Class: 344742

Estimating, CAD, Build Data Roundtrip – The future of Data Analytics for Contractors

Featured Speaker
Greg Davis
Owner - founder
Dolphin Consulting – AnalyticsMEP
Greg@DolphinConsulting.org

www.AnalyticsMEP.com | www.DolphinConsulting.org

Learning Objectives

- Understand What Data can be used for by the MEP Contractor
- Review industry changes that impact existing Estimating and Job Costing Procedures
- Understand at a basic level: descriptive statistics versus inferential statistics.
- Understanding Statistical Correlation and misunderstandings about Statistical Causation
- Understand the Benefits of developing systems to align, store and retrieve data
- Understand at a high level, how to use data to get better insights and make better **decisions** leading to improved margins (Analytics Data – When and Where can you use it).
- Review major sources of misalignment in methods and storage of data between Estimating takeoff (model, revit families, fabrication items, other 3rd Party Estimating Systems), estimating reporting, CAD Model to Fabrication data and Job Cost Data.
- Review and discuss cost code requirements and processes in your company
- Review requirements to align data between Estimating, CAD and actual reported costs (Final cost codes).
- Review Autodesk Fabrication Data Fields and Alignment potential
- Discuss forms of data, its usefulness and reporting methods

Class Description

How do contractors adapt their estimating and project management processes to the new world of increased modular manufacturing and prefabrication? Because contractors want to take advantage of CAD model data and manage projects with more precise estimates, we need methods to better align data to process. Discover the next generation of Data Analytics – hidden in your data. This class reviews the power and benefits of using model-based estimates to predict and create future estimates more accurately and manage projects more concisely. Explore the challenges and obstacles to aligning field provided data, including cost code structures and automation of cost code creation, along with data alignment and the requirements to implement estimating of fabrication ITMs and families. Examine how to practically proceed in building process around tracking data, collecting then organizing data to use in the next generation of tools for predictive estimate values and project sales tools like CRM.

Speaker: Greg Davis

With a background as an engineer and Project Manager for an MEP contractor, Greg is passionate about revolutionary technology in the construction industry. As co-founder of TSI (Sysque/ Building Data), Greg helped drive the industry to adopt BIM and integration. Consulting with clients today, Greg helps companies implement integrated CAD-Estimating-Manufacturing processes (using Autodesk Fabrication parts, Families and other components) and is developing the next generation software to make a long sought connection between the data in CAD models and Estimating systems. Because manufacturing processes are changing, the old models for estimating accurately and managing projects will be passed by with new modern data analytics-driven analysis based methods making use of all real installation data as a real-time feedback loop.

Understand What Data can be used for by the MEP Contractor

Big Data Analytics is growing widely in use across many industries. This data offers incredible potential to better understand and run your contracting business.

In order to understand the potential use of data, first look at the sources of data available and the implications of what data is readily available and what is not. The Construction industry, and specifically subcontractors have long been what is called a “fragmented” industry, with few companies with a national presence and large market shares. Instead, contractors have often been privately owned companies that have represented a lot of opportunities for dedicated and entrepreneurial people to grow. In the last 20-30 years, large suppliers have consolidated within the industry to gain a less fragmented stake, such as Ferguson in piping and plumbing supply.

Over the last 5-10 years we are now starting to see consolidation within the contractors themselves, as well as information sharing within multi-company consolidators. This has been driven by efficiencies, generally from software-based technology in administrative tasks, project management, and fabrication (CAD-Fab/ CAM) with pre-fabrication.

So, in general, we have a relatively short history of information sharing potential.

Sources of Information

One of the challenges faced in gathering industry data, is that contractors are competitive and private, and their transactional information is not shared. Contractors, in general track, Material Costs for items, Labor Hours and Costs for Fabrication and Installation of items, other project costs, Schedules for projects and successful adherence to schedules. Often contractors track, loosely, the productivity of different crews or foreman versus others, and incentivize productivity.

Although much emphasis is placed on tracking cost data on projects, a higher level of accuracy is potentially available with improved data which requires using information from Both the original Estimate AND the CAD Model (and other takeoffs if necessary) inclusive of all fabricated and installed items, along with the resultant job costs. Connecting the data fully eludes contractors because of the challenges in aligning all of the data and reporting it fully.

Instead, for some contractors, which represents a successful first step, they are running estimates on their CAD models and comparing Estimates, Models and Job costs manually. This leads to an improved intuitive understanding and certainly better decisions without learning from this type of review.

The next step represents the use of Data Analytics.

Uses for Data

The exciting potential for Data Analytics for MEP Contractors is improved decision making in major aspects of the contracting business. This is becoming critical as industry veterans retire and competition is driving more projects, more work with less management time, less veteran skills and less labor. Our industry is evolving to a real manufacturing operation, and the industry has done a great job in leveraging CAD models and Autodesk Fabrication software to increase pre-fabrication levels.

Evaluating decisions in this new environment will be required as we attempt to understand how far we can take combined skids and result in real savings as the cost of skid frames and combining objects offsite becomes more complex. Additionally, we will continue to need to understand and make decisions about larger and larger investments in manufacturing machinery (and eventually Robots) and the real ROI versus other approaches.

Competitive pressures also drive us to make hiring and promotion decisions to get the best people and put them in the best situations for success.

Traditionally these decisions have relied on inferential decision making requiring a lot of experience. In reality, without more broad mathematical feedback, misunderstandings can be reinforced leading to less than optimal solutions.

Decision-making using Analytics is rapidly growing in many areas, even where traditional experience is highly relied upon and highly valued. At a minimum, analytics provides confirmation of decisions and supplements knowledge of leaders that are very experienced. Moving further many industries, businesses and forms of entertainment are rapidly changing based on new information uncovered through analytics.

One example outside of our industry (which is at its infancy in analytics) that many are familiar with is in sports. For those familiar with the book, movie and story of “Moneyball”, Billy Beane revolutionized Major League Baseball by introducing Analytics in the game. His work has inspired Analytics implementation in other areas, including both other professional sports and industries in general. For those that follow sports, Analytics is driving decisions not only on player selection (HR), but also on Play Selection in football and shot selection in basketball (See the change in NBA 3 Point shots over the last few years). And it doesn’t stop there, many industries are seeing an emphasis on using analytics, including in the construction industry, and specifically in MEP, where suppliers are investing in solutions to mine data to order and stock material more accurately.

In fact, emphasis and interest in MEP construction Analytics seems to be reaching a crescendo, as Billy will be the Keynote speaker at Construct and Connect during Autodesk University this year.

Let’s look at a high level of what Billy accomplished, and specifically what questions he tried to answer with Analytics, along with what information he had to start from.

Billy worked to improve the accuracy of selecting players (hiring) and predicting those that would be successful. He was able to do so because from their college and minor league playing days, players (potential hires) had numerous statistics on their performance available in **public**.

Billy went on to evaluate statistics like on-base percentage that were ignored by other teams and found players that were under-valued in the marketplace.

Fast forward to thinking about MEP Contracting. Would you like to know which potential employees you should hire for your labor force? Would it benefit you to better understand which employees you currently have are performing the best, which are performing well, and which are underperforming?

What can we measure? … What can we not?

Well, one challenge is that there are no public figures on potential hires. And the data that you use now to evaluate current staff could be better, and in fact may not be fair at all.

Fabrication Shop Labor Team Members and Installation Team Members are often evaluated today based upon an estimate that does not exactly match the model and items actually fabricated or installed. To be sure, it is close. It is better than nothing. But CAD models can vary significantly from Estimated Models. So the data we need to fairly evaluate people and make those hiring and promotion decisions can improve greatly.

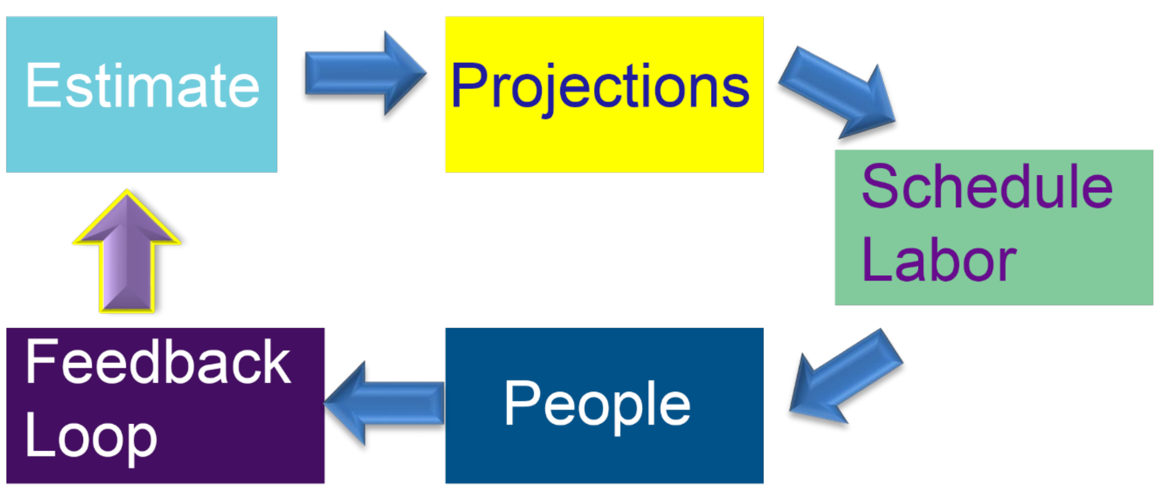
Construction Labor Team Members are more likely to respond well to fair, data-driven evaluation, than inferred guesses which include biases and favoritism that don't benefit the company as a whole.

Additionally, estimators and Project Managers, need a fairer evaluation as well. And to improve they need a feedback loop.

The good news is that the information in that Feeback loop, the **Estimating-CAD-Build Data Roundtrip**, can be very valuable to your company, allowing you to estimate better and make better projections to manage projects better.

Further, with the improved feedback loop data, *improved personnel decisions* are just a step away.

WHAT CAN YOU USE DATA FOR – MAJOR CATEGORIES AND THE FEEDBACK LOOP



Industry Changes Driving Move to analytics

Contractors often lament the challenges of finding good people. As we know, good people are key to performance in a labor-driven industry. Interest in construction careers, however are not matching supply. At the same time, because of efficiencies found in consolidation and adoption of technology, the competitive pressure aren't making competing in the MEP contracting space easier. And, perhaps as always, cash-flow management is a critical component to maintaining a healthy MEP Contracting business.

All of these key challenges require solutions that can benefit from Analytics.

Finding Good People

MEP-Driven Analytics may be a few steps away from providing insights for most contractors. It will take some time for data to be gathered with an analytics solution outcome to be realized, and in that process, the feedback loop on people, or specifically performance may take the longest.

There are other non-MEP specific industry solutions in HR for analyzing talent.

Attracting Talent

The modern workforce thinks differently than they did in the past. Culture, Lifestyle, Freedom, Transparency, and the Nature of the work Experience are highly valued today. Developing a system of measurement and that is consistent is one tool for transparency and providing a sense of both empowerment and control that will attract employees to your company. Freedom, or flexibility in work environment and schedule may be more possible in Virtual Construction and as Virtual Construction takes on a higher share of the work load. In general, even those in Fabrication and Field positions today are attracted to implementation of technology and new systems as a compelling part of the role and job.

Grooming Good People

Because the supply of experienced staff is limited and dwindling, an analytics-based system will help you make better decisions about staff development, including who can be coached in certain areas to resolve problems and who may not be a good fit for the role they are in.

Better Financial Performance

Cash Flow Management and financial performance are the key to your ability to sustain your organization as you adopt and evolve. Improving estimates and job performance are the key areas that Analytics can help address for MEP Contractors…And within the process of developing the EST-CAD-Build Roundtrip and Establishing an Analytics-driven system, it is a result that comes in early, while evolving and getting more refined as well in time.

Descriptive Versus Inferential Statistics

When we develop data sets, we can find out different information based on the nature of the data relative to the nature of all the possible data. Although, perhaps we never are really collecting all of the data, we use the Descriptive Statistics approach when we have a set of data that we want to look at and of that set determine the total items in the set, the average value of the items, and how “dispersed” the data is, meaning how different is one result from another.

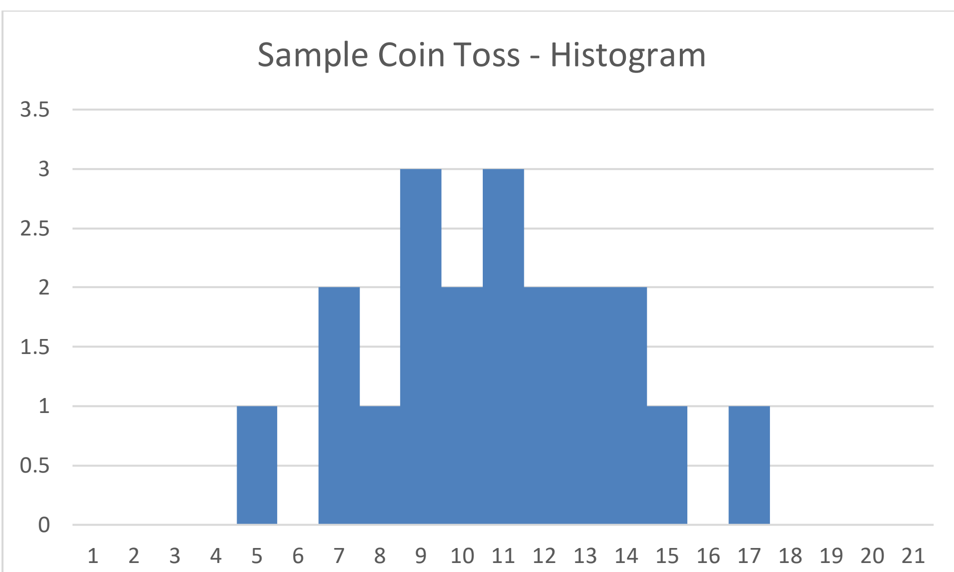
Inferential Statistics starts from an understanding that we have only sampled partial data, and we draw conclusions from that limited sample. Historical knowledge by industry experts, leaders, senior project managers, etc., though often vast, typically falls in this category. Based on experience, and a lot of it, people can infer information about data they have seen. However, even inferential statistics requires more complete data collection than we find in the industry today, in order to make use of the data outside of people’s intuitive sense (which requires a vast supply of very experienced staff to make all decisions).

Descriptive Statistics

People are often first introduced to statistics through a descriptive statistical example. Imagine a toss of a coin. Let’s assume the coin toss analytically should in fact land on heads 50% of the time. The larger the data set, the closer our average will in fact get to 50%.

In this coin toss experiment, if we toss the coin 20 times, and count the number of times that it landed on heads, we will one data point (in this case the average should eventually be 10).

In practice we may get results like 8,11,6,9,16,13,4,10,12,7,13,8,6,10,9,12,14,11,10,8 if we toss the coin 20 times per round and perform 20 rounds.

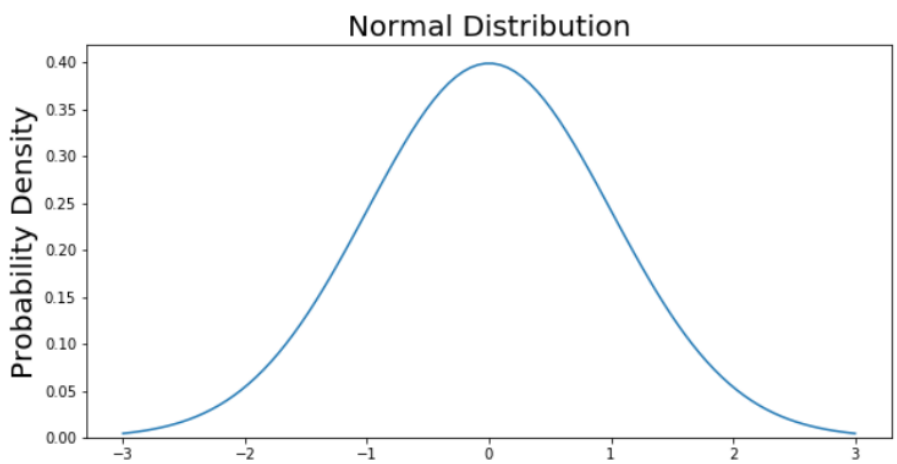


Sample Results in Histogram Form

The total out of 20 rounds would be 197, with an average of 9.85, which is very close to perfect information. But, how risky is it to assume you will get 10, or even a number within some range of 10. Las Vegas is built on similar games that explore (and understand on behalf of the house) these risks.

Normal Distribution

You may recall learning about normal distributions in statistics in school. They look like the this:



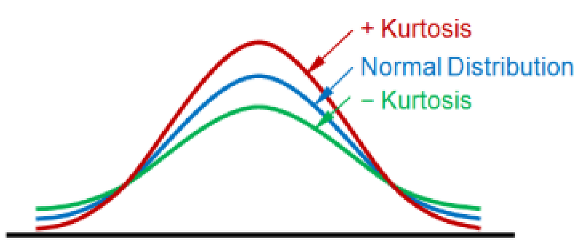
Normal Distribution Curve

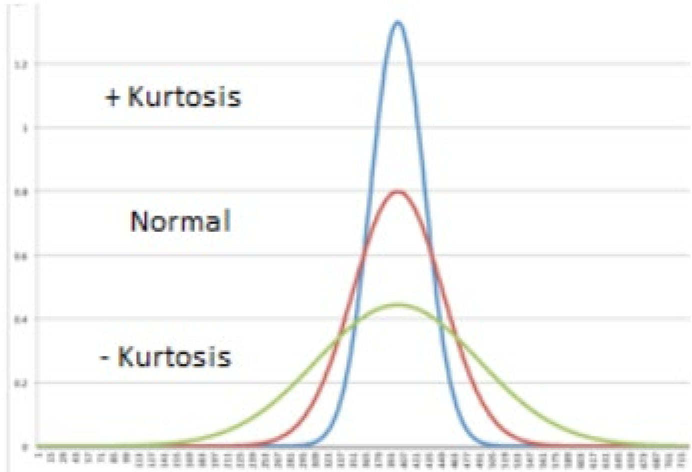
A normal distribution is what you should get (eventually with enough rounds) from the coin toss experiment, or any perfect 50/50 odds split.

Kurtosis

Although it may not look like it, our coin toss example is very much like a normal distribution, as it should be. First, our average is in the middle (in this case near 10 at 9.85) A distribution can be low "risk", and reliable when it follows the curve in expected ways:

Kurtosis – peaked-ness – measures variability of data (often representing risk)





Kurtosis for a normal distribution is 3, and higher for + Kurtosis

The Kurtosis of a Normal distribution is 3, but often, like in the excel formula can be expressed as “excess Kurtosis” which starts at 0 for a normal distribution. In our data set example, the excess Kurtosis was -0.237146432, which is near zero, as expected. So, again it may not look it to the naked eye, but the example of our random experiment is following a normal distribution as expected and if we predict results based on reliable processes like this, we will have good results. It is important to understand from a statistical standpoint, that if you estimate jobs or look at data, and are alarmed by the variation (without running numbers, aka without Analytics) you might over-react to under or over-performance. This example was normal, but who would let it ride when they got a 4 out of 20, expecting 10 to be the average, or when the result was 16. Descriptive statistics helps you uncover a lot of hidden answers that you wouldn’t intuitively expect.

Skewed Distribution

A distribution can be low “risk”, and not look normal. Our process by design could be to have something that hits 75% of the time. This would represent a “skewed” distribution, which could otherwise be normal, with zero excess Kurtosis.

Inferential Statistics

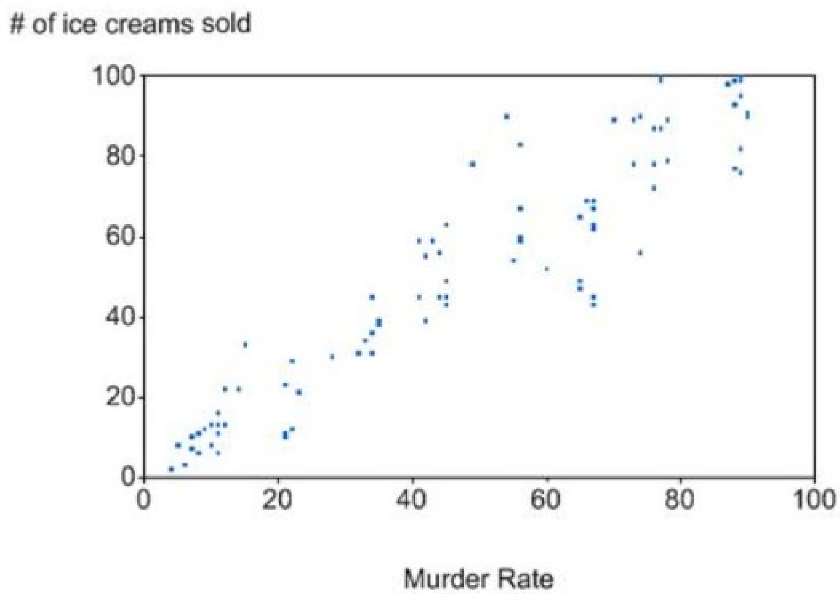
When discussing the practical implementation of Analytics systems with contractors, concerns arise about the use of limited data sets. It is true that the first data results built around an analytics system could be less accurate than the collective wisdom of your senior people. The real question is, at what point do you have enough data to infer something.

Regression and Correlation analysis are tools used to infer information from data samples, and to provide some context to there usefulness. The regression part is a fancy way to say “Do all the math.” The correlation part is the result you are aiming for. And correlation is often misunderstood for causation.

Correlation

A correlation means, that when there is a lot of one thing going on in the data, there is a lot of another thing going on. Sometimes that is enough information to more accurately infer a result (predict) than what we might pull out of our heads.

One cautionary example cited in statistical explanation is the example of rates of murder and rates of ice cream. For example, it has been found that murder rates increase when ice cream sales increase. It would be wrong to conclude that eating ice cream makes you into a murder, though.



Map of Ice Cream Sold against Murder Rate in plot points – the basis for correlation analysis

Instead, this correlation appears useless. Well, if it was all the data you had, it might actually not be. If you owned an ice cream truck, you could monitor murder stories or other murder data, and stock more ice cream as those rates increased. You wouldn't necessarily be wrong, even if your reasoning is wrong.
If that was all, the data you had, it wouldn't be possible to do further study to see why the ice cream sales went up at the same time murder rates did. You could just sell more ice cream and move on, hopefully not too traumatized by the hypothetical, which in no way is meant to be insensitive, but it meant to illustrate the concepts. Further study might reveal that Ice Cream sales went up when the days were hotter, or perhaps longer, and murder rates might have appeared to follow that trend for those reasons. Isolating those factors however, eventually would make you more accurate to focus on the factors that were better "Causes", like what could be the case of hotter days.

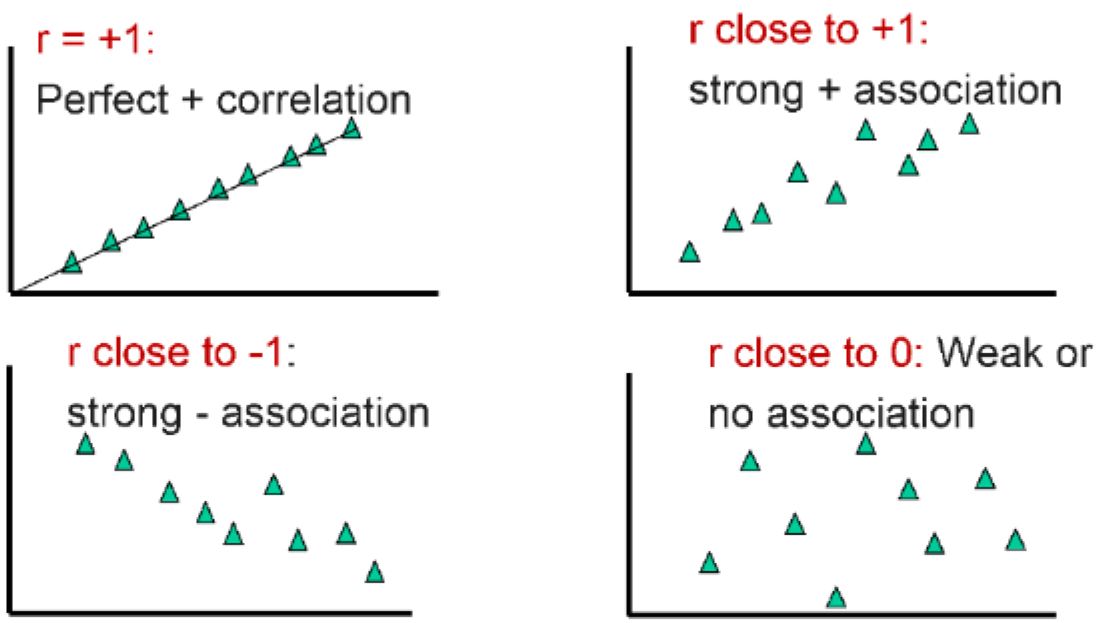
Examples in MEP of possible correlations between one factor and another that you probably use in one form or the other today include:

- More labor when more Feet of pipe.
- More labor when higher percentage of Welded versus copper.
- More labor when more fittings versus straights of duct.
- More labor when more branches or coil connections versus mains.
- More labor when more high floor versus low floor,
- More labor when more mechanical rooms versus general areas.

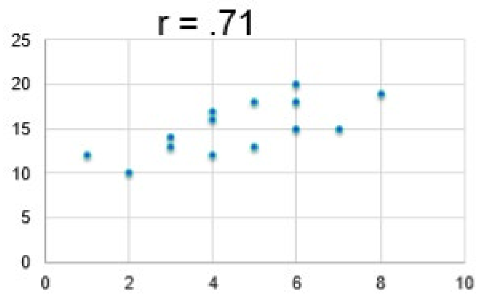
Causation

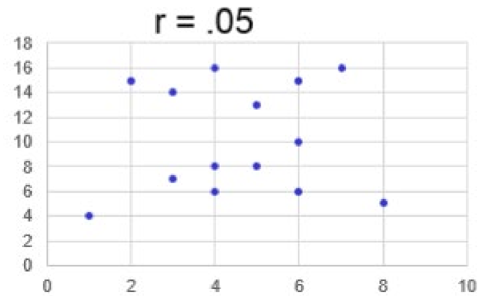
To truly determine, through statistics, that one factor causes another, you have to run experiments that isolate factors, and have many samples of each option compared. If for example, usually sales were good on hot days, that doesn't full make it a "cause" of ice cream sales to go up, statistically, unless when isolating and holding all the other factors considered, the sales are always better with Hot weather than not. Maybe a large concert in the area happens during warm weather usually, but in fact the concert (or several different dates of concerts throughout the year) are the real cause.

For those not experienced in Analytics, the tendency is to draw the wrong conclusions be assuming causation, or to discount the value of the correlations that they don't agree with (or weren't their idea). The key take-away, is that analytics is value agnostic, and people agnostic. And the results of the statistics give you measurements. In statistics, the Correlation Coefficient (r) describe how correlated one thing is to another in your data.



Scenarios for r – value in Correlation





The left graph shows the when the x-axis goes up, the y-axis goes up – and indicates the value of the correlation. On the right, the x and y values don't relate and have a low correlation (r) value for r – value

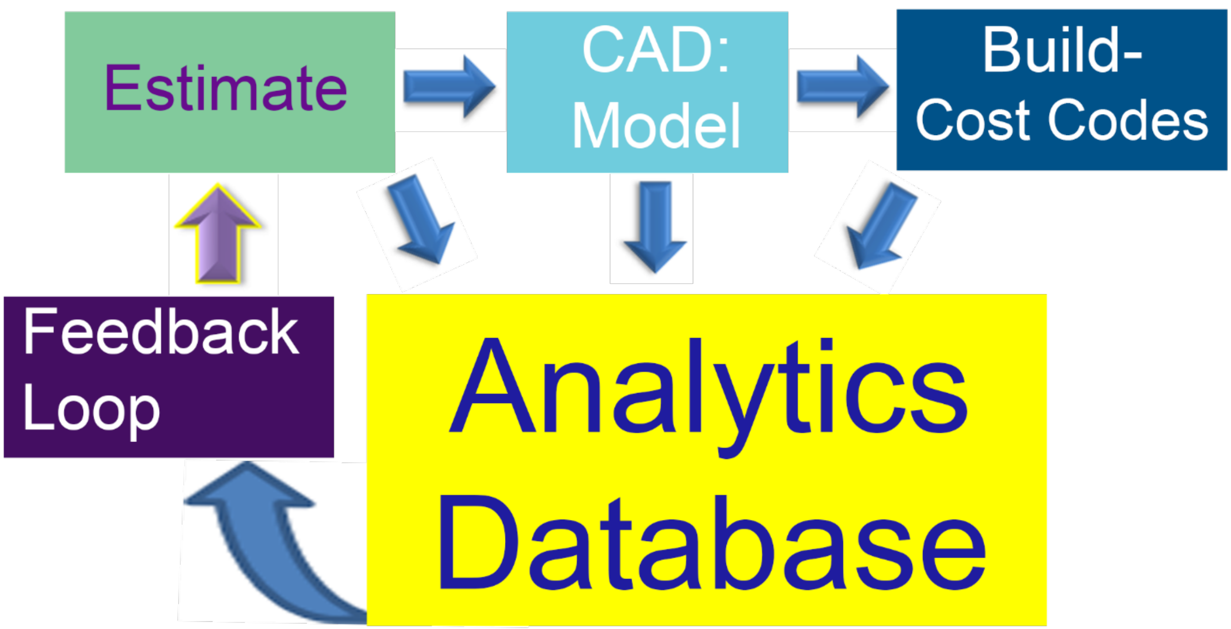
Misalignment of Data sets in MEP Contractor Processes

Data inputs that would be useful and necessary for analytics in contracting come from many sources, which presents two challenges. One, systems need to be built to automate the connection from one information source to the other to bring them all together. And Two, the data is stored in different forms, with different fields, different groupings of fields, different data structure of those fields, and different units.

This data will need to be aligned to connect together useful analytics data. The feedback loop requires three Major Parts (data sources or Data Source Stages), and one database to collect the data, report the data and through analytics calculation provide predictive modelling within the other systems.

EST-CAD-Build Data Loop

This data will need to be aligned to connect together useful analytics data. The feedback loop requires three Major Parts (data sources or Data Source Stages), and one database to collect the data, report the data and through analytics calculation provide predictive modelling within the other systems.



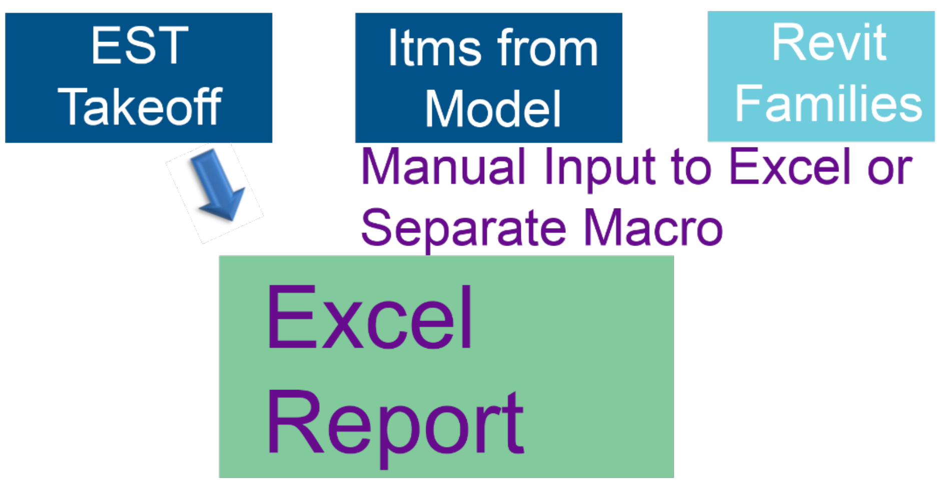
Analytics MEP – EST-CAD-Build Round Trip Elements

Estimate Information and Disconnected Sources

The future of the MEP Industry and Construction industry in general is to extend the use of Revit and extend out the connections to that model, or to extend the abilities of Addons to use the Revit Model and its data. That being said, it should be noted that some Estimates are still done:

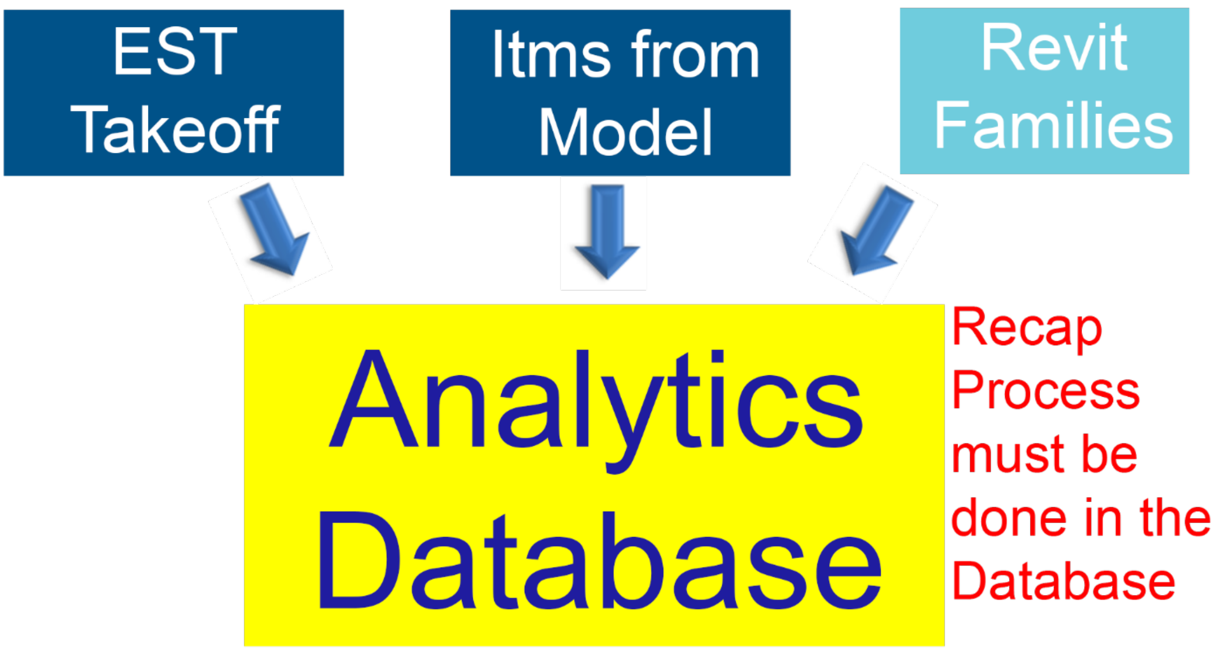
- via pdf takeoff using Fabrication itms (ESTmep)
- via QuickPen, Quote Software and other tools that don't connect to fabrication items
- via model-driven Service Takeoff, in which all items in the generic engineered model are re-drawn with a combination of Revit Families and Fabrication items (.itms).
- Via conversion of the model using design to fabrication and some combination of the Fabrication parts or Revit Families.

Just a brief review of these options clearly shows that the data comes in different forms. Even in the most advanced implementation scenario, the full conversion of the model, still leaves us with some Fabrication Items, and Some Revit Families (Grilles, Flex, Pumps, Fixtures, Equipment, etc.). On top of that, most contractors then leave the project model environment or Estimating program's environment, and complete the remainder of their bid adjustments in excel.



Estimating Data Sources only semi-connect. Once in Excel, Connection to the Model & the Data is lost.

By Connecting These Sources, we enable the ability to align to an Analytics Approach. And in order to connect and line up data from different sources, you need a database program that has routines to convert the data in one form to match the other. You also need to make all adjustments to the bid for final recap in an interfaces that updates all of the items in takes to store their original Default Material and Labor Costs (usually MCAA and SMACNA for Labor) as well as their Adjust Material and Labor Cost values (hours and Dollars).



New Model does all storage of data and adjustments connected to a the AnalyticsMEP Database

CAD Model Information and Disconnected Sources

Like Estimating, we do have models in many forms. Consistent with most contractors implementations, we will limit our discussion to the use of Fabrication items and Revit Families, and not other CAD Systems, which have limited market share.

Because of the high level of adoption of Fabrication Items, the alignment process is generally more simply at a basic level. However, the requirement to align Revit Families for those items in models that aren't available as items, means that this new data model needs to support both. And by doing so, it makes possible the use of other systems, either from third party Revit Drawing add-ins or systems built in Revit Families in general.

We will discuss alignment issues that remain after Families and Items are brought into a single database, based on data implementation inconsistencies that are common to contractors and different process requirements in the build stage, as we continue.

Build Phase and Cost Codes – another source of disconnect

On the surface, cost codes and their data provide little complexity to the EST-CAD-Build Model. But their simplicity creates the normal alignment issue and common disconnects.

Cost Codes are built with different requirements than the Estimating Process, though we tend to think they are the same. When you estimate, you are effectively counting up a bunch of stuff. For example, if you do a takeoff from an equipment schedule which is not drawn within the model, you may not need to, or find it to be a cost effective use of your time to indicate in the takeoff, the floor, section, service, etc. that each and every equipment item will belong to. Certainly, for major equipment, we can expect the field to be able to reasonably mark or enter their time for those types of equipment. But that doesn't apply to all equipment and accessory items. The person in the field usually only can easily differentiate between time spent within an area, floor, or system, not necessarily on the individual parts. This disconnect can go in both directions, as the takeoff person may provide area details on items, and once prefabricated as spools field installation may be too quick to separate by the area or section.

In order to align data, we need to align cost codes back to the items. To do this requires a completely consistent approach to cost codes, which in practice means an automated cost code creation system. That way you can apply, as well as possible, labor in the field back (through a database interface) to the items themselves in the takeoffs.

Fabrication Item Alignment Challenges

To make analytics a success, you need to resolve fabrication alignment issues. These issues come down to two categories:

- Item Setup – Database Issues
- Alignment to Cost Code and Estimating Process Issues

Item Setup and Database Issues

The Fabrication Database, which on the one hand was designed for maximum flexibility, is made up of item files that are not connected to one another. The database is not a fully relational database. A relational database when implemented well, indexes fields, and doesn't store all of the data on each item. So, for example, if you change the name of a material in the database, then all items which had that material will know their new "Material Name". This can be broken in Fabrication very easily with materials, material abbreviations, connectors, service names, service abbreviations, etc. Other fields have ambiguous meanings and have been implemented in multiple ways, such as specification. For duct, it may be clear that the specification drives all of the parameters of the ductwork pieces. For pipe, while some information can be driven by the spec, other information can be driven by the material, and there isn't a clear reason to have multiple entries for specification. However, because there are opportunities to do so, some contractors use this fields effectively as an additional reporting category.

Alignment to Cost Code and Estimating Processes

Because MEP Contractors have different setups in the shop and the field, the cost codes and groupings (correctly) tend to make the given process.
While the goal of the estimate is to count all the stuff in the form that it comes in through, the breakdown in the pre-fabrication shop is done based on the types of things that use the same processes. These can include welding stations, Copper Fabrication stations, Hanger Fabrication stations, Coil Line Duct fabrication stations, Spool assembly stations, etc.

In the field, often these disconnected items (via spools) are brought back together, like Welded Carbon Steel spools for one section joined with Copper for that section, while Welded and Copper for another section are staged (and of course installed) somewhere else.

The Fabrication Database doesn't always support the breakdowns actually used by contractors.
For example:

- Contractors may takeoff (estimate) and group for install cost codes, Low Pressure Duct and Medium Pressure Duct (regardless of the specific service or specific Low or Medium Pressure number 2",3"4" wg, etc.)
- Contractors may break up all fittings and straights in the Fab Shop, but not strictly group Low Pressure and Medium Pressure items at fabrication time.

- Contractors may wish to group all items that are "Copper"-like, such as copper, brass and bronze. While they do want to report the specific material at times, and in their material grouping, they want to group for pre-fab and install all of these together somehow.
- Also, they may wish to group these together by material and manufacturer in purchasing and then split up the parts into different groups in the fabrication shop based on the parts that will be assembled together.

So how do we ever align the data to store and retrieve the information which has different categories for different purposes? First, you need to establish different cost codes for Material, Fab and Field. Next, you need routines that can find all of these different things and know that they belong together. The answer lies in creating data tables within a database that relate the object together. For example, if Copper, Bronze and Brass go to a Copper Material Group, their entry in a Material Group table might be "Copper" or "Copper Alloys". By using a relational database tool, you can then group these parts together.

Likewise, you could use a table to list services and the Pressure Group that they belong to (LP or MP)

Cost Code Automation

In order to make the EST-CAD-Build Feedback loop work, you need to align data from Estimate and Models (Items, Families and possibly other takeoffs) to actual Cost Code data applied by the field, shop and job administration (be that via a foreman, project manager, etc. for material).

And to make alignment of cost codes work, you need a system that is consistent and automated. Much work has been done with real contractors on this process to this point, and the first concern is often a fear of a lack of flexibility to meet the requirements of different jobs when cost code automation is used.

And we certainly know that we receive feedback in different ways at different times. To make the automation meet the needs of different jobs is not an easy task. However, we can do it with a series of routines that allow seletion of a cost code breakdown scheme desired, or simply a set of rules that selects the breakdown level based on the rules.

For example, if you want to end up with at least 200 hours in a given field cost code, and no more than 2000 hours, the automation routine can loop through to determine which break downs to use.

Possible candidates for breakdown include things like, Pressure Group, Material Group, Service, Material, Section, Service Type.

Each company has their own rules for how things are group together, like service type. Also, each company has some for of unique need for prioritizing these types of fields, and an automation routine can process this.

Cost Code Routine Example

Using our example, with a goal of no more than 2000 hours in a given cost code (and for simplicity only discussing a few fields), let say that our Mechanical Piping (Pipe Valves and Fittings) had a total of 8,600 hours. Clearly being over 2,000, it needs to be broken down. Let's say in our company, our first breakdown priority for Mechanical Piping is "Material Group". And to simplify, lets say we have Copper, Welded Steel, Grooved Ductile Iron as our categories. Let's say we have 1,800 hours of Copper, 1,800 hours of Grooved, and the remaining 5,000 hours in the welded steel. Now lets say we use section as our next priority breakdown field. Let's say we have L1, L2, L3, and L1M (for L1-Mechanical Room). In those sections, we will say we have 1,700; 1,300; 800; and 1,200 hours in those sections. We would now have the following breakdowns:

Material Group	Section	Hours
Copper		1,800
Ductile Iron (Grv)		1,800
CS Welded	L1M	1,200
CS Welded	L1	1,700
CS Welded	L2	1,300
CS Welded	L3	800

We might use one of many forms of cost code name. Lets say our group is 30 for Pipe Field. Our codes might be:

Material Group	**Section**	**Hours**	**FieldCostCode**
Copper		1,800	30.1
Ductile Iron (Grv)		1,800	30.2
CS Welded	L1M	1,200	30.3
CS Welded	L1	1,700	30.4
CS Welded	L2	1,300	30.5
CS Welded	L3	800	30.6

Or we may include some designations that aren't numeric that tie to our breakdown fields, like:

Material Group	**Section**	**Hours**	**FieldCostCode**
Copper		1,800	30.C
Ductile Iron (Grv)		1,800	30.D
CS Welded	L1M	1,200	30.W.L1M
CS Welded	L1	1,700	30.W.L1
CS Welded	L2	1,300	30.W.L2
CS Welded	L3	800	30.W.L3

Analytics Database and Alignment of EST-CAD-Build

We can illustrate how the data connects in practice through a simplified example. Remember the goal is to have information that helps MAKE a DECISION at the Place and Time of the DECISION.

We first start by recognizing that, although you may have a Project Management System containing Job Records, or Job Cost Accounting System, you likely are not storing in 1 place all of the Bids in a database. If you are, we want to look at Bid information (Project Information) that will be useful in comparing types of jobs. This can include the Project Type (Hospital or Data Center, etc), or the General Contractor, Engineer, etc.

Next we recognize that your estimate needs to allow for more than one takeoff. Some companies will do multiple takeoffs in ESTmep, batch them together and run a report. Others need tools to group them together. Regardless, as we discussed, the future requires at a minimum a takeoff sources for families and items (or other takeoff program data). So a job can have multiple takeoffs.

Next at the item level, all of the items are stored, regardless of which takeoff they apply to, and all of their item data. Each of the items, from the onset of the original Bid Recap, need to automatically have their cost codes applied, and to store these cost codes (unique for material, fab and field).

Analytics Data and Process

Data Tables

Bids

BidID	Job Type
1001	Hospital
1002	Office
1003	Hospital
1004	Retail

Takeoffs

TakeoffID	Takeoff Type	BidID
1	EST	1001
2	EST	1001
3	CAD	1001
4	EST	1002
5	BUILD	1001

Items

Item	Weight	Field Hrs	Adj.Field Hrs	Cost Code	Line Item Code	TakeoffID
Elbow	55	1.2	1.0	20	11.1	1
Straight	65	0.4	0.4	20	11.1	2
Elbow	50	1.0	0.916	20	11.1	3
Elbow	50	1.0	.0727	20	11.1	5

Basic Data Tables – EST-CAD-Build Roundtrip

When we bring in the takeoffs from the estimate and use the database itself to create and store markups and adjustments to complete our final bid, we now have information that potentially can align to the next phases. Likewise, we bring in the CAD model in a similar fashion, and the Cost Code data from our Job Cost Accounting System. The result, when the data is lined up properly, is the ability to see in real time the actual Factor that if applied to takeoff items would have collectively matched our result on our jobs.

Analytics Data and Process

Data Flow

Items	Item	Weight	Field Hrs	Adj.Field Hrs	Cost Code	Line Item Code	TakeoffID
EST	Elbow	16.5	3.6	3.0	20	11.1	1
	Straight	19.5	1.2	1	20	11.1	1
	Straight	19.5	1.2	1.2	20	11.2	1
	Straight	19.5	1.2	1.2	20	11.3	2
Totals	**(20)**	**75.0**	**7.2**	**6.4**	**Factor CAD by EST Adj Field Hrs / Field Hrs**		
CAD	Elbow	**15.0**	3.3	**2.75**	20	**11.1**	3
	Straight	19.5	1.2	**1.0**	20	**11.1**	3
	Straight	19.5	1.2	1.2	20	11.2	3
	Straight	19.5	1.2	1.2	20	11.3	3
Totals	**(20)**	**73.5**	**6.9**	**6.15**			

BUILD

Original EST Field Hrs (Job Cost) 7.2 — **Final Adj Field Hrs Build / Field Hrs EST = 4.92 / 7.2 = 0.68333**

Updated Forecast to Completion (Job Cost) 6.15 — **CAD-Build Field Hrs / EST Field Hrs = 6.9/7.2 = 0.958**

Actual Job Costs 4.92 — **Final Adj (Actual) Field Hrs Build / CAD-Build Field Hrs = 4.92 / 6.9**

Factor to Apply to Items = .713

Data Alignment – EST-CAD-Build Data – True Predictive Factor Created

From the base factor, additional factors can be mathematically determined. Keeping our discussion of statistics in mind, there are many possibilities for how this factor would be calculated, factored and applied. In this simple example, the math is shown to just factor the table once. However, as the quality of the data set grows over time, you need have an analysis tool to validate whether or not there is an adequate correlation to apply a prediction. For example, you may have enough data to predict Ductwork in a Hospital Mechanical Room versus a Data Center Mechanical Room, but not have adequate correlations to factor in the floor.

Analytics Data and Process

Cost Code 20 Data Summary

Item	EST Weight	Adj. Weight	EST Field Hrs (TO)	EST Field Hrs-CAD	Actual Field Hrs
Elbow	16.5	15.0	3.6	3.0	2.353
Straight	58.5	58.5	3.6	3.4	2.567
(20)	**75.0**	**73.5**	**7.2**	**6.4**	**4.92**

Overall

- Lbs/ Hr target (EST) = 75 / 4.92 = **15.24**
- Lbs/Hr target (CAD) = 73.5 / 4.92 = **14.93**
- Lbs Escalation Factor = 73.5 / 75.0 = **0.98**

Breakdown

	Elbow	Straight
Lbs/ Hr target (EST)	7.01	22.8
Lbs/Hr target (CAD)	6.37	22.8
Lbs Escalation Factor	0.91	1.0

These breakdowns are simplified examples. They could be anything because of the flexibility of the database.

Data Alignment – EST-CAD-Build Data – True Predictive Factor Created

Examples of Viewing and Using the Data – At the Point of a DECISION

The goal in developing a system and process for Data Analytics in your company should be to directly impact specific decisions that you make.

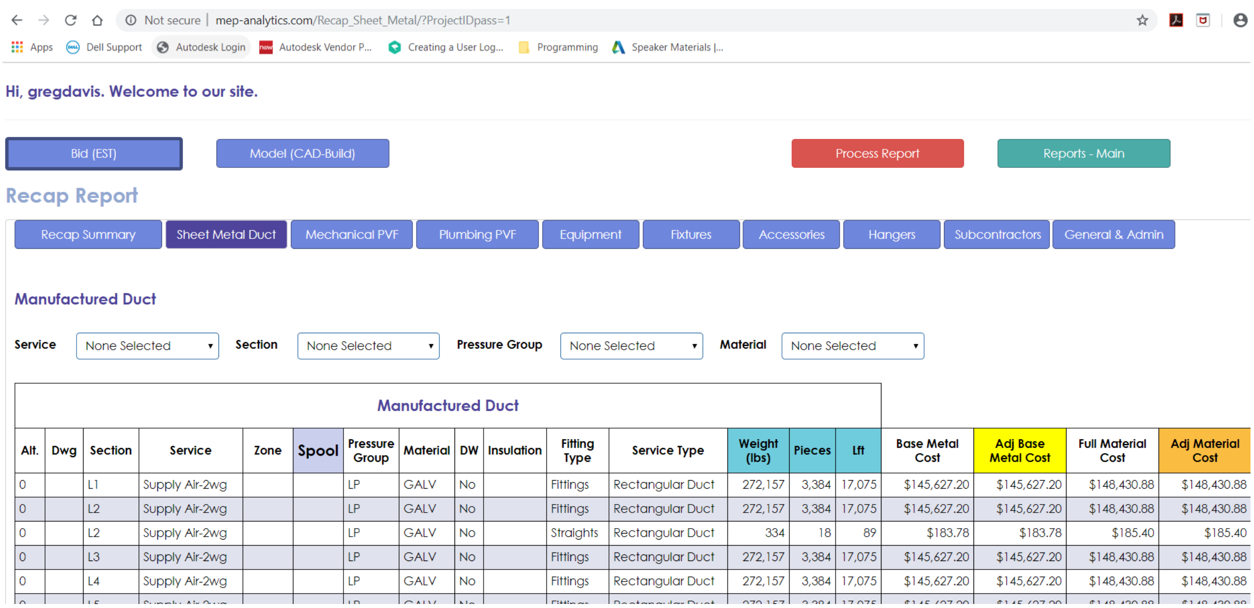
The following are specific decision points that Analytics MEP could provide analytics data:

- What factor of MCAA do I bid in a Hospital, 3rd Floor Mechanical Room for a Chilled Water Main?
- What Pounds/Hour Factor should I bid for a Data Center, Medium Pressure Supply Air System with 35% fittings?
- How many Hours do I need to schedule for the 4th Floor of the Office Tower project, and how many people do I need?
- How soon do I need the prefab of duct done for the 4th floor of the Office Tower?
- How soon do I need the Copper spools done, and what day should I ship them to get the most productivity out of my labor?
- Are we over-building our Mechanical Room Skids. Can we save costs by using less framing, and doing spool connections with flanges and Victaulic couplings on site?

The key point, once again, is that Data is useful if it is readily accessible when and where you need to make a decision. If you are never asked to ballpark the Dollars per SQFT for a Hospital and always do a takeoff, there isn't much benefit to a table predicting the Dollars per SQFT in a hospital. It might feel great to know, but it only becomes useful, which it may, if you can use it to sell jobs by providing that figure. If that figure is meant to be a ballpark guess that you aren't held to, then your priority is to have a system that gets you close enough for a ballpark guess, and to focus on details in other areas where you need the detail to make DECISIONS.

AnalyticsMEP – Apply Cost Adjustments to your bid with Predictive Analytics

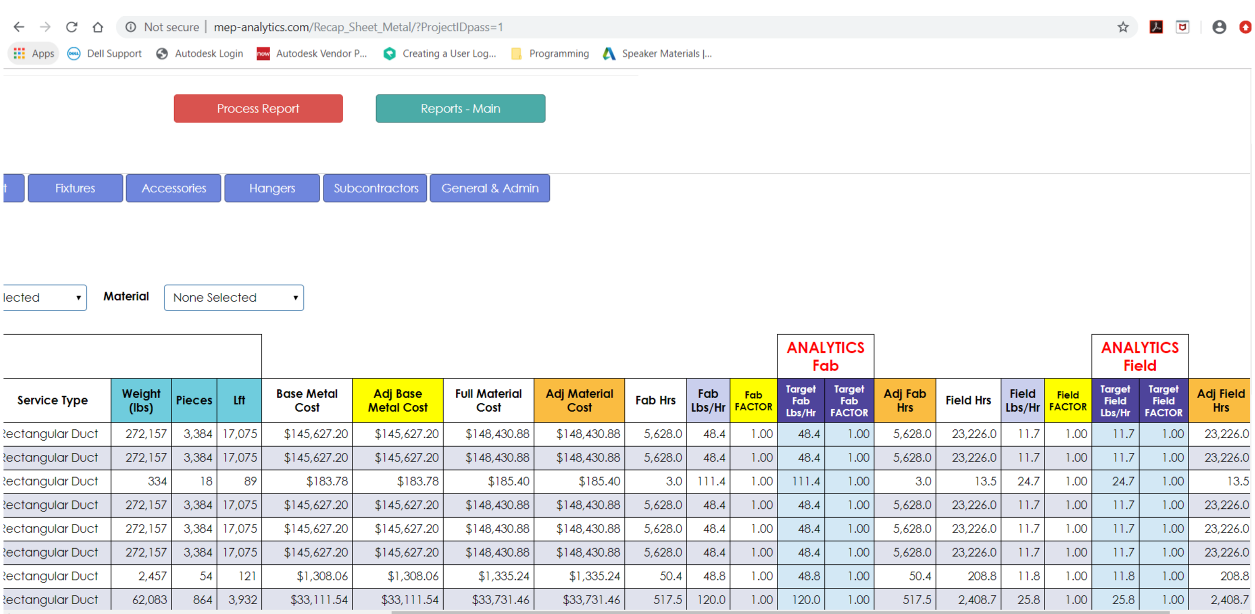
Because AnalyticsMEP is all online and database-driven, all adjustments write back to the real database in real time. Breakdowns of the Job into major category grouping (Duct displayed here) allow for different processes and grouping for each category.



Alt.	Dwg	Section	Service	Zone	Spool	Pressure Group	Material	DW	Insulation	Fitting Type	Service Type	Weight (lbs)	Pieces	Lft	Base Metal Cost	Adj Base Metal Cost	Full Material Cost	Adj Material Cost
0		L1	Supply Air-2wg			LP	GALV	No		Fittings	Rectangular Duct	272,157	3,384	17,075	$145,627.20	$145,627.20	$148,430.88	$148,430.88
0		L2	Supply Air-2wg			LP	GALV	No		Fittings	Rectangular Duct	272,157	3,384	17,075	$145,627.20	$145,627.20	$148,430.88	$148,430.88
0		L2	Supply Air-2wg			LP	GALV	No		Straights	Rectangular Duct	334	18	89	$183.78	$183.78	$185.40	$185.40
0		L3	Supply Air-2wg			LP	GALV	No		Fittings	Rectangular Duct	272,157	3,384	17,075	$145,627.20	$145,627.20	$148,430.88	$148,430.88
0		L4	Supply Air-2wg			LP	GALV	No		Fittings	Rectangular Duct	272,157	3,384	17,075	$145,627.20	$145,627.20	$148,430.88	$148,430.88

AnalyticsMEP – real time bidding with storage to an Analytics Database

In the Table, adjustments are entered in the orange columns. The purple columns give you real predictive factors to use – Updated live as your database changes. (At the time of publishing, no adjustments to the dataset have been made. Future updates to be published at analyticsmep.com/Sample.php



										ANALYTICS Fab						ANALYTICS Field		
Service Type	Weight (lbs)	Pieces	Lft	Base Metal Cost	Adj Base Metal Cost	Full Material Cost	Adj Material Cost	Fab Hrs	Fab Lbs/Hr	Fab FACTOR	Target Fab Lbs/Hr	Target Fab FACTOR	Adj Fab Hrs	Field Hrs	Field Lbs/Hr	Field FACTOR	Target Field Lbs/Hr	Target Field FACTOR	Adj Field Hrs
Rectangular Duct	272,157	3,384	17,075	$145,627.20	$145,627.20	$148,430.88	$148,430.88	5,628.0	48.4	1.00	48.4	1.00	5,628.0	23,226.0	11.7	1.00	11.7	1.00	23,226.0
Rectangular Duct	272,157	3,384	17,075	$145,627.20	$145,627.20	$148,430.88	$148,430.88	5,628.0	48.4	1.00	48.4	1.00	5,628.0	23,226.0	11.7	1.00	11.7	1.00	23,226.0
Rectangular Duct	334	18	89	$183.78	$183.78	$185.40	$185.40	3.0	111.4	1.00	111.4	1.00	3.0	13.5	24.7	1.00	24.7	1.00	13.5
Rectangular Duct	272,157	3,384	17,075	$145,627.20	$145,627.20	$148,430.88	$148,430.88	5,628.0	48.4	1.00	48.4	1.00	5,628.0	23,226.0	11.7	1.00	11.7	1.00	23,226.0
Rectangular Duct	272,157	3,384	17,075	$145,627.20	$145,627.20	$148,430.88	$148,430.88	5,628.0	48.4	1.00	48.4	1.00	5,628.0	23,226.0	11.7	1.00	11.7	1.00	23,226.0
Rectangular Duct	272,157	3,384	17,075	$145,627.20	$145,627.20	$148,430.88	$148,430.88	5,628.0	48.4	1.00	48.4	1.00	5,628.0	23,226.0	11.7	1.00	11.7	1.00	23,226.0
Rectangular Duct	2,457	54	121	$1,308.06	$1,308.06	$1,335.24	$1,335.24	50.4	48.8	1.00	48.8	1.00	50.4	208.8	11.8	1.00	11.8	1.00	208.8
Rectangular Duct	62,083	864	3,932	$33,111.54	$33,111.54	$33,731.46	$33,731.46	517.5	120.0	1.00	120.0	1.00	517.5	2,408.7	25.8	1.00	25.8	1.00	2,408.7