Parking Meter Price Optimization: How Big Data Can Increase City Revenue

Jeffrey D. Buck, PMP, MBA

***Abstract*—This article explores how big data technologies can optimize City parking meter systems using San Diego as a case study. The article begins by reviewing the history of parking meters, the current state of the market, and recent trends in dynamic pricing strategies. The article will then explore the current state of big data analytics, current approaches, and details around the Hadoop big data eco-system. Parking meter data from San Diego was used to analyze trends and insights into parking meter locations, configurations, transactional volumes, and time studies to determine if big data applications could have a meaningful impact. A detailed analysis of this data is included. The article will close out by reviewing the key observations from the data and presents recommended next steps in utilizing these insights to commercialize big data in optimizing parking meter pricing.**

***Index Terms*—Parking meters, dynamic pricing strategy, surge pricing, San Diego, Hadoop, big data, Hive, ParkDetroit, Uber, NoSQL.**

I. Introduction

The North American parking management market is projected to grow from $4,189.9 million in 2013 to $7,430.1 million by 2018. This is a compounded annual growth rate of over 12%. This growth correlates with the expected continued growth in car sales and increased urban densities. This creates both a problem and an opportunity for city planners. Do nothing, and cities will suffer from continued traffic congestion. According to industry data, over 30% of traffic congestion is caused by drivers who are unable to find parking spaces. However, given the growth in smart parking system technologies and the proliferation of smart phones and the internet of things, city planners can capitalize on this problem by upgrading parking systems to help reduce congestion as well as increase revenue[1].

The iconic parking meter has been around since the mid-1930's. At the time, city infrastructures were designed to support horses and carriages. As cars became more prevalent by the average citizen, they became the preferred mode of travel to and from work every day. This caused an increased in traffic volume and coincidentally the problem of parking. Most drivers would park along the edges of the street in front of stores for the duration of the day which caused uproar from shop owners who did not have adequate parking for shoppers. Oklahoma City was the first city to formally address this issue by holding a local contest to design a device to solve the issue. The result was the 'Black Maria', a mechanical device that required a nickel per hour. The first meter was installed on the southeast corner of First Street and Robinson Avenue on July 16th, 1935 where they were spaced at 20 foot intervals[2].

Since the 1930's parking meter technologies have not progressed significantly. The original mechanical design of coin and timer are still seen throughout the US. Recently, however, cities such as Detroit have started to challenge the status quo of traditional parking meters. In 2015, Detroit launched a ParkDetroit campaign to modernize city parking. This campaign was spearheaded by a new app that users could download to phones which interfaced with new parking meters. The app was complete with a payment and alert system that enabled users to pay for parking from their smartphone and receive alerts when the meter is at risk of expiring[3].

Recent developments in pricing strategies have seen wide spread acceptance of dynamic pricing models. Dynamic pricing models adjust the price of a product or service based on the relationship of supply and demand. If demand exceeds supply prices increase and vice versa if supply exceed demand. Uber is the most popular deployment of this type of strategy. Uber utilizes a real time dynamic algorithm to determine the ratio of supply to demand in zones. When demand exceeds supply, the algorithm will adjust pricing (P_B) structures to 'surge' pricing (S_M) where the base price will be multiplied by a multiplier variable (M_O) to determine the new price.

$$P_B \times M_o = S_M$$

In most cities these multipliers are regulated by city governments to ensure that prices do not get out of control. This is known as the greatest surge multiplier[4]. Parking meters could be an ideal model for utilizing a dynamic pricing strategy.

The integration of parking meters with smartphones provide an opportunity for further advancement by utilizing big data technologies to optimize the parking experience for both drivers

[1] Kevin Anchi Author, "Parking Management Market Worth $7.43 Billion 2018," in Academia.edu 2016.
[2] Amanda Erickson Author, "A Brief History of the Parking Meter," Popular Mechanic Apr 3, 2012.
[3] Author Unknown, "City of Detroit Launches New ParkDetroit Mobile App", City of Detroit Press Release July 31, 2015.
[4] M. Keith Chen, Michael Sheldon Author, "Dynamic Pricing in a Labor Market: Surge Pricing and Flexible Work on the Uber Platform". December 11, 2015.

and city planners. By integrating big data technologies such as Hadoop and a dynamic pricing strategy, drivers can be given automated recommendations of parking opportunities, city planners can better parking supply and demand analysis, and city financial planners can deploy a dynamic payment system to adjust parking fees relative to demand. This article will start with an overview of big data and how it is different than traditional technologies. The article will then give an overview of the different big data technologies, ecosystem, and advantages of using the big data ecosystem. Finally, the article will do an in-depth analysis of the San Diego parking meter system to identify potential opportunities for cities to incorporate big data techniques to increase revenue from parking meters.

II. BIG DATA – OVERVIEW

Big Data can be described "as volumes of data available in varying degrees of complexity, generated at different velocities and varying degrees of ambiguity, that cannot be processed using traditional technologies, processing methods, algorithms, or any commercial off-the-shelf solutions". This can include data generated from websites, social media, sensor networks, airplane engines, consumer data driven, financial data, etc[5].

A. Characteristics The characteristics of big data can be summarized by the three V's, volume, velocity, and variety. In addition, ambiguity, viscosity, and varality are often used to describe the benefits of big data ecosystems. Data volume can be defined as the amount or quantity of data that is continuously being generated. For example, a text post on Facebook is a few kilobytes, a video post is a few megabytes; the log of individuals that viewed your Facebook post or the clickstream history resulting from your post could be in gigabytes. Data volume is generally measured in the number of bytes. Traditional technologies are capable of affordably storing giga and terabytes of data due to the underlying data storage approach[6]. Big data technologies are capable of storing petabytes of data or more at extremely low costs.

Data velocity can be defined as the speed at which the data is received. For example, Amazon receives millions of clicks a second. Data velocity is generally measured by an action such as clicks, measurements, posts, etc. over a period of time. Traditional approaches have processed this data in batches. A specific amount of data would be processed over a given time horizon, transformed, stored in a data warehouse, and then made available for reporting and analysis. This approach was effective in an un-scalable infrastructure but caused the data to become stale and unresponsive. Valuable insights were only discovered after a sale was complete forcing organizations to be reactive instead of proactive[7]. Big data ecosystems enable a more scalable and real time processing at any data velocity. This enables insights to be available almost immediately.

Data variety can be defined as the format or type of data that is received. For example, a CRM system that is needed to intake Tweets, Facebook messages, voicemails, and text messages. This data may arrive in at the warehouse in structure, unstructured, and semi-structured formats with little to no meta-data. In a traditional ecosystem, the data would have to do through multiple transformations to apply the meta-data or overlay a schema to enable downstream data-mining and analysis. Big data ecosystems do not require complex transformations or even schemas. The data can be immediately stored and either analyzed as is or have a schema applied later in the process which enables more rapid ingestion and analysis.

B. Opportunities. The full potential of big data has yet to be realized. Every industry is currently exploring ways to employ big data ecosystems. Below are some use case examples that depict the opportunities for big data.

A wind energy company explored using big data platforms to collect weather data at a more granular level to help drive wind turbine placement decisions. Traditional technologies only afforded the energy company the ability to process data within a 17 x 17 mile area, across only a couple dozen parameters, and within batches of information every 24 hours. By transitioning to a big data platform, the energy company was able to capture weather data within a 32 x 32 foot area, across 178 different parameters, every 3 hours. This enabled the engineers to process 18-24 petabytes of data a day and optimizes the placement of wind turbines[8].

A government security manager required a surveillance system that could detect, classify, locate, and track potential threats to a perimeter and border area by analyzing streaming surveillance data. Traditional data processing platforms were not capable to supporting the velocity and variety of data nor were they capable of processing them in a timeframe that could afford the security manager the ability to respond to potential threats, significantly reducing the probability of their effectiveness. A big data platform was implemented and successfully integrated over 1,000 devices and enabled the simultaneous monitoring and analysis of the information. In addition, the analysis response time was reduced from hours to fractions of a second enabling security potential to respond to threats real time. The deployed solution was capable of processing over 40 terabytes of data a day[9].

The opportunities for big data are endless. The above examples only illustrate a few potential use cases and the scale at which big data can be effective. Two and a half exabytes of data are created every day and that number doubles every month[10]. Ninety percent of the world's currently stored data has been created in the last two years[11]. The need and opportunities for big data are only going to grow further. Retail, churches, sports, politics, healthcare, academia, banking, and

[5] Krish Krishnan, "Introduction to Big Data," in Data Warehousing in the Age of Big Data, 1st ed. Massachusetts.
[6] Krish Krishnan, "Introduction to Big Data," in Data Warehousing in the Age of Big Data pg 19, 1st ed. Massachusetts.
[7] Krish Krishnan, "Introduction to Big Data," in Data Warehousing in the Age of Big Data pg 20-22, 1st ed. Massachusetts.
[8] Krish Krishnan, "Introduction to Big Data," in Data Warehousing in the Age of Big Data pg 102-103, 1st ed. Massachusetts.
[9] Krish Krishnan, "Introduction to Big Data," in Data Warehousing in the Age of Big Data pg 105-108, 1st ed. Massachusetts.
[10] Website: http://hbr.org/2012/10/big-data-the-management-revolution
[11] Website: http://www.01.ibm.com/software/data/big data

science are some of the top industries predicted to be the largest growth areas for big data[12].

III. BIG DATA – ECOSYSTEM

Big data encompasses numerous ecosystems, technologies, and applications. It is important to understand these definitions and variations before embarking on an in-depth analysis. The following section will review the conceptual components of big data and some examples of currently available ecosystems. Finally we will review the specific distributions that are available for the Hadoop ecosystem.

Components. Big data can be split into four primary components; data sources and collection, data storage, data transformation and performing analytics, and data visualization. Data sources and collection serve as the inputs to the big data ecosystem. These are the tools that will be used to interact with source data bases such as legacy SQL databases, system logs, emails, social media posts, sensor data, call logs, etc. and land the data into the eco-system. Data storage and computing infrastructure is made up of the technologies that will store, distribute, manage, and replicate the landed data and make it available to the other components. Data transformation and analytics consist of the family of technologies that enable a user to transform structured, unstructured, and semi-structured data in a consistent manner and prepare the data for analysis. This component provides the querying and machine learning algorithms to analyze the data. Data visualization and delivery serves as the output technologies to the ecosystem. . This component will also provide visualization tools to graph outputs.

IV. BIG DATA - TECHNOLOGIES

Since the early 80's, it was recognized that the constraints inherent in a SQL RDBMS would eventually hinder an architectures capabilities in the future. A series of working sessions were organized to review alternate designs that cope with the growing need to manage exponential growths in data. Over the course of two decades various approaches taken with two dominate approaches emerged. Big data has two primary ecosystems, NoSQL and Hadoop, and a third smaller emerging ecosystem known as textual ETL processing. These approaches purposely deviate from traditional data management approaches.

A. NoSQL. NoSQL has become the industry term to describe a family of non-relational databases built for specific workloads. These have broadly been categorized into four broad categories. Table 1, lists the four categories and a brief description of each.

Category	**Description**
Key-value Pairs	Utilize a hash table where there is a unique key and point to a particular item of data creating a key-value pair.
Column Family Stores	An extension of the key-value architecture with columns and column families, the overall goal was to process distributed data over a pool of infrastructure.
Document Databases	Data is stored as a document and is represented in JSON or XML formats.
Graph Databases	Based on the graph theory, this class of database supports the scalability across a cluster of machines.

Table 1: NoSQL Categories[13]

B. Hadoop. Hadoop has emerged as the dominate big data platform at the time of writing this article. Hadoop ecosystem is built on top of what is known as the foundation core components; Hadoop distributed file system, key value store (HBase), job scheduling/execution system (MapReduce), serialization (Avro), and coordination (Zookeeper). The ecosystem was built as an open-source system and designed to operate as a scalable, distributed file system that operates on commodity hardware[14]. Hadoop will be the focus of the remaining sections of this article.

C. Textual ETL Processing. NoSQL and Hadoop ecosystems consist of highly technical technologies that require, in most cases, developers to implement, query, and analyze data. Business users have asked for a more user friendly way of interrogating unstructured data without the need for technical developers. Textual ETL processing was introduced to meet this use case. Textual ETL processing enables a user to apply various algorithms and modeling techniques to unstructured data in a user friendly interface. The outputs of the analysis are exported into a RDMS for further analysis. The components of this ecosystem consist of a textual ETL rules engine, user interface, taxonomies, and an output database[15].

D. Hadoop Distributions. Hadoop continues to be an open source big data platform provided through Apache however, commercial organizations have adopted the base platform and applied additional tools and references. These various incarnations of the Hadoop platform are known as distributions. The leading distribution offerings are provided by organizations such as Cloudera, HortonWorks, MapR, Intel, and PivotalHD, and Cloudera at the time of writing this article were considered to be the leading providers of Hadoop distributions due to their ability to track changes in the base Hadoop offering and their additional install and administration tools that come in both open source and premium cost models.

V. HADOOP TECHNOLOGIES

A. Foundation Core Technologies. Hadoop it built on top of what is known as foundational core technologies that provide the core functionality that enables Hadoop to provide a scalable,

[12] Website: https://datafloq.com/read/8-industries-benefit-big-data-infographic/402

[13] Krish Krishnan, "Introduction to Big Data," in Data Warehousing in the Age of Big Data pg 87, 1st ed. Massachusetts

[14] Krish Krishnan, "Introduction to Big Data," in Data Warehousing in the Age of Big Data pg 54, 1st ed. Massachusetts

[15] Krish Krishnan, "Introduction to Big Data," in Data Warehousing in the Age of Big Data pg 97-98, 1st ed. Massachusetts.

distributed data management platform. These core foundational technologies are listed in Table 2:

Component	Technology
Coordination	Zookeeper
Distributed Processing	MapReduce
NoSQL Database	Hbase
Distributed Storage	Hadoop Distributed File System (HDFS)

Table 2: Hadoop Foundational Core Technologies

B. Supporting Technologies. In addition to the foundational core technologies, Hadoop ecosystem provides additional capabilities to support various components of the ecosystem. Table 3 below provides a summary of those components and technologies.

Component	Technology
Query	Hive
Scripting	PIG
Machine Learning	Mahout
Workflow & Scheduling	Oozie
Management & Monitoring	Ambari
Data Integration	Sqoop REST ODBC

Table 3: Supporting Hadoop Technologies

VI. ADVANTAGES OF HADOOP

The Hadoop platform offers significant advantages over traditional relational databases. These advantages include scalability, cost effectiveness, fault tolerance, speed, and flexibility. These are reviewed in detail below.

Scalable. The scalability of the Hadoop platform is one of if not the biggest advantage. Hadoop takes advantage of the growth in hardware as a commodity in recent years. Traditionally an organization purchased additional racks and hardware when in need of additional storage. This is known as vertical expansion, Hadoop enables organizations to leverage hardware/infrastructure as a service and purchase additional nodes and rack space. This is known as horizontal expansion. This enables an organization to store terabytes and petabytes of data across thousands of nodes with unlimited scalability.

A. Cost Effective. A byproduct of Hadoop's scalability is the cost effectiveness of this approach. Traditional relational databases require a non-linear growth in storage capacity. Meaning, as an organization creates more bytes of data it requires a disproportional growth in the amount of storage capacity required to process it. Based on traditional infrastructure costs, the cost to store 300 TB on an typical RDBMS platform would be approximately $9m as compared to only $300k on a Hadoop platform. .

B. Fault tolerance. Hadoop is able to achieve a high level of fault tolerance through its replication, heartbeat, and block reporting features. Fault tolerance in Hadoop architecture is critical because given its distributed nature it is highly probably that a machine will fail from time to time if not often. To address this risk, Hadoop architecture includes multiple safe guards. First, the block replication follows a strict placement policy to ensure that blocks are efficiently distributed across multiple and separate racks in case one fails. Secondly, the NameNode will periodically receive a 'Heartbeat' for each of the datanodes to ensure that all the machines are operable. Lastly, the NameNode will also receive a block report which lists the location of all the blocks on a given data node. If the NameNode does not receive a Heartbeat from a datanode, the block list is then used to locate the replicated blocks that need accessing[16].

C. Speed. Hadoop's unique distributed storage system enables the platform to maintain a mapping of all the data across the hundreds or thousands of 'clusters' in which it is stored. In addition, the tools for processing the data are often on the same machines that the data is stored which vastly simplifies the architecture. When processing large sets of unstructured data this enables a user to process terabytes of data in seconds and petabytes of data in hours[17]

D. Flexibility. Hadoop's capability to process structured, unstructured, and semi-structured data separates itself from traditional architectures and makes it uniquely flexible. The ability to combine these various types of data into a single architecture with the tools to query, analyze, visualize, model, and apply machine learning algorithms gives organizations and paralleled ability to quickly identify insights that would otherwise go unseen. This positions Hadoop as an extremely flexible platform to conduct robust business analytics.

VII. DATA

Two sets of data were used for this analysis:

- San Diego Parking Meter Locations[18]
- San Diego Parking Meter Transactions[19]
- San Diego Parking Meter Transactions (Days)

A. San Diego Parking Meter Location. This data set lists all of the parking meters in San Diego in 2015. The data for this report was downloaded on November 23, 2016. It was downloaded as a .csv file and imported as a Metastore table into a Cloudera node. The data contains 4,653 parking meter locations. Table 4 contains the data dictionary for the dataset.

Field	Description
zone	The parking districtÂ to which theÂ meter belongs
area	The neighborhood the meter resides in
sub_area	The block number and street of the meter
pole	Unique pole id
config_code	Configuration code
config_name	Configuration description

[16] Krish Krishnan, "Introduction to Big Data," in Data Warehousing in the Age of Big Data pg 50-56, 1st ed. Massachusetts

[17] Michele Nemschoff Author. "Big Data: 5 Advantages of Hadoop". IT Proportal.com. December 20th, 2013.

[18] City of San Diego Datasets: http://data.sandiego.gov/dataset/parking-meter-locations

[19] City of San Diego Datasets: http://data.sandiego.gov/dataset/parking-meter-transactions

longitude	meter location longitude (may not be accurate)
latitude	meter location latitude (may not be accurate)

Table 4: San Diego Parking Meter Locations Data Dictionary

Appendix 1 depicts screenshots of this data and examples.

B. San Diego Parking Meter Transactions. This dataset lists all parking meter transactions in San Diego in 2015. The data for this report was downloaded on November 23, 2015. It was downloaded as a .csv file and imported as a Metastore table in Cloudera node. The data contains 9,068,312 parking meter transactions. Table 5 depicts the data dictionary for the dataset.

Field	Description
meter_type	Type of meter
pole_id	Unique Meter ID
trans_amt	Transaction amount in cents
pay_method	Method of payment
trans_start	Transaction start date and time
meter_expire	Meter expiration date and time
smartcard_id	Smart card serial number

Table 5: San Diego Parking Meter Transactions Data Dictionary

VIII. DATA PREPERATION

To properly prepare the data three major steps were conducted:

- Load Data
- Join Tables
- Time Data Transformations

A. Load Data. The source data was downloaded from the City of San Diego website as a .csv. All tables were uploaded into Hadoop cluster via the Hue file browser. The files were then loaded as tables utilizing Metastore and saved in a new database. Appendix 1 depicts screenshots of the tables created.

B. Join Tables. The location of the parking meters and the transaction history of each parking meter were split between two tables. A transformation was required to join both tables. Appendix 3 contains a screen shot of this transformation in Hive.

C. Time Data Transformations. The time data provided in the tables was in the following format:

DD-MM-YYYY HR:MIN:SEC

In order to conduct the time data analysis, a new column was added and the time data parsed.

IX. DATA ANALYSIS & OBSERVATIONS

The following tools were used to conduct this analysis:

- Hive
- Microsoft Excel
- ThinkCell

Hive was used to conduct all of the queries and some of the visualizations in the analysis. Microsoft Excel and ThinkCell were utilized to assist in the visualization of the data.

An initial exploratory analysis of the data was conducted to answer four primary questions:

- Meter Configuration: Does San Diego already use a dynamic pricing strategy on its parking meters?
- Meter Distribution: Which areas of the city have the highest number of parking meters?
- Meter Transactions: Which areas of the city have the highest transactional volume on the parking meters?
- Meter Time Study: Are their peak usage times for the parking meters in the high volume areas?

A. Meter Configuration. San Diego parking meters are capable of being configured by:

- Maximum time limit
- Hourly Rate
- Hours that Meter Applies
- Days of the Week that Meter Applies

In 2015, San Diego had 53 potential configurations for a parking meter. 77% of these meters are the 6569 configuration which is set to a parking limit of 2 hours at $1.25 an hour, Monday through Saturday. Table *6 lists* the top five configurations in the dataset. Appendix 4 depicts the queries and results.

Description	Count
2 Hr Max - $1.25	3,518
9 Hr Max - $.50	233
4 hr Max - $1.00	121
15 Min Max - $1.25	63
1 Hr Max - $1.25	19

Table 6: Top Five Meter Configurations

This confirms that San Diego parking meters do not currently follow a dynamic pricing strategy and the dominate static pricing structure utilized is a 2 hour max, $1.25 an hour configuration.

B. Meter Distribution. Parking meters are distributed across San Diego by Zone, Area, and Sub-Area. San Diego has four zones; Downtown, Uptown, Mid-City, and City. The majority of the marking meters are located within the Downtown and Uptown zones. Diagram 1 depicts this zone distribution. Appendix X depicts Hive queries and additional results.

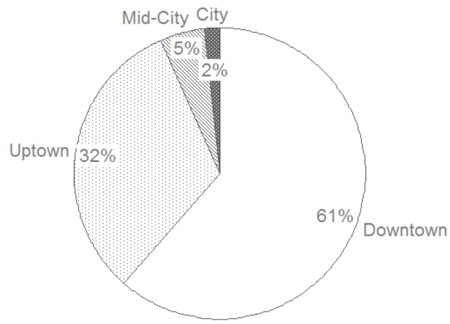


Figure 1

The Zones are further broken down into 23 areas. Eighty percent of the city's parking meters reside in the top six areas with East Village being the largest area with 1,095 parking meters. Table 7 provides a list of the top 10 areas.

Area	Count
East Village	1,096
Bankers Hill	650
Hillcrest	637
Core-Columbia	414
Cortez Hill	391
Gaslamp	331
Little Italy	317
Marina	292
North Park	116
Mission Hills	116

Table 7: Top 10 Parking Meter Areas

Appendix 7 depicts the Hive queries and additional detail.

The zones and areas are even further divided down into 540 sub-areas. The majority of the sub-areas are broken down into the specific streets with approximately 10-35 meters per street. However, over 400 parking meters are in what is known as a 'Hospitality Zone' sub-area. These are locations that the city has identified as either key local or tourist locations in the city that receive a large volume of traffic. Unlike the other sub-areas, the Hospitality Zone is located in multiple areas and sub-areas. Table 8 depicts the top five meters by sub-area. Appendix 5 depicts Hive queries and additional results.

Sub-Area	Count
Hospitality Zone	417
3600 Fifth Ave	35
1400 E. Street	33
3900 Fifth Ave	29
3900 Fourth Ave	26

Table 8: Top 5 Meter Count by Sub-Area

This confirms that the density of parking meters and transactional volume is focused in the 'Downtown' zone and the 'Hospitality Zones' sub-areas.

C. Meter Transactions. Whenever a driver pays for time on a parking meter a transaction is logged. Utilizing this data joined with the meter locations, an analysis of the transaction densities can be conducted. An initial analysis was conducted to understand the locations of the most transactions by area and sub-area. East village had the highest level of transactions followed closely by Gaslamp, Core-Columbia, Little Italy, and Marina. Table 9 depicts the top five areas by transaction volume:

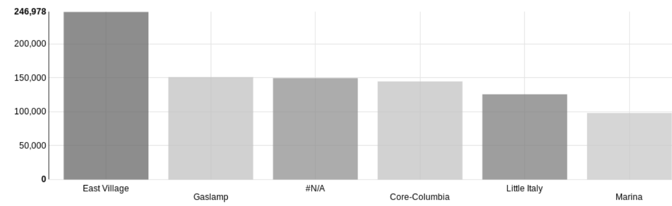


Figure 2: Top 5 Areas by Transaction Volume

Utilizing these results, additional analysis was conducted on the top transactional volume areas to determine which sub-areas had the highest transactions by sub-area. Not surprising, the highest transactional volume came from the hospitality zones. Table 9 provides the list of the top four sub-areas by transactional volume.

Area	Sub-Area	Count
Gaslamp	Hospitality Zone	68,652
Marina	Hospitality Zone	42,076
East Village	Hospitality Zone	34,728
Core-Columbia	Hospitality Zone	29,594

Table 9: Top 4 Sub-areas by Transactions

Appendix 8 and 9 depict the Hive queries and additional detail.

D. Meter Utilization over Time. Utilizing the transaction time and count data in conjunction with the meter count data, additional analysis was conducted to understand the utilization of the parking meters over the course of the day. Multiple queries were written to count the number of transactions by the time of the day across the top four sub-areas by transactional volume to determine peak utilization times and the variability in demand. Core-Columbia's Hospitality Zone utilization spans from 6:00 am to 6:00 pm with peak utilization around 7:30 am, 10:00 am, and 5:30 pm. East Village's, Gaslamp's, and Marina's Hospitality Zones all have utilization spans generally from 7:00 am to 8:00 pm with peak utilization at 10:00 am. In addition, each Hospitality zone predominantly use the same meter configuration, 6569 - 2 hour max at $1.25 per hour. Figure 3 depicts a depiction of this utilization. Appendix 10 depicts Hive queries and additional details.

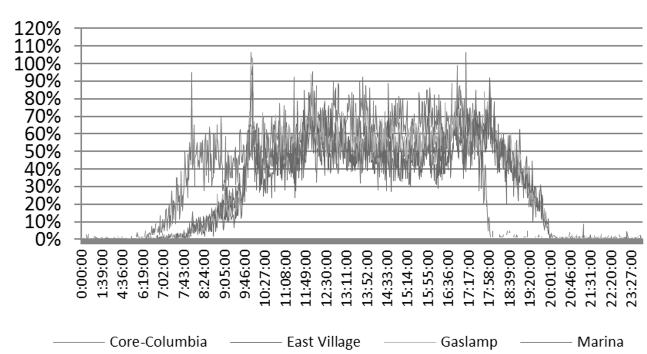


Figure 3: Top Four Utilization Sub-Areas

This confirms that San Diego parking meters do follow similar utilization patterns with periods of high demand (i.e. 100% utilization).

X. RESULTS/OBSERVATIONS

Based on the analysis the following conclusions were drawn:

- The Downtown, Hospitality Zones are high volume areas that would be ideal for a dynamic pricing strategy prototype.

San Diego does not currently pursue a dynamic pricing strategy on its parking meters.

XI. CONCLUSION/NEXT STEPS

Major cities could positively impact its parking meter revenue by incorporating big data. Utilizing similar technologies employed by the City of Detroit that integrates parking meters with smartphones and utilizing a dynamic pricing strategy similar to that utilized by Uber could enable cities to optimize the parking experience while simultaneously increasing revenue. As seen in the analysis of the San Diego parking meter data, high volume parking areas follow utilization patterns throughout the course of a day. Big data in conjunction with a pricing algorithm that monitors this utilization could identify sub-areas that are experiencing high levels of parking utilization and switch the parking meter configurations to a surge pricing model. For example, when a sub-area has a parking meter utilization that exceeds 90%, the meters could automatically shift to a surge pricing of 1.5 times the base price. When utilization approaches 95%, surge pricing could shift to 2.0 times the base price.

Before deploying such a strategy additional work should be conducted. Additional analysis should be conducted to further understand he seasonality of parking meter utilization. This article only analyzed the utilization based on time. Seasonal, daily, monthly, and major event days should be analyzed to identify additional insights.

Select cities should be identified to conduct public opinion surveys to determine if the shift in pricing strategy could cause public outrage. It would be critical to understand the public's perception of this new approach and understand if certain demographics, city zoning, use cases, etc. are more accepting of this approach.

Additional analysis of general parking pricing adjacent to the Hospitality Zones should be conducted to understand the generally accepted market rate for parking in the area. This will help to further understand what pricing would be accepted by drivers in areas and model surge pricing scenarios.

Finally, a business case should be developed to understand the total cost of ownership for the creation and deployment of a new parking meter application and infrastructure to support this model. In addition, a value model should be constructed to determine how much revenue could be increased. The combination of these two components would be critical to understanding if this was a positive investment.

APPENDIX

APPENDIX 1: Table Schemas

Tables

Databases > parking_meters

STATS

No comment.

TABLES

Search for a table... View

	Table Name
	meter_combined
	meter_locations
	meter_time
	meter_transactions2

Parking Meter Transaction Data

Databases > parking_meters > meter_transacti

Add a description...

Overview **Columns (7)** Sample Details

	Name	Type
1	meter_type	string
2	pole_id	string
3	trans_amt	smallint
4	pay_method	string
5	trans_start	string
6	meter_expire	string
7	smartcard_id	string

Parking Meter Location Data

Databases > parking_meters > meter_locations

Add a description...

Overview **Columns (10)** Sample Details

	Name	Type
1	p	string
2	area	string
3	sub_area	string
4	pole	string
5	config_code	smallint
6	config_name	string
7	longitude	float
8	latitude	float
9	pole_concat	string
10	pole_id_conc	string

Parking Meter Combined Data w/ Concatenation

Databases > parking_meters > meter_time

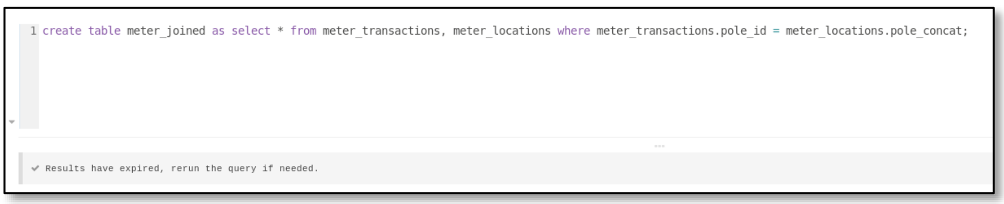
Add a description...

Overview **Columns (16)** Sample Details

	Name	Type
1	meter_type	string
2	pole_id	string
3	trans_amt	smallint
4	pay_method	string
5	trans_start	string
6	start_date	string
7	start_time	string
8	meter_expire	string
9	end_date	string
10	end_time	string
11	smartcard_id	string
12	p	string
13	area	string
14	sub_area	string
15	config_code	smallint
16	config_name	string

APPENDIX 2: SAN DIEGO PARKING METER TRANSACTIONS QUERY

APPENDIX 3: HIVE JOIN CODE



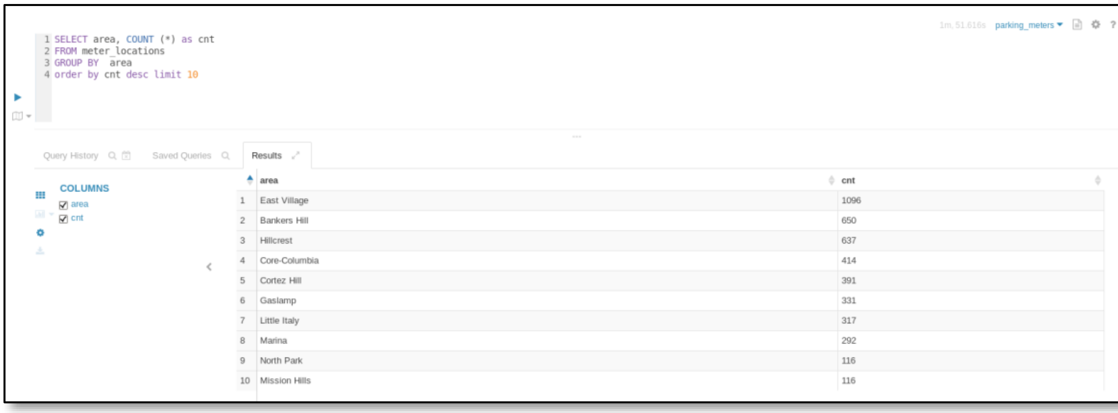
APPENDIX 4: PARKING METER CONFIGURATION QUERY

APPENDIX 5: Parking Meter Zone Analysis

```
SELECT p, COUNT (*) as cnt
FROM meter_locations
GROUP BY  p
order by cnt desc limit 20
```

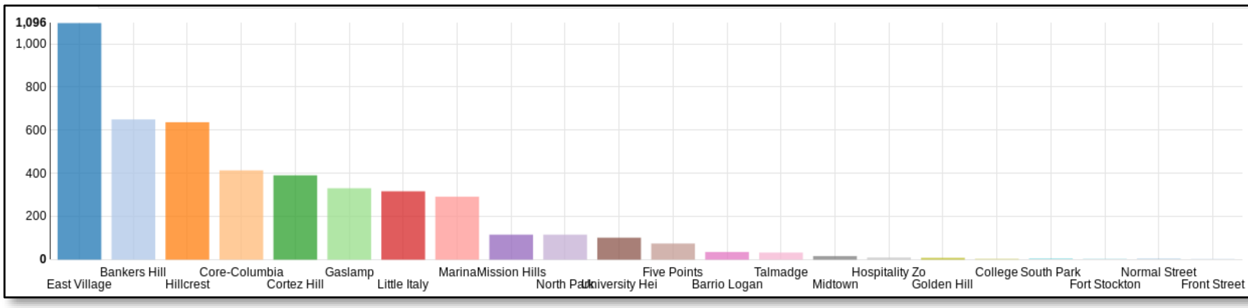
	p	cnt
1	Downtown	2851
2	Uptown	1489
3	Mid-City	240
4	City	73

APPENDIX 6: Parking Meter Area Analysis

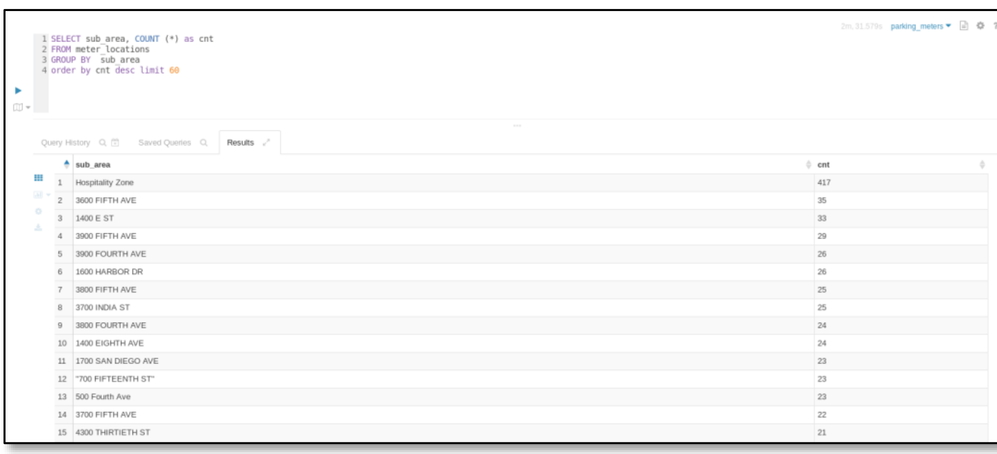


```
SELECT area, COUNT (*) as cnt
FROM meter_locations
GROUP BY  area
order by cnt desc limit 10
```

	area	cnt
1	East Village	1096
2	Bankers Hill	650
3	Hillcrest	637
4	Core-Columbia	414
5	Cortez Hill	391
6	Gaslamp	331
7	Little Italy	317
8	Marina	292
9	North Park	116
10	Mission Hills	116



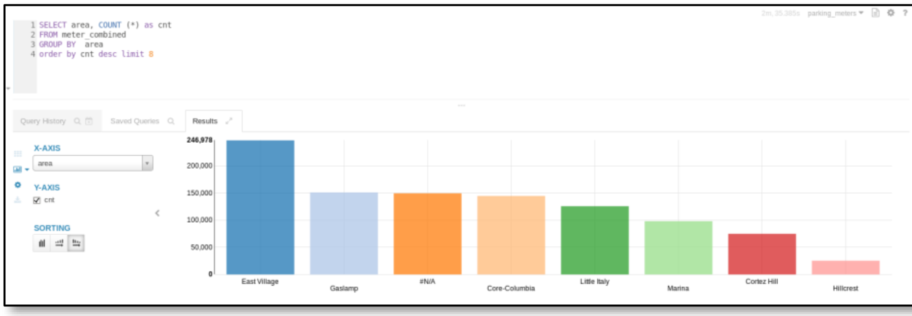
APPENDIX 7: Parking Meter Sub-Area Analysis



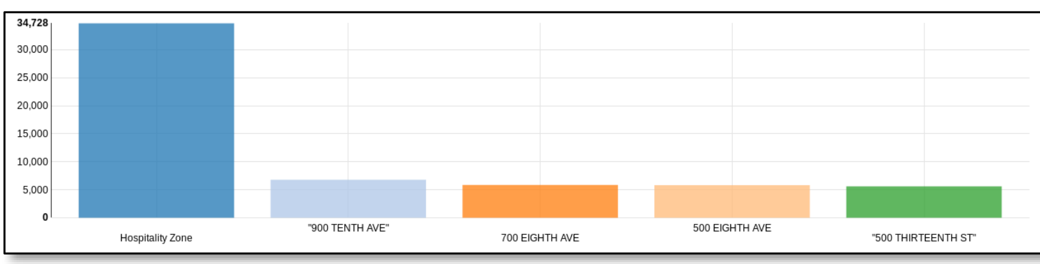
```
SELECT sub_area, COUNT (*) as cnt
FROM meter_locations
GROUP BY  sub_area
order by cnt desc limit 60
```

	sub_area	cnt
1	Hospitality Zone	417
2	3600 FIFTH AVE	35
3	1400 E ST	33
4	3900 FIFTH AVE	29
5	3900 FOURTH AVE	26
6	1600 HARBOR DR	26
7	3800 FIFTH AVE	25
8	3700 INDIA ST	25
9	3800 FOURTH AVE	24
10	1400 EIGHTH AVE	24
11	1700 SAN DIEGO AVE	23
12	"700 FIFTEENTH ST"	23
13	500 Fourth Ave	23
14	3700 FIFTH AVE	22
15	4300 THIRTIETH ST	21

APPENDIX 8: Transaction Volume by Area



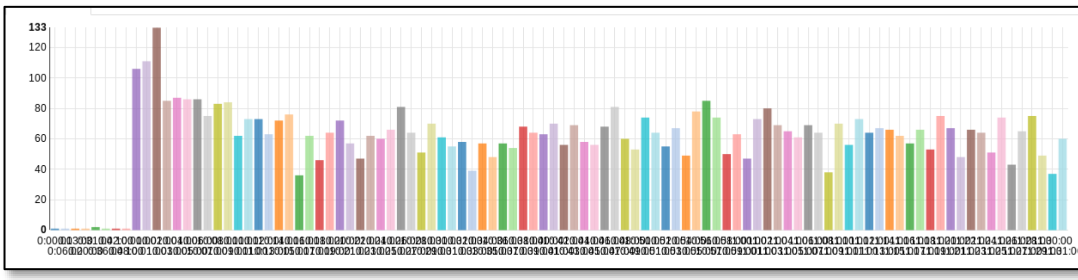
APPENDIX 9: Transaction Volume by Area



APPENDIX 10: Transaction Volume by Area

```
SELECT start_time, COUNT (*) as cnt
FROM meter_time
WHERE area = "Marina" AND sub_area = "Hospitality Zone"
GROUP BY start_time
```

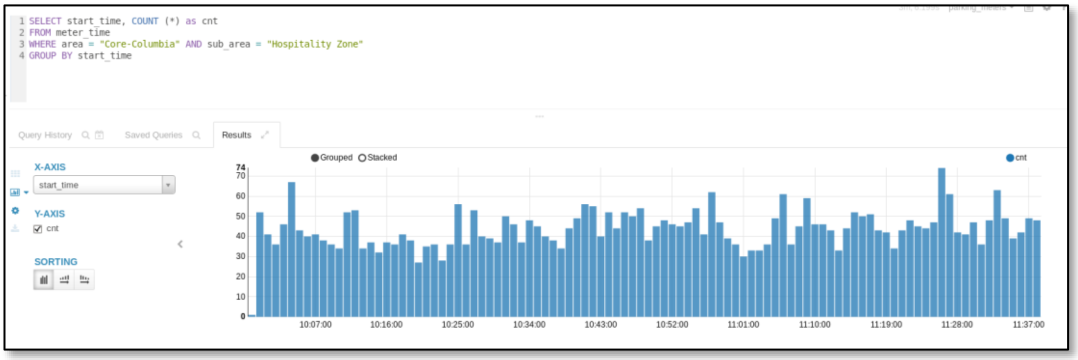
	start_time	cnt
1	0:00:00	1
2	0:06:00	1
3	0:13:00	1
4	0:20:00	1
5	0:31:00	2
6	0:36:00	1
7	0:42:00	1
8	0:48:00	1
9	10:00:00	106
10	10:01:00	111
11	10:02:00	133
12	10:03:00	85
13	10:04:00	87
14	10:05:00	86
15	10:06:00	86



```
SELECT start_time, COUNT (*) as cnt
FROM meter_time
WHERE area = "Marina" AND sub_area = "Hospitality Zone"
GROUP BY start_time
```

	start_time	cnt
1	0:00:00	1
2	0:06:00	1
3	0:13:00	1
4	0:20:00	1
5	0:31:00	2
6	0:36:00	1
7	0:42:00	1
8	0:48:00	1
9	10:00:00	106
10	10:01:00	111
11	10:02:00	133
12	10:03:00	85
13	10:04:00	87
14	10:05:00	86
15	10:06:00	86

```
SELECT start_time, COUNT (*) as cnt
FROM meter_time
WHERE area = "East Village" AND sub_area = "Hospitality Zone"
GROUP BY start_time
```


SELECT start_time, COUNT (*) as cnt
FROM meter_time
WHERE area = "Core-Columbia" AND sub_area = "Hospitality Zone"
GROUP BY start_time
Query History
Saved Queries
Results
X-AXIS
start_time
Y-AXIS
cnt
SORTING
Grouped
Stacked
cnt
74
70
60
50
40
30
20
10
0
10:07:00
10:16:00
10:25:00
10:34:00
10:43:00
10:52:00
11:01:00
11:10:00
11:19:00
11:28:00
11:37:00