date is ready to be mapped.
2	14	Map the Crime data.	Using ArcGis Pro software the crime locations in Hamilton were mapped.	Map of crime locations in Hamilton.
Test Result Summary: It was obtained the map of crime locations in Hamilton.				

Test Goal: Occupation density and average occupation for all Lots.				
Step	Dataset ID	Activity	Description	Result
1	Terminal occupation and Parking lots occupation datasets	Calculate the average daily occupation by lot and the occupation density for the parking lots.	It was written a Python script to calculate daily average occupation and density for the parking lots.	Data for Average occupation and occupation density by lot.
2	Terminal occupation and Parking lots	Map occupation data.	Using ArcGIS Pro software it was mapped the daily average occupation by lot for the Cale terminals and the Occupation	Map of lots of occupations.

	occupation datasets		density for the Passport Parking lots.	
3	Map Crime and occupation.	Map crime and occupation to identify correlation or trends in the data.	Maps of occupation and crime distribution in Hamilton were merged using the ArcGIS Pro software.	Map of Crime and Parking lots occupation - Hamilton
Test Result Summary: It was obtained the map of parking occupation and crime locations in Hamilton.				

Evaluation

Evaluate Results

With the data analysis, we were able to identify underused parking lots based on population density and the patterns contributing to some lots being used more than others. Also, we identified a subpopulation of the user that could be catered to with a new permit, such as a weekly pass, weekend pass, or three-day pass at a discounted rate that will optimize the user experience.

The models approved during this project were the fee per transaction for Cale and Passports datasets, Average Daily Occupation for Cale Terminals data, and Daily Average Occupation density for Passports dataset. All the models proved to be correct and were important tools to obtain insights into the use of parking lots and terminals in the city of Hamilton.

Further, the data obtained in the models were plotted in a map, showing the occupation for the Parking lots and Terminals, the criminal activity around the city, and important locations in the city as hospitals and municipal service centers that helped us to have a better understanding of the city parking lots locations and its use.

The criminal data analysis did not show any relevant trends to rates of occupation. We were expecting to find an important correlation between criminal activity and underused lots, however, criminal data did not provide important insight into the use of parking lots on Hamilton.

Review Process

The model that was built used all data that was available in the dataset, along with a few other variables that were created in order to complete the parking lot fee per transaction. The model addresses the primary objectives of the project, all underused and high-traffic parking lots were determined using a population density and a weighted average formula. A subpopulation was identified of users that frequently use the parking lots on a weekly basis throughout the year. The city of Hamilton should use the python code that was generated in order to determine the number of chargeable minutes for a transaction.

Determine Next Steps

A list of possible next actions includes:

- Completing the poster and submitting the results to the judges to be evaluated
- Implementing the changes through the business to optimize the user experience

The team has decided to complete the poster and submit the final poster and video to the competition for the judges to evaluate and assess.

Deployment

Deployment Plan

The plan for deployment would be to look into getting permits to build bigger parking structures to meet the demand for the lots that were closed, particularly in the areas in Dundas and near Hamilton General Hospital. As well as look into creating the weekly parking pass that will meet the needs of many of the users and how they can implement it. For instance, would they make it available on the Passport App or implement a paper copy that can be purchased at a terminal and displayed on a windshield or would they make it available for purchase online.

Additionally, we assume the City of Hamilton has deployed resources into the creation of the Passport App in order to aid citizens of Hamilton in parking at the Municipal Parking lots. Our findings hopefully support the city in that Passport usage still has room for growth and lead them to integrating it with CALE Terminals as it can see an increase in benefits and ease of parking for those using Municipal Parking Lots.

Maintenance Plan

Maintenance after deployment would have to be through further usage of data generated by both the Passport app and CALE Terminals. As our analysis looked at average usage throughout the day compared to lot size (parking spaces) to identify lots that were heavily used throughout the course of the day the same methodology would have to be employed throughout the life cycle of Municipal Parking Lots. Should usage of lots change, the City of Hamilton should respond as well. For example, if usage increases in lots previously unused plans may change and considerations should be given to expand lots in an area or not. Furthermore, results of expansion and additional lots that were a result of initial deployment should be heavily monitored to see if residents take advantage of the initial space and that they relieve congestion in the area in particular.

Further integration of the Passport Parking app for lots that exclusively use CALE Terminals would also require maintenance in the same way

Maintenance can be implemented by monitoring the usage of parking lots as well as weekly parking pass permit purchases. It could be that as time goes on, the subpopulation could split or a new one could be created. Meaning, there might be more of a demand for a new parking pass such as a three day pass or a weekend pass. Continually monitoring the implementation of the

weekly pass and attempting to determine if a new subpopulation would benefit from a new package is important for the business.

Close Report

Project Summary

Project scope covered the ability to analyze the parking data generated by both the Passport App and Cale Terminals to describe the problem of parking in Hamilton and prescribe data-backed solutions to improve it while identifying new trends and/or opportunities for parking in the city of hamilton. Performing a demographic focused analysis of parking lot users and assessing street parking were, and continue to remain, out of scope. However, during the course of the project we added an analysis of crime data from 2020-2021 provided by the Hamilton Police Department as well to figure into our analysis of parking lot usage which was a change in scope. Another change of scope was doing a density analysis of the municipal parking lots and of crime data using ARCGIS. Lastly, identifying the subpopulation of weekly users was a change in scope. Originally, we wanted to look at the top users of the parking lots and see if we observed any behaviors and we identified these users that are using the lots at a high rate and we pivoted our project.

Project Success Criteria

Objective/Deliverables	Met?	Comments
Clean and prepare datasets	Yes	Once we decided as a group what data we needed, we deleted the others and prepped the data for analysis.
Create and analyze graphs	Yes	We decided to use Tableau, MS Excel and ArcGIS to create visualizations. We used a pie chart, tree map, bar chart and ArcGIS map as our visualizations.
Prepare a poster and presentation of parking analysis	Yes	We used MS PowerPoint to create the poster and record our video presentation.

Statement of Sponsor and Stakeholder Feedback

According to the stakeholders, the objectives of the project which was to see the different behaviors of the lots was fully met and the information obtained and analyzed was correct. Some recommendations from stakeholders included improving the legends and axis titles in the graphs and using timestamps to calculate the hourly occupation.

The feedback on the poster from our sponsor included excellence in data effectiveness, overall impression, abstract, and background quality. Things that were good but can be improved to get to excellence were the flow, the writing, reference, and discussion quality, and the conclusions and recommendations.

Project Challenges

1. Anonymized data - The crime data we received from the police department was anonymized and it made our analysis more tedious. We imputed the values to de-

anonymize the data to make it useful for analysis. For example, “7XX Barton St.” was changed to “700 Barton St.” the value XX was imputed as 00.

2. Deleting data that we were unable to use because the address was given as an intersection. This was a difficult decision for us to make because originally we saw that a lot of crime was happening on Barton St. and the action of deleting these data points took away from that original analysis as a lot of the intersections given were on Barton St. We believe this data was random street crime data points as well, meaning crime that did not occur at a residence or commercial address but random crime that did not have an exact address.
3. Utilizing the crime data - Initially wanted to test the assumption that more crime in the downtown core implied less people utilizing parking lots in that area. Analysis showed that wasn’t the case and there was a point we thought our analysis was for nothing. Ultimately it nullified the original hypothesis to show that people parked in high crime areas and steps should be taken for increased security measures in these lots.
4. Learning to use ARCGIS as an app to develop a map visual and provide context for density analysis of both parking lot usage and of crime reported to have happened in the City in 2020-2021. This was a new skill to learn to utilize it for the most benefit and had problems with making a web app with it in terms of activating interactivity to the poster
5. Calculating parking lot fees - Was tricky to figure out an Excel formula to calculate parking lot fees generated from citizens paying for municipal parking while factoring in hours of operation, particularly 8am-2am. Ultimately took steps to use Python functions to do this. The 8am-2am parking lots still presented a problem even with a python library calculating the chargeable minutes of each transaction. In the end, we shifted back all the times by three hours in order to keep all the entry times, exit times, opening hour of operation and closing hour of operation within the same day in order to eliminate the 2AM closing time. For example, lots that had operating hours of 8am-2am now had operating hours of 5am-11pm with the shift taken into account.
6. We also weren’t too sure how the parking system really worked, so we went out into Hamilton to clarify some questions we had. After looking at a few signs, visiting a few parking lots and making a few calls to the city parking lot line, we were made aware of the fact that regardless of duration time, minimum minutes determined the fee per transaction. Meaning, the maximum minutes of the parking lot fees is always a multiple of the minimum minutes. This was key to finding the fees because it allowed us to express the fees per transaction as a rate for minimum minutes and the daily fee as a rate for minutes for daily rate.
7. Another problem with determining the fee per transaction was with transactions that lasted for more than one day. These were also creating some anomalies with the Excel function that was implemented. A little tweak to the function that accounted for these transactions gave us an accurate representation of the fees.
8. Parking data included, this was in fact not true as some of the terminals that had missing information were on the street. It was also determined that street terminals have a maximum time limit of three hours. Meaning, that you could not park on the street for longer than 3 hours or risk a parking ticket. For this reason, the CALE terminals were excluded from the definition of the subpopulation of weekly users as all of these transactions will have a

duration of three hours or less.

9. Defining the subpopulation of weekly users also presented many issues. At first the entire population was being taken into account to determine what percentage of users were behaving as weekly users. We then eliminated many users by applying four rules: only transactions that took place in the year 2021 were accounted for, only users that had used the parking lot for at least the daily duration, only transactions that involved parking lots and only users that had used the lots at least 4 times of the week at least one week out of the year.
10. Accounting for users that were using the parking lots outside of the traditional work week of Monday-Sunday. For instance, there are quite a few people that would use the parking lots say, from a Wednesday-Tuesday at least four times during that seven day period but the original code was not taking them into account. We found a clever way to take the people into account by creating a window of time and looking at the difference between the number of days between each entry date of each user. This accounted for all users that were entering within 4, 5, 6 and 7 days regardless of the weekday and regardless of the week of the year.
11. Determining a measure to distinguish between high-traffic parking lots and low-usage parking lots. We wanted to take into account the size of the parking lots and for this reason we used a weighted average to distinguish between the lots. As well, we were not sure if we should use the number of users that entered the parking lot, the average duration of stay per user of each parking lot or a population density. It was determined that population density is the most accurate representation of the usage per parking lot.

Lessons Learned

1. When we got the data from Hamilton Police, it was clear that the data was not showing us what we wanted it to show, which in the moment felt like a major failure. This particular situation is one that we may face many times in the future. The lesson learned is that even though the data doesn't show what we're expecting, there is no reason to label it as wasted effort because it could be showing something that is still interesting and useful. In our case, the data nullified our original hypothesis and gave us a small piece of information that we didn't expect.
2. Data cleaning, creating visualizations, finding the right angle, obtaining secondary datasets, and even finding the time to work together are all tasks that take a significant amount of time and we ran around in circles a few times. That being said, we were relentless in our effort to complete the project to the best of our ability and it led to our success. In the future, experience with more projects will help us work around these issues. Also, in a place of work, it is a lot easier to dedicate time to a project because we are being well compensated for our time.
3. Data analysis and creation of visualizations is just as important as data preparation and modeling. We must leave enough time for these two activities, since even if we have done a great job modeling, if we do not analyze and present the results well, we will not have the best results.

Project Wins

1. Using python to determine the fee per transaction expanded our knowledge of python coding and helped us all bond as a team working together toward the same goal and experiencing challenges together. It also gave us the ability to show how much each of the weekly users were spending on average which was a powerful tool to drive our reason for a weekly pass.
2. The ARCGIS Map to capture a visualization and understand parking lot usage density and crime density in the city was a win for us as it allowed us to really contextualize and visualize the problem in reference to the city as a whole and make use of the crime data we requested.
3. Working on a team always presents many issues and we were always able to overcome our issues and work together to complete the project. There were approximately 50 teams that registered to compete in this data competition and of the 50 teams only 20 teams submitted a poster and presentation and of those twenty we ranked first in Phase I. This was the biggest win of all. We were able to make it out of the entire competition in the top three.

Outstanding Items

Outstanding Tasks/Issues	Assigned To	Target Completion Date
HEAD Competition - Phase II	Entire Team	March 25, 2022.
Project Close Report	Entire Team	March 28, 2022.
Final Report	Entire Team	April 4, 2022.

Out of scope – Future items

1. Performing a demographic analysis of parking lot users, data on users is not available. This is in reference to the fact that while the data in both the Passport Mobile App dataset and the Cale Terminal dataset had a vehicle number, further demographic information was unknown. Can be addressed if the Passport app has a register of some demographic data of the user using the app. Might be more difficult to collect with CALE Terminal as that seems to only collect payment and not user information.
2. Assessing the level and usage of “street parking” was out of scope. This is in reference to parking meters found on the side of streets. Data was not available for this and this would have been more beneficial to analyze which streets were busy and parking as this can help analyze potential problem streets that were busy. It potentially isn’t feasible to collect this data, but if there was some sort of mobile integration to pay by phone for parking meters on the side of the street in the future - it could be a start.
3. Performing an analysis on high-traffic days of the year for parking lots that are close to arenas. We are still interested to see how sports events impact the usage of municipal parking lots when there are events such as “Disney on Ice” or concerts or hockey games or football games or soccer games. This analysis was not performed as no dataset to compare events to parking lot traffic was obtained.

Action Items and Considerations		
Priority	**Action Item**	**Department/Service Area**
High	Allocate enough time to analyze the data modeled and to produce the visualizations and the final presentation.	Data Analysis Team
High	Effective and descriptive visualizations in the presentation to enhance storytelling and audience understanding	Data Analysis Team
Medium	Spend less time cleaning the data to merge into the dataset. Some errors were made that required going back and re-cleaning the data	Data Analysis Team

Conclusion

The results of our project during PHASE I found that the proportion of users using a parking lot at least 4 times a week is a significant proportion. On average they are spending $165.00 a year on average compared to all other users spending an average of $5.28. These people are the City of Hamilton's top users and deserve to be compensated for their loyalty. We also found that our original hypothesis that people would use parking lots less in areas of high crime within the City of Hamilton was false according to the data and by the Density Map of Parking Lots we created with a heat map of Crime overlaid on top.

People who need to use a municipal parking lot will park regardless and crime activity did not influence that decision. This was the result of our cleaning of the datasets provided by the City of Hamilton for the HEAD Competition that contained payment information from the PASSPORT Mobile App and CALE Terminals located on site in the Municipal Parking Lots. Furthermore, we found that the areas of Parking Lots clustered in the Dundas area have high occupancy rates and the city would have to carefully plan how they utilize this land and may consider restructuring. A similar pattern was noted on Barton Street E near the General Hospital with Parking Lot 82.

This data was cleaned and merged using Excel with additional details to derive revenue calculated using a Python code. Visualizations were made using Tableau to present our findings and additionally by ARCGIS Online to create a map that showed the density of parking lots and a crime heat map using data supplied by the City of Hamilton Police Department upon our request.

Recommendations

The recommendations we made based on our results in PHASE I of the HEAD Competition are as follows:

- Considering 20% of the current users are eligible for a weekly pass, it would be great investment.
- The Dundas Area, specifically Lot 9DU, sees a high amount of occupancy density and adding a new lot in the Dundas area, or more street parking meters are options to be considered
- Lot 82 on Victoria Ave N, close to Hamilton General, sees a high amount of density and opportunity to be explored
- We also recommend that given that users park in these lots anyway even if there is crime activity nearby that security measures should be taken to protect Hamilton residents such as surveillance.

References

1. Competition Dataset provided by the City of Hamilton: https://datacompetition.mohawkcollege.ca/2022-data-links/
2. Location of Municipal Car Parks in Hamilton. Obtained from: https://open.hamilton.ca/datasets/d56d996d4725499da2a5555aa5e5b651_5/about
3. Data requested from the Hamilton Police Department website: https://hamiltonpolice.on.ca/how-to/file-freedom-information-request
4. Census Data for Hamilton, obtained from: https://www.hamilton.ca/moving-hamilton/community-**profile/census-data-hamilton**