

Big Data!



Majed Al-Ghandour, PhD, PE, CPM
Division of Planning and Programming
NCDOT
2016 NCAMPO
Conference- Greensboro, NC
May 12, 2016





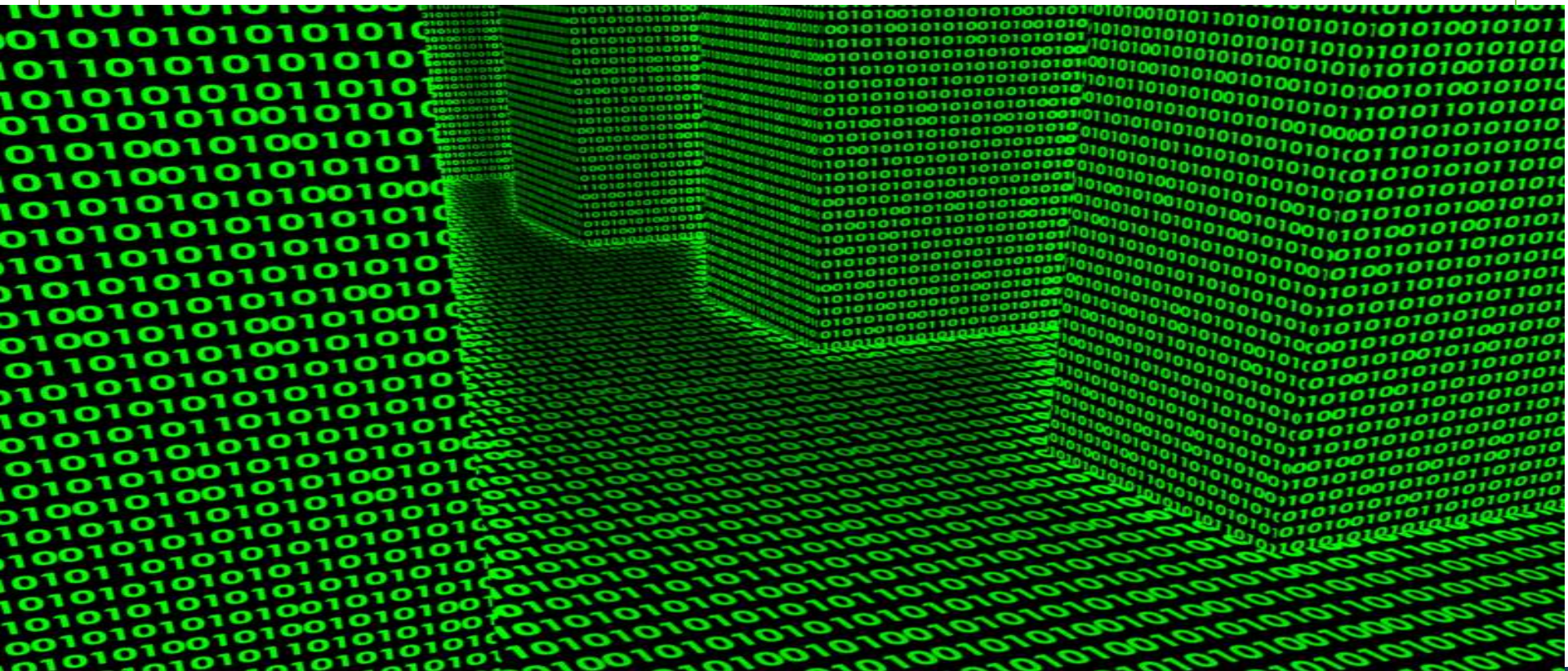
Outline

- **Introduce Big Data (BD)**
- **List the components of BD**
- **Visualization and Analytics Tools**
- **Demo**



Data

- Transportation Data increasingly drives the decision-making process in business, nonprofits, and the government.



What good is all of your data?

- ...**Data** as raw unorganized facts, is in and of itself worthless.

- **Information**... potential valuable pieces based on data.

	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
Total Lane Departure Crashes	61,239	57,520	56,243	56,435	54,999	56,475	55,498	50,022	51,853	53,615
Fatal Crashes	823	796	846	927	803	737	708	629	671	666
A Injury Crashes	1,744	1,663	1,554	1,386	1,237	1,126	981	1,011	983	893
B Injury Crashes	9,880	9,596	9,236	9,481	9,082	8,444	8,186	7,880	8,107	7,681
C Injury Crashes	13,831	12,787	12,741	12,807	12,285	12,598	12,204	11,464	11,922	11,787
Fatalities	917	871	939	1,040	871	805	766	693	718	737
A Injuries	2,304	2,161	2,025	1,780	1,556	1,426	1,264	1,259	1,230	1,137
B Injuries	13,145	12,444	12,003	12,252	11,651	10,950	10,439	10,091	10,294	9,706
C Injuries	20,640	19,034	18,677	18,893	18,031	18,446	17,968	16,990	17,673	17,189

What is Big Data (BD)?

Name

- Kilobyte (kb) 10^3
- Megabyte (MB) 10^6
- Gigabyte (GB) 10^9
- Terabyte (TB) 10^{12}
- Petabyte (PB) 10^{15}
- Exabyte (EB) 10^{18}
- Zettabyte (ZB) 10^{21}
- Youttabyte (YB) 10^{24}

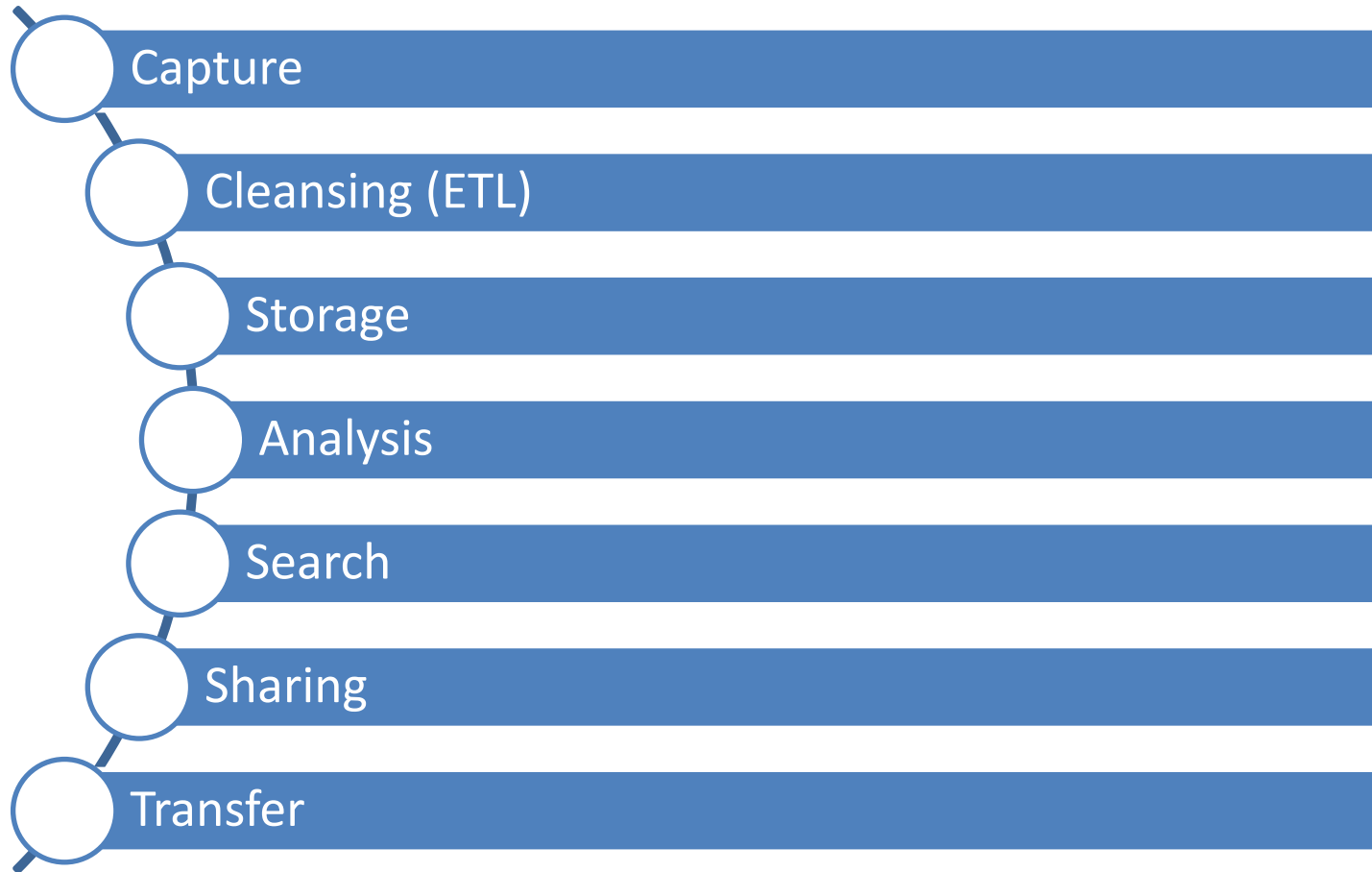
1 YB

= 1000^8 bytes = 10^{24} bytes = 1000000000000000000000000

bytes = 1000zettabytes = 1trillionterabytes

BD is so large data that it becomes difficult to process it using the traditional system.

Area of Challenges for DATA



Introduce Data: Making Sense of Data

- **Data is being generated at increasingly rapid rates in many disciplines such as:**
 - Internet of Things (IoT)
 - Healthcare
 - Banking and Financial Services
 - Insurance
 - Agriculture
 - Pharmaceutical
 - Manufacturing
 - Businesses
 - Logistics
 - Academic (Education)
 - Transportation
 - Planning
 - Assets Managements
 - Retail Sales (supply chain, consumer products, logistics..)
 - Inventory
 - Procurement(Electronic Commerce)
 - Order Management
 - Accounting
 - Customer Relationship Management
 - Human Resources Management
 - Telecommunications
 - GIS GeoSpatial
 - Social Media
 - Engineering, Energy, Aerospace



BIG DATA

By the Numbers

2.8 Zetabytes

The amount of information stored worldwide in 2012

140 Zetabytes

**The amount of information estimated to be stored
worldwide in 2020**

0.5%

Percentage of all the world's data currently analyzed

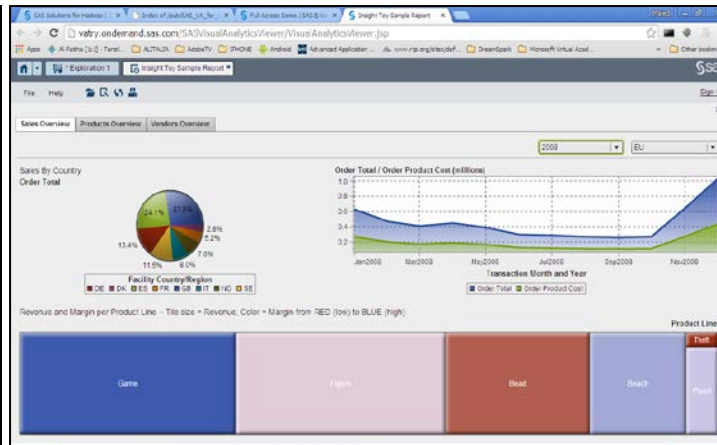
33%

**Percentage of existing data in the world that could be
useful if properly tagged and analyzed**

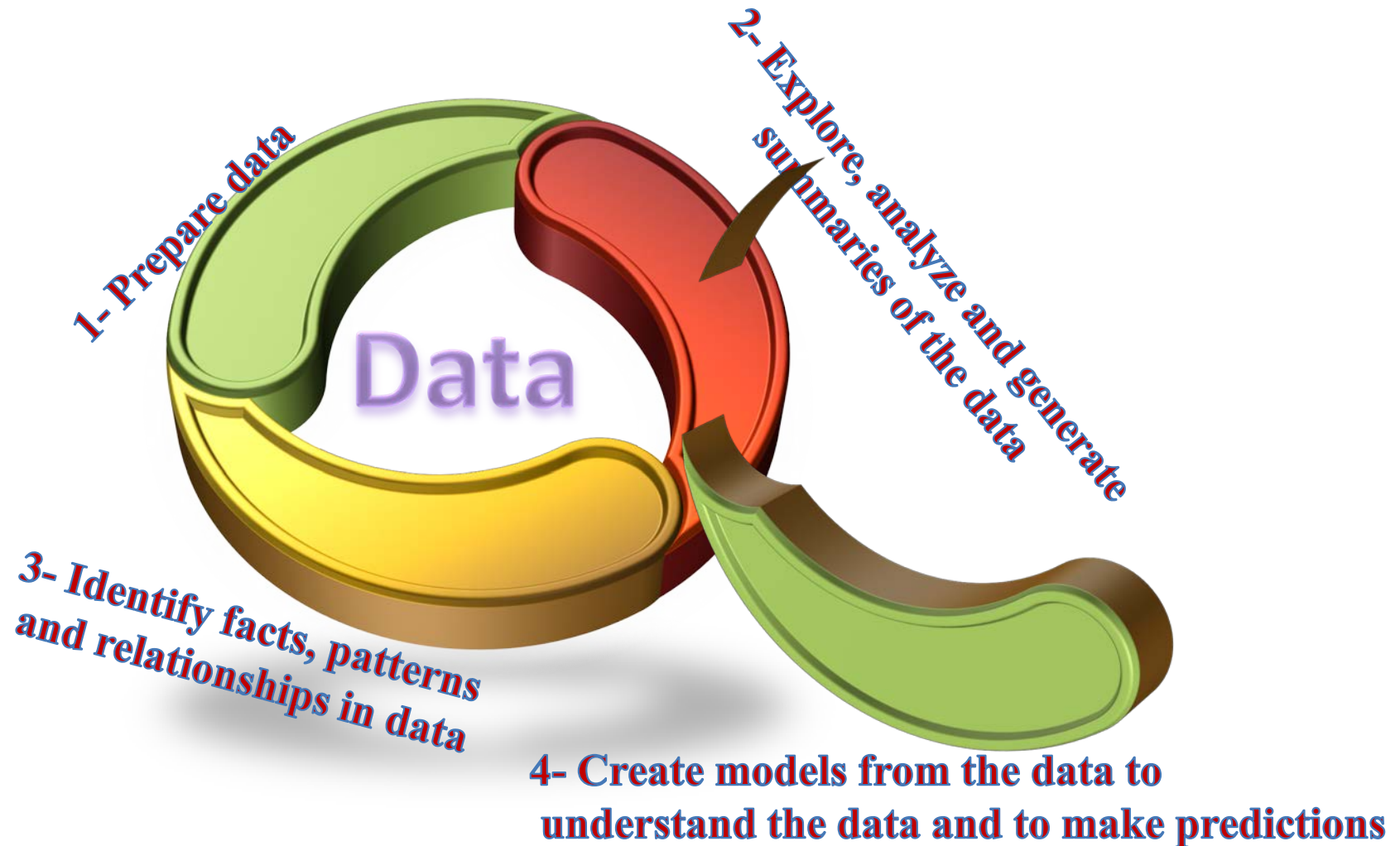
Over 5,000 miles

**The estimated cumulative size of all the world's data
centers in 2016**

It is all about making
business decisions based
on analysis of DATA.



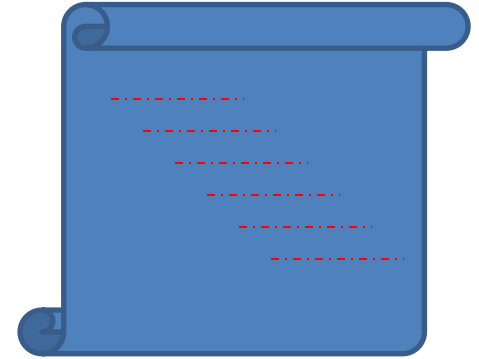
Making Sense of Data: from 3,000 ft.



Hot Topics in Business

- **Big Data Analysis**
- **Hadoop**
- **Business Intelligence**
- **Business Analytics**
- **Advanced Business Analytics**
- **Data Management**
- **Geospatial**
- **Dimensional Models and Data warehousing**
- **Predictive Analysis, Forecasting, and Time Series**
- **Mining**
- **Modeling**
- **Data Analytics**
- **Data quality, security, and privacy**

Sources of Data



- **Observations**
- **Surveys**
- **Polls**
- **Interviews**
- **Questionnaires**
- **Experiments (to understand cause-and-effect events by controlling some factors)**
- **Business transactions**
- **Planning, Designing, Operational**
- **Payrolls, Funding**
- **CRM (Customer Relationship Management databases)**
- **ERP (Enterprise Resource Planning databases)**
- **HRM (Human Resource Planning databases)**

Introduce Big Data (BD)

- **Big data** is a popular term used to describe the exponential growth and availability of data, both structured and unstructured.
- Measures by the Three **Vs** of big data:
volume, velocity and variety.
- Other Two Dimensions about Big Data:
Variability and Complexity.

Traditional and BD

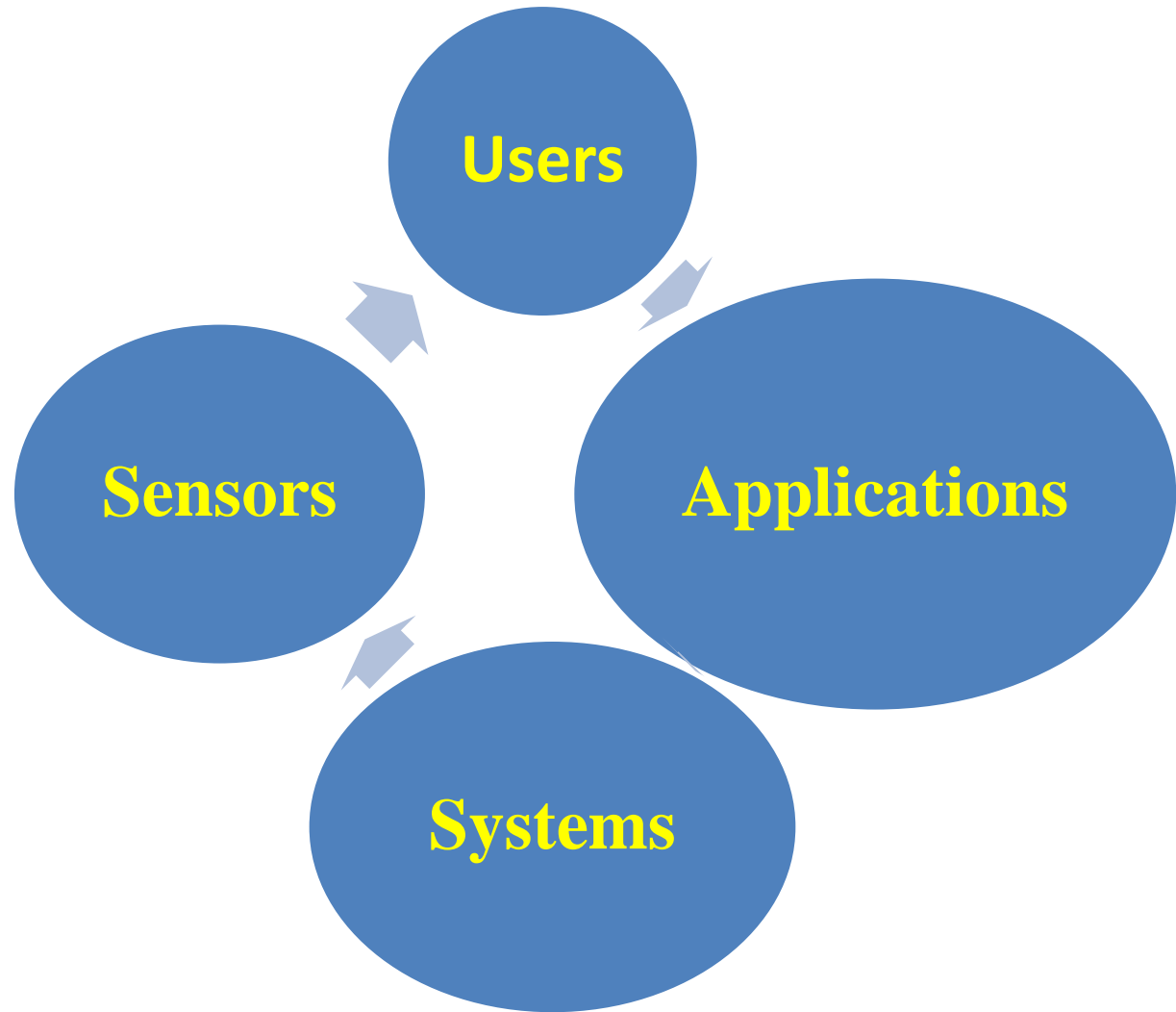
Traditional Data

- Documents
- Finances
- Personal File
- Payroll
-

Big Data

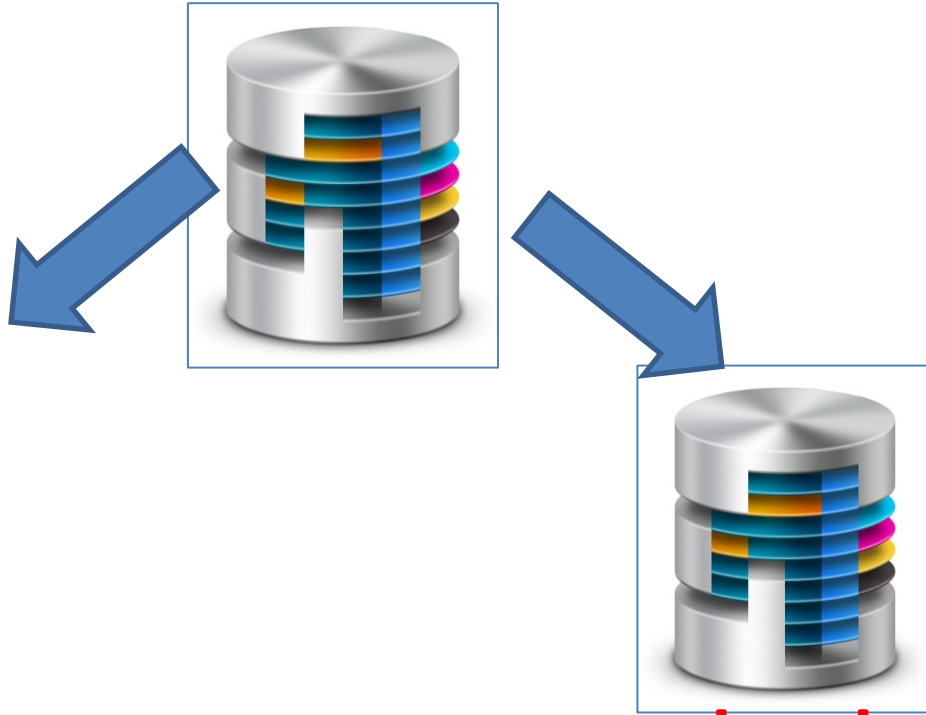
- 3D models
- Audio & Video
- Simulations
- Images
- Location Data (GeoSpatial)

Big Data Sources



Corporation Data

Corporate Data



structured

Transaction,
Payments,
Sales activity,
ATM,
Medical records

unstructured

Emails,
warranty claims,
Call center,
Market research

Analog processing,
Telephone call
records



repetitive



nonrepetitive

Corporation Data



unstructured



Analog processing,
Telephone call records

repetitive



Emails,
warranty claims,
Call center,
Market research

nonrepetitive



Corporation Data

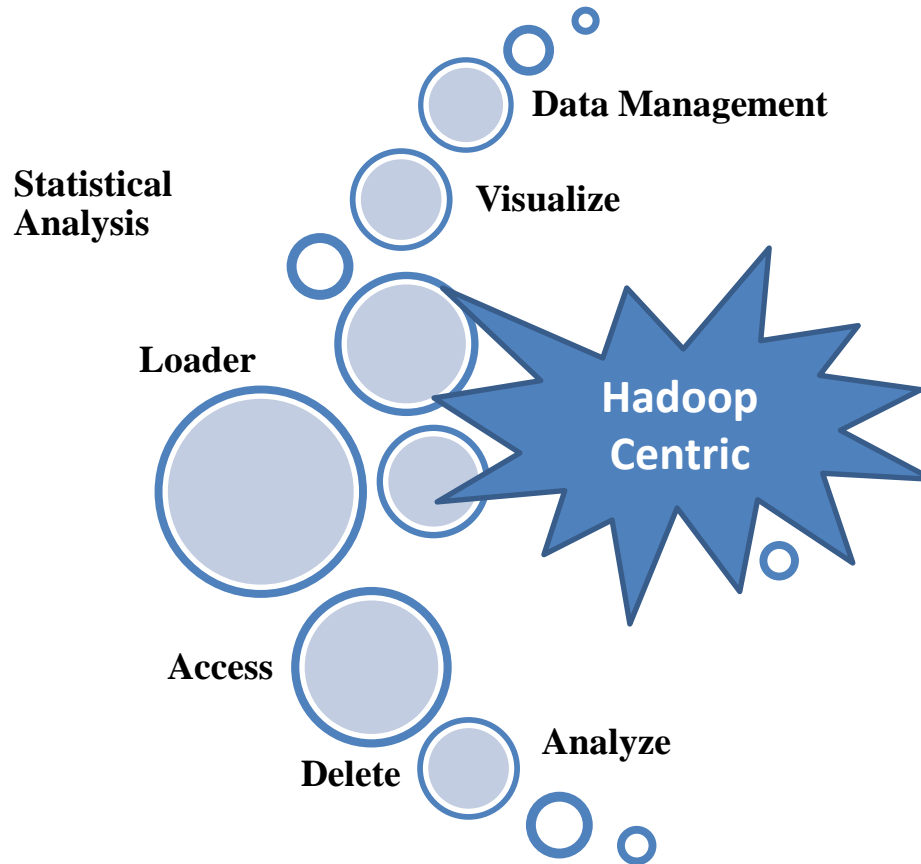
BD tools



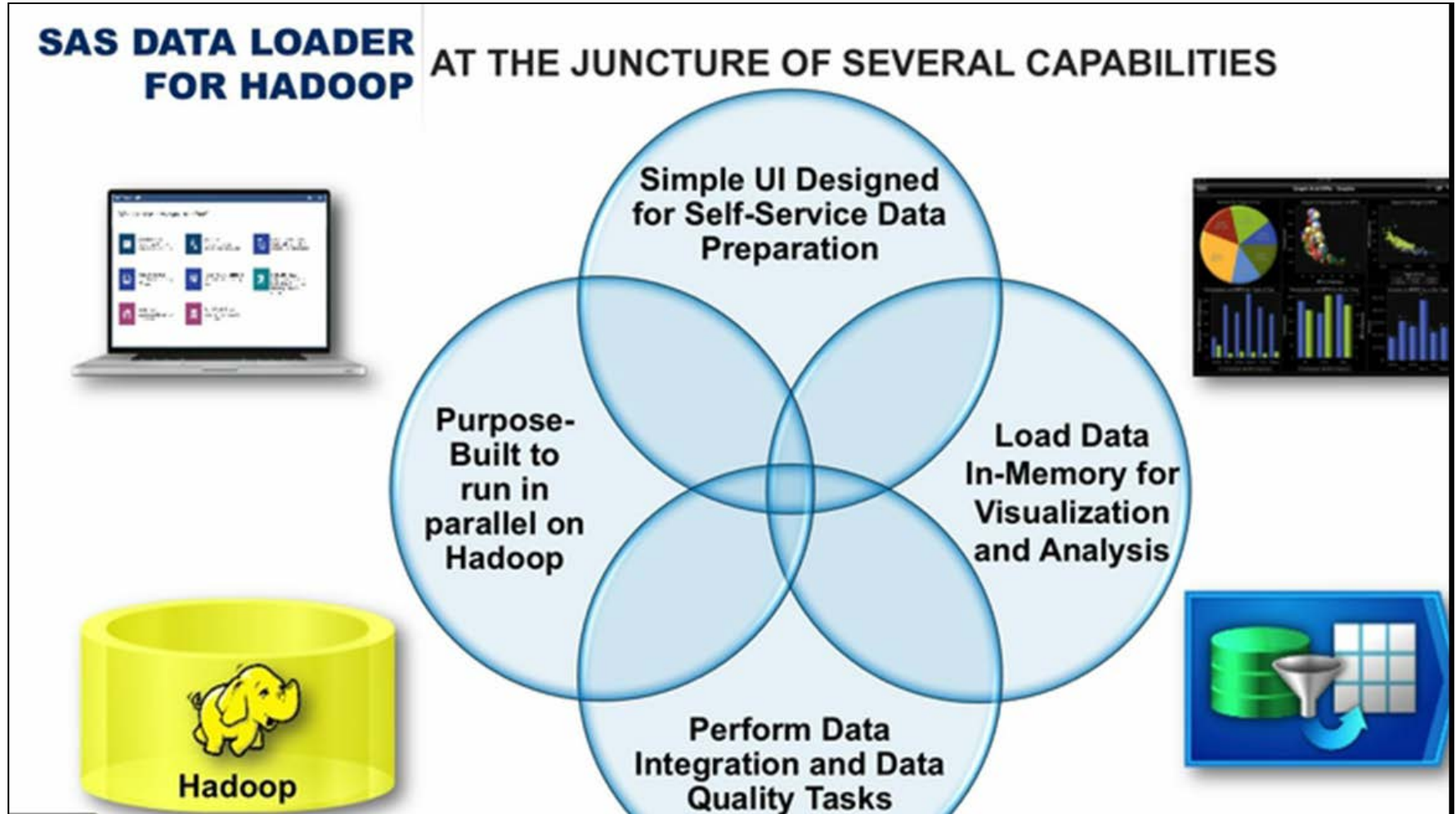
- **Hadoop** is the primary storage mechanism for Big Data with the following characteristics:
 - Managing very large volumes of data
 - Managing data on less expensive storage
 - Managing data by the “Roman Census” methods
 - Storing data in an unstructured manner.
 - Open source software framework can store large amounts of data and run applications on clusters of commodity hardware.

Corporation Data

- **Technology Services around Hadoop**



SAS Data Loader for Hadoop



SAS Data Loader for Hadoop

SAS® DATA LOADER FOR HADOOP

SELF-SERVICE DATA PREPARATION, DATA INTEGRATION
AND DATA QUALITY FOR HADOOP



**Business
Analyst/Data
Scientist**



BD Analytics

- **B**ig data analytics is the process of examining big data to uncover hidden patterns, unknown correlations and other useful information that can be used to make better decisions.”
- With big data analytics, data scientists and others can analyze **huge volumes** of data that conventional analytics and business intelligence solutions can't touch.
- Consider this; it's possible that your organization could accumulate (if it hasn't already) billions of rows of data with hundreds of millions of data combinations in multiple data stores and abundant formats.
- High-performance analytics is necessary to process that much data in order to figure out what's important and what isn't.

BD Users



BD Architecture



http://www.sas.com/en_us/insights/big-data/what-is-big-data.html

Why is BD Analytics Important?



Why Big Data?

- Data Growth is Huge.
- All that Data is Valuable.
- Data won't fit on a single computer distributed Data.
- Distributed Data= Faster Computation.
- Image Recognition
- Statistical Data Analysis
- Machine learning and more....

How BD works and what are the Key Technologies?

There's no single technology that encompasses big data analytics.
Here are the biggest players:

Data management

Data needs to be high quality and well-governed before it can be reliably analyzed.

With data constantly flowing in and out of an organization, it's important to establish repeatable processes to build and maintain standards for data quality.

Once data is reliable, organizations should establish a master data management program that gets the entire enterprise on the same page (Agile project management).

How BD works and what are the Key Technologies?

Data mining

Data mining technology helps you examine large amounts of data to discover patterns in the data – and this information can be used for further analysis to help answer complex business questions.

With data mining software, you can sift through all the chaotic and repetitive noise in data, pinpoint what's relevant, use that information to assess likely outcomes, and then accelerate the pace of making informed decisions.

How BD works and what are the Key Technologies?

Hadoop

This open source software framework can store large amounts of data and run applications on clusters of commodity hardware.

It has become a key technology to doing business due to the constant increase of data volumes and varieties, and its distributed computing model processes big data fast.

An additional benefit is that Hadoop's open source framework is free and uses commodity hardware to store large quantities of data.

How BD works and what are the Key Technologies?

In-memory analytics

By analyzing data from system memory (instead of from your hard disk drive), you can derive immediate insights from your data and act on them quickly.

This technology is able to remove data prep and analytical processing latencies to test new scenarios and create models; it's not only an easy way for organizations to stay agile and make better business decisions, it also enables them to run iterative and interactive analytics scenarios.

How BD works and what are the Key Technologies?

Predictive analytics

Predictive analytics technology uses data, statistical algorithms and machine-learning techniques to identify the likelihood of future outcomes based on historical data.

It's all about providing a best assessment on what will happen in the future, so organizations can feel more confident that they're making the best possible business decision.

Some of the most common applications of predictive analytics include fraud detection, risk, operations and marketing.

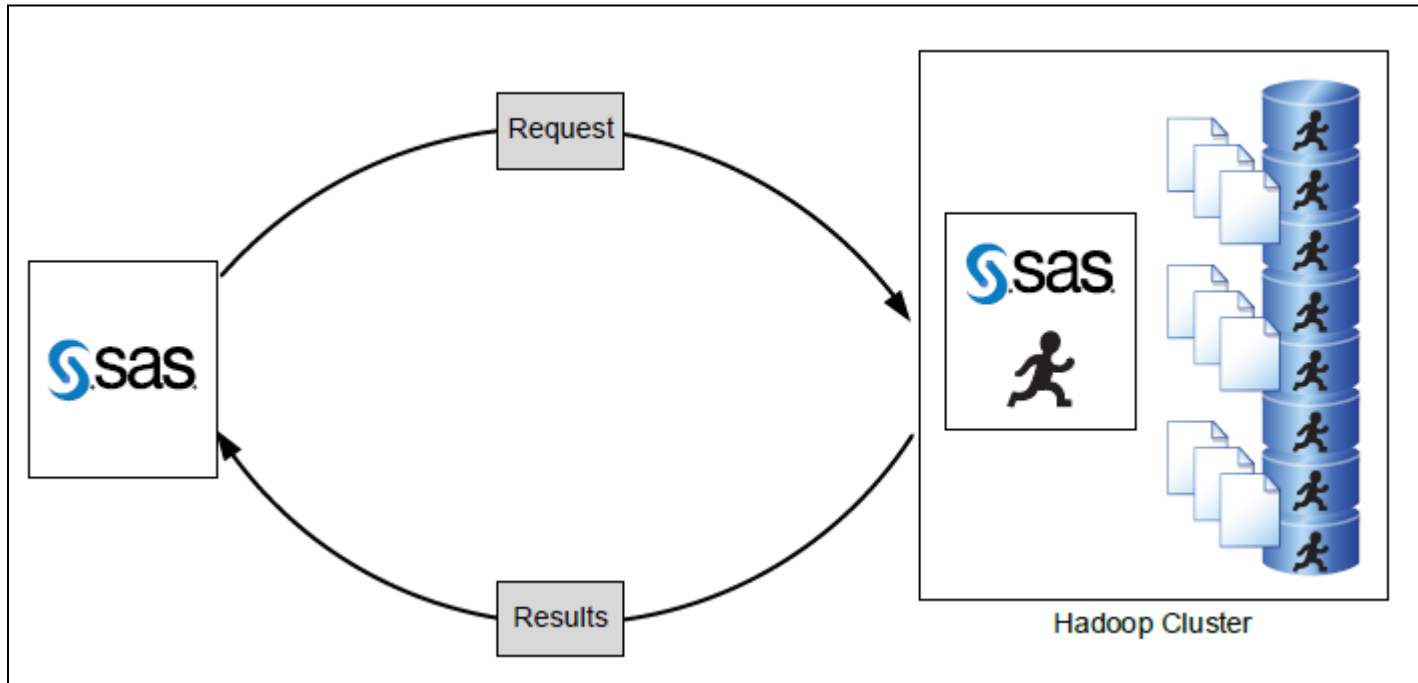
How BD works and what are the Key Technologies?

Text mining

With text mining technology, you can analyze text data from the web, comment fields, books and other text-based sources to uncover insights you hadn't noticed before.

Text mining uses machine learning or natural language processing technology to comb through documents – emails, blogs, Twitter feeds, surveys, competitive intelligence and more – to help you analyze large amounts of information and discover new topics and term relationships.

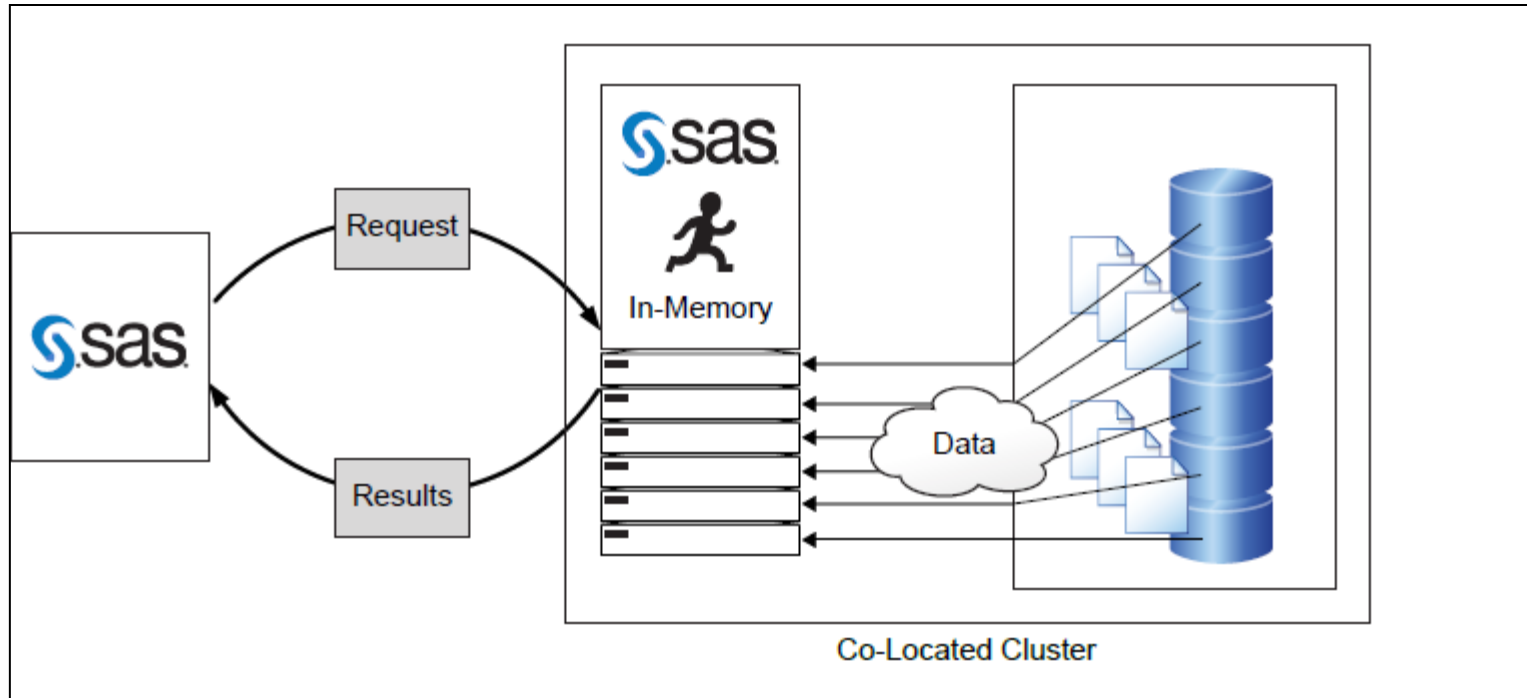
Processing in the Hadoop Cluster



SAS Data Loader and SAS In-Database Technology can process data in the Hadoop cluster.

SAS/ACCESS Interface to Hadoop can pass SQL code to the Hadoop cluster, the SAS Scalable Performance Data (SPD) Engine can submit data sub setting to the Hadoop cluster, and PROC HADOOP enables you to submit MapReduce programs for further processing by Hadoop.

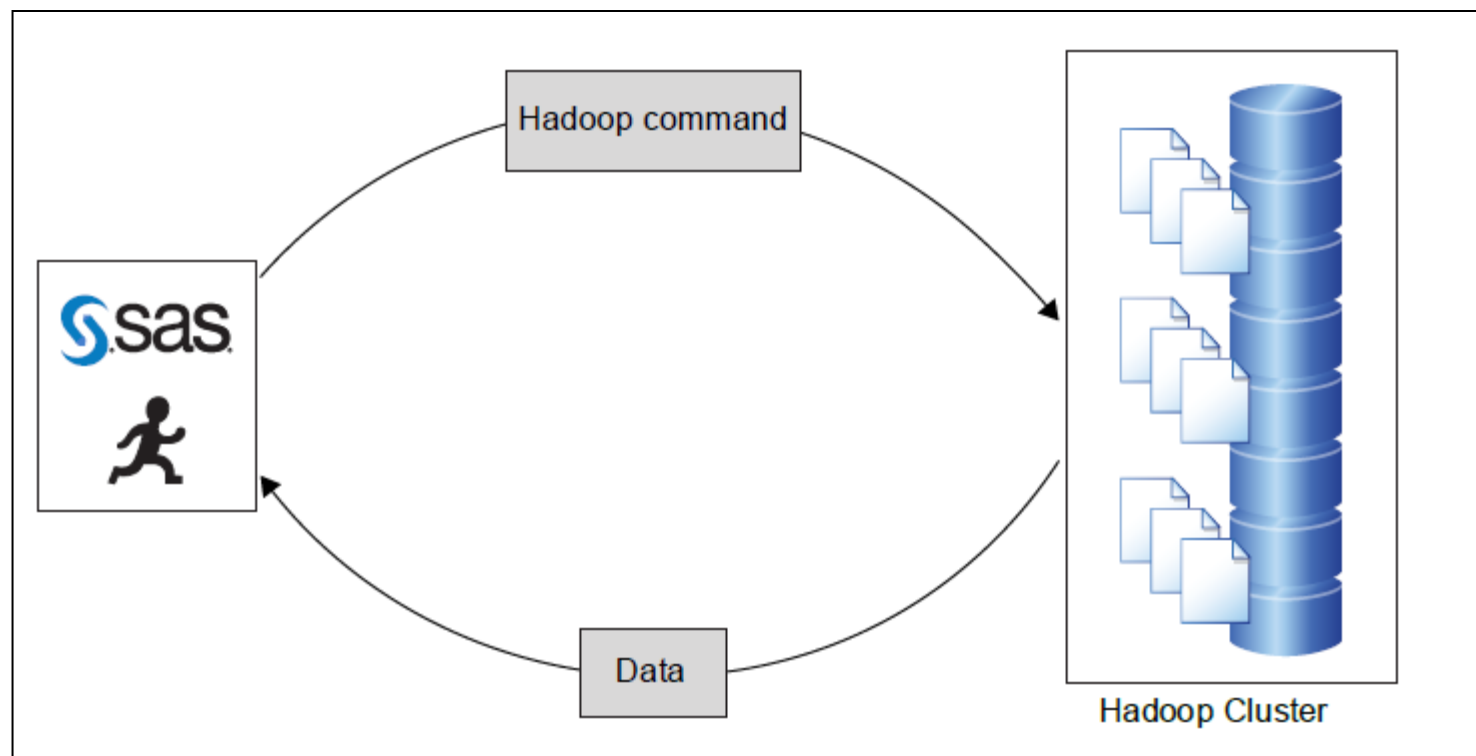
Processing in a SAS In-Memory Environment



The in-memory environment is on a separate set of machines from the Hadoop cluster.

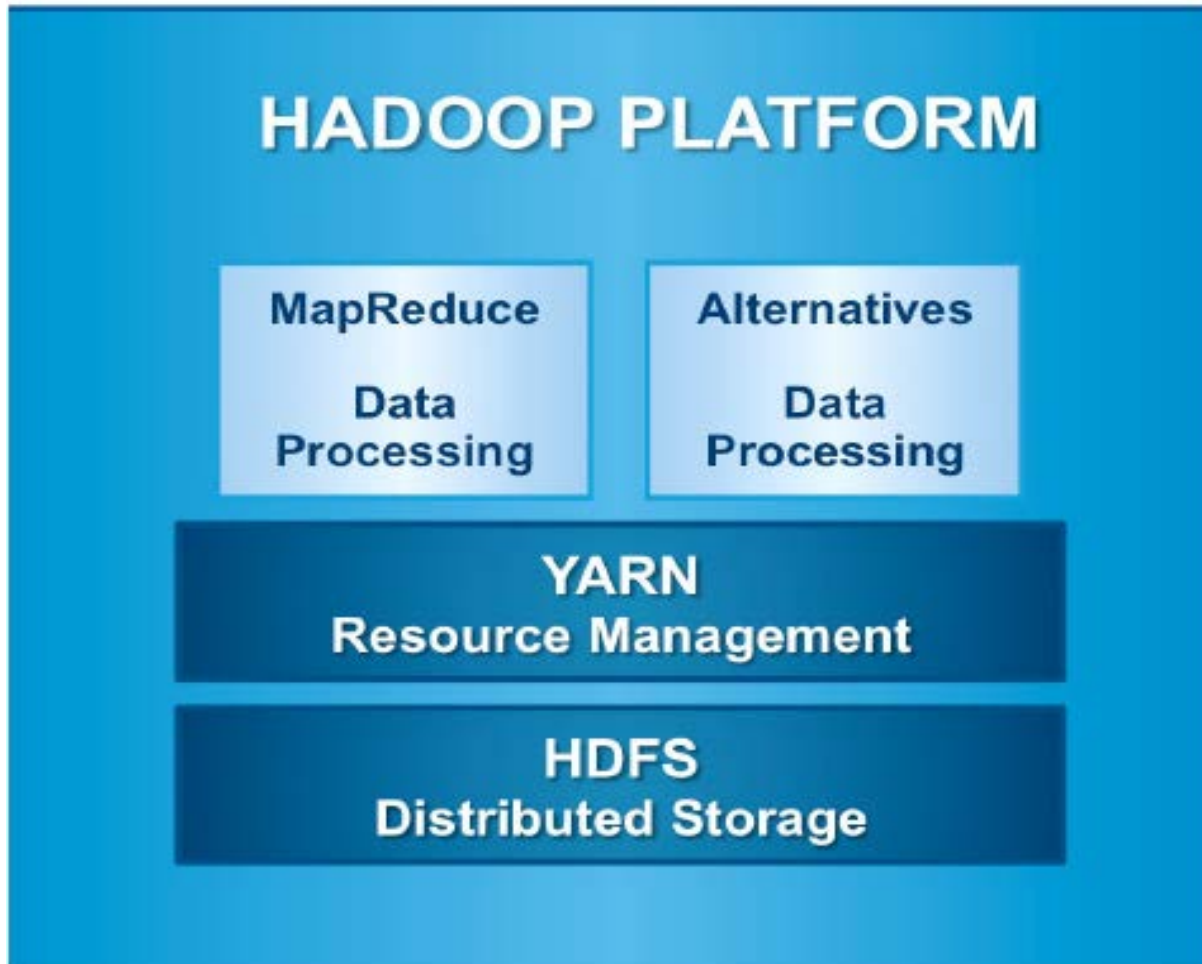
The SAS server or client connects to the analytics cluster that is remote from the Hadoop cluster, submits a request, loads the Hadoop data to the in-memory environment, processes the request, and then sends only the results back to SAS.

Processing in a SAS In-Memory Environment



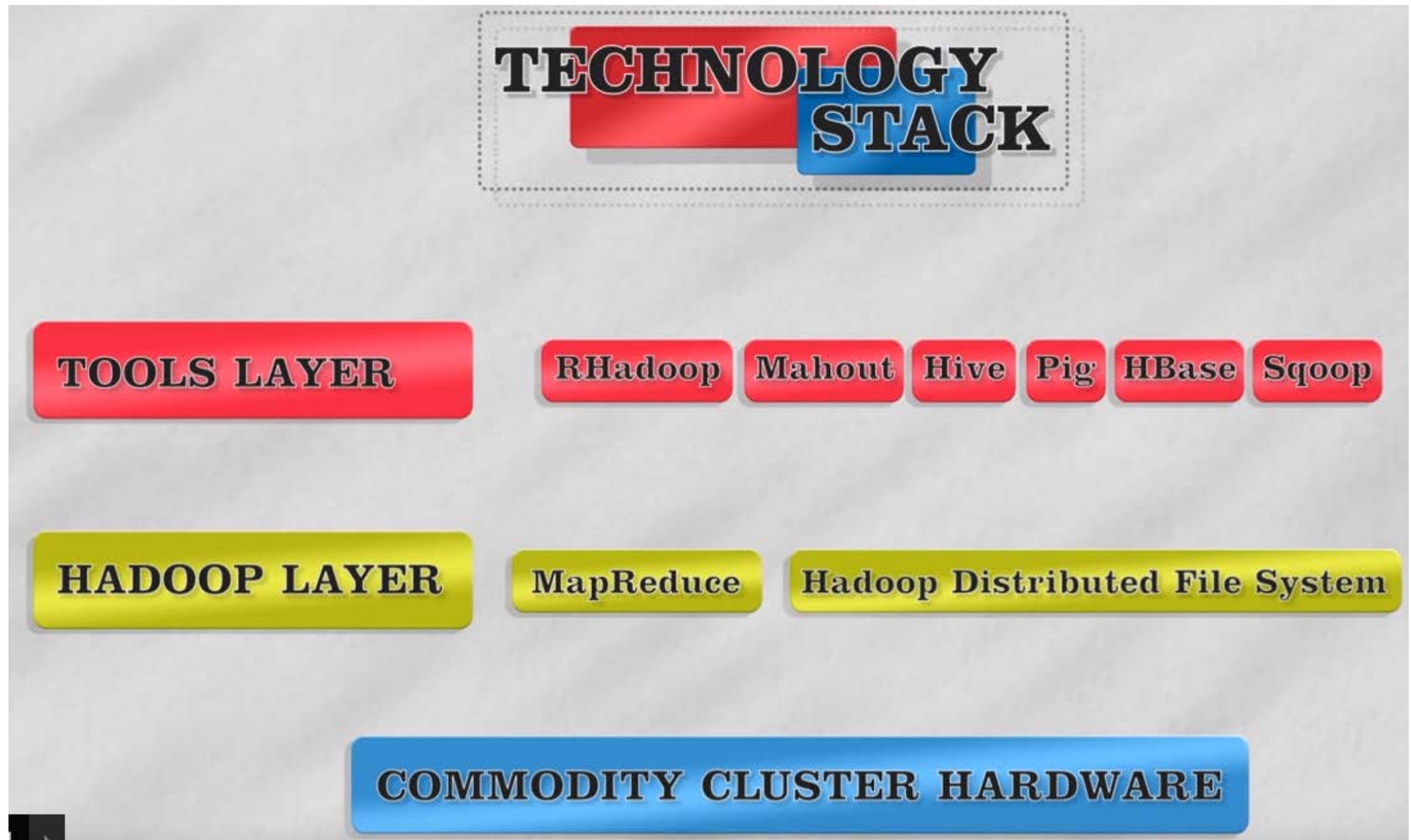
SAS Visual Analytics, SAS In-Memory Statistics, SAS High-Performance Analytics products (such as SAS High-Performance Data Mining, SAS High- Performance Econometrics, SAS High-Performance Optimization, SAS High- Performance Statistics, and SAS High-Performance Text Mining), SAS High- Performance Risk, and SAS Visual Scenario Designer can process Hadoop data in an in-memory environment.

The Hadoop Platform



- Hadoop Distributed File System (HDFS).
- Hadoop YARN – a resource-management platform responsible for managing compute resources in clusters and using them to schedule users' applications.

The Hadoop Platform



Explaining Big Data



BIG DATA, WHAT'S THE BENEFIT FOR TRANSPORTATION?

Asset Management

Big data presents opportunities to identify problems, analyze and reduce project costs.

Real Time Traffic Management

Big data provides new opportunities to predict congestion before it happens, using a combination of real time information, historical trends and clever algorithms.

BIG DATA, WHAT'S THE BENEFIT FOR TRANSPORTATION?

Asset Management

Big data presents opportunities to identify problems, analyze and reduce project costs.

Real Time Traffic Management

Big data provides new opportunities to predict congestion before it happens, using a combination of real time information, historical trends and clever algorithms.

Planning

An understanding of origins and destinations of trips – why, who, how and when they are made.

Funding

Big data presents funding, expenditures, and cash models.

Data Security

Big data helps you discover the 'who, what, when, where and why' of any security event.

Open Government

Big data analytics can help governments improve the way they allocate funds by finding and eliminating potential inefficiencies in spending.

Transit

Transit authority officials can use ticketing data correlated with GPS data to analyze rider habits, traffic schedules and other factors that influence how people get around town.

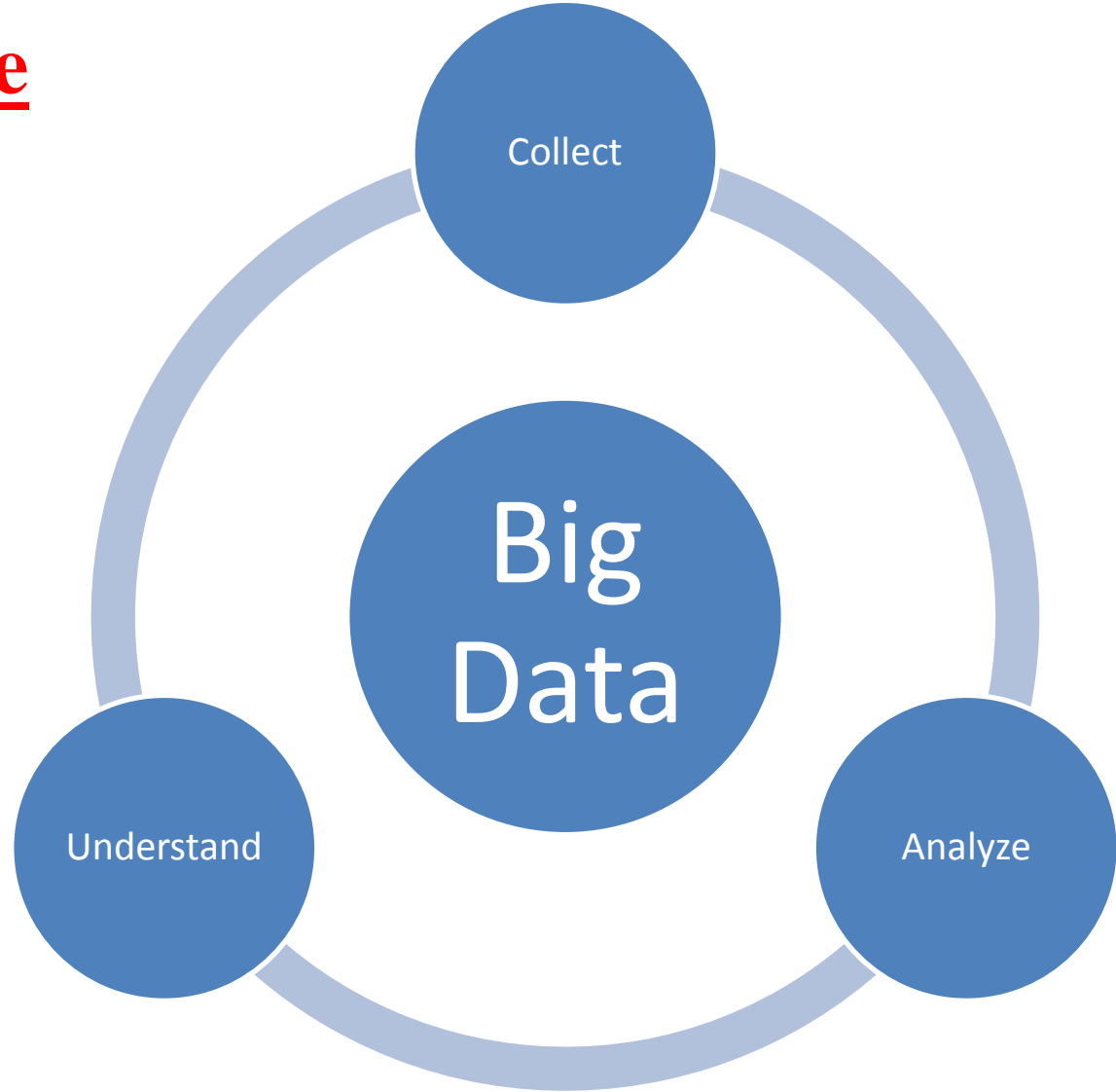
....

State and Local Government



Study BD Challenge

- Big Data Management
- Data Visualization
- Big Data Analytics
- Statistical Methods for Data Science
- Google BigQuery



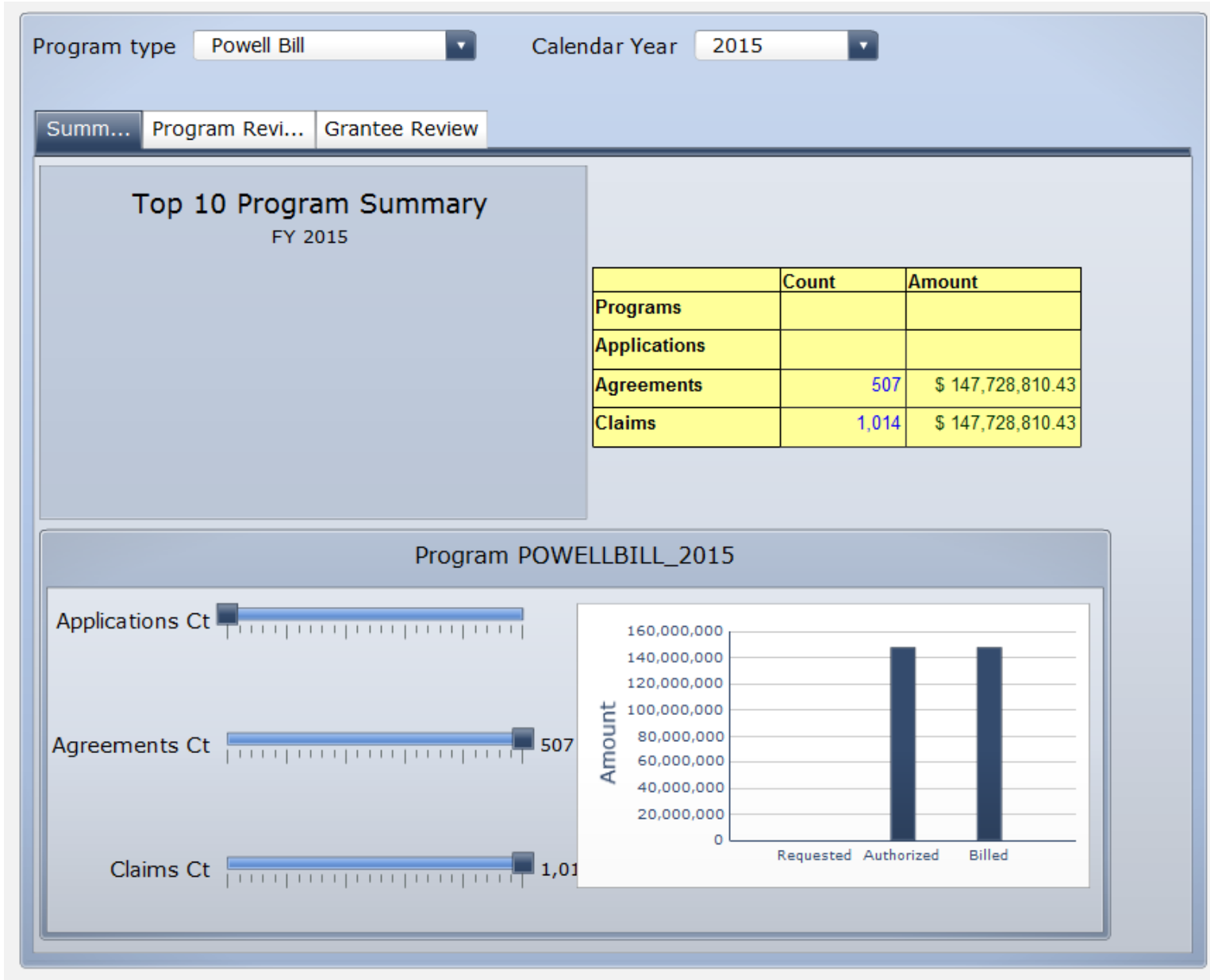
North Carolina Government Data Analytics Center (GDAC)

Data & Analytics

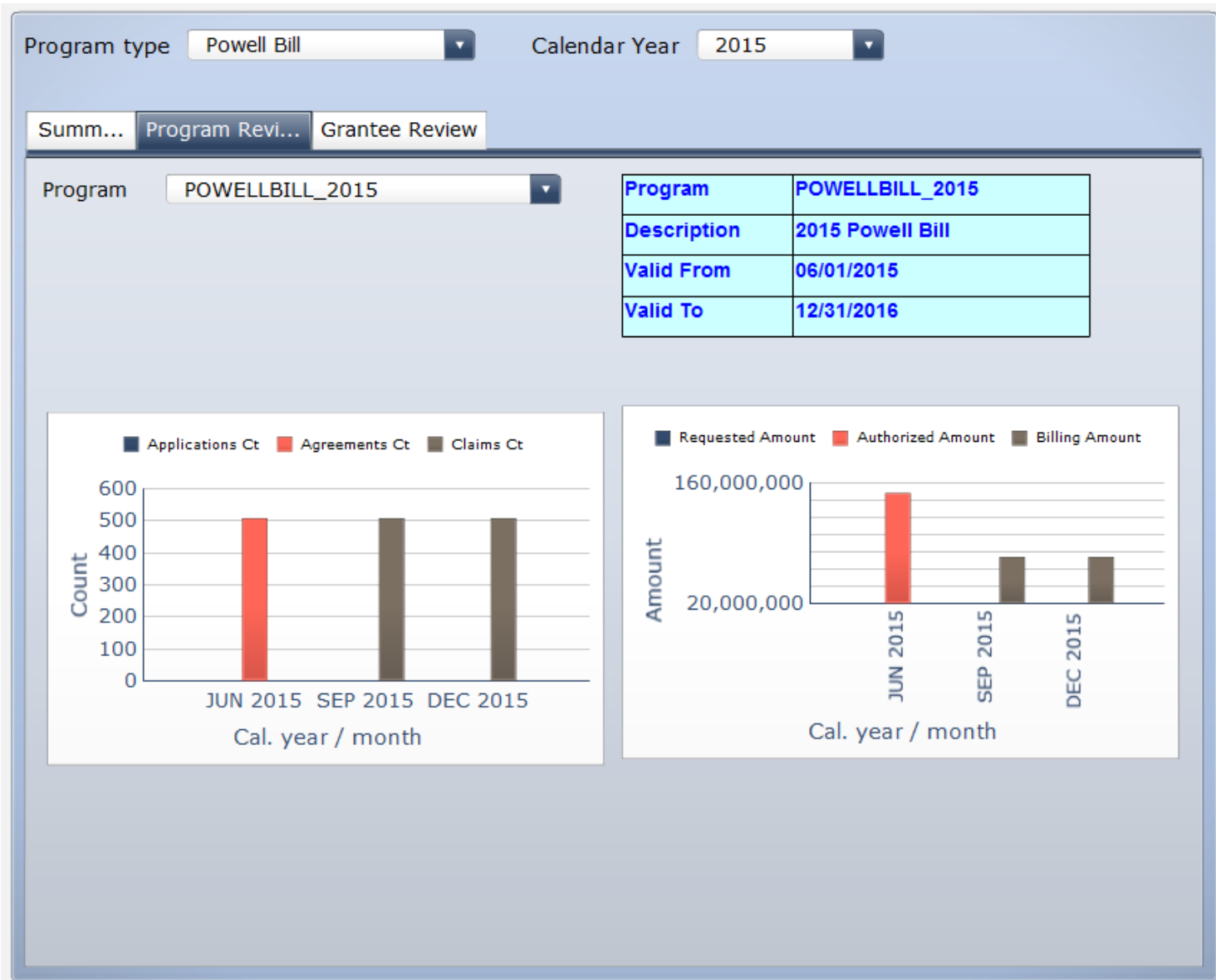


<http://it.nc.gov/nc-gdac>

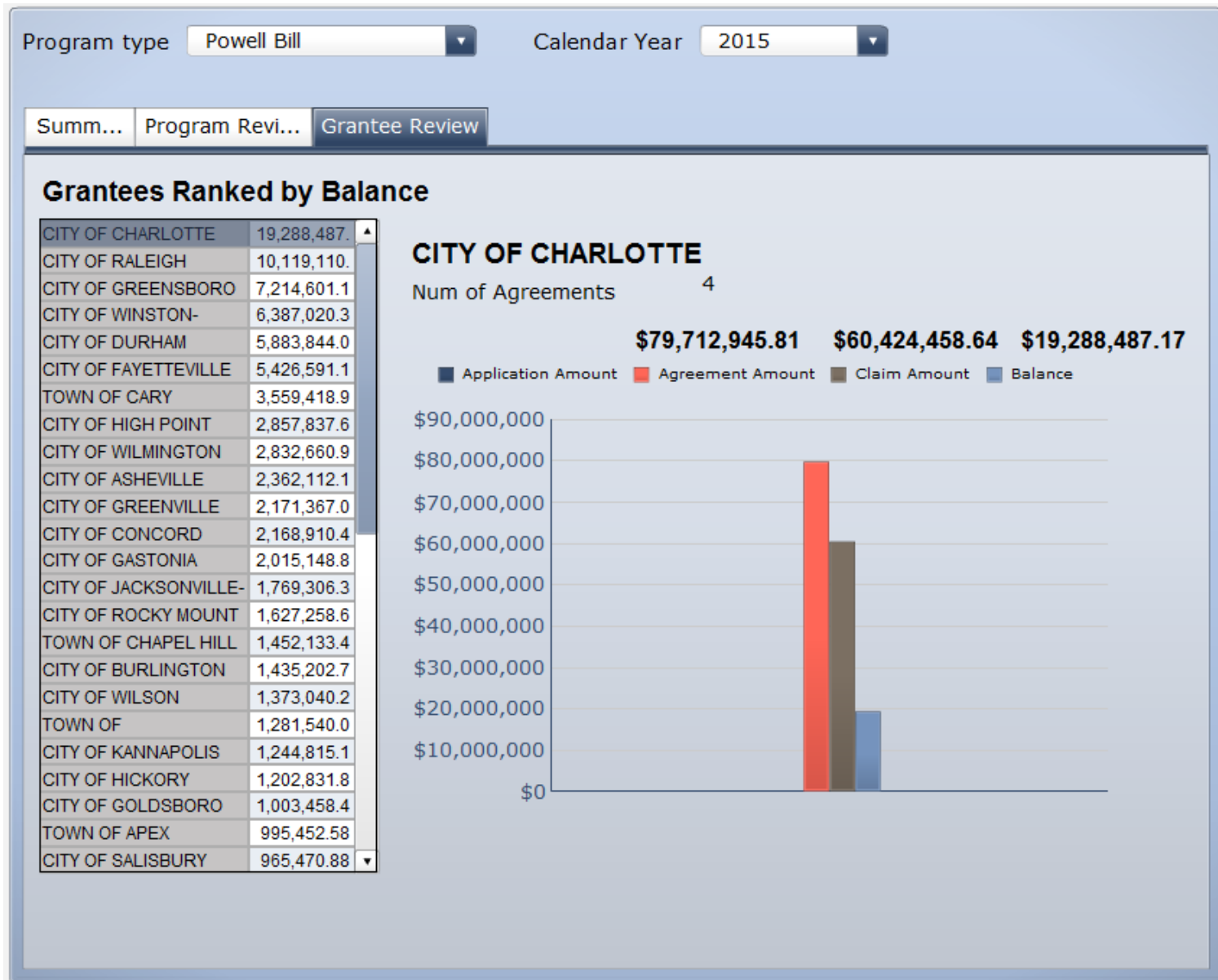
Dashboards SAP Business Objects



Dashboards SAP Business Objects

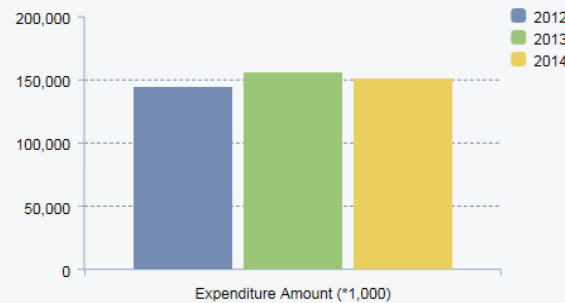
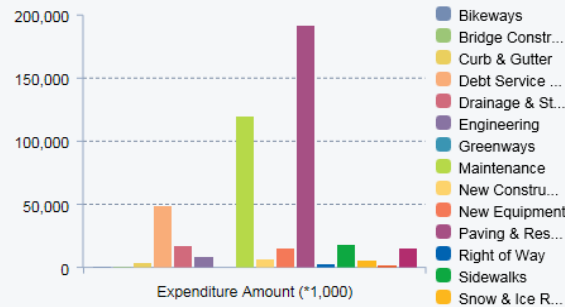


Dashboards SAP Business Objects



Graph Table

Grantee	Expenditure Amount 000 \$
CITY OF ALBEMARLE	1,728
CITY OF ARCHDALE	895
CITY OF ASHEBORO	2,574
CITY OF ASHEVILLE	7,199
CITY OF BELMONT	935
CITY OF BESSEMER CITY	417
CITY OF BOILING SPRING LAKES	866
CITY OF BREVARD	538
CITY OF BURLINGTON	4,340
CITY OF CHARLOTTE	70,206
CITY OF CLAREMONT	176
CITY OF CLINTON	775
CITY OF CONCORD	5,439
CITY OF CONOVER	793
CITY OF CREEDMOOR	241
CITY OF DREXEL	256
CITY OF DUNN	1,077
CITY OF DURHAM	18,267
CITY OF EDEN	1,434
CITY OF ELIZABETH CITY	2,490
CITY OF FAYETTEVILLE	14,796
CITY OF GASTONIA	6,022
CITY OF GOLDSBORO	3,002
CITY OF GRAHAM	1,182
CITY OF GREENSBORO	21,323
CITY OF GREENVILLE	7,781
CITY OF HAMLET	586
CITY OF HAVELOCK	1,911
CITY OF HENDERSON	1,156
CITY OF HENDERSONVILLE	1,406
CITY OF HICKORY	3,694
CITY OF HIGH POINT	8,712
CITY OF HIGH SHOALS	64
CITY OF JACKSONVILLE-FINANCE DEPT	5,074
CITY OF KANNAPOLIS	4,087
CITY OF KING	633
CITY OF KINGS MOUNTAIN	1,469
CITY OF KINSTON	1,891
CITY OF LAURINBURG	1,271
CITY OF LENOIR	1,717



Graph

Table

Cancel

Apply

Program Year

- >

Expense Type

- >

Grantee

- >

Key Figures

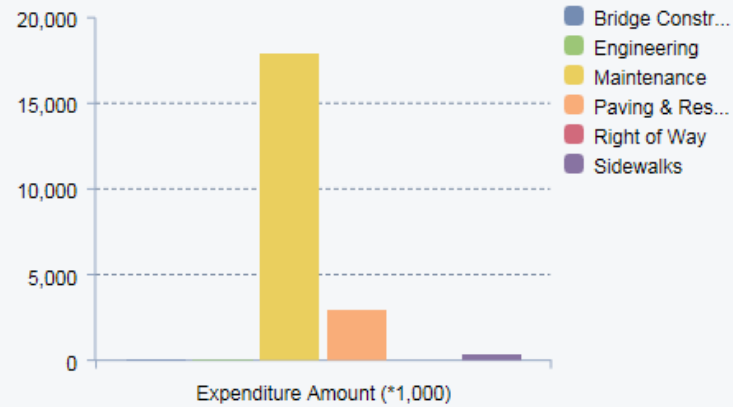
- >

Clear Filters

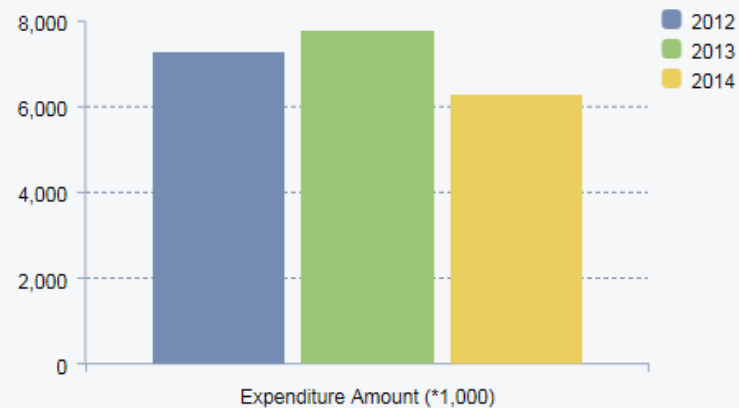
	Expenditure Amount			
Program Year	2012 ₪	2013 ₪	2014 ₪	Overall Result ₪
Grantee Δ	* 1,000 \$	* 1,000 \$	* 1,000 \$	* 1,000 \$
CITY OF ALBEMARLE	626	558	544	1,728
CITY OF ARCHDALE	253	390	252	895
CITY OF ASHEBORO	782	891	901	2,574
CITY OF ASHEVILLE	2,362	2,400	2,436	7,199
CITY OF BELMONT	267	271	396	935
CITY OF BESSEMER CITY	141	152	124	417
CITY OF BOILING SPRING LAKES	283	247	337	866
CITY OF BREVARD	166	271	101	538
CITY OF BURLINGTON	1,437	1,450	1,453	4,340
CITY OF CHARLOTTE	19,700	25,566	24,941	70,206
CITY OF CLAREMONT	161	11	4	176
CITY OF CLINTON	256	258	260	775
CITY OF CONCORD	1,803	2,388	1,249	5,439
CITY OF CONOVER	213	327	253	793
CITY OF CREEDMOOR	21	104	117	241
CITY OF DREXEL	191	59	5	256
CITY OF DUNN	277	352	447	1,077
CITY OF DURHAM	5,946	6,094	6,227	18,267
CITY OF EDEN	467	485	482	1,434
CITY OF ELIZABETH CITY	365	1,257	868	2,490
CITY OF FAYETTEVILLE	5,507	5,474	3,815	14,796
CITY OF GASTONIA	1,970	1,998	2,054	6,022
CITY OF GOLDSBORO	1,004	1,001	997	3,002
CITY OF GRAHAM	390	398	394	1,182
CITY OF GREENSBORO	7,266	7,790	6,266	21,323
CITY OF GREENVILLE	2,212	2,983	2,586	7,781
CITY OF HAMLET	203	163	220	586
CITY OF HAVELOCK	867	495	550	1,911
CITY OF HENDERSON	503	310	343	1,156
CITY OF HENDERSONVILLE	350	671	385	1,406
CITY OF HICKORY	1,215	1,239	1,240	3,694
CITY OF HIGH POINT	2,858	2,913	2,941	8,712
CITY OF HIGH SHOALS	2	48	13	64
CITY OF JACKSONVILLE-FINANCE DEPT	2,024	916	2,133	5,074
CITY OF KANNAPOLIS	1,134	1,283	1,670	4,087
CITY OF KING	198	216	219	633
CITY OF KINGS MOUNTAIN	476	399	594	1,469
CITY OF KINSTON	626	636	629	1,891
CITY OF LAURINBURG	470	428	373	1,271
CITY OF LENOIR	570	572	576	1,717

Grantee	Expenditure Amount (\$1,000)
CITY OF ALBEMARLE	1,728
CITY OF ARCHDALE	895
CITY OF ASHEBORO	2,574
CITY OF ASHEVILLE	7,199
CITY OF BELMONT	935
CITY OF BESSEMER CITY	417
CITY OF BOILING SPRING LAKES	866
CITY OF BREVARD	538
CITY OF BURLINGTON	4,340
CITY OF CHARLOTTE	70,206
CITY OF CLAREMONT	176
CITY OF CLINTON	775
CITY OF CONCORD	5,439
CITY OF CONOVER	793
CITY OF CREEDMOOR	241
CITY OF DREXEL	256
CITY OF DUNN	1,077
CITY OF DURHAM	18,267
CITY OF EDEN	1,434
CITY OF ELIZABETH CITY	2,490
CITY OF FAYETTEVILLE	14,796
CITY OF GASTONIA	6,022
CITY OF GOLDSBORO	3,002
CITY OF GRAHAM	1,182
CITY OF GREENSBORO	21,323
CITY OF GREENVILLE	7,781
CITY OF HAMLET	586
CITY OF HAVELOCK	1,911
CITY OF HENDERSON	1,156
CITY OF HENDERSONVILLE	1,406
CITY OF HICKORY	3,694
CITY OF HIGH POINT	8,712
CITY OF HIGH SHOALS	64

CITY OF GREENSBORO

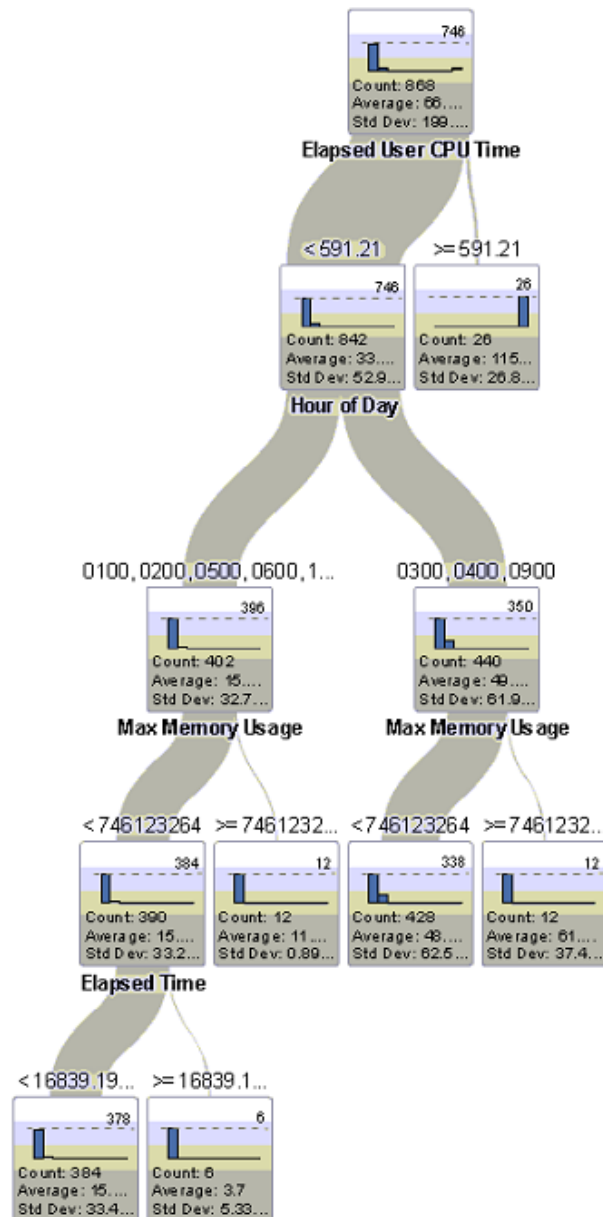


All Expense Types



Decision Tree with SAS® Visual Analytics

Decision Tree of Elapsed Max CPU Time



Demo:

NCDOT

SAS Model & SAS VA

Dashboards with SAS® Visual Analytics



Construction Dashboard provides detailed forecasts for active and future construction projects; monthly or fiscal year views are supported. Payout curves can be viewed from the aggregate down to the individual project level. A breakdown of project let delays is also provided.

Click [Here](#) to view **Revenue Dashboard**

Standard Forecast

Setup

[Data Refresh](#): Select source files to refresh the reports

Available Reports

[Overall Forecast](#): Overall construction expenditure forecast by month and year

[Fiscal Year Breakdown](#): Overall fiscal year construction expenditure forecast

[Individual Project Payouts](#): Payout curves by project

[Let Delay Breakdown](#): Various summarizations of project let delays

Custom and What If Forecast

Custom Setup

[Project Setup](#): Modify parameters and refresh custom construction expenditure forecast

Available Reports

[Overall Forecast](#): Custom Overall construction expenditure forecast by month and year

[Fiscal Year Breakdown](#): Custom Overall fiscal year construction expenditure forecast

What If Setup

[Download Parameter Form](#): Allows for customization – Modify project cost, let delay, or remove a project

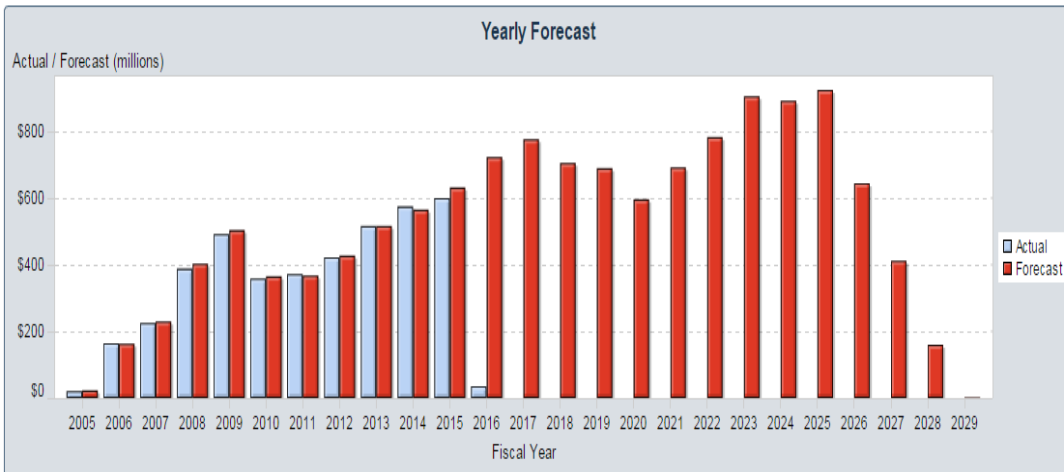
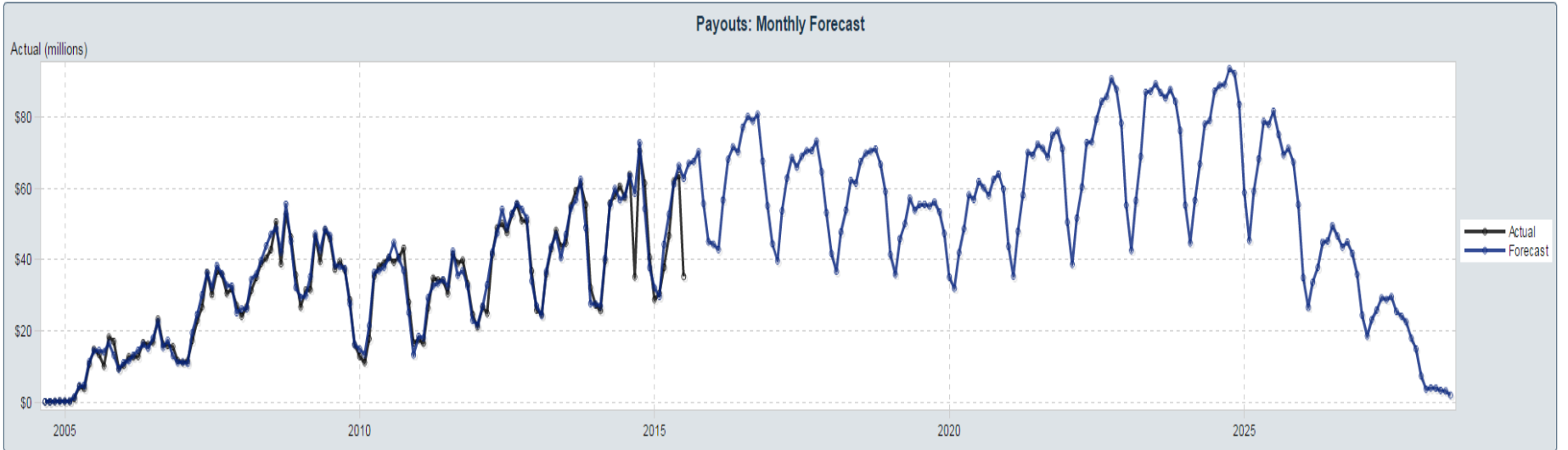
[Upload Parameter Form](#): Upload parameter form(automatically updates all what-if forecast reports)

Available Reports

[Forecast - Fiscal Year](#): What-if Overall fiscal year construction expenditure forecast

[Forecast - Monthly](#): What-if Overall construction expenditure forecast by month and year

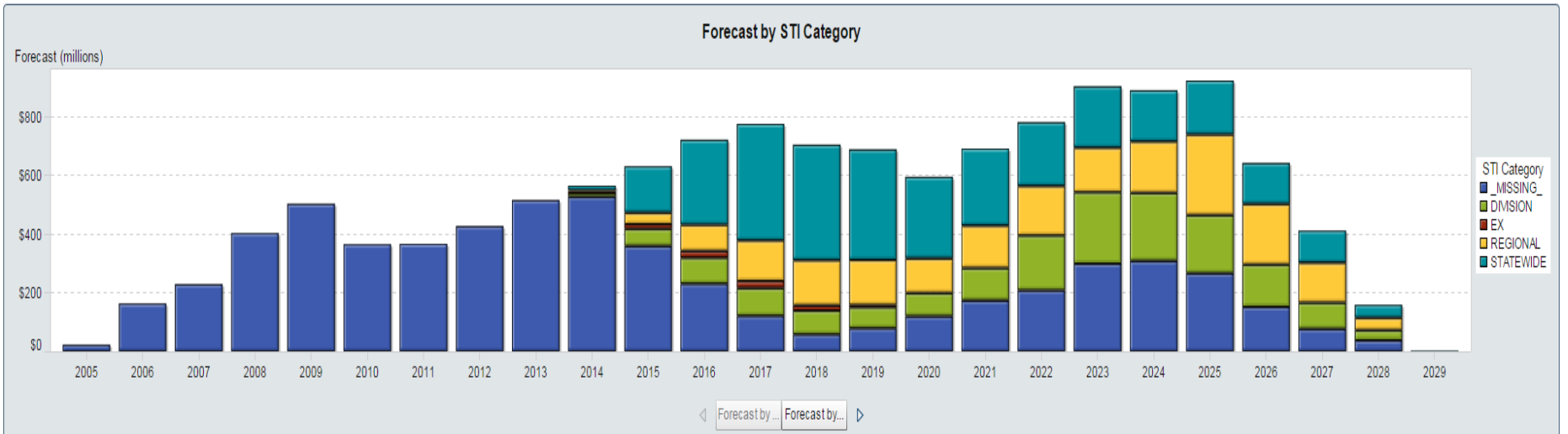
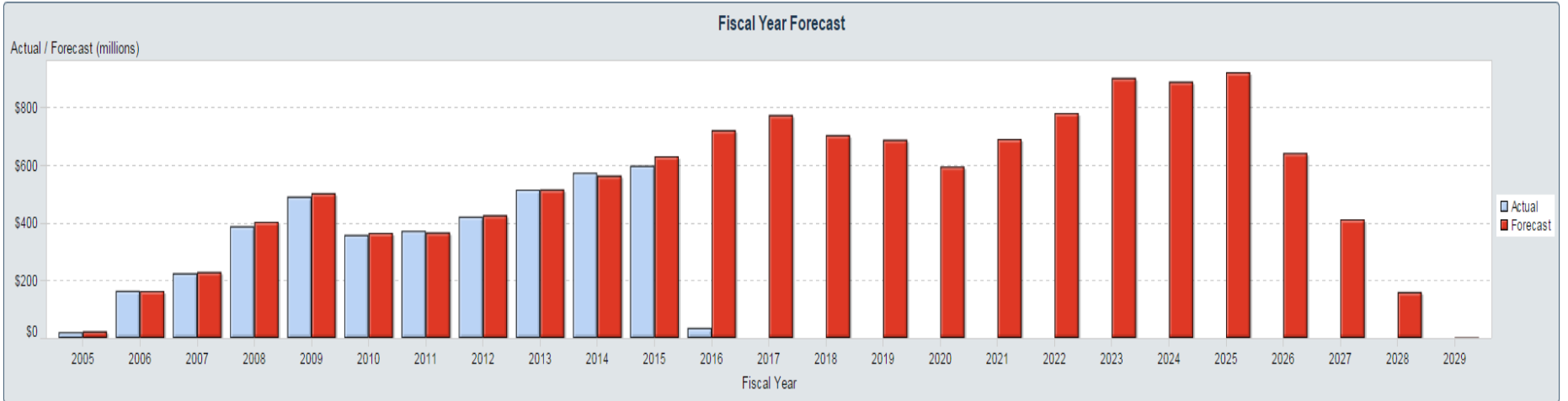
Upload Date 14Sep2015	Fiscal Year 2005 ————— 2029 2005 2029	STIP No <input type="checkbox"/> A-0009B <input type="checkbox"/> A-0010AA <input type="checkbox"/> A-0011BB <input type="checkbox"/> A-0011C <input type="checkbox"/> B-0682	STI Category <input type="checkbox"/> _MISSING_ <input type="checkbox"/> DIVISION <input type="checkbox"/> EX <input type="checkbox"/> REGIONAL <input type="checkbox"/> STATEWIDE	Project Type <input type="checkbox"/> A <input type="checkbox"/> B <input type="checkbox"/> BK <input type="checkbox"/> EB <input type="checkbox"/> F	Sub Type <input type="checkbox"/> _MISSING_ <input type="checkbox"/> BRIDGE <input type="checkbox"/> CAPACITY <input type="checkbox"/> FERRY <input type="checkbox"/> HIGHWAY	Let Type <input type="checkbox"/> _MISSING_ <input type="checkbox"/> BDRL <input type="checkbox"/> DBL <input type="checkbox"/> DDL <input type="checkbox"/> DDRL	Bid Status <input type="checkbox"/> AWARDED <input type="checkbox"/> NOT AWA <input type="checkbox"/> NOT AWARDED	Bid Fund <input type="checkbox"/> (missing values) <input type="checkbox"/> F <input type="checkbox"/> S
---------------------------------	--	---	--	---	---	---	---	--



Yearly Forecast Summary Table

Fiscal Year ▲	Actual	Forecast
2005	\$20,868,113	\$23,103,491
2006	\$164,402,730	\$163,615,192
2007	\$225,407,316	\$229,841,988
2008	\$388,955,282	\$403,549,509
2009	\$491,557,264	\$504,159,266
2010	\$359,089,547	\$365,554,578
2011	\$372,265,994	\$367,256,739
2012	\$422,007,791	\$427,994,364
2013	\$516,184,949	\$516,858,747
2014	\$575,050,128	\$564,869,947
2015	\$599,098,088	\$631,590,002
2016	\$655,000,000	\$700,000,000
2017	\$750,000,000	\$750,000,000
2018	\$700,000,000	\$700,000,000
2019	\$680,000,000	\$680,000,000
2020	\$580,000,000	\$580,000,000
2021	\$680,000,000	\$680,000,000
2022	\$780,000,000	\$780,000,000
2023	\$880,000,000	\$880,000,000
2024	\$880,000,000	\$880,000,000
2025	\$880,000,000	\$880,000,000
2026	\$640,000,000	\$640,000,000
2027	\$400,000,000	\$400,000,000
2028	\$150,000,000	\$150,000,000
2029	\$0	\$0

Upload Date 14Sep2015	Fiscal Year 2005 ————— 2029 2005 2029	STIP No <input type="checkbox"/> A-0009B <input type="checkbox"/> A-0010AA <input type="checkbox"/> A-0011BB <input type="checkbox"/> A-0011C <input type="checkbox"/> B-0682	STI Category <input type="checkbox"/> _MISSING_ <input type="checkbox"/> DIVISION <input type="checkbox"/> EX <input type="checkbox"/> REGIONAL <input type="checkbox"/> STATEWIDE	Project Type <input type="checkbox"/> A <input type="checkbox"/> B <input type="checkbox"/> BK <input type="checkbox"/> EB <input type="checkbox"/> F	Sub Type <input type="checkbox"/> _MISSING_ <input type="checkbox"/> BRIDGE <input type="checkbox"/> CAPACITY <input type="checkbox"/> FERRY <input type="checkbox"/> HIGHWAY	Let Type <input type="checkbox"/> _MISSING_ <input type="checkbox"/> BDRL <input type="checkbox"/> DBL <input type="checkbox"/> DDL <input type="checkbox"/> DDRL	Bid Status <input type="checkbox"/> AWARDED <input type="checkbox"/> NOT AWA <input type="checkbox"/> NOT AWARDED	Bid Fund <input type="checkbox"/> (missing values) <input type="checkbox"/> F <input type="checkbox"/> S
---------------------------------	--	---	--	---	---	---	---	--



Let Date

15Dec200501Jan2030

Year

☐ 1 Year
☐ 2 Years
☒ 3 Years

STIP No

☐ A-0009B
☐ A-0010AA
☐ B-1037
☐ B-1382
☐ B-2146

STI Category

☐ _MISSING_
☐ DIVISION
☐ EX
☐ REGIONAL
☐ STATEWIDE

Project Type

☐ _MISSING_
☐ BPOC
☐ CITY
☐ DB
☐ DDL

Sub Type

☐ _MISSING_
☐ BICYCLE AND PEDESTRIAN
☐ BRIDGE
☐ CAPACITY
☐ HIGHWAY

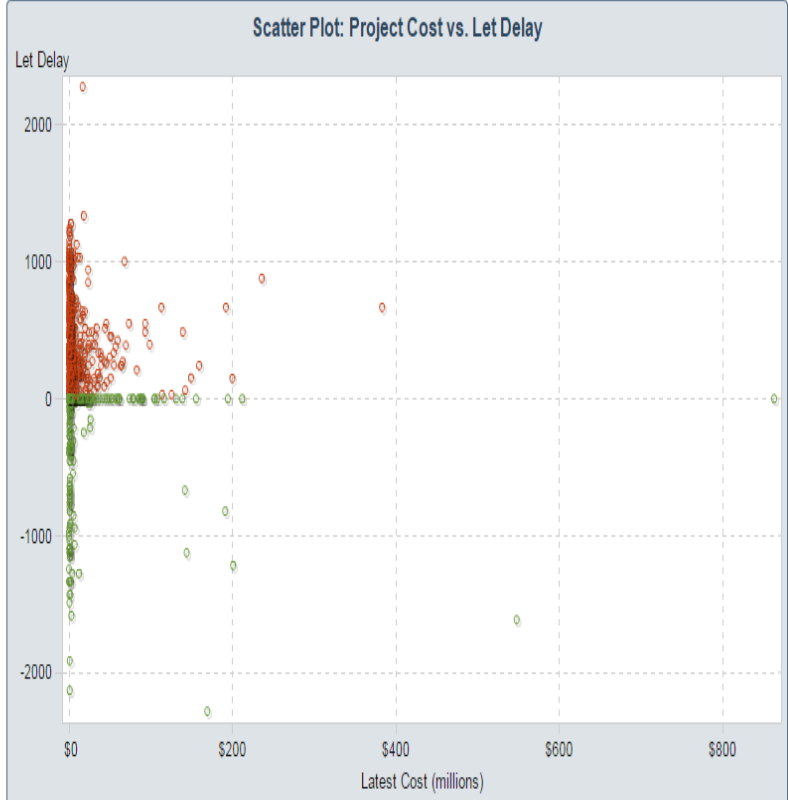
Project Type

☐ B
☐ BD
☐ I
☐ Other
☐ U

Division

☐ 1
☐ 10
☐ 11
☐ 12
☐ 13

Let Data								
STIP No	Description	Bid Status	Latest Cost	Design Build	Let Date (current)	Let Date (hist)	Let Delay	
A-0009B	US 74 FROM US 129 AT ROBBINSVILLE TO NC 143 NORTH OF CH...	NOT AWARDED	\$78,000,000	NO	01Jul2017	01Jul2017	0	
A-0010AA	NORTH OF I-240 IN ASHEVILLE TO US 25.	NOT AWARDED	\$89,000,000	NO	01Jan2022	01Jan2022	0	
B-1037	SOUTH FORK NEW RIVER. REPLACE BRIDGE NO. 39 (COORDINAT...	AWARDED	\$2,450,000	NO	01Jan2010	17Feb2009	318	
B-1382	BRIDGE 36 OVER BLACK RIVER OVERFLOW ON NC 44 SW OF WAR...	AWARDED	\$4,100,000	NO	01Aug2008	01Dec2008	600	



Let Date
15Dec2005 01Jan2030
15Dec2005 01Jan2030

Year
☐ 1 Year
☐ 2 Years
☒ 3 Years

STIP No
☐ A-0009B
☐ A-0010AA
☐ B-1037
☐ B-1382
☐ B-2146

STI Category
☐ _MISSING_
☐ DIVISION
☐ EX
☐ REGIONAL
☐ STATEWIDE

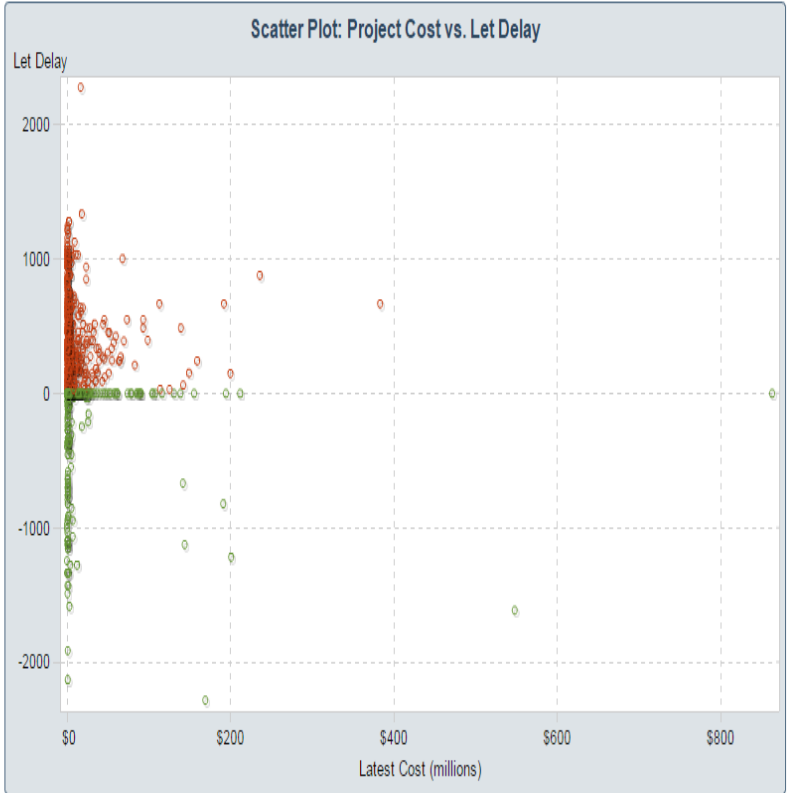
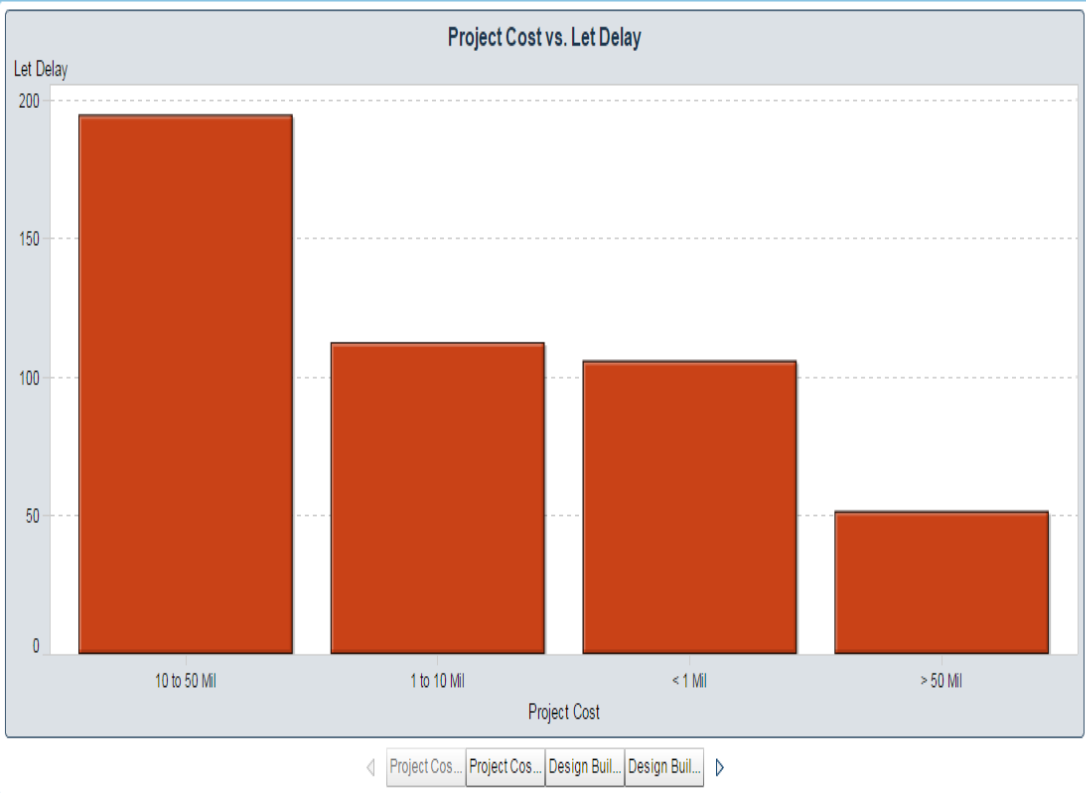
Project Type
☐ _MISSING_
☐ BPOC
☐ CITY
☐ DB
☐ DDL

Sub Type
☐ _MISSING_
☐ BICYCLE AND PEDESTRIAN
☐ BRIDGE
☐ CAPACITY
☐ HIGHWAY

Project Type
☐ B
☐ BD
☐ I
☐ Other
☐ U

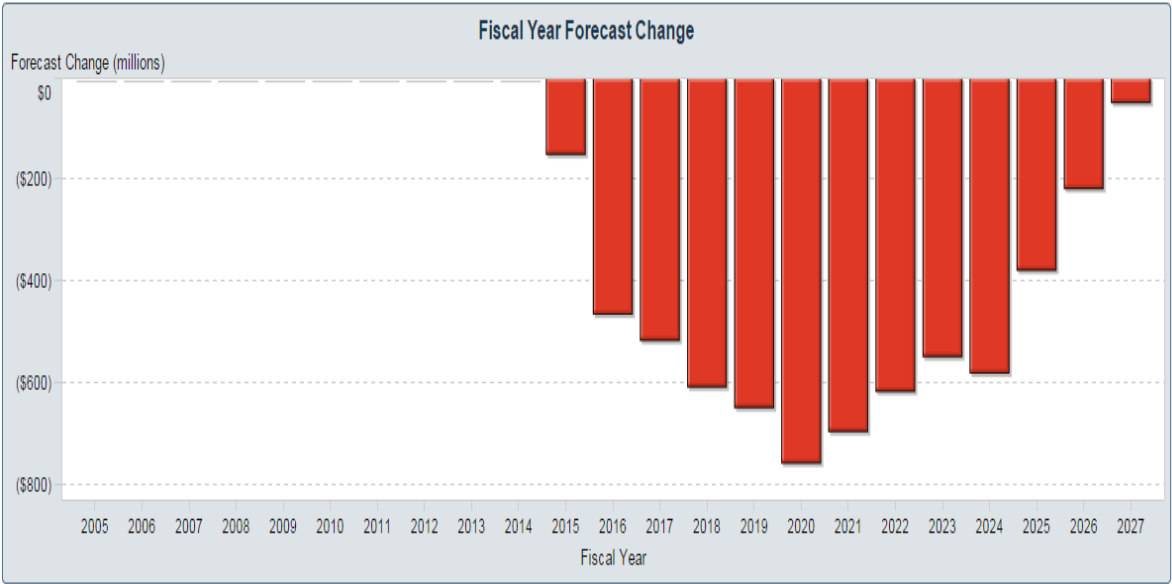
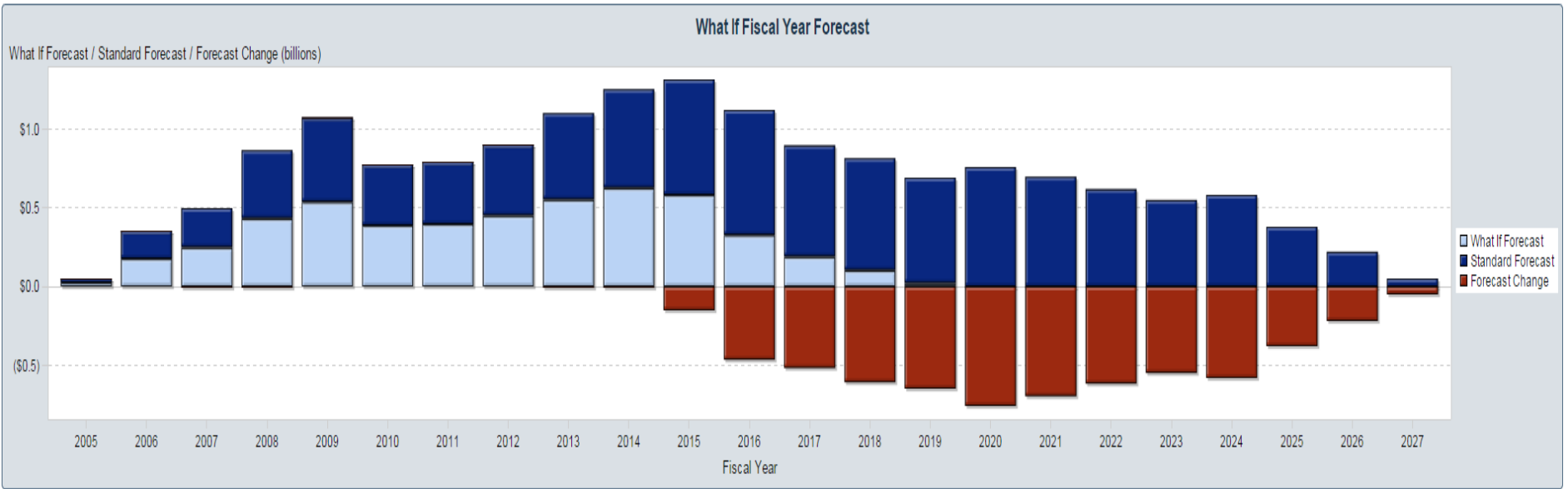
Division
☐ 1
☐ 10
☐ 11
☐ 12
☐ 13

Let Data								
STIP No	Description	Bid Status	Latest Cost	Design Build	Let Date (current)	Let Date (hist)	Let Delay	
A-0009B	US 74 FROM US 129 AT ROBBINSVILLE TO NC 143 NORTH OF CH...	NOT AWARDED	\$78,000,000	NO	01Jul2017	01Jul2017	0	
A-0010AA	NORTH OF I-240 IN ASHEVILLE TO US 25.	NOT AWARDED	\$89,000,000	NO	01Jan2022	01Jan2022	0	
B-1037	SOUTH FORK NEW RIVER. REPLACE BRIDGE NO. 39 (COORDINAT...	AWARDED	\$2,450,000	NO	01Jan2010	17Feb2009	318	
B-1382	BRIDGE 26 OVER BLACK RIVER OVERFLOW ON NC 14 SIMS CREEK	AWARDED	\$4,100,000	NO	01Aug2009	01Dec2008	600	



20052027

20052027



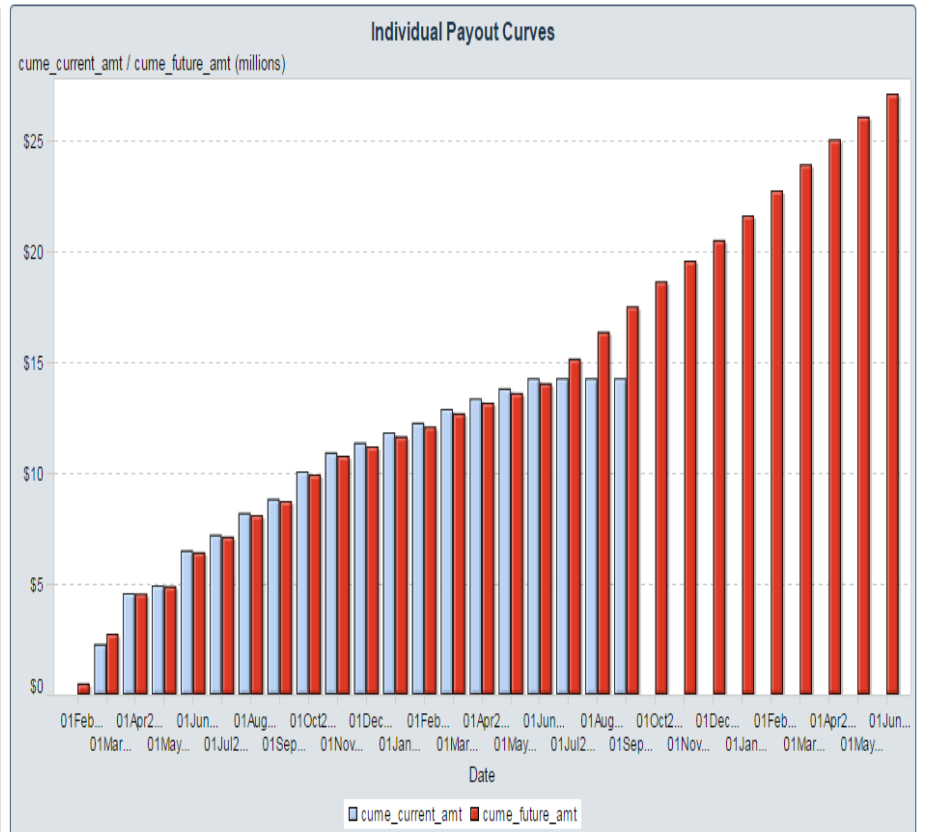
What If Fiscal Year Summary Table

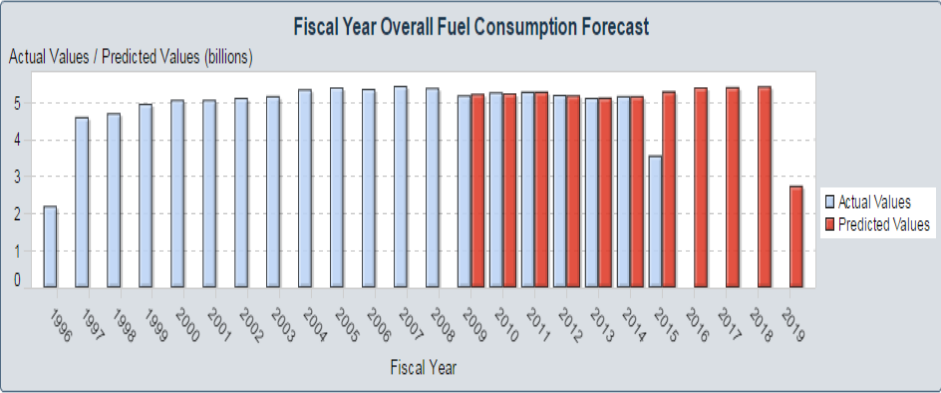
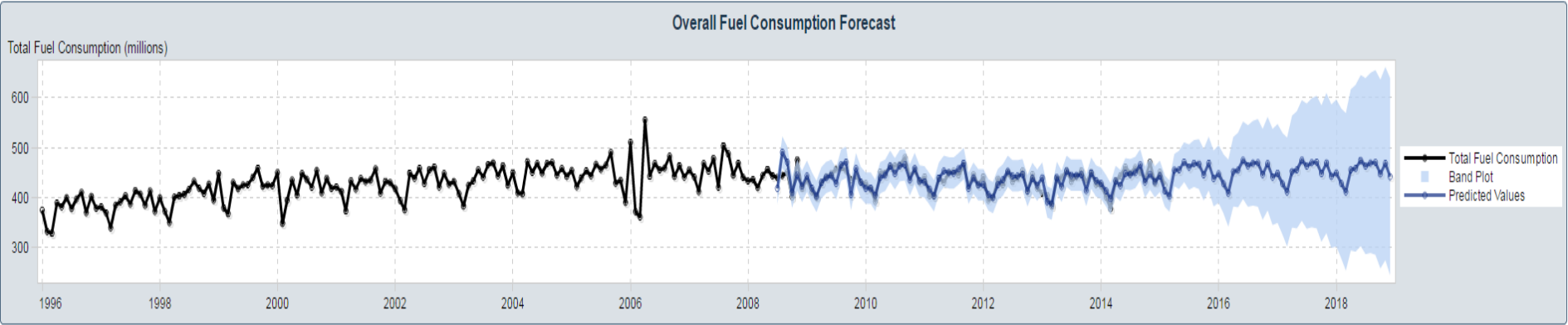
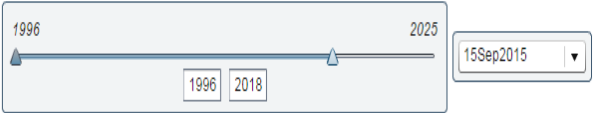
Fiscal Year ▲	Standard Forecast	What If Forecast	Forecast Change	
2005	\$25,042,538	\$25,042,538	\$0	▲
2006	\$177,347,208	\$177,347,208	\$0	
2007	\$249,113,206	\$248,885,280	(\$227,926)	
2008	\$434,840,347	\$433,633,514	(\$1,206,833)	
2009	\$538,315,822	\$538,315,822	\$0	
2010	\$388,622,781	\$388,622,781	\$0	
2011	\$396,666,513	\$396,666,513	\$0	
2012	\$452,188,018	\$452,188,018	\$0	
2013	\$552,453,449	\$552,375,328	(\$78,121)	
2014	\$630,046,815	\$626,659,951	(\$3,386,864)	
2015	\$735,301,191	\$582,179,238	(\$153,121,952)	
2016	\$795,450,885	\$327,934,322	(\$467,516,563)	
2017	\$709,276,071	\$189,763,596	(\$519,512,475)	
2018	\$740,707,300	\$400,070,404	(\$340,636,896)	▼

Upload Date 23Sep2015	Fiscal Year 2005 2016 2029	STIP No B-2500A	STI Category <input type="checkbox"/> _MISSING_	Project Type <input type="checkbox"/> B	Sub Type <input type="checkbox"/> _MISSING_	Let Type <input type="checkbox"/> _MISSING_	Bid Status <input type="checkbox"/> AWARDED	Bid <input type="checkbox"/> F
---------------------------------	--------------------------------------	---------------------------	---	---	---	---	---	--

STIP No	Program Code	Division	County	STI Category	Let Type	Sub Type	Bid Status	Latest Let Date	Latest Cost
B-2500A	B	1	DARE	_MISSING_	_MISSING_	_MISSING_	AWARDED	01Nov2013	\$3,602,235,000

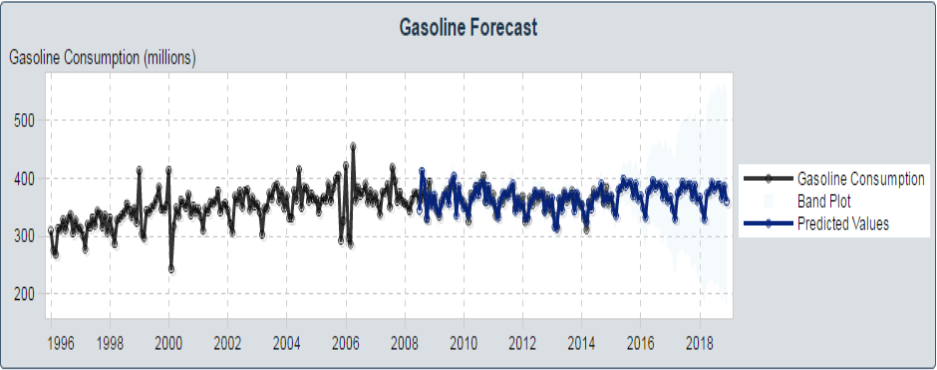
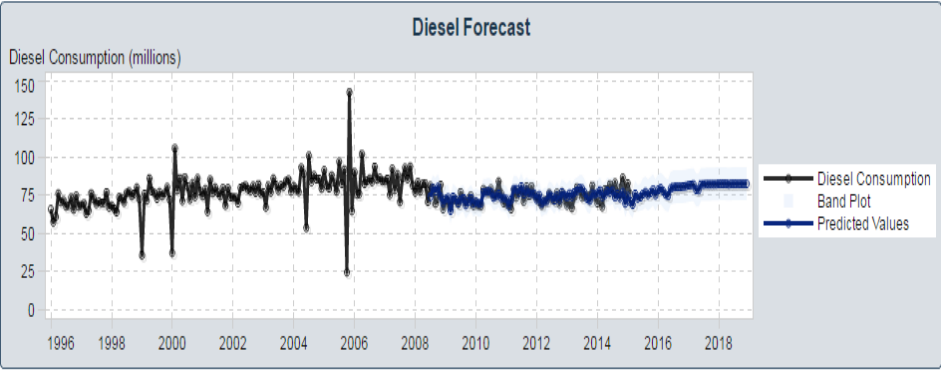
Payout Table					
STIP No	Fiscal Year	Date	Forecast	Actual	
B-2500A	2014	01Feb2014	\$507,485	.	
B-2500A	2014	01Mar2014	\$2,249,785	\$2,284,140	
B-2500A	2014	01Apr2014	\$1,806,231	\$2,294,722	
B-2500A	2014	01May2014	\$343,292	\$355,033	
B-2500A	2014	01Jun2014	\$1,535,747	\$1,591,134	
B-2500A	2015	01Jul2014	\$702,315	\$704,735	
B-2500A	2015	01Aug2014	\$959,096	\$985,302	
B-2500A	2015	01Sep2014	\$645,359	\$637,259	
B-2500A	2015	01Oct2014	\$1,205,671	\$1,236,350	
B-2500A	2015	01Nov2014	\$838,303	\$866,858	
B-2500A	2015	01Dec2014	\$428,803	\$435,000	
B-2500A	2015	01Jan2015	\$453,282	\$460,000	
B-2500A	2015	01Feb2015	\$431,982	\$434,664	
B-2500A	2015	01Mar2015	\$603,353	\$616,824	
B-2500A	2015	01Apr2015	\$478,687	\$488,421	
B-2500A	2015	01May2015	\$439,478	\$445,306	
B-2500A	2015	01Jun2015	\$448,396	\$471,433	
B-2500A	2016	01Jul2015	\$1,108,782	.	
B-2500A	2016	01Aug2015	\$1,209,523	.	
B-2500A	2016	01Sep2015	\$1,142,920	.	
B-2500A	2016	01Oct2015	\$1,137,079	.	





Fiscal Year Summary Table

Fiscal Year	Actual Values	Predicted Values
1996	2,207,642,531	-
1997	4,609,555,837	-
1998	4,713,614,748	-
1999	4,967,575,356	-
2000	5,078,651,771	-
2001	5,073,809,631	-
2002	5,130,405,695	-
2003	5,178,307,319	-
2004	5,366,350,040	-



Using Tableau for Decision

The NCDOT distributed more than **\$147.7 million** in State Street Aid, also known as the Powell Bill fund, to **507 municipalities** across the state per population and mileages.

Tableau used to visualize the initial allocations.

Simple Dashboard helps to identify which city had received the largest allocations from 507 **municipalities**.

\$ 147,728,810.43



Data

Analytics

export (2014 Allocations and B...

Dimensions

Abc F11
 # Grantee ID
 Abc Grantee Name
 Abc Measure Names

Measures

2014 Allocation Amount
 # Agreement
 # Beginning Balance 7/1/14
 # Ending Balance 6/30/15
 # Mileage
 # Mileage Amount
 # Population
 # Population Amount
 -# Number of Records
 # Measure Values

Pages

Columns

Rows

Filters

Marks

Circle



Color



Size



Label

Detail

Tooltip



SUM(Population A..



SUM(Mileage Am..



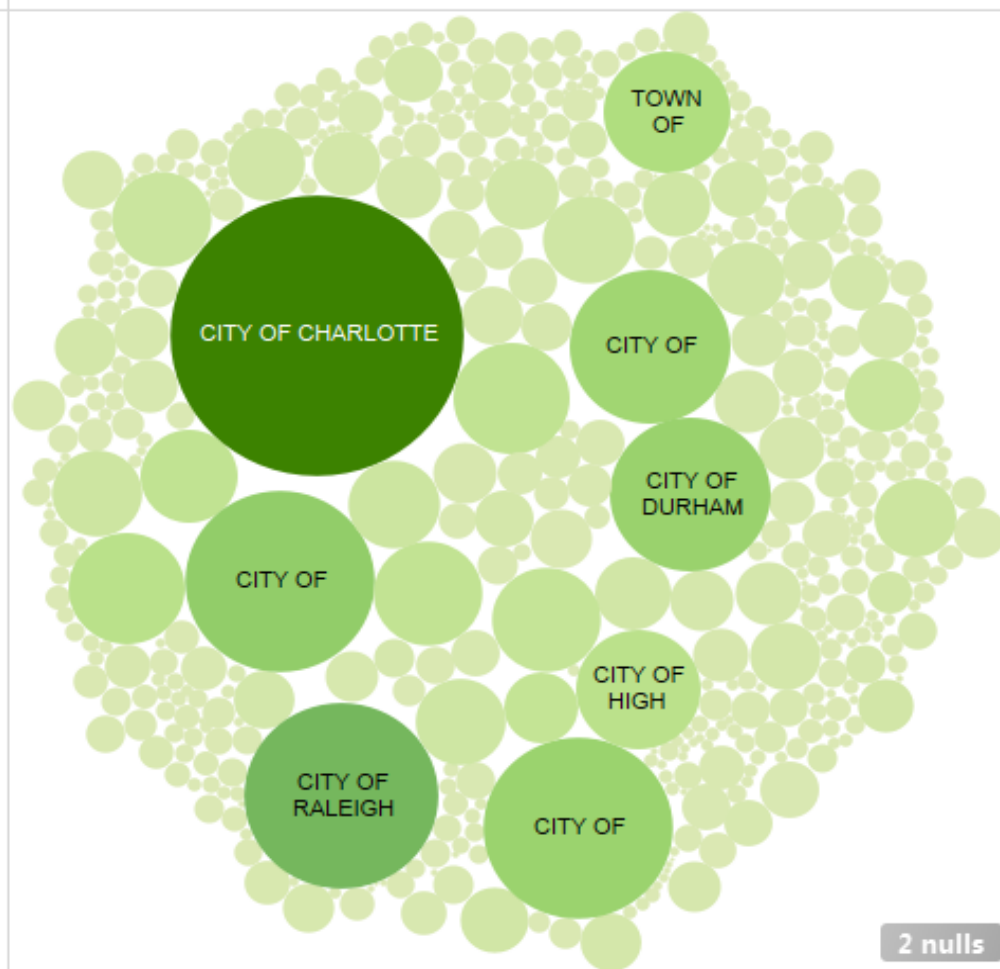
Grantee Name

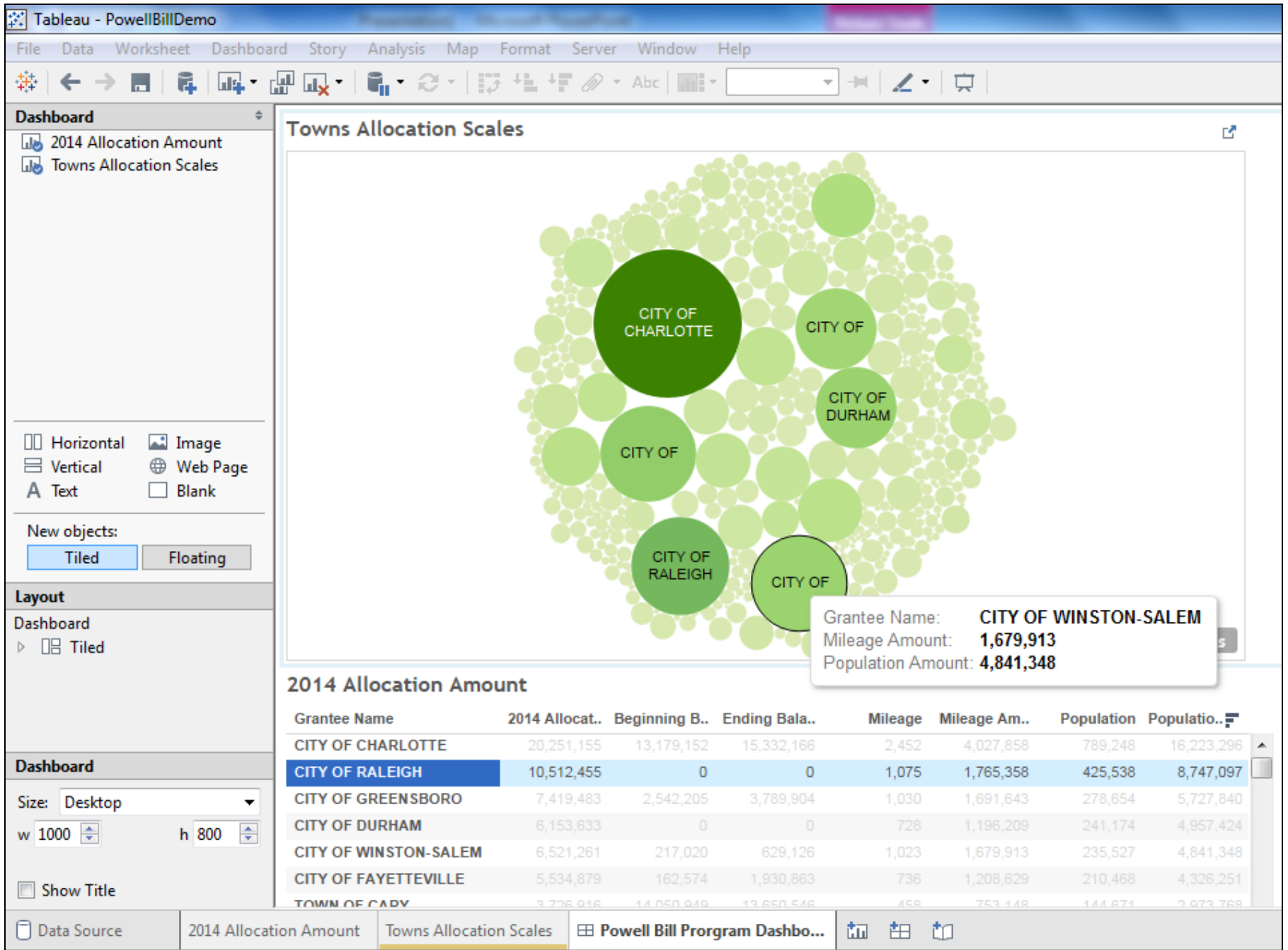
SUM(Population Amount)



411

16,223,296





Conclusion

Using Big Data does come with its own set of challenges, including data security and cluster of networks that turn big data into ‘Smart Information’.



Contact Information & Questions

Dr. Majed Al-Ghandour, PhD, PE, CPM, M.ASCE
Manager

**Division of
Planning and Programming**

malghandour@ncdot.gov

919-707-4620

