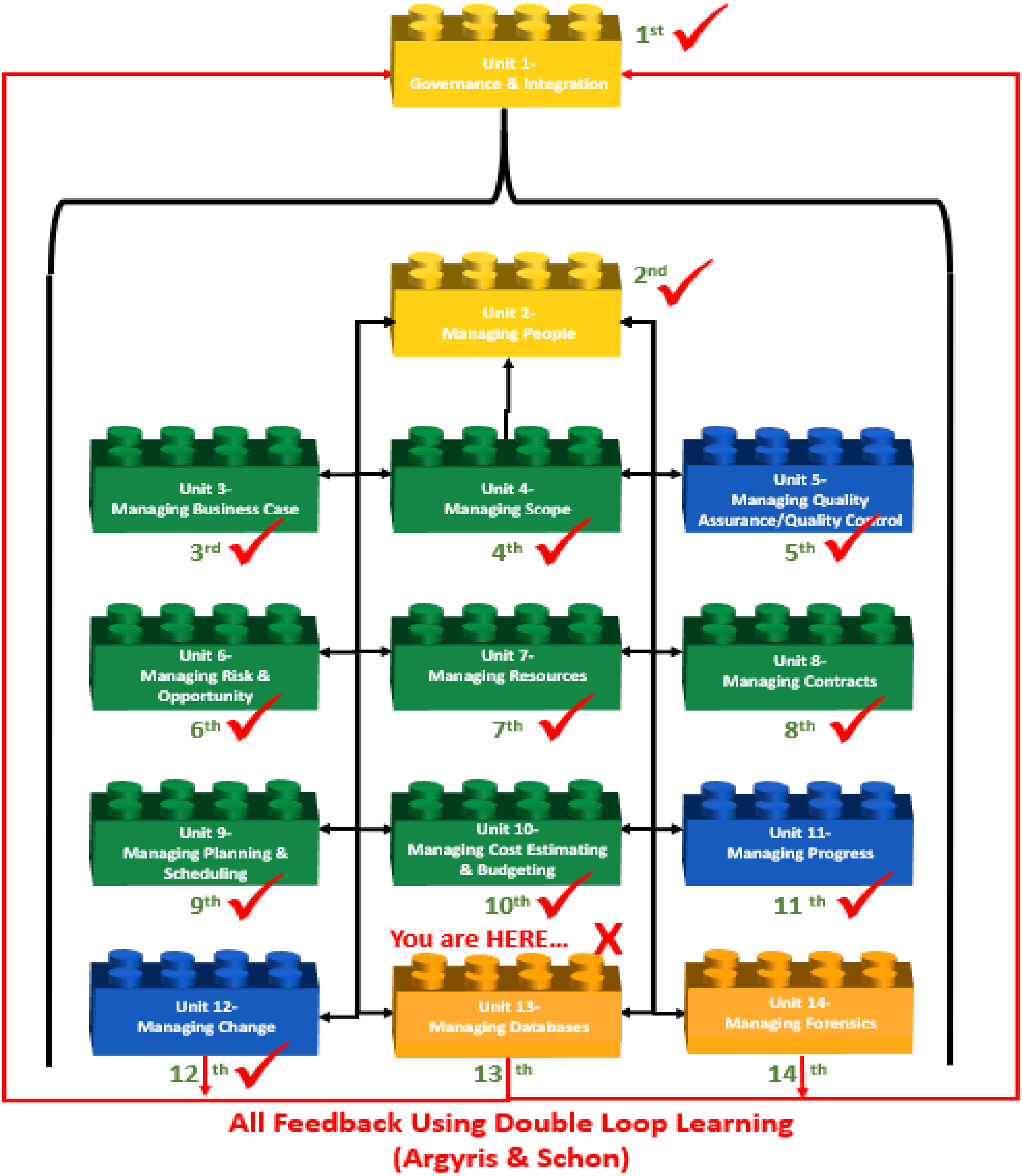
1 UNIT 13- MANAGING DATABASES



2 Figure 1- High-Level Process Map Showing Progress

RECAP OF UNITS 1-12

Keeping in mind that this book is designed to be a "How To Do It" or "Cookbook" of "best tested and proven" "recipes" derived or based on 90+ years of experience and solid research, here is a quick review of what previous Units have covered. While it is unnecessary to follow this sequencing, it is highly recommended to start with.

- ✓ **Unit 1- Governance and Integration**: We researched, documented, and shared the infrastructure or framework to enable project management to function as an asset delivery system. Treating "projects" as being one of several "delivery systems" for organizations to "create, acquire, update, expand, repair, maintain and eventually dispose of ORGANIZATIONAL ASSETS" has been tested and proven to work for 65 years at least. Trying to set up an "Integrated Asset, Portfolio, Program and Project Management ("IA3PM") methodology or system without the supporting infrastructure and framework makes it almost impossible.
- ✓ **Unit 2- Managing People**: We researched, documented, and shared the core "Individual, Management and Organizational" competencies necessary for project practitioners regardless of job title with the "skills, knowledge, attitudes and aptitudes" to function in a project environment, either as a CONTRACTOR'S or an OWNER'S organization. We also explored how to identify and score stakeholders to determine which ones are the "most" important.
- ✓ **Unit 3- Managing Business Case**: We contributed to or showed you how to facilitate the Business Case development, regardless of whether you are working for an OWNER or CONTRACTOR organization, and regardless of the sector you are working in or for.
- ✓ **Unit 4- Managing Scope**: We contributed to or facilitated the development of STANDARDIZED, MULTI-DIMENSIONAL WBS/CBS coding structures that enable us to present project information in a way that makes sense to ALL stakeholders.
- ✓ **Unit 5- Managing QA-QC**: We identified the Quality Assurance and Quality Controls tools & techniques developed for general use and demonstrated how to adopt or adapt them for use in an "IA3PM" (Integrated Asset, Portfolio, Program, and Project Management) environment. We also added a few NEW or DIFFERENT tools and techniques that we have found to add value as project practitioners.
- ✓ **Unit 6- Managing Risk & Opportunity**: We identified the tools & techniques associated with Risk and Opportunity Management, but most importantly, we not only identified both POTENTIAL Risks and Opportunities, but we have also identified people responsible for making both STRATEGIC and TACTICAL decisions about those risks and opportunities and have formally EMPOWERED them with authority to ACT, either to protect us against the impacts of potentially NEGATIVE outcomes and to exploit or enhance the probability presented by an OPPORTUNITY.
- ✓ **Unit 7- Managing Resources**: The logic or rationale behind doing **Unit 6- Managing Risk & Opportunity** BEFORE we did Unit 7 is, does it NOT make sense that the risk or opportunity RESPONSES will impact what resources we need when we need them, and how are we going to obtain or secure them? First, we had to identify our Single Points of Contact (SPOCs) and then empower them to make decisions through the Delegation of Authority process but supposing they need more than that? If we need an oil spill response team, won't they require not only people but

machines, tools, and training? Likewise, "Managing Resources" becomes a PREREQUISITE to **Unit 8- Managing Contracts**. Why? Because only when we know what RESOURCES or are and are NOT available, only then can we decide what are the appropriate STRATEGIC responses, which is, do we do this project with our own people "in-house" ("insource") and what needs to be contracted out or "outsourced?"

- ✓ **Unit 8- Managing Contracts:** Moving forward, the reason we started the MANAGING CONTRACTS process AFTER **Unit 7- Managing Resources** was the availability and competency of the available resources has a major impact on the decision of what to "insource" or do with our own people and what to "outsource" to one or more contractors or vendors. And the reason we do the Planning and Scheduling AFTER we've made the "Make (insource) or Buy (outsource)" decision is IF we have outsourced, then the detailed or Level 4-5 SCHEDULING will be done by the CONTRACTOR, not the OWNER. Now, if the decision were made to insource, then creating the Level 4 or Level 5 schedule would or SHOULD be the responsibility of our in-house project controls team.
- ✓ **Unit 9- Managing Planning & Scheduling:** The reason we held off exploring Managing Planning and Scheduling until AFTER we know what we were going to do in-house and what we wanted to outsource (contract out) is because that decision determines who is going to be responsible for developing the Level 4, 5 or even Level 6 CPM Schedule. (The "Execution" Schedule) Likewise, we do the Schedule BEFORE we do the cost budgeting (loading or allocating the cost estimates into the schedules using Activity Based Costing) because WHEN we do an activity may have a major impact on the COSTS of that Activity, the classic example being the cost of placing concrete in the winter vs. the cost of placing concrete in the summer or doing excavation or other civil activities during the rainy season. Again we cannot emphasize enough the importance of "APPLIED COMMON SENSE" in the sequencing of these Units, provided we recognize and accept the impact of those pesky FEEDBACK LOOPS.
- ✓ **Unit 10- Managing Cost Estimating and Budgeting:** In **Unit 9- Managing Planning & Scheduling**, we created the CPM Schedule and using "Start-to-Start" logical relationship, we took the cost ESTIMATES, and by spreading them over time by allocating those budgets over the duration of the activities and then plotting the Early Date S-Curves with the activities constrained to start as EARLY as possible and the Late Date S-Curve by constraining the activities to start as LATE as possible, we were able to produce a profile against which to start monitoring PHYSICAL PROGRESS against that plan.
- ✓ **Unit 11- Managing Progress:** In **Unit 9- Managing Planning & Scheduling** and **Unit 10- Managing Cost Estimating and Budgeting**, we created (or at least SHOULD have created) as realistically as we possibly could, a "model" showing WHEN and HOW we planned on allocating our scarce or limited resources to deliver the ASSET the project was undertaken to "create, acquire, update, expand, repair, maintain or dispose of." This took the form of an S-Curve showing BOTH the early and late dates. This was (or should have been) required by the owner as a prerequisite for them issuing the "Notice to Proceed." (NTP). This cost and resource loaded CPM schedule and the associated Early and Late Date S-Curve is known as the PERFORMANCE MEASUREMENT BASELINE (PMB) and is the basis against which all progress is measured. IF there are any CHANGES (and we know of no project in HISTORY that has not had SOME changes), then the PMB needs to be ADJUSTED accordingly.

- ✓ **Unit 12- Managing Change:** In **Unit 11- Managing Progress**, we created the performance measurement baseline and recognizing that change WILL happen, and as Integrated Asset, Portfolio, Program, and Project Management practitioners, we have a moral and ethical, if not a legal, obligation to manage those changes professionally.

INTRODUCTION TO UNIT 13- MANAGING DATABASES

At the same time, we are managing change because the programs and projects are now being executed, with data being generated daily, we also have a professional obligation to be capturing that data as close to real-time as possible, understanding that especially for contractors, are using today's cost and productivity data to bid tomorrows work. And for OWNERS to be able to realize the full value from Earned Value Management, they too need to start by knowing the FAIR MARKET VALUE of the goods and services they are buying. This means it is up to us, as professional project controllers and PMO practitioners, to ensure that the data remains ACCURATE, PRECISE, and RELIABLE. Unit 13 also becomes critical as, right now, a lack of up-to-date cost and localized productivity and cost databases is one of the constraints or impediments to implementing Building Information Modeling (BIM), specifically for 4D and 5D Apps. For any of you with an entrepreneurial spirit, this provides a multitude of LOCAL opportunities.

WHAT IS THE PURPOSE OF MANAGING PROJECT DATABASES?

The purpose of the Managing Databases is to introduce the tools, techniques, and methodologies, deemed appropriate to designing, creating, updating, and otherwise managing databases, that have been identified as being "best tested and proven" practices and which have been found to work on "most projects, most of the time"; provide a logical or rational sequence showing when those tools or techniques would normally and customarily be used and in selected instances, show how to use those tools/techniques and/or where to find additional information on how to use or apply them.

In terms of the change management processes, there is not any major or significant difference between how owners and contractors design, create, update, or otherwise manage databases, and, in the case of commercial databases (i.e., RS Means, Richardsons, Compass, etc.) the same database can be used as least as a starting point by both owners and contractors.

So what is a database? A database is a collection of information either in written or numeric form, which is stored for a specific purpose and organized to allow its contents to be easily accessed, managed, and updated. Although this definition includes stored data collections such as libraries, file cabinets, and address books, when we talk about databases, we almost invariably mean collecting data stored on a computer.

There are two basic categories of the database. The most commonly encountered category is the transactional database, used to store dynamic data, such as inventory contents, which is subject to change on an ongoing basis. The other category is the analytical database, used to store static data, such as geographical or chemical test results, which is rarely altered. For project control professionals, the classic example of a transactional database is the cost and productivity databases, while most "lessons learned" databases tend to be more static.

Strictly speaking, a database is just the stored data itself, although the term is often used erroneously to refer to a database and its management system (DBMS).

Given that we see more "planning and scheduling" and "cost estimating" being automated through the use of computer software and building information modeling (BIM), the role of the "project control professional" of the future is more likely going to be less focused on producing quantity take-off's, bills of material or bills of quantities or even on creating CPM schedules, but more on the creation and expansion of the coding structures as well as the development, maintenance and expansion of the databases of information necessary to enable this automation to be possible.

- **For Planners / Schedulers, we will have to focus on capturing, analyzing, and coding productivity rates, learning curves,** and procurement lead times than we are actually creating the activities and the logic. As practitioners, we often maintain "libraries" of data that often consist of previous project schedules, production rates, build and procurement times, sample fragnets, reports, presentations, procedures, and narratives, etc. All of which we use and utilize as part of our planning and scheduling duties.
- **For Cost Estimators, it means** that instead of spending our time doing quantity take-offs and producing bills of materials or bills of quantities, more of our time is going to be spent keeping the cost estimating databases current and updated, especially the need to develop location factors for different cities or regions around the world.

And even our Forensic Analysts need to be able to access and use the same databases as the planners, schedulers, and cost estimators, plus they need to be able to access, use and understand the various legal databases, such as Lexus/Nexus.

All of the above are kept in databases of various forms; in its simplest form, it could simply be a coordinated suite of folders on a computer holding reports and useful information, an excel spreadsheet which simply has many "rows" storing the information, or even a more complex excel database capable of being filtered and sorted. Taken to the next level, we can use databases from bespoke and off the shelf software solutions to build complex databases into which we can place either historic or current project scheduling and cost data and complete "relational" or "interlinked databases" which integrate project time, cost and accounting content.

Summarized, the future of "project controls" is more than likely to shift to more emphasis on data collection, data analysis and normalization, data codification, and data mining, all of which require a complete understanding of database systems database management.

WHAT ARE THE PROCESS MAPS FOR MANAGING PROJECT DATABASES?

At the 1,000 meter level of detail, the process flow chart looks like managing change. At the same time, this process applies equally to both the Owner and Contractor organizations there are subtle but important differences which is why we show separate process maps for each.

With the exception of the acquisition of the initial database, which is normally done by owners and to a lesser extent by contractors, the process of designing, creating, updating, and otherwise managing databases is largely an internal process, although the levels of detail between an owner's database and that of a contractor is likely to be different.

Thus once the initial database template is purchased or created, it is updated to reflect the scope of work normally and customarily performed by the owner (Unit 4- Managing Scope) and is (or at least should be) based on actual performance (in terms of costs and productivity) coming from progress on real projects. (Unit 11- Managing Progress). As the actual cost and productivity data come in from the field, the owner's project control analyzes this productivity, normalizes and adjusts it, and ideally, that "real-time" data is used to estimate the cost and duration of tomorrow's projects.

In addition to the more obvious cost and productivity data, which provides key inputs to Unit 9- Managing Planning & Scheduling and Unit 10- Managing Cost Estimating and Budgeting, consistent with the PTMC's advocating the use of Double Loop Learning, it is expected that the owner's project control team will also assume responsibility to research, analyze and share "lessons learned" from previous projects as part of the risk/opportunity management process (Unit 6- Managing Risk & Opportunity) and selection of the best or most appropriate contracting method and type (Unit 8- Managing Contracts), with the objective to minimize or mitigate claims and disputes (Unit 14- Managing Forensics).

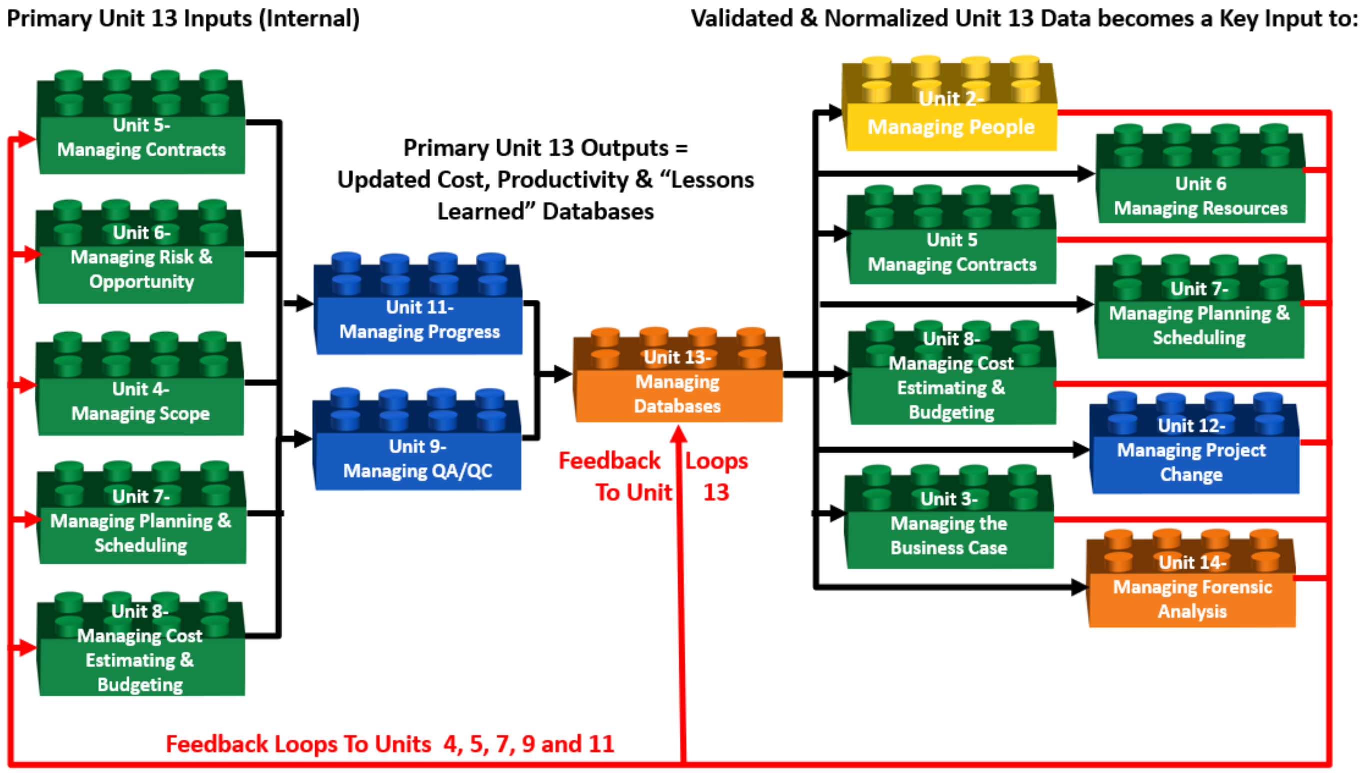


Figure 2- 1,000 Meter Level Process Flow Chart for Unit 13- Managing Databases from BOTH OWNER'S and CONTRACTOR'S PERSPECTIVE
Source: PTMC Team

While in terms of the PROCESS elements, their relationships, and sequencing, there is very little difference between the owner and contractor's perspectives an important difference which is worth noting is the fact that scope definition from the OWNER'S perspective derives from Unit 4- Managing Scope in the form of the Work Breakdown Structure (WBS) for a CONTRACTOR, scope definition derives from the CONTRACT, meaning that the key input for CONTRACTORS is the contract documents, which form the basis of the Contractual Work Breakdown Structure (CWBS).

Another important difference worth noting is that because of the highly competitive nature of contracting, combined with the fact that most contractors are working on single-digit EBIT margins, contractors very rarely use commercial databases with significant modifications to them, both in terms of crew sizes and allocations as well as the productivity calculations. As these important nuances give contractors a real or perceived competitive advantage in the marketplace, while the commercial databases are often purchased to provide the standardized coding structures, the actual cost and productivity numbers are almost sure to be modified. This is why for contractors, "lessons learned" databases are essential as even small improvements to the processes yield a significant competitive advantage.

Thus, while owner organizations can get away with less detail for use in "top-down" estimating methods, contractors generally require more detail. They are producing cost and duration estimates using "bottom-up" methods. Also, as CONTRACTOR'S are bidding in highly competitive markets, where single-digit EBIT and Net Margins are the norms, they cannot be bidding using data that is not PRECISE, RELIABLE, and ACCURATE.

PROCESS MAPS FROM 100 METERS

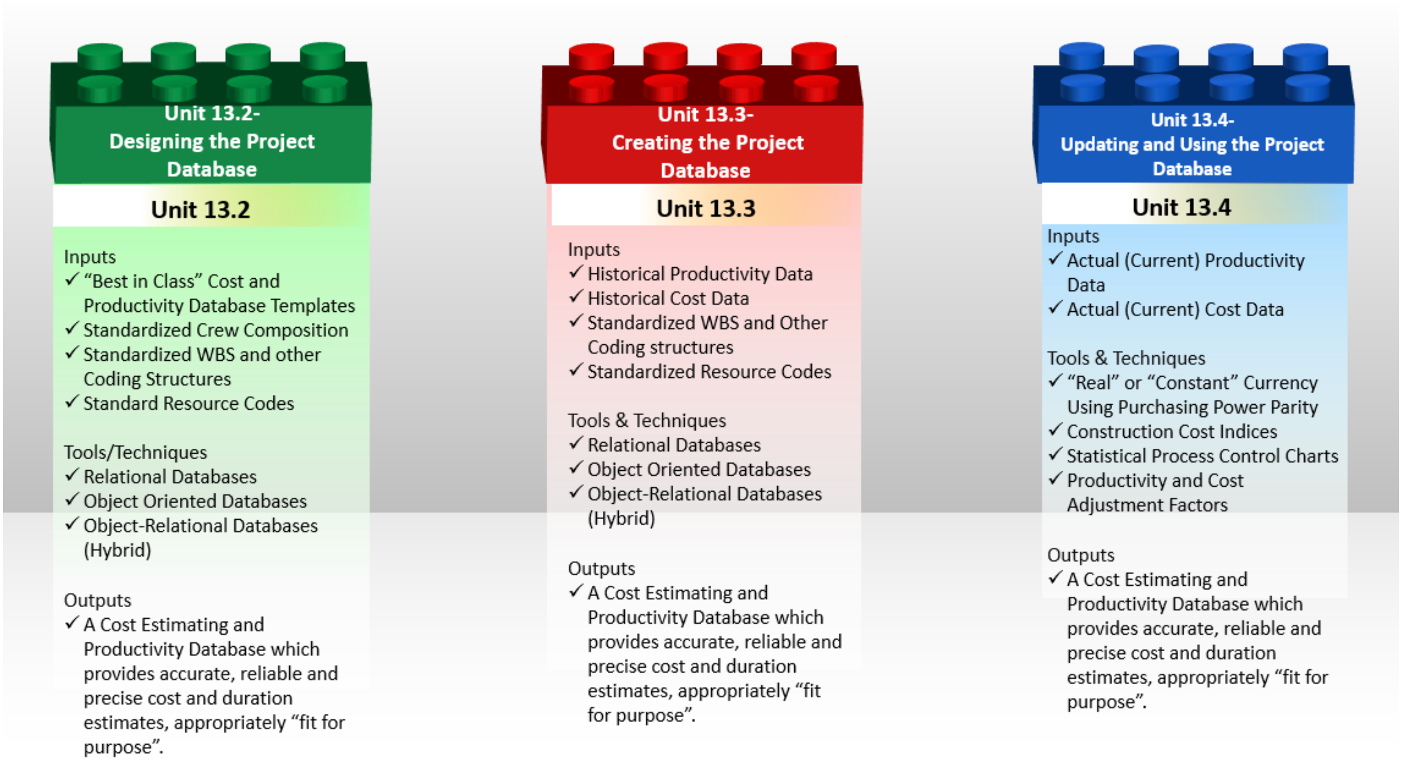


Figure 4- 100 Meter Level Process Flow Chart for Managing Change, from both the OWNER'S and CONTRACTOR'S ORGANIZATION PERSPECTIVE

Source: PTMC Team

Database development is a four-step process:

- ✓ **In the first step**, it is important to create a policy and procedures manual for database management that is easy to understand and follow. Failure to do this will result in people creating their own modifications, which may or may not work the way the management needs or wants. Developing SoP was covered in **Unit 1- Governance and Integration** rather than repeating it for each follow-on unit.
- ✓ **In the second step**, which is where we start this unit, you should create the logical design for the database, based solely on the data you want to store, rather than thinking of the specific software used to create it or the types of reports created. Therefore, many owners and contractors begin the process by purchasing an existing "Commercial Off the Shelf" (COTS) database and then CUSTOMIZING it to be "fit for purpose." This step defines tables and fields and establishes primary and foreign keys and integrity constraints. In the event your organization chooses to create your own, the PTMC Team has included templates, which have been "tested and proven to work over many years of actual use in the marketplace" rather than just some theoretical design.
- ✓ **In the third step**, you implement your plan within the database software program, which for owners includes making adjustments for location, currency fluctuations, or inflation, and for contractor's means optimizing crew sizes and compositions and enhancing productivity wherever possible.
- ✓ **In the fourth and final step**, you develop the end-user application that will allow your user(s) to interact with the database, including the most critical responsibility for project control professionals, which is to ensure that the database is continually updated with "real-time" (current) cost and productivity information as well as capturing "lessons learned." This last step where the use of STANDARDIZED CODING STRUCTURES (WBS, CBS, CREW, and RESOURCE DICTIONARY IDs) becomes of critical importance, especially if the project has been designed using Building Information Modeling. (BIM) Failure to adopt the standardized coding structures which are pre-loaded with each object in the design will require the project control team who has not adopted the standardized coding structures to write translator programs to enable their "homegrown" or "ad hoc" coding structures to "talk" to or exchange data with the BIM software packages.

While the 100 Meter level of detail provides a more granular look at the processes and how they interact than the 1,000 Meter view, there is yet another deeper level of detail which the PTMC calls the "ground" or "working-level." It is the next level deeper which contains the explanation for each of the Units shown above, telling more about what inputs are required, including providing some examples; what tools, techniques are typically used, including providing examples or templates, and in selected instances, specific step by step instructions or links to additional resources, showing how to use each of these tools or techniques consistent with the PTMC's commitment to identify and advocate "best tested and proven" practices.

BACKGROUND INFORMATION FOR MANAGING PROJECT DATABASES

Given the rapid proliferation of Building information modeling, it is becoming increasingly obvious that the "project control professional" of tomorrow is going to be less involved in doing quantity take-offs and cost estimates or creating schedules, which already can be done faster and arguably enough, more accurately, by computer software than doing it by hand, which means that the real "added value" services which still require a human to perform are the creation, populating, updating and maintaining the cost and productivity databases which the 4D, 5D and 6D BIM require in order to produce "realistic" durations, cost estimates and cost budgets. Without these databases developed and updated in real-time, results in the old "Garbage In/Garbage Out" paradigm.

For this reason, the PTMC Team has included a separate Unit on this topic as this is likely to become one of the most important responsibilities we as project controllers have in the very near future.

- ✓ **What is a Database?**

According to Oracle, a database

> *"is an organized collection of structured information, or data, typically stored electronically in a computer system. A database is usually controlled by a database management system (DBMS). Together, the data and the DBMS and the applications associated with them are referred to as a database system, often shortened to just "database."*
>
> *Data within the most common types of databases in operation today is typically modeled in rows and columns in a series of tables to make processing and data querying efficient. The data can be easily accessed, managed, modified, updated, controlled, and organized. Most databases use structured query language (SQL) for writing and querying data."*

There are two basic categories of databases. The most encountered category is the **transactional database**, used to store dynamic data, such as inventory contents, which is subject to change on an ongoing basis. The other category is the **analytical database**, used to store static data, such as geographical or chemical test results, which is rarely altered.

Strictly speaking, a database is just the stored data itself, although the term is often used erroneously to refer to a database and its management system (DBMS).

Given that we see more "planning and scheduling" and "cost estimating" being automated through the development of "Artificial Intelligence" (AI), "Machine Learning" (ML), and Building Information Modeling (BIM), the role of the "project control professional" of the future is more likely going to be less focused on producing quantity take off's, bills of material or bills of quantities or even on creating CPM schedules, but more on the creation and expansion of the coding structures as well as the development, maintenance and expansion of the databases of information necessary to enable this automation to be possible.

- ✓ **For Planners / Schedulers**, we will have to focus on capturing, analyzing, and coding productivity rates, learning curves, and procurement lead times than we are actually creating the activities and the logic. As practitioners, we often maintain "libraries" of data which often consist of previous project schedules, production rates, build and procurement times, sample fragnets, reports, presentations, procedures, and narratives, all of which we use and utilize as part of our planning and scheduling duties.
- ✓ **For Cost Estimators**, it means that instead of spending our time doing quantity take-offs and producing bills of materials or bills of quantities, more of our time is going to be spent keeping the cost estimating databases current and updated, especially the need to develop location factors for different cities or regions around the world.

- ✓ **And even our Forensic Analysts** need to be able to access and use the same databases as the planners, schedulers, and cost estimators, plus they need to be able to access, use and understand the various legal databases, such as Lexus/Nexus.

All of the above are kept in databases of various forms; in its simplest form, it could simply be a coordinated suite of folders on a computer holding reports and useful information, an excel spreadsheet which simply has many "rows" storing the information, or even a more complex excel database capable of being filtered and sorted. Taken to the next level, we can use databases from bespoke and off the shelf software solutions to build complex databases into which we can place either historic or current project scheduling and cost data and complete "relational" or "interlinked databases" which integrate project time, cost and accounting content.

Summarized, the future of "project controls" is more than likely going to shift to more emphasis on **data collection**, data **analysis** and **normalization**, data **codification**, and data **mining**, all of which require a complete understanding of "**Database Management."**

Before getting started, we need to ensure that all know and understand database terminology.

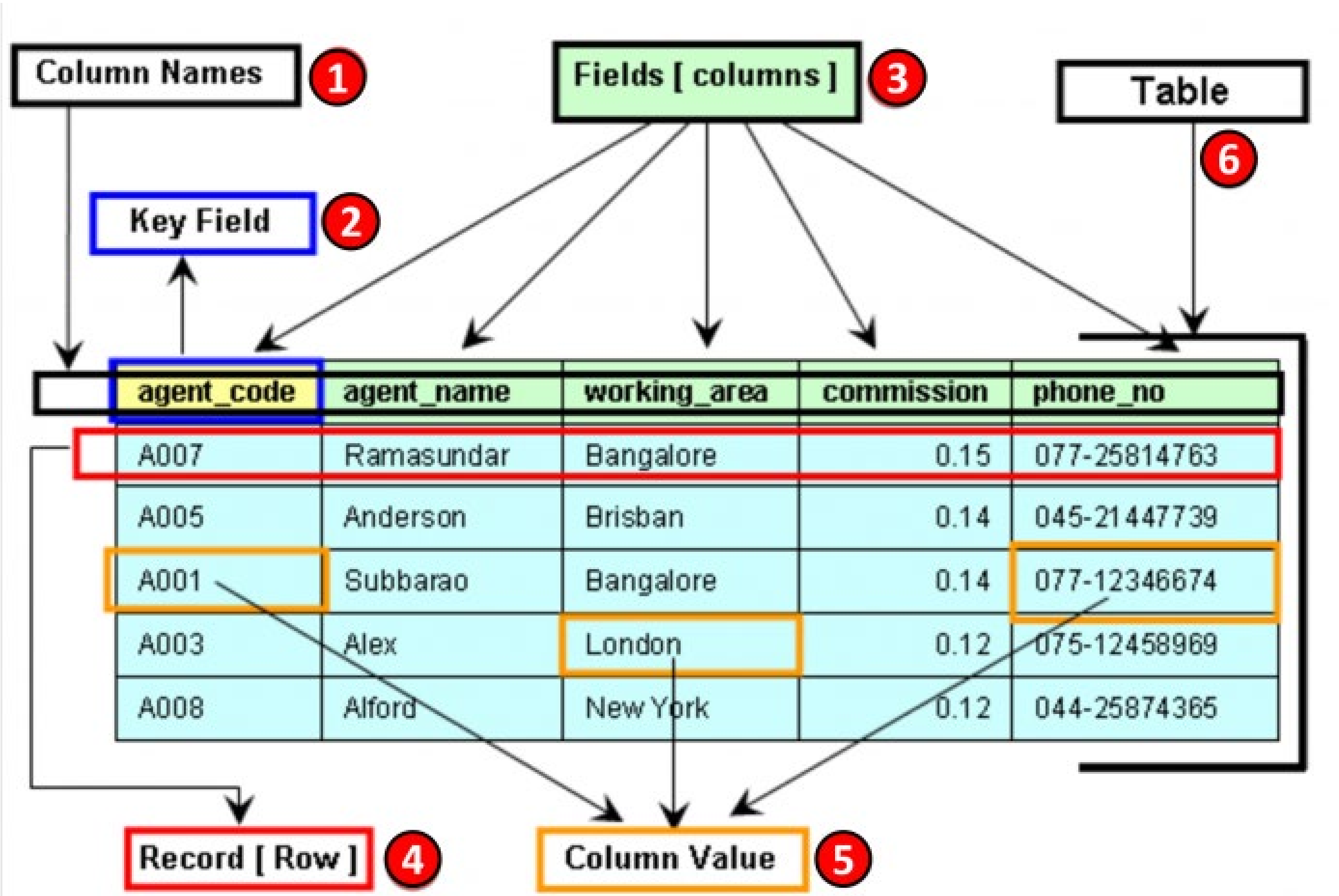


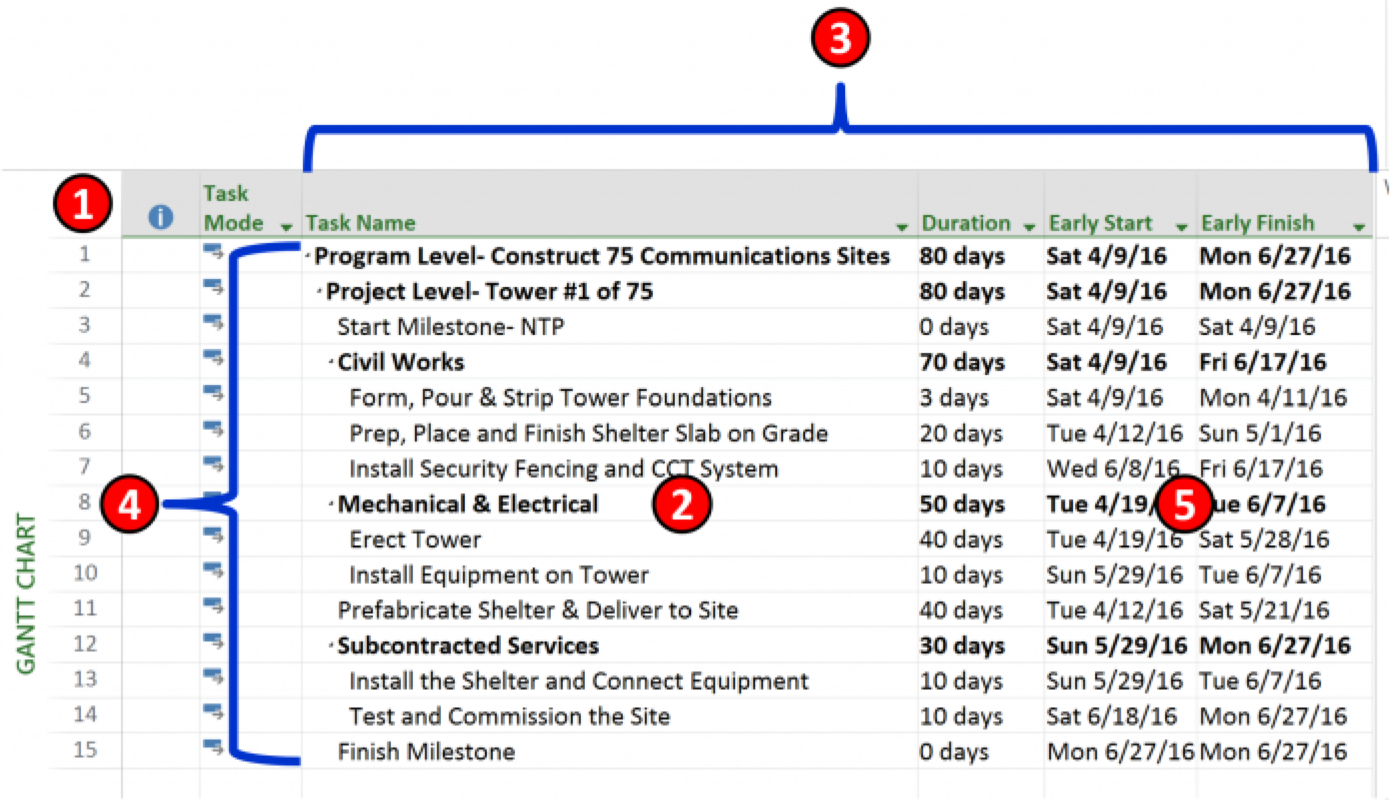
Figure 5- Basic Database Nomenclature

Source: W3Resource - Components of a table (of a database)

- ✓ **(5.1) Column Names**- All columns need to be named, and in any table, no two columns can have the same name. All column names need to be unique

- ✓ **(5.2) Key Fields**- There need to be one or more "key fields" which enable different tables (databases) to share or exchange information. Examples of Key Fields would be the Activity ID in Primavera or the Omniclass Tables or Norsok Z-014 Tables
- ✓ **(5.3) Fields**- Information in a table is relevant to a specific column header or heading. Fields are also called "Attributes."
- ✓ **(5.4) Record**- Each ROW of data is called a RECORD which contains all of the available information. For project control professionals, the "record" we most commonly deal with is an ACTIVITY. Records are also known as Tuples.
- ✓ **(5.5) Column Value**- These are the specific pieces of information for a given record that are relevant or appropriate for each field. A column value can be left blank, but if left blank will not be able to sort or filter using that field.
- ✓ **(5.6) Table**- A table is a set of one or more RECORDS, and it takes one or more tables to form a database. Explained another way, a table is the mass storage of information cross-referenced by RECORD (Row) and FIELD (Column)

As explained above, "databases" can come in many forms, and **the CPM Schedule** is just one of those:



	Task Mode	Task Name	Duration	Early Start	Early Finish
1		**Program Level- Construct 75 Communications Sites**	**80 days**	**Sat 4/9/16**	**Mon 6/27/16**
2		**Project Level- Tower #1 of 75**	**80 days**	**Sat 4/9/16**	**Mon 6/27/16**
3		Start Milestone- NTP	0 days	Sat 4/9/16	Sat 4/9/16
4		**Civil Works**	**70 days**	**Sat 4/9/16**	**Fri 6/17/16**
5		Form, Pour & Strip Tower Foundations	3 days	Sat 4/9/16	Mon 4/11/16
6		Prep, Place and Finish Shelter Slab on Grade	20 days	Tue 4/12/16	Sun 5/1/16
7		Install Security Fencing and CCT System	10 days	Wed 6/8/16	Fri 6/17/16
8		**Mechanical & Electrical**	**50 days**	**Tue 4/19/**	**ue 6/7/16**
9		Erect Tower	40 days	Tue 4/19/16	Sat 5/28/16
10		Install Equipment on Tower	10 days	Sun 5/29/16	Tue 6/7/16
11		Prefabricate Shelter & Deliver to Site	40 days	Tue 4/12/16	Sat 5/21/16
12		**Subcontracted Services**	**30 days**	**Sun 5/29/16**	**Mon 6/27/16**
13		Install the Shelter and Connect Equipment	10 days	Sun 5/29/16	Tue 6/7/16
14		Test and Commission the Site	10 days	Sat 6/18/16	Mon 6/27/16
15		Finish Milestone	0 days	Mon 6/27/16	Mon 6/27/16

Figure 6 - Database Example from CPM Schedule Software

Source: Giammalvo, Paul D (2015) Course Materials Contributed Under Creative Commons License BY v 4.0

- ✓ **(6.1)** We see the "Key Field" is the Activity Number which there can be one and only one with that single unique identifier.
- ✓ **(6.2)** This is a single ROW
- ✓ **(6.3)** These are examples of FIELDS or ATTRIBUTES that are associated with each ROW

- ✓ **(6.4)** This is the TABLE containing all the data
- ✓ **(6.5)** Here are examples of VALUES, some of which are entered manually (i.e., Duration) or others that are calculated. (i.e., early and late finish dates)

Having made certain everyone knows the vocabulary and understands how it applies in the world of project controls, we can walk you through creating and maintaining your database.

The following introduction was copied in its entirety from Learn IT- The Power of the Database consistent with the PTMC belief that there is no need to "reinvent the wheel," we believe this is a great explanation of what a database is and how to use it in the context of project control databases.

- ✓ **A Brief History of the Database:**

The first attempts at computer databases arose around the mid-twentieth century. Early versions were file-oriented. A database file became known as a table because its structure was the same as a paper-based data table. For the same reason, the columns within a table were called fields, and the rows were called records. Computers were evolving during that same period, and their potential for data storage and retrieval became recognized.

The earliest computer databases were based on a flat-file model, in which records were stored in text format. In this model, no relationships are defined between records. Without defining such relationships, records can only be accessed sequentially. For example, if you wanted to find the record for the fiftieth customer, you would have to go through the first 49 customer records in sequence first. The flat-file model works well for situations where you want to process all the records but not for situations in which you want to find specific records within the database.

The hierarchical model, widely used in mainframe environments, was designed to allow structured relationships to facilitate data retrieval. Within an inverted tree structure, relationships in the hierarchical model are parent-child and one-to-many. Each parent table may be related to multiple child tables, but each child table can only be related to a single parent table. Because table structures are permanently and explicitly linked in this model, data retrieval was fast. However, the model's rigid structure causes some problems. For example, you can't add a child table that is not linked to a parent table: if the parent table was "Doctors" and the child table was "Patients," you could not add a patient record independently. That would mean that if a new patient came into a community's health care system, under the system, their record could not be added until they had a doctor. The hierarchical structure also means that if a record is deleted in a parent table, all the records linked to it in child tables will be deleted.

Also, based on an inverted tree structure, the next approach to database design was the network model. The network model allowed more complex connections than the hierarchical model: several inverted trees might share branches, for example. The model connected tables in sets, in which a record in an owner table could link to multiple records in a member table. Like the hierarchical model, the network model enabled very fast data retrieval. However, it also had many problems. For example, a user would need a clear understanding of the database structure to get information from the data. Furthermore, if a set structure were changed, any reference to it from an external program would have to be changed as well.

In the 1970s, the relational database was developed to deal with data in more complex ways. The relational model eventually dominated the industry and has continued to do so through to the present day. We'll explore the relational database in some detail in the next segment. For more, refer to Learn IT- The Power of the Database.

✓ **Related Links:**

- SearchDatabase offers a selection of resources for Database Backgrounders and General Information.
- Selena Sol's interesting and informative article, "What is a Database?" explores the historical development of databases.
- The Database Journal offers Ian Gilfillan's "Introduction to Relational Databases."

✓ **What is a Relational Database?**

In the relational database model, data is stored in relations, more commonly known as tables. Tables, records (sometimes known as tuples), and fields (sometimes known as attributes) are the basic components. Each individual piece of data, such as the last name or a telephone number, is stored in a table field, and each record comprises a complete set of field data for a particular table. In the following example, the table maintains customer shipping address information. Last_Name and other column headings are the fields. A record, or row, in the table, comprises the complete set of field data in that context: all the address information that is required to ship an order to a specific customer. Each record can be identified and accessed through a unique identifier called a primary key. In the Customer_Shipping table, for example, the Customer_ID field could serve as a primary key because each record has a unique value for that field's data.

	Typical Database Components							
Field Names	ID	First	Last	Apt.	Address	City	State	Zip
Records	101	John	Smith	147	123 1st Street	Chicago	IL	60635
	102	Jane	Doe	13 C	234 2nd Street	Chicago	IL	60647
	103	June	Doe	14A	243 2nd Street	Chicago	IL	60647
	104	George	Smith	N/A	345 3rd Street	Chicago	IL	60625

Figure 7 - Showing Data for the Example Above- Customer Shipping
Source: Learn IT- The Power of the Database.

The term relational comes from set theory rather than the concept that relationships between data drive the database. However, the model does, in fact, work through defining and exploiting the relationships between table data. Table relationships are defined as one-to-one (1:1), one-to-many (1:N), or (uncommonly) many-to-many (N: M):

- If a pair of tables have a one-to-one relationship, each record in Table A relates to a single record in Table B. For example, in a table pairing consisting of a table of customer shipping addresses and a table of customer account balances, each single customer ID number would be related to a single identifier for that customer's account balance record. The one-to-one relationship reflects that each customer has a single account balance.
- If a pair of tables have a one-to-many relationship, each individual record in Table A relates to one or more records in Table B. For example, in a table pairing consisting of a table of

university courses (Table A) and a table of student contact information (Table B), each single course number would be related to multiple records of student contact information. The one-to-many relationship reflects that each course has multiple students enrolled.

- Suppose a pair of tables have a many-to-many relationship. In that case, each individual record in Table A relates to one or more records in Table B, and each individual record in Table B relates to one or more records in Table A. For example, in a table pairing consisting of a table of employee information and a table of project information, each employee record could be related to multiple project records, and each project record could be related to multiple employee records. The many-to-many relationship reflects that each employee may be involved in multiple projects and that each project involves multiple employees.

✓ **Where did the Relational Model Come From?**

The relational database model developed from the proposals in "A Relational Model of Data for Large Shared Databanks," a paper presented by Dr. E. F. Codd in 1970. Codd, a research scientist at IBM, explored better ways to manage large amounts of data than were currently available. The hierarchical and network models of the time tended to suffer from problems with data redundancy and poor data integrity. By applying relational calculus, algebra, and logic to data storage and retrieval, Codd enabled developing a more complex and fully articulated model than had previously existed.

One of Codd's goals was to create an English-like language that would allow non-technical users to interact with a database. Based on Codd's article, IBM started their System R research group to develop a relational database system. The group developed SQL/DS, which eventually became DB2. The system's language, SQL, became the industry's de-facto standard. In 1985, Dr. Codd published a list of twelve rules for an ideal relational database. Although the rules may never have been fully implemented, they have provided a guideline for database developers for several decades.

✓ **Codd's Rules:**

- **The Information Rule**: Data must be presented to the user in table format.
- Guaranteed Access Rule: Data must be reliably accessible by referencing the table name, primary key, and field name.
- **Systematic Treatment of Null Values**: Fields that are not primary keys should be able to remain empty (contain a null value).
- **Dynamic On-line Catalog Based on the Relational Model**: The database structure should be accessible through the same tools that provide data access.
- **Comprehensive Data Sublanguage Rule:** The database must support a language that can be used for all interactions (SQL was developed from Codd's rules).
- **View Updating Rule:** Data should be available in different combinations (views) that can also be updated and deleted.
- **High-level Insert, Update and Delete:** It should be possible to perform all these tasks on any set of data that can be retrieved.
- **Physical Data Independence:** Changes to the database's architecture should not affect the user interface.

- **Logical Data Independence:** If the logical structure of a database changes, that should not be reflected in how the user views it.
- **Integrity Independence:** The language used to interact with the database should support user constraints to maintain data integrity.
- **Distribution Independence:** If the database is distributed (physically located on multiple computers), that fact should not be apparent to the user.
- **Non-subversion Rule:** It should not be possible to alter the database structure by any other means than the database language.

✓ **Related Links:**

- ITWorld goes into more detail about Codd's 12 Rules.
- The DB Group provides "A Brief History of Databases."
- The NAP Reading Room offers a chapter on "The Rise of Relational Databases" from the book Funding a Revolution.

✓ **What Other Types of Databases are there?**

Although the relational model is by far the most prevalent one, several other models are better suited to particular data types. Alternatives to the relational model include:

- **Flat-File Databases:** Data is stored in files consisting of one or more readable files, usually in text format.
- **Hierarchical Databases:** Data is stored in tables with parent/child relationships with a strictly hierarchical structure.
- **Network Databases:** Similar to the hierarchical model, but allows more flexibility; for example, a child table can be related to more than one parent table.
- **Object-Oriented Databases:** The object-oriented database model was developed in the late 1980s and early 1990s to deal with data types that the relational model was not well-suited for. Medical and multimedia data, for example, required a more flexible system for data representation and manipulation.
- **Object-Relational Databases:** A hybrid model, combining features of the relational and object-oriented models.

✓ **Related Links:**

- Phil Howard's article on SearchDatabase explores "A proliferation of database types."
- Ryan Stephens and Ronald Plew offer a tip on "Alternatives to the relational database" from their book Teach Yourself Database Design.

✓ **What "languages" are used to Interact with Databases?**

- **SQL (Structured Query Language)** is by far the most common language used to interact with relational databases. Originally developed for use with IBM's DB2, the standard -- often pronounced "sequel" -- is promoted in various formats by both the American National Standards Institute (ANSI) and the International Standards Organization (ISO).
- SQL commands are fairly straightforward and easy to understand. For example, suppose you wanted a list of all your customers within a specific zip code area. In that case, the following

command (based on the table in response to question #2, above), for example, will return that information, which in this case would be "George Smith."

- Select First_Name, Last_Name from Customer_Shipping where Zip = '60625';

Although many use proprietary extensions specific to their own products, most databases use SQL.

✓ **Related Links:**

- SearchDatabase.com has more in-depth information in their Learning Guide: SQL.
- You can browse through hundreds of questions answered by SearchDatabase SQL expert Rudy Limeback or ask him about something you don't see answered here.
- Search400 offers a selection of Best Web Links for SQL and Query.
- SQLCourse.com is a free, interactive SQL tutorial with a beginner's level followed by more advanced sections.

✓ **How can I Ensure a Good Database Design?**

Hands down, the most important thing you can do to ensure a successful database design is to **put enough resources into the planning stage.** The proliferation of off-the-shelf databases and database applications has led many people to a number of erroneous conclusions, such as:

- **Off-the-shelf databases can be easily customized.**

In fact, although there are ready-made databases available for any number of applications, their design typically differs significantly from the ideal model for your specific needs. And tailoring them to fit is often more complicated than starting from scratch.

Anyone can create a perfectly functional database.

In fact, almost anyone could create a perfectly functional database -- if they took the time to learn what they needed to know before they started to develop.

You can jump right into the development process, adjusting as you go along.

You could build a database without a carefully constructed plan. You could also build a house in that manner -- but it's not advisable. Databases are complicated constructions. Whether or not major problems rear their ugly heads through the development phase, they are bound to pop up in implementation. Fixing those problems can be difficult, time-consuming, and expensive. Furthermore, because of the intricate ways that data is connected in a database, a problem in one area can affect data in other areas in surprising ways.

Databases came into being because of the computer, and the two have enjoyed a mutually beneficial symbiotic relationship ever since, each helping the other grow by leaps and bounds. Somewhat ironically, however, the best way to start a plan for database development is to take out a paper, a pencil -- and a big eraser.

Database development is a three-phase process. In the first phase, you should create the logical design for the database, based solely on the data you want to store, rather than thinking of the specific software used to create it or the types of reports created. This phase defines tables and fields and establishes primary and foreign keys and integrity constraints. In the second phase, you implement your plan within the database software program, and in the third phase, you develop the end-user application that will allow your user(s) to interact with the database.

- ✓ **What are the most important things to keep in mind during the design phase?**
  - Create your design on paper first, as completely as possible.
  - Eliminate as much redundancy of data as possible.
  - Start from scratch -- don't try to use parts of a database with structural problems.
  - Make sure that each table represents a single subject.
  - Assign a primary key whose value clearly identifies each record and only a single record.
  - Ensure that each field represents a single value.
  - Take the time to be certain of data integrity.

- ✓ **Related Links:**
  - Michael J. Hernandez has a handy tutorial on Database Design Tips.
  - SearchDatabase offers a selection of resources for Database Languages and Development.
  - Hernandez's Database Design for Mere Mortals is available from the TechTarget Bookstore.
- ✓ **What is normalization, and why do I need to know about it?**

  - **In short:**

Well normalized data makes programming (relatively) easy and works very well in multi-platform, enterprise-wide environments. Non-normalized data leads to heartbreak. -- Steve Litt

Normalization is a guiding process for database table design that ensures, at four levels of stringency, increasing confidence that results of using the database are unambiguous and as intended. Basically, a refinement process, normalization tests a table design for the way it stores data so that it will not lead to the unintentional deletion of records, for example, and that it will reliably return the data requested.

Normalization degrees of relational database tables:

- **First normal form (1NF)**
  - This is the "basic" level of normalization and generally corresponds to the definition of any database, namely:
  - It contains two-dimensional tables with rows and columns corresponding to records and fields.
  - Each field corresponds to the concept represented by the entire table: for example, each field in the Customer Shipping table identifies some component of the customer's shipping address.
  - No duplicate records are possible.
  - All field data must be of the same kind. For example, in the "Zip" field of the Customer_Shipping table, only five consecutive digits will be accepted.
- **Second normal form (2NF)**

- In addition to 1NF rules, each field in a table that does not determine the contents of another field must itself be a function of the other fields in the table. For example, in a table with three fields for customer ID, the product sold, and price of the product when sold, the price would be a function of the customer ID (entitled to a discount) and the specific product.

- **Third normal form (3NF)**
  - In addition to 2NF rules, each field in a table must depend on the primary key. For example, using the customer table just cited, removing a record describing a customer purchase (because of a return perhaps) will also remove the fact that the product has a certain price. In the third normal form, these tables would be divided into two tables so that product pricing would be tracked separately. The customer information would depend on the primary key of that table, Customer_ID, and the pricing information would depend on the primary key of that table, which might be Invoice_Number.
- ✓ **Domain/key normal form (DKNF)**
  - In addition to 3NF rules, a key, which is a field used for sorting, uniquely identifies each record in a table. A domain is the set of permissible values for a field. By enforcing key and domain restrictions, the database is assured of being freed from modification anomalies. DKNF is the normalization level that most designers aim to achieve.

Related Links:

- Steve Litt's Normalization tutorial explains the practice and the processes involved.
- Tom Russell and Rob Armstrong's SearchDatabase article describes "13 reasons why normalized tables help your business."
- Database Journal has Ian Gilfillan's tutorial on Database Normalization.
- Database Words-to-Go Glossary: Browse through database vocabulary in a printable glossary.
- **Database development is a Three-Phase Process -** in the **first phase**, you should create the logical design for the database, based solely on the data you want to store, rather than thinking of the specific software that will be used to create it or the types of reports that will be created from it. This phase defines tables and fields and establishes primary and foreign keys and integrity constraints. In the **second phase**, you implement your plan within the database software program, and in the **third phase**, you develop the end-user application that will allow your user(s) to interact with the database.

UNIT 13.2- DESIGNING THE ORGANIZATION'S DATABASE

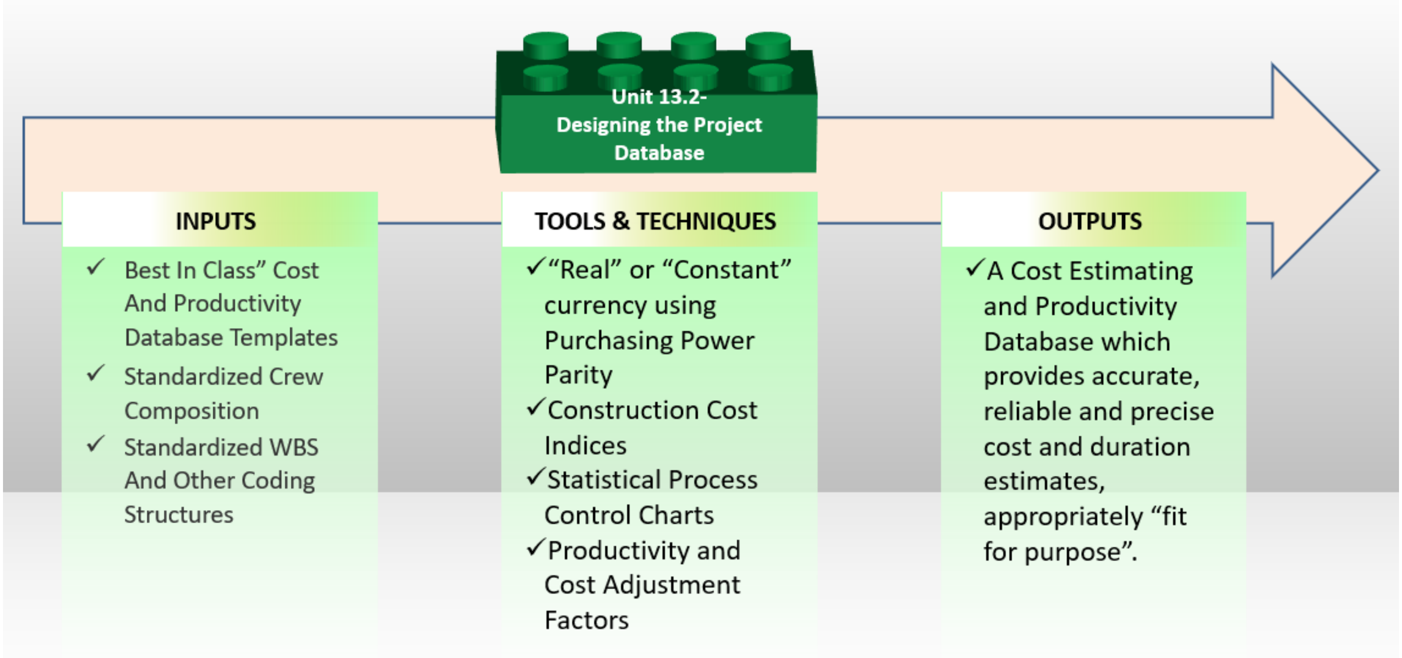


Figure 8- The Designing the Project Databases Process Map

Source: PTMC Team

INTRODUCTION

Designing the Project Database is perhaps one of the most important steps in creating a useable (complex) project database. Given we are moving towards the use of Building Information Modelling and more sophisticated and demanding projects, the three most logical options would be to choose from the following database types:

- **Relational Databases**: Computer databases in which all data is stored in relation (to the user) are tables with rows and columns. Each table is composed of records (called Tuples), and each record is identified by a field (attribute) containing a unique value. Every table shares at least one field with another table in 'one to one,' 'one to many,' or 'many to many relationships. These relationships allow the database user to access the data in almost unlimited ways and combine the tables as building blocks to create complex and very large databases.
- **Object-Oriented Databases**: A database specifically designed to work in an object-oriented programming environment, where data of various types may be stored, including text, graphics, sound, and video, and it provides database management system capabilities to objects (3) created by object-oriented programming languages. Its abbreviation is OODB. Also called object database.
- **Object-Relational Databases**: A hybrid model, combining features of the relational and object-oriented models.

INPUTS

- ✓ Best In Class" Cost And Productivity Database Templates
- ✓ Standardized Crew Composition
- ✓ Standardized WBS And Other Coding Structures

TOOLS & TECHNIQUES

- ✓ **Designing Database Codes**

As we know from our exploration of database elements, every item must have a unique identifier known as **a Key Field**.

Meaning one of the first challenges we face is establishing a coding structure, preferably one which has been standardized and ideally been done so across your industry and is not just unique to your organization. As we are using R.S.Means and the CSI/ISO Omniclass Tables as representing what the PTMC Team believes to be "best in class" examples, we will use those for our case studies. **However, there may be other coding structures, especially in other industries, which may be used instead of these examples**.

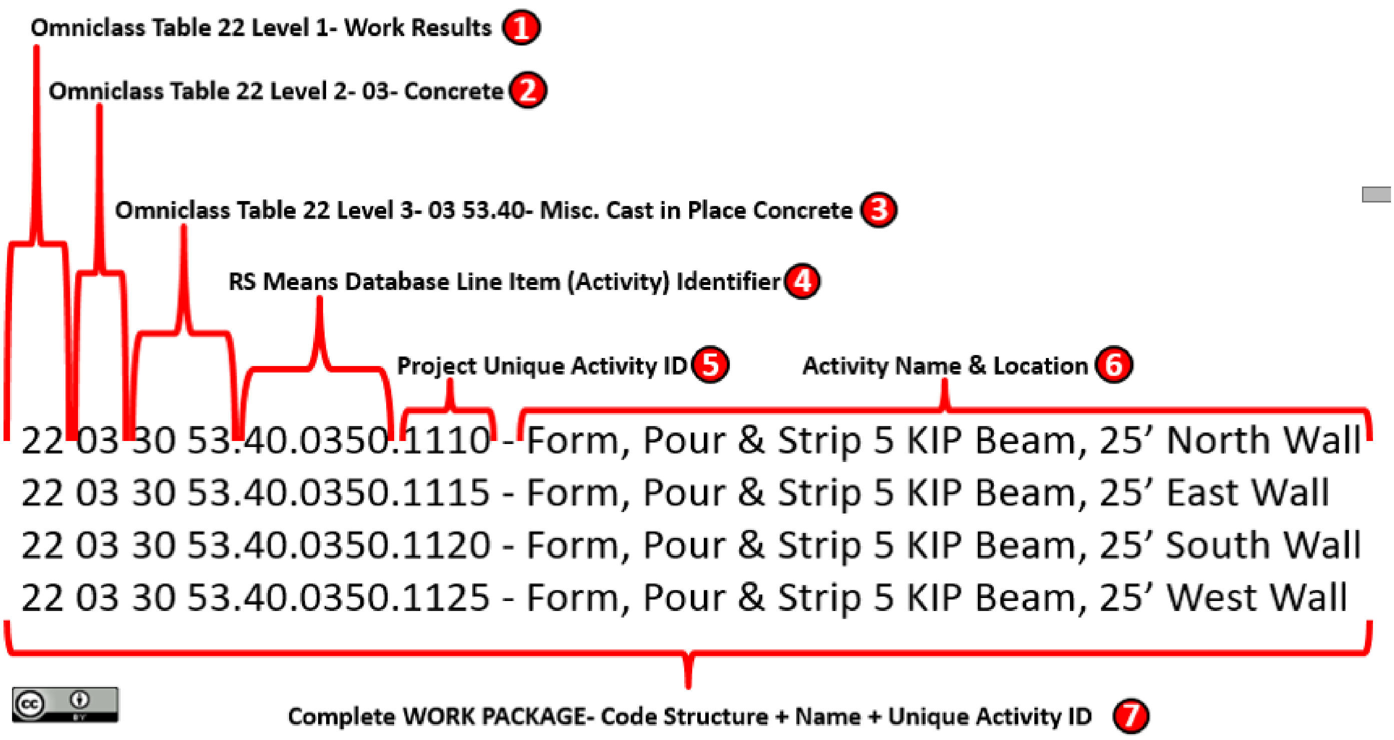


Figure 9- R.S Means Database Coding Structure Explained Showing KEY FIELDS
Source: Giammalvo, Paul D (2015) Course Materials Adapted from R.S. Means 2008 Facility Cost Estimating Database. Contributed Under [Creative Commons License BY v 4.0]

The example above illustrates the levels of detail that most CONTRACTORS would normally develop and maintain for their bids, estimates and projects, and databases which is why these coding structures are critical if we want to be able to maintain Vertical Integration capabilities so that the owners can "roll-up" the contractor's detailed cost estimates for their own use down to Level 3 minimum or more ideally, Level 4.

- **(9.1) 22** is CSI's Level 1 WBS. Using CSI's Masterformat or Omniclass Table 22, we can see ALL the project's deliverables (work results).
- **(9.2) 03** is CSI/Omniclass Level 2. For this example, 03 is ALL concrete on the project.
- **(9.3) 03 30** is CSI/Omniclass Level 3. In this example, 03 30 covers ONLY cast in place concrete. This is the MINIMUM level of detail that OWNERS should be providing to CONTRACTORS in the contract documents.
- **(9.4) 03 30 53.40** is the CSI/Omniclass heading covering Miscellaneous Beam Concrete in Place. This is the IDEAL or RECOMMENDED Level of Detail that OWNERS should be providing to CONTRACTORS in the contract documents.
- **(9.5) 03 30 53.40.1110-** This is the Activity level of detail that the contractor would be developing from the contract documents provided by the owner. Now, if there were more than ONE of these same activities (in other words, if there were several different places in the project which required the Forming (0.0010); Installing Rebar (0.0020) and Placing and Finishing Concrete (0.0050) 25' Span Beams then the contractor would add another level of code, let's say there are four places on the project where we need these Beams, North, South, East and West Walls.
- **(9.6) 03 30 53.40.0350.1110** would be the coding structure that the contractor would likely use to identify the first Activity on the North Wall, **03 30 53.40.1115.**, would be the coding structure that the contractor would likely use to identify the second location and **03 30 53.40.1120** would be the coding structure that the contractor would likely use to identify the third and **03 30 53.40.1125** and fourth and final location.
- **(9.5) 03 30 53.40.1110 to 03 30 53.40.1125** would identify the entire **WORK PACKAGE.**

✓ **Designing Database Structures**

Having explained the coding structure, let's explore what a well-established time, cost, and productivity database looks like.

03 Concrete (1)										
03 30 Cast in Place Concrete (2)										
03 30 53 Miscellaneous Cast in Place Concrete (3)										
03 30 53.40 Concrete in Place (4)										
0.0010 0.0020 0.0050	Including Forms (4 uses), reinforcing steel, concrete placement and finishing, unless otherwise indicated.	Crew Type	Daily Output per Unit	Labor Hours per Unit	Unit of Measure	Material Costs	Labor Costs	Equip-ment Costs	Total Costs per Unit	Total Price/Unit Including OH&P
0.0300	Beams- 5 kip per lineal foot, 10' long spans	C14-A	15.62	12.8	Cubic Yard (CY)	$340.00	$645.00	$58.00	$1,043	$1,475
0.0350	Beams- 5 kip per lineal foot, 25' long spans	"	18.55	10.78	CY	$355.00	$545.00	$49.00	$949	$1,325
(5)		(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)

Source- RS Means Facilities Construction Costs 2018- Page 99

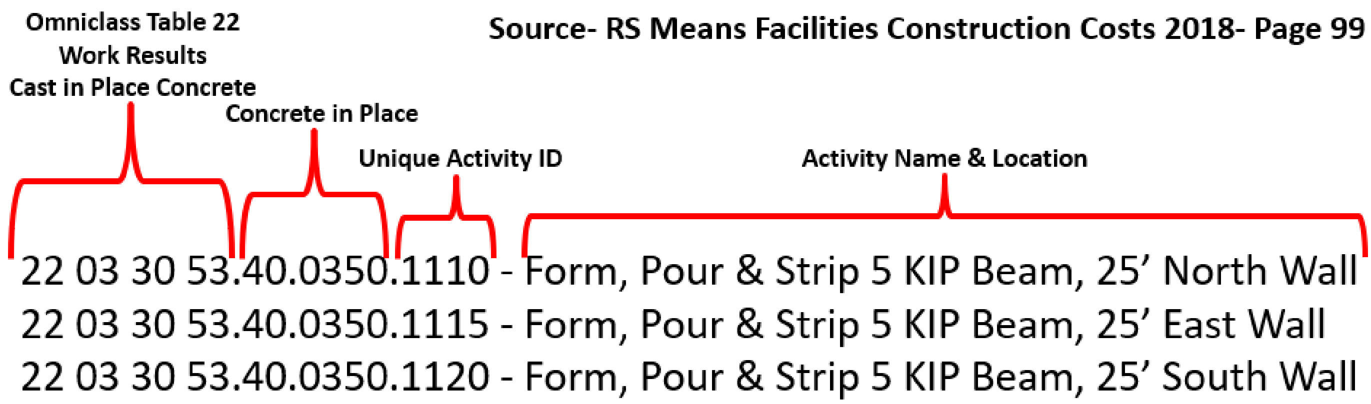


Figure 10- Case Study Using Commercial Database Information

Source: Giammalvo, Paul D (2015) Course Materials Adapted from R.S. Means 2018 Facility Cost Estimating Database. Contributed Under Creative Commons License BY v 4.0

In the example above, which shows a classic example of Activity Based Costing (ABC) at Level 4, 5, or 6 commonly used by contractors but also shows how using, in this case, Masterformat or Omniclass Table 22, the activity costs can be "rolled up" to whichever level is deemed appropriate or necessary for use in "Rolling Wave Planning." (For more on the importance of these coding structures, see also Units 7.10, Managing Horizontal and Vertical Traceability (Time) and **Unit 10- Managing Cost Estimating and Budgeting 10**.8- Managing Horizontal and Vertical Traceability (Costs):

- **(10.1) 03 Concrete** is CSI Masterformat or CSI Omniclass Table 22 2nd Level WBS Structure
- **(10.2)** 03 30 Cast In Place Concrete is CSI Masterformat/Omniclass Table 22 3rd Level WBS Structure. (Level 1 in Table 22, Level 2, 03 is Concrete and Level 3 is 03 30 is Cast in Place Concrete)
- **(10.3) 03 30 53 Miscellaneous Cast in Place Concrete** is CSI Masterformat/Omniclass Table 22 4th Level WBS (Level 1, Table 22, Level 2 is 03, Concrete, Level 3, 03 30 is Cast in Place Concrete)
- **(10.4) 03 30 53.40- Concrete in Place** is CSI Masterformat/Omniclass Table 22 Level 5 WBS, which from this level down becomes not only a WBS but a Cost Breakdown Structure (CBS) and Productivity Breakdown Structure (PBS) as well. One coding structure serving three purposes. Worth noting is Level 3 of the WBS is the MINIMUM level of detail that an owner should be

providing to a contractor, assuming the owner wants to minimize claims and disputes while at the same time obtaining highly competitive bids.

- **(10.4) 03 30 53.40.0010, 03 30 53.40.0020 and 03 30 53.40.0050 is an ACTIVITY** consisting of Forming (0.0010); Installing Rebar (0.0020) and Placing and Finishing Concrete (0.0050)
- **(10.4)** There are two types of Activity "03 30 53.40.0300 is based on 10' (foot) long spans" and Activity **"03 30 53.40.0350 is based on 25' (foot) long spans"**. Depending on the number of scenarios, you could create other cost and productivity calculations for as many different combinations as you are likely to use.
- **(10.5)** For the remaining example, we will be using **Activity 03 30 53.40.0350 to form, reinforce, place concrete and strip 25' long, 5 kip (5,000 PSI concrete) beams.** Notice that the only information MISSING from this is the Quantity Take-Off or Bill of Materials? When we start to use Building Information Modelling (BIM), this information will be coming to us using these coding structures. If we do not use BIM, we have to do it the old-fashioned way- doing quantity take-offs using manual methods.
- **(10.6)** This is where we know the composition of the crew that was used to calculate the productivity and costs. For this example, we will use **Crew C-14A**. See below for a more detailed look at what Crew C-14A is made of.
- **(10.7)** This Is simply the daily output **Crew C-14A** can produce ON AVERAGE. (P50) value. It is not adjusted for any risks. This is one of the most important pieces of information that the Planner/Scheduler needs as this is how we calculate the DURATION, and the Cost Estimator needs to know and understand to estimate the COSTS. Crew C-14A can produce, on average, 18.55 Cubic Yards (CY) per day for this particular activity. Different activities will have different productivity rates.
- **(10.8)** Is the number of **Crew Labor Hours per Cubic Yard (CY)** of beam concrete. This is another useful piece of information for all project control professionals but especially planners/schedulers, as many projects are not tracked based on money but on person-hours expended or earned vs. planned Person-hours. This was covered in Unit 11- Managing Progress. On average, it takes 10.78 labor hours per Cubic Yard (CY) in place for this particular activity. Different activities will have different productivity rates.
- **(10.9)** This is the Unit of Measure. In this case, it **is Cubic Yards (CY),** but it could have been Cubic Meters (M3) or any other fast and reliable way to measure physical progress.
- **(10.10)** This is the Field where we enter **Material Costs.** As material costs tend to be location-specific, the professional cost estimate needs to keep this updated and adjust these values for different locations. The Material Costs for this activity is $355.00 per Cubic Yard of Concrete in place. It is ESSENTIAL that the project controller validate this number.
- **(10.11) Labor Costs-** As with Material costs, labor costs are highly variable and need to be checked and validated by the cost estimator/project controller for each location and each trade. This is the weakest part of any of these systems as the data has to be VERY localized down to each major city or metropolitan area. These should be updated not more than quarterly and in the event of high inflation (which is already starting in many countries) needs to be done monthly, weekly or in extreme cases like Zimbabwe or Venezuela) daily. The Labor Costs for this activity is $545.00 per Cubic Yard of Concrete in place. This means the crew composition needs to be checked and

validated, as does the labor costs of each person on the crew, as it is unlikely they are all being paid the same hourly rate.

- **(10.12) Equipment Costs**- Tend to be less variable than material or labor costs and must be updated at least semi-annually and preferably quarterly. Equipment productivity tends to be relatively stable and predictable. The Equipment Cost for this activity is $410.00 per Cubic Yard of Concrete in place
- **(10.13) Total Costs is the sum of 11, 12, and 13.** The Total Cost per Cubic Yard of Concrete in Place is $949.00.
- **(10.14) Marked Up Costs (Contractors Selling Price)** is shown in this column and, as we can see, ranges from about 44% to 47% to cover Project Overhead, Home Office Overhead and all the other items shown above in Figure XX R.S. Means 2008 Facility Cost Estimating Database Back Cover Showing Labor Rate Markups. The Fair Market Value or Contractors Selling Price for each Cubic Yard of Concrete in Place is $1,325, including a markup for OH&P of ~40% (15% Home Office Overhead + 15% for Project Overhead + 10% for Contractors Gross Profit = 40%)

This means that once we have the STANDARDIZED the Records and Fields, it becomes relatively easy to enter and update the actual costs and productivity to fit local conditions.

Knowing the base composition of each gang or crew makes it easy to adjust the crew composition to fit local practices, labor laws, or union agreements.

			Subcontractors Bare Costs		Including Subcontractors OH & P		Prime Contractors Cost Per Labor Hour	
(1) #	(2) Crew C-14A	(3) Standard Occupational Code (SOC)	(4) Hourly	(5) Daily	(6) Hourly	(7) Daily	(8) Bare Costs	(9) Billing Rate Including OH & P
1	Carpenter Foreman	47-1011	$52.70	$421.60	$85.55	$684.40	$50.54	$81.64
16	Carpenters	47-2030	$50.70	$6,489.60	$82.30	$10,534.40		
4	Rodmen	47-2171	$54.65	$1,748.80	$87.30	$2,793.60		
2	Laborers	47-2061	$39.85	$637.60	$64.70	$1,035.20		
1	Cement Finisher	47-2051	$47.55	$380.40	$75.20	$601.60		
1	Equipment Operator	47-2073	$53.75	$430.00	$84.80	$678.40		
1	Gas Engine Vibrator			$25.80		$28.16		
1	Concrete Pump (Small)			$881.60		$969.76	$4.54	$4.99
200	Total Daily Labor Hours	25 X 8 = 200 Hrs @100% Efficiency		$11,015.40		$17,325.52	$55.08	$86.63
(10)		(11)		(12)		(13)	(14)	(15)

Figure 11 - Case Study Demonstrating Crew Composition Details

Source: Giammalvo, Paul D (2015) Course Materials Adapted from R.S. Means 2008 Facility Cost Estimating Database. Contributed Under Creative Commons License BY v 4.0

- ✓ **(11.1), (11.2),** and **(11.3)** Crew C-14A consists of the 8 Labor and Equipment items; 1 carpenter foreman plus 16 carpenters, 4 rodmen, 2 laborers, and 1 equipment operator.
- ✓ **(10.10)** This crew of 25 people equals 25-man days of labor, and assuming they are working an 8 hour day = 200 Person-hours of labor per crew working day.

- ✓ **(10.12)** The bare COST of this crew is $11,015.40 per day (bare costs are wages and fringe benefits for the labor and the EXPENCED costs of the equipment)
- ✓ **(10.13)** This is what the PRIME CONTRACTOR has to CHARGE for this crew to work one day
- ✓ **(10.14)** This is the additional amount the PRIME CONTRACTOR has to add to cover his/her Overhead and Profit on his Foreman and the Concrete pump the prime contractor provided. Thus the PRIME CONTRACTOR would have to take the $17,325.52 and ADD the OH&P of $86.63 for his Foreman and $55.08 for his Pump for a total of $17,467.23 per 8 hour working day. **(10.15)** And don't forget that is a P50 value that ASSUMES working at 100% productivity of 18.55 CY per day or 10.78 Crew Hours Per Cubic Yard. If YOUR crews are not meeting these PRODUCTIVITY ASSUMPTIONS, you need to adjust your bid price. (Could be UP if your productivity is LOWER, or you can REDUCE the price if your productivity is HIGHER.)

Having established and kept current a **PROJECT CONTROLS DATABASE** when we have created a schedule, then we have to draw from this database to create our **RESOURCE POOL** or **RESOURCE ASSIGNMENTS** or **RESOURCE DICTIONARY** to justify costs, budgets, and durations, etc.

OUTPUTS

- ✓ A Cost Estimating and Productivity Database Which Provides Accurate, Reliable And Precise Cost And Duration Estimates, Appropriately "Fit For Purpose."

UNIT 13.4- CREATING THE PROJECT DATABASE(S)

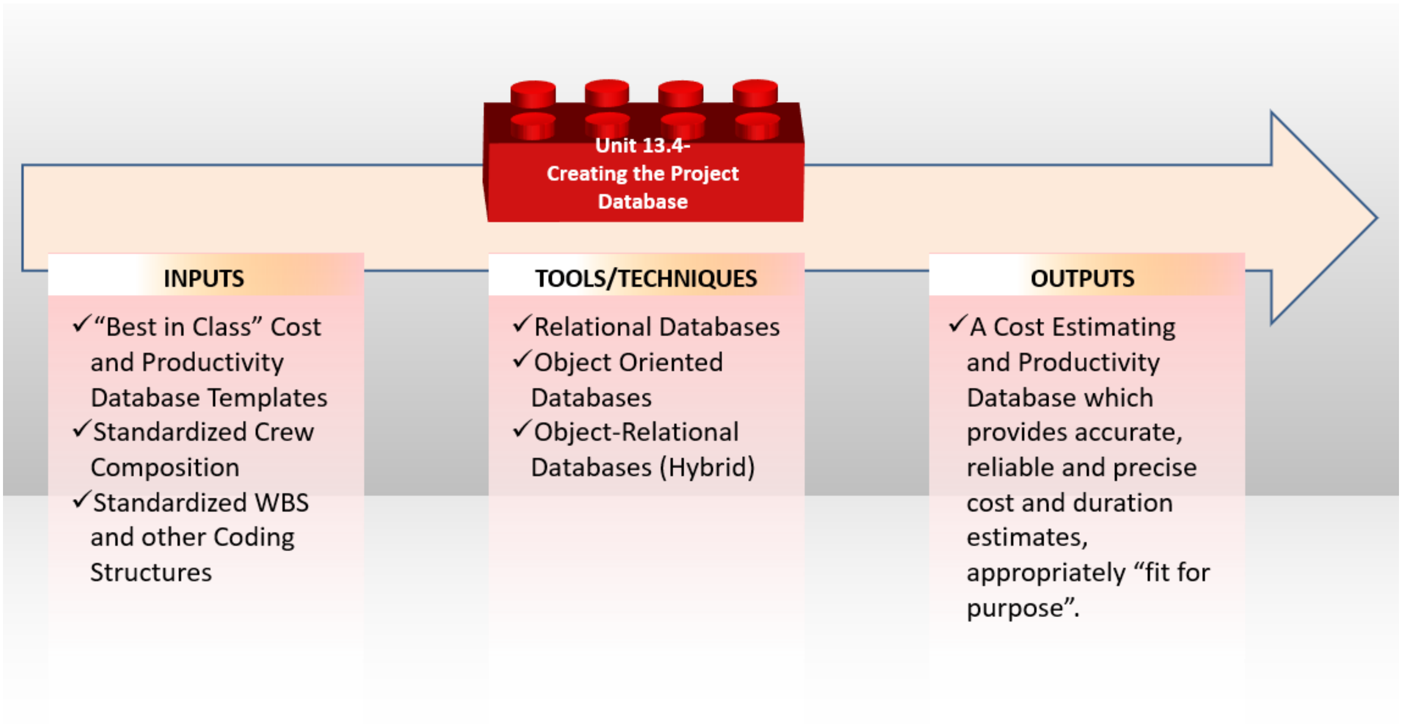


Figure 12 - The Creating the Project Databases Process Map
Source: PTMC Team

INTRODUCTION

The best way to demonstrate how to create a project database is by showing a real example of one created in Excel, understanding that knowing the Records and Fields required can start with this template and modify or adapt it to suit your needs. Here is the URL to download this template, and you can use it to follow along - Class A Cost Estimate Template

INPUTS

- ✓ "Best In Class" Cost And Productivity Database Templates
- ✓ Standardized Crew Composition
- ✓ Standardized WBS And Other Coding Structures
- ✓ Historical Productivity Data
- ✓ Historical Cost Data

TOOLS & TECHNIQUES

- ✓ **Cost & Productivity Database Home Page/Demographics**

As seen from the example below, we have many fields that should be included in our own internal databases. This is important because when we are tendering / budding for new projects, we need to know not only the location but the year of construction, time of year the construction was being done, and any other information which will enable us to select comparable projects which are as close to the new project as possible.

By including as much demographic data as possible, the objective is to quickly sort through what can grow to be a very large database and find as many projects as possible, similar to the one you are currently estimating. Consistent with fundamental statistical theory, the larger the sample population you can find, the smaller the variance is likely to be. The smaller the variance, the more ACCURATE, PRECISE, and RELIABLE the cost estimate you produce is likely to be.

This should include a brief narrative, the names of the project manager, foreperson, or other key people on the project, as well as any other keywords that might help people who may not be familiar with the project to be able to come as close as possible to matching it with the project they are now bidding. The more comparable the projects from the database can match the project being bid, the more likely you will not only win the bid but be able to make money on it.

The database should also include the "Lessons Learned" on each project. Things that went RIGHT and things that went WRONG.

For those who want to include photos, scanned documents, audio or video files, these too can be embedded into the spreadsheet. This is also where you would put the KEYWORDS if you have set up your database to be searched and filtered based on keywords.

There are many cost estimating templates for both owners and contractors available either in "hard copy" (paper-based) or, more commonly, spreadsheets.

The best examples for both owner and contractor that the PTMC has been able to locate in our research are the templates provided at no cost and under open source licensing are those offered by the US Parks Department.

PROJECT INFORMATION		
Project:	Project Name	
Park:	Park Name	
Park Apha:	Park Alpha Code	
PMIS:	TBD or PMIS number if known	
Estimate Da	Estimate Date	
Prepared By	Estimator Name	
Company:	Company	
Address:	Address	
City, State	City	
Phone:	Phone	
BACKGROUND SUPPORTING MATERIAL (Scope of Work):		
Briefly describe scope of work included in the estimate, plan dates, exclusions, etc. List any Government Furnished Property (GFP)		
SOURCE OF COST DATA:		
Document all sources of cost information used in the estimate. (Attach additional information if necessary)		
ESTIMATE ASSUMPTIONS:		
Describe any assumptions made to prepare estimate and highlight areas needing clarification for future estimates.		
MAJOR CHANGES FROM PREVIOUS ESTIMATE:		
Describe any major changes in scope of work, materials, systems, assumptions, etc. from the previous estimate version.		
DESCRIPTION OF MARK-UP & ADD-ONS:		
Design Contingency:	0.00%	Explain & Justify
Standard. General Condition	0.00%	Explain & Justify
Government General Conditi	0.00%	Explain & Justify
Contractor Overhead:	0.00%	Explain & Justify
Contractor Profit:	0.00%	Explain & Justify
Contracting Method Adjustn	0.00%	Describe anticipated contract method
Annual Inflation Escalation F	0.00%	Projected annual inflation rate.
Time Until Project Midpoint	0	Number of months from estimate (or
OTHER COMMENTS:		
Provide any additional information, qualifications, etc.		

Basis of Estimate | Project Cost Summary | Bid Ite

Figure 13- Basis For Estimate Home Page (For a completed sample, go here)
Source: US National Park Service Cost Estimating Handbook (2011)

United States Department of the Interior
National Park Service
Class A Construction Cost Estimate

PROJECT COST SUMMARY

Project:	Oso Comida Trailhead Improvements:	Estimate By:	YIB
Park:	Bear Arbor NRA	Date:	01/12/11
Park Alpha:	BEAR	Reviewed By:	BBB
PMIS Number:	XXXXXX	Date:	01/17/11

Bid Item No.	Bid Item Description		Total Material Cost	Total Labor Cost	Total Equipment Cost	Total Direct Construction Costs	Design Contingency	General Conditions	General Contractor Overhead	General Contractor Profit	Contracting Method Adjustment	Inflation Escalation	Bid Item Total
							2.00%	3.00%	8.50%	10.00%	15.00%	APR 3.60% Month 32	
Bid Item: 1	Replace Pit Toilets with New Comfort Station						TOTAL VALUE OF GOVERNMENT FURNISHED PROPERTY (if any):						$ 46,000.00
	A10	Foundations	$ 30,028	$ 33,082	$ 7,293	$ 70,403							
	A20	Basement Construction	$ -	$ -	$ -	$ -							
	B10	Superstructure	$ 15,622	$ 13,198	$ 460	$ 29,280							
	B20	Exterior Enclosure	$ 35,992	$ 29,477	$ -	$ 65,469							
	B30	Roofing	$ 18,471	$ 8,706	$ -	$ 27,177							
	C10	Interior Construction	$ 25,573	$ 9,308	$ -	$ 34,881							
	C30	Interior Finishes	$ 4,476	$ 13,424	$ -	$ 17,900							
	D20	Plumbing Systems	$ 26,655	$ 16,121	$ -	$ 42,776							
	D30	HVAC	$ 1,269	$ 1,170	$ -	$ 2,439							
	D50	Electrical	$ 8,753	$ 9,366	$ -	$ 18,119							
	F20	Selective Building Demolition	$ 463	$ 1,990	$ 3,862	$ 6,315							
	G10	Site Preparation	$ 2,188	$ 4,362	$ 6,952	$ 13,502							
	G20	Site Improvements	$ 8,900	$ 7,300	$ -	$ 16,200							
	G30	Site Mechanical	$ 86,213	$ 32,582	$ 44,542	$ 163,337							
	G40	Site Electrical	$ 5,000	$ -	$ -	$ 5,000							
	XX	Standard General Conditions	$ 31,900	$ 101,200	$ 18,610	$ 151,710							
Total - Bid Item 1	Replace Pit Toilets with New Comfort Station		$ 301,503	$ 281,286	$ 81,719	$ 664,508	$ 12,370	$ 18,926	$ 55,233	$ 64,980	$ 122,403	$ 92,813	$ 1,031,234
Bid Item: 2	Construct New Parking Lot & Site Utilities						TOTAL VALUE OF GOVERNMENT FURNISHED PROPERTY (if any):						$ -
	G10	Site Preparation	$ 2,500	$ 11,711	$ 19,776	$ 33,987							
	G20	Site Improvements	$ 143,581	$ 36,335	$ 43,670	$ 223,586							
	G30	Site Mechanical	$ 12,153	$ 14,232	$ 4,241	$ 30,626							
	XX	Standard General Conditions	$ 12,925	$ 8,350	$ 6,500	$ 27,775							
			$ -	$ -	$ -	$ -							
Total - Bid Item 2	Construct New Parking Lot & Site Utilities		$ 171,159	$ 70,628	$ 74,187	$ 315,974	$ 6,319	$ 9,669	$ 28,217	$ 33,196	$ 59,006	$ 44,742	$ 497,123
Bid Item: 3	Picnic Area & Trailhead Improvements						TOTAL VALUE OF GOVERNMENT FURNISHED PROPERTY (if any):						$ -
	G10	Site Preparation	$ -	$ 11,860	$ 4,845	$ 16,705							
	G20	Site Improvements	$ 59,448	$ 25,960	$ 12,270	$ 97,678							
	G30	Site Mechanical	$ 2,125	$ 2,275	$ 330	$ 4,730							
	XX	Standard General Conditions	$ 5,775	$ 7,550	$ 2,500	$ 15,825							
			$ -	$ -	$ -	$ -							
Total - Bid Item 3	Picnic Area & Trailhead Improvements		$ 67,348	$ 47,645	$ 19,945	$ 134,938	$ 2,699	$ 4,129	$ 12,050	$ 14,177	$ 25,199	$ 19,107	$ 212,299
		Total Bid Items 1-3	$ 540,010	$ 399,559	$ 175,851	$ 1,115,420	$ 21,388	$ 32,724	$ 95,500	$ 112,353	$ 206,608	$ 156,662	$ 1,740,656

Figure 14 - Summary Level Cost Estimating Template (Owners)

Source: US National Park Service Cost Estimating Handbook (2011)

The example above shows what any cost estimating database should contain for information, whether Owner or Contractor:

- ✓ **(14.1) Activity Name-** Below are some examples of well-written Activity Names: Below are some examples of well-written Activity Names:

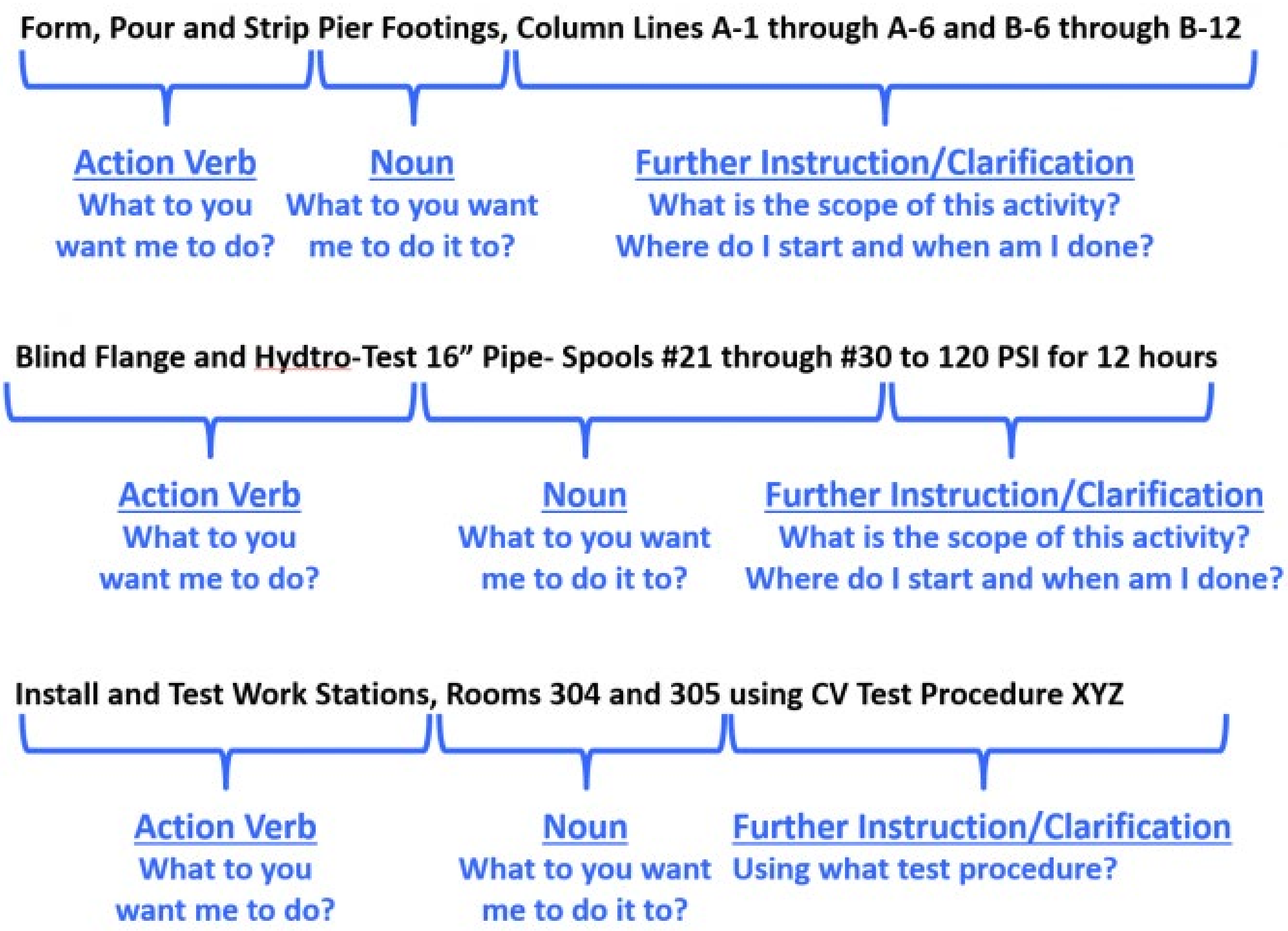


Figure 15 - Well Written Activity Names
Source: Giammalvo, Paul D (2015) Course Materials Contributed Under [Creative Commons License BY v 4.0]

However, in the example shown above, because this is an owner's cost summary, they rolled it up by bid items rather than by the more detailed activity names likely to be used by contractors or owners for any "self-performed" work.

- ✓ **(14.2) Coding Structure Sub Sort**- In this case, the US Parks Department has opted to use CSI's Uniformat as the basis to "roll up" their project costs. For more examples of CSI, Uniformat download this reference, which takes you to Level 3 of Uniformat for each of the main headings, or you can download the Omniclass Table 21, which takes you to Level 5 for all headings.

Worth noting is that OWNER's generally like using Uniformat/CSI Table 21 as it enables them to develop databases that are useful in the early phases of the project to develop cost estimates, while on the other hand, contractors tend to prefer using Masterformat/CSI Table 22 as it provides for a much greater level of detail (Activity) than does Uniformat, which identifies the components or elements of an asset.

- ✓ **(14.3) Material Cost-** This should be self-explanatory, but for cost estimators, we need to be sure to check as materials prices vary significantly depending on location. The more remote the site, generally the more expensive materials are because of shipping and storage costs.
- ✓ **(14.4) Labor Costs-** Another largely self-explanatory heading, but again, many factors go into calculating labor costs, not the least of which is the actual productivity.
- ✓ **(14.5) Equipment Costs-** Self-explanatory with the note that getting equipment to and from any given site (mobilization and demobilization costs) can often be significant, and that has to be factored into the equipment costs along with the daily or hourly rental fee or cost of ownership.
- ✓ **(14.6)** Total Direct Costs- Simply the sum of 3, 4, and 5 above. In this example above, because the contract is a cost-plus type, the owner has every right to request and receive this information. However, if the contract were being let on a "firm fixed price" basis, the owner would never see this level of detail. However, the contractor should have gone through the same process.

(14.7) Design Contingency- This is a RISK ALLOWANCE to cover the probability that the contractor cannot provide exactly what was specified or what was specified cannot be obtained at the price it was bid at.

- ✓ **(14.8) General Conditions-** This is what is known as the "Project Indirect" costs and covers things like the fencing/hoarding around the project, the site offices, electricity, fuel, QA/QC, Safety, Protective Equipment, etc. and other items identified in CSI Division 1/CSI Table 22 "General Conditions."
- ✓ **(14.9) Contractors Home Office Overhead-** This is a very real yet often contentious expense, which covers the salaries and facilities associated with the contractor's home office. As this is generally considered a fixed expense, the percentage allocated to any project can vary, depending on the volume of work.
- ✓ **(14.10) Contractors Profit Margin-** As noted previously, single-digit EBIT margins are the norm for contractors worldwide. So even if you go in with a 10% target, every mistake, error, or omission the contractor makes comes out of that amount. This is why we explain that for a contractor, his/her profit margin is the "Management Reserve."
- ✓ **(14.11) Contracting Method Adjustment-** This too is a "risk contingency" adjustment applied at the project level (as opposed to activity level), which covers such risk events as remote site construction, labor shortages/inefficiencies, or working in adverse climates, either very hot and/or humid or very cold and dry. Again, while it is unusual to see an owner organization recognizing this, if you are an owner's project control professional, you need to recognize that this adjustment is or should be made by your contractors to put together their cost estimate for bidding.
- ✓ **(14.12) Inflation Adjustment Factor-** Again, self-explanatory with the caution that we tend to underestimate what it really is. Given most governments lie about what the real or true inflation rate is in their country (the US underestimates inflation by a factor of 50%), the competent cost or project controls practitioner will take material and labor prices over a period of time and use those to project into the future what the real or true inflation rate is likely to be. Also, for those who are working on International projects, don't forget to factor in the exchange rate fluctuations. Those often have a far worse impact than inflation, especially in today's global marketplace.
- ✓ **(14.13) Marked Up "Selling" price-** This is a summation of the direct costs **(14.6)** plus the adjustments **(14.7-14-11)** to give us the CONTRACTORS SELLING PRICE, which, when the work has been done and is billed by the contractor, becomes the OWNERS ACTUAL COST OF the WORK PERFORMED (ACWP or AC)

- ✓ **Description of Mark-Ups and Contingency (Cost Reimbursable Contract)**

As for owners, a project is a cost or investment center, and they do not mark up the price for profit; however, they DO need to mark up the quote the contractor submits to cover:

- Owner's Project Management Overhead
- Owner's Home Office Overhead (i.e., Finance charges)
- Owner-Supplied Equipment/Materials or Services
- Owner Contingency (NOT Management Reserve as that does not belong to the project unless asked for and approved by management) Keep in mind that for a contractor, the profit margin is their "management reserve."

In the example below, we see a Level 3 Cost Estimate. This is the level contractors would normally provide to the owner under, and this is the level of detail they would normally report their progress against as well as bill against.

OmniClass defines an Element to be "a major component, assembly, or "construction entity part which, in itself or in combination with other parts, fulfills a predominating function of the construction entity" (ISO 12006-2). Predominating functions include, but are not limited to, supporting, enclosing, servicing, and equipping a facility. Functional descriptions can also include a process or an activity. A Designed Element is an "Element for which the work result(s) have been defined." (ISO 12006-2)."

Assuming we are using a relational, object-oriented, or hybrid database, we can assign MORE than one code, thus enabling multiple sorts and/or combinations of sorts.

Using the example shown in Figure 15 below, we could FILTER, SORT, and SUMMARIZE by any of the Fields shown across the top PLUS those fields shown in the ROWS:

- Bid Item
- Uniformat or Omniclass Table 21- Elements
- Masterformat or Omniclass Table 22- Work Results

Bid Item No.	Bid Item Description		Total Material Cost	Total Labor Cost	Total Equipment Cost	Total Direct Construction Costs	Design Contingency 2.00%	General Conditions 3.00%	General Contractor Overhead 8.50%	General Contractor Profit 10.00%	Contracting Method Adjustment 15.00%	Inflation Escalation (APR 3.60%, Month 32)	Bid Item Total
Bid Item: 1	Replace Pit Toilets with New Comfort Station						TOTAL VALUE OF GOVERNMENT FURNISHED PROPERTY (if any):						$ 46,000.00
	A10	Foundations	$ 30,028	$ 33,082	$ 7,293	$ 70,403							
	A20	Basement Construction	$ -	$ -	$ -	$ -							
	B10	Superstructure	$ 15,622	$ 13,198	$ 460	$ 29,280							
	B20	Exterior Enclosure	$ 35,992	$ 29,477	$ -	$ 65,469							
	B30	Roofing	$ 18,471	$ 8,706	$ -	$ 27,177							
	C10	Interior Construction	$ 25,573	$ 9,308	$ -	$ 34,881							
	C30	Interior Finishes	$ 4,476	$ 13,424	$ -	$ 17,900							
	D20	Plumbing Systems	$ 26,655	$ 16,121	$ -	$ 42,776							
	D30	HVAC	$ 1,269	$ 1,170	$ -	$ 2,439							
	D50	Electrical	$ 8,753	$ 9,366	$ -	$ 18,119							
	F20	Selective Building Demolition	$ 463	$ 1,990	$ 3,862	$ 6,315							
	G10	Site Preparation	$ 2,188	$ 4,362	$ 6,952	$ 13,502							
	G20	Site Improvements	$ 8,900	$ 7,300	$ -	$ 16,200							
	G30	Site Mechanical	$ 86,213	$ 32,582	$ 44,542	$ 163,337							
	G40	Site Electrical	$ 5,000	$ -	$ -	$ 5,000							
	XX	Standard General Conditions	$ 31,900	$ 101,200	$ 18,610	$ 151,710							
Total - Bid Item 1	Replace Pit Toilets with New Comfort Station		$ 301,503	$ 281,286	$ 81,719	$ 664,508	$ 12,370	$ 18,926	$ 55,233	$ 64,980	$ 122,403	$ 92,813	$ 1,031,234

Bid Item Direct Costs SUMMARIZED using CSI Uniformat or Omniclass Table 21 Elements.

We can sort and summarize by Multiple Levels of DETAIL and also by CSI Masterformat/Omniclass Table 22 Work Results

Figure 15 - Cost Summary Based on BID ITEM sub-sorted by CSI Uniformat Coding Structure (OmniClass Table 21 Elements)

Source: US National Park Service Cost Estimating Handbook (2011)

Figure 16 below shows us the flexibility we have when we use relational databases to create Pivot Tables enabling us to view our cost and productivity data (and any other data) in multiple ways depending on how the stakeholders want to see it.

Bid Item Number	Asset / Project Element / Description	Size/Count	Units
BID ITEM 1	Construct Wet Comfort Station	1000	SF

Item No.	WBS	Description	Material Cost/Unit	Total Material Cost	Labor Cost/Unit	Total Labor Cost	Equipment Cost/Unit	Total Equipment Cost	Direct Cost/Unit	Total Direct Costs	NET Cost/Unit	Total NET Costs
1	A10	Foundations	$ 30.03	$ 30,028	$ 33.08	$ 33,082	$ 7.29	$ 7,293	$ 70.40	$ 70,402	$ 109.26	$ 109,256
2	A20	Basement Construction - INC. IN FOUNDATION	$ -	$ -	$ -	$ -	$ -	$ -	$ -	$ -	$ -	$ -
3	B10	Superstructure	$ 15.62	$ 15,622	$ 13.20	$ 13,198	$ 0.46	$ 460	$ 29.28	$ 29,279	$ 45.44	$ 45,438
4	B20	Exterior Enclosure	$ 35.99	$ 35,992	$ 29.48	$ 29,477	$ -	$ -	$ 65.47	$ 65,469	$ 101.60	$ 101,600
5	B30	Roofing	$ 18.47	$ 18,471	$ 8.71	$ 8,706	$ -	$ -	$ 27.18	$ 27,176	$ 42.17	$ 42,174
6	C10	Interior Construction	$ 25.57	$ 25,573	$ 9.31	$ 9,308	$ -	$ -	$ 34.88	$ 34,881	$ 54.13	$ 54,131
8	C30	Interior Finishes	$ 4.48	$ 4,476	$ 13.42	$ 13,424	$ -	$ -	$ 17.90	$ 17,900	$ 27.78	$ 27,778
9	D20	Plumbing Systems	$ 26.65	$ 26,655	$ 16.12	$ 16,121	$ -	$ -	$ 42.78	$ 42,776	$ 66.38	$ 66,383
10	D30	HVAC	$ 1.27	$ 1,269	$ 1.17	$ 1,170	$ -	$ -	$ 2.44	$ 2,438	$ 3.78	$ 3,784
11	D50	Electrical	$ 8.75	$ 8,753	$ 9.37	$ 9,365	$ -	$ -	$ 18.12	$ 18,118	$ 28.12	$ 28,117
12	F20	Selective Building Demolition	$ 0.46	$ 463	$ 1.99	$ 1,990	$ 3.86	$ 3,863	$ 6.32	$ 6,315	$ 9.80	$ 9,800
13	G10	Site Preparation	$ 2.19	$ 2,188	$ 4.36	$ 4,362	$ 6.95	$ 6,952	$ 13.50	$ 13,502	$ 20.95	$ 20,953
14	G20	Site Improvements	$ 8.90	$ 8,900	$ 7.30	$ 7,300	$ -	$ -	$ 16.20	$ 16,200	$ 25.14	$ 25,140
15	G30	Site Mechanical Utilities	$ 86.21	$ 86,213	$ 32.58	$ 32,582	$ 44.54	$ 44,543	$ 163.34	$ 163,337	$ 253.48	$ 253,479
16	G40	Site Electrical Utilities	$ 5.00	$ 5,000	$ -	$ -	$ -	$ -	$ 5.00	$ 5,000	$ 7.76	$ 7,759
17	XX	General Conditions	$ 31.90	$ 31,900	$ 101.20	$ 101,200	$ 18.61	$ 18,610	$ 151.71	$ 151,710	$ 235.44	$ 235,435
18	XX		$ -	$ -	$ -	$ -	$ -	$ -	$ -	$ -	$ -	$ -
19	XX		$ -	$ -	$ -	$ -	$ -	$ -	$ -	$ -	$ -	$ -
		Subtotal Direct Construction Costs	$ 301.50	$ 301,501	$ 281.28	$ 281,284	$ 81.72	$ 81,720	$ 664.50	$664,505	$ 1,031.23	$1,031,229
		Total Value of Government Furnished Property (GFP) Inc. in Direct Cost		$46,000.00		$ -		$ -	$ 46,000	$46,000	In most cases GFP is normally zero - see footnote-	
		Direct Cost Subtotal without GFP		$ 255,501		$ 281,284		$ 81,720		$618,505		
		Design Contingency	2.00%							$12,370		
		Total Direct Construction Costs								$676,875		
		Standard General Conditions	0.00%	Applied to Total Direct Construction Cost less GFP						$0		
		Government General Conditions	3.00%	Applied to Total Direct Construction Cost less GFP						$18,926		
		Subtotal NET Construction Cost								$695,801		
		Overhead	8.50%							$55,233		
		Profit	10.00%							$64,980		
		Estimated NET Construction Cost								$816,014		
		Contracting Method Adjustment	15.00%							$122,402		
		Inflation Escalation	32	Months	Annual Rate =	3.60%				$92,812		
		Total Estimated NET Cost of Construction								$1,031,229		

Notes & Comments:
Building only direct cost = $308.44/sf
Building total NET cost = $474.28/sf
GFP Septic King Treatment System Pre-Purchased by Government = $46,000

917 **Figure 16 - Bid Item Detail Showing Detail by CSI Uniformat Coding Structure (OmniClass Table 21 Elements)**
918 Source: US National Park Service Cost Estimating Handbook (2011)

In this example, you can see that the level of detail is quite extensive, which is the same level of detail that a contractor would be expected to create for their in-house SUMMARY level database. (Level 4 of the WBS)

03 Concrete										
03 30 Cast in Place Concrete										
03 30 53 Miscellaneous Cast in Place Concrete										
03 30 53.40 Concrete in Place		1	2	3		4	5	6	7	8
0.0010 0.0020 0.0050	Including Forms (4 uses), reinforcing steel, concrete placement and finishing, unless otherwise indicated.	Crew Type	Daily Output per Unit	Labor Hours per Unit	Unit of Measure	Material Costs	Labor Costs	Equip-ment Costs	Total Costs per Unit	Total Price/Unit Including OH&P
0.0300	Beams- 5 kip per lineal foot, 10' long spans	C14-A	15.62	12.8	Cubic Yard (CY)	$315.00	$490.00	$48.50	$853.50	$1,225.00
0.0350	Beams- 5 kip per lineal foot, 25' long spans	"	18.55	10.78	CY	$325.00	$415.00	$40.50	$780.50	$1.100.00

Figure 17 - Modifying the US Park Cost Estimating Database for Use in Scheduling Databases
Source: Giammalvo, Paul D (2015) Course Materials Contributed Under Creative Commons License BY v 4.0

For those wanting to adopt/adapt this model, you could follow the R.S. Means template shown above, and adding fields named "Crew," "Daily Output," and "Labor Hours" to the Excel template would enable you to link this to your schedule software database.

Project:	Project Name	Estimate By:	:stimator Name
Park:	Park Name	Date:	Estimate Date
Park Alpha:	Park Alpha Code	Reviewed By:	
PMIS:	TBD or PMIS number if known	Date:	

Summary Iten C10 Interior Construction — Total Cost: $0

Uniformat II WBS Code	Description	Quantity	Unit	MATERIAL		LABOR		EQUIPMENT		TOTALS	
				Material Cost/Unit	Total Material Cost	Labor Cost/Unit	Total Labor Cost	Equipment Cost/Unit	Total Equipment Cost	Total Cost/Unit	Total Cost
C1010	INTERIOR PARTITIONS										
MF-2004 Code	Description	0	Unit	$ -	$0	$ -	$0	$ -	$0	$ -	$0
MF-2004 Code	Description	0	Unit	$ -	$0	$ -	$0	$ -	$0	$ -	$0
MF-2004 Code	Description	0	Unit	$ -	$0	$ -	$0	$ -	$0	$ -	$0
MF-2004 Code	Description	0	Unit	$ -	$0	$ -	$0	$ -	$0	$ -	$0
MF-2004 Code	Description	0	Unit	$ -	$0	$ -	$0	$ -	$0	$ -	$0
MF-2004 Code	Description	0	Unit	$ -	$0	$ -	$0	$ -	$0	$ -	$0
MF-2004 Code	Description	0	Unit	$ -	$0	$ -	$0	$ -	$0	$ -	$0
MF-2004 Code	Description	0	Unit	$ -	$0	$ -	$0	$ -	$0	$ -	$0
SUBTOTAL	INTERIOR PARTITIONS	0	Unit	#DIV/0!	$0	#DIV/0!	$0	#DIV/0!	$0	#DIV/0!	$0

Uniformat II WBS Code	Description	Quantity	Unit	MATERIAL		LABOR		EQUIPMENT		TOTALS	
				Material Cost/Unit	Total Material Cost	Labor Cost/Unit	Total Labor Cost	Equipment Cost/Unit	Total Equipment Cost	Total Cost/Unit	Total Cost
C1020	INTERIOR DOORS										
MF-2004 Code	Description	0	Unit	$ -	$0	$ -	$0	$ -	$0	$ -	$0
MF-2004 Code	Description	0	Unit	$ -	$0	$ -	$0	$ -	$0	$ -	$0
MF-2004 Code	Description	0	Unit	$ -	$0	$ -	$0	$ -	$0	$ -	$0
MF-2004 Code	Description	0	Unit	$ -	$0	$ -	$0	$ -	$0	$ -	$0
MF-2004 Code	Description	0	Unit	$ -	$0	$ -	$0	$ -	$0	$ -	$0
MF-2004 Code	Description	0	Unit	$ -	$0	$ -	$0	$ -	$0	$ -	$0
MF-2004 Code	Description	0	Unit	$ -	$0	$ -	$0	$ -	$0	$ -	$0
MF-2004 Code	Description	0	Unit	$ -	$0	$ -	$0	$ -	$0	$ -	$0
SUBTOTAL	INTERIOR DOORS	0	Unit	#DIV/0!	$0	#DIV/0!	$0	#DIV/0!	$0	#DIV/0!	$0

Summary Iten C10 Interior Construction

Uniformat II WBS Code	Description	Quantity	Unit	MATERIAL		LABOR		EQUIPMENT		TOTALS	
				Material Cost/Unit	Total Material Cost	Labor Cost/Unit	Total Labor Cost	Equipment Cost/Unit	Total Equipment Cost	Total Cost/Unit	Total Cost
C10	Interior Construction	0	Unit	#DIV/0!	$0	#DIV/0!	$0	#DIV/0!	$0	#DIV/0!	$0

Figure 18 - Showing a Level 5 WBS Cost Estimating Detail
Source: US National Park Service Cost Estimating Handbook (2011)

Figure 18 provides an example of a Level 5 Cost Estimate, using Activity Based Costing (ABC). Notice that in the first column (field), while the costs have been summarized using Uniformat/OmniClass Table 21, that at the individual activity level, we see Masterformat/OmniClass Table 22 being used. Rarely would owners go to this level of detail; however, for work being done "in-house" by your teams, this would be the recommended level of detail if you are serious about project management and project controls as a core competency.

As with the previous examples, modifying this Excel spreadsheet (Access database) for use with scheduling databases requires the addition of fields named "Crew," "Daily Output," and "Labor Hours" along with the appropriate data.

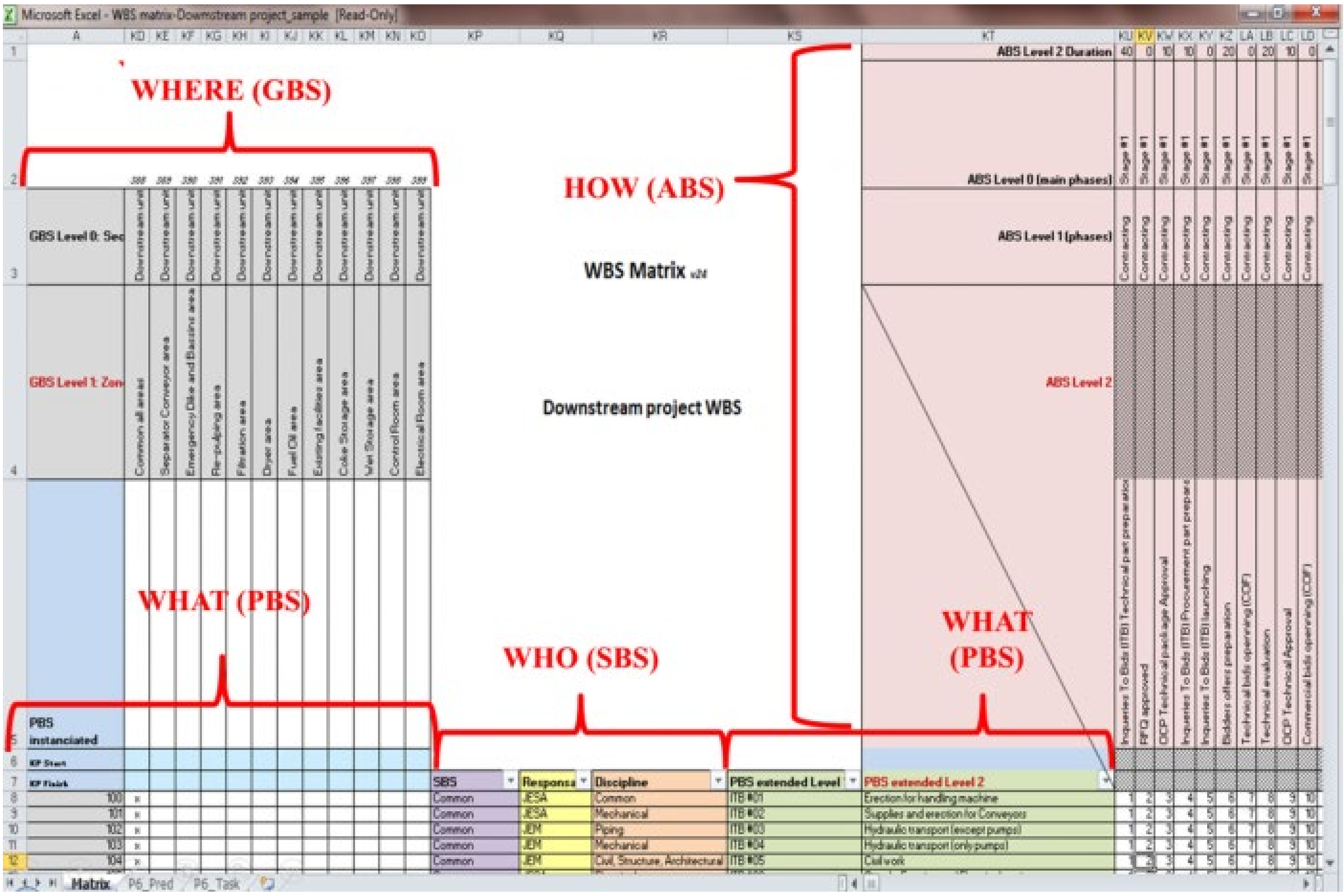


Figure 19 - Example of an Excel Table Designed to enable the schedule data to be entered into MS Excel and then be imported into Primavera P6 or MS Project

Source: Moine, Jean Yves, Leynaud, Xavier and Giammalvo, PD (2015) Creating and Using Multi-Dimensional WBS Structures

In the example above, you can see an Excel template (Access Database) set up to import directly into Primavera's P6 or MS Project 2013. In addition to the Durations (See Row 1 ABS Level 2 Duration), we can also import the costs, crew assignments, and other fields created in the database. However, the key to this is determining which coding structures your stakeholders need and want, and then instead of creating "home-built" coding structures, adopt one of the standardized coding structures, such as OmniClass or Norsok Z-014. For those organizations or sectors that do not yet have a standardized WBS, Resource Code, etc., there is a great opportunity for the more entrepreneurial people out there to create one.

The Figure below, also taken from R.S. Means 2008 Facilities Cost Estimating Database is typical for the USA. While other countries will undoubtedly vary, the concept remains the same. For cost estimators who are preparing costs for projects in countries other than their own need to check to find out what the mark-up requirements are for Labor especially.

These database values need to be updated annually and/or whenever a new labor agreement is signed if working with unionized labor forces or whenever the relevant ministry of manpower issues new regulations.

			A		B	C	D	E	F		G	
1 Resource Code	2 SOC Code	3 Resource Name	Base Rate Including Fringe Benefits 4		Workers Comp. 5	Project Overhead 6	Home Office Overhead 7	Profit Margin 8	Total Overhead and Profit (Colums B through E) 9		Rate With Overhead and Profit (Columns A + G) 10	
			Hourly	Daily	%	%	%	%	%	Amount	Hourly	Daily
SkWkr	47-2000	Skilled Workers	$52.35	$418.80	11.8%	18.3%	16.0%	15.0%	61.10%	$31.99	$84.34	$674.69
Hlpr	47-3000	Helper/Apprentice	$37.80	$302.40	15.0%	18.3%	16.0%	15.0%	64.30%	$24.31	$62.11	$496.84
Suprv	47-1000	Foreman	$54.35	$434.80	11.8%	18.3%	16.0%	15.0%	61.10%	$33.21	$87.56	$700.46
Lab	47-2060	Laborer	$39.85	$318.80	11.8%	18.3%	16.0%	15.0%	61.10%	$24.35	$64.20	$513.59

Figure 20 - R.S. Means 2018 Facility Cost Estimating Database Back Cover Showing Labor Rate Markups
Source: R.S. Means 2008 Facility Cost Estimating Database Back Cover Showing Labour Rate Markups
Explaining Figure 10 above-

- ✓ **(20.1) CODING STRUCTURE**- as defined in the RESOURCE DICTIONARY. As with all other coding structures, it needs to be standardized to as great an extent as possible, not only within an organization but within an industry.
- ✓ **(20.2) STANDARD OCCUPATIONAL CODE (SOC)** This is the RESOURCE CODING STRUCTURE developed by the US Department of Labor, Bureau of Labor Statistics. This is FREE OF CHARGE and makes an excellent LABOR RESOURCE DICTIONARY as it includes a well-written description
- ✓ **(20.3) RESOURCE NAME-** This could be generic, or it could be real people's names
- ✓ **(20.4) BASE RATE-** including fringe benefits (i.e., vacation, insurance) This is the taxable income as shown on your weekly or monthly pay stub.
- ✓ **(20.5) WORKER COMPENSATION INSURANCE-** this insurance is to cover your expenses in the event you are hurt while working on the job.
- ✓ **(20.6) PROJECT OVERHEAD-** are all the indirect costs directly attributable to the project but NOT identifiable to any single activity or work package. This includes the project manager's salary, site offices, fuel for the vehicles, temporary heat, electricity, and water. Basically, any of Division 1 (General Requirements) on the project. In accounting terms, these are often known as "above the line" or "Cost of Goods Sold."
- ✓ **(20.7) HOME OFFICE OVERHEAD-** this is the owner's salary and payroll for accounting, legal, and the bidding team, the rent, heat, electricity, and water for the home office. In accounting terms, these are known as "below the line" costs or General, Sales, and Administrative expenses (GS&A)
- ✓ **(20.8) PROFIT MARGIN-** which, as has been noted, is normally targeted at 10% but often ends up less as for a contractor, this is his/her "management reserve." If there are any "unknown-unknown" risk events that there was no budget or contingency allocated, the cost comes out of the profit margin.
- ✓ **(20.9) TOTAL OVERHEAD and PROFIT-** % is the sum of Columns 4-7 while
- ✓ **(20.10) TOTAL OH&P AMOUNT-** is the total % (8) X the Hourly Base Rate (3)
- ✓ **(20.10) HOURLY BILLING RATE-** is the amount from 9 plus the hourly billing rate from 3

- ✓ **(20.10)** DAILY BILLING RATE- is the Hourly Rate from 9 X 8-hour working day

- ✓ **"Lessons Learned" Databases**

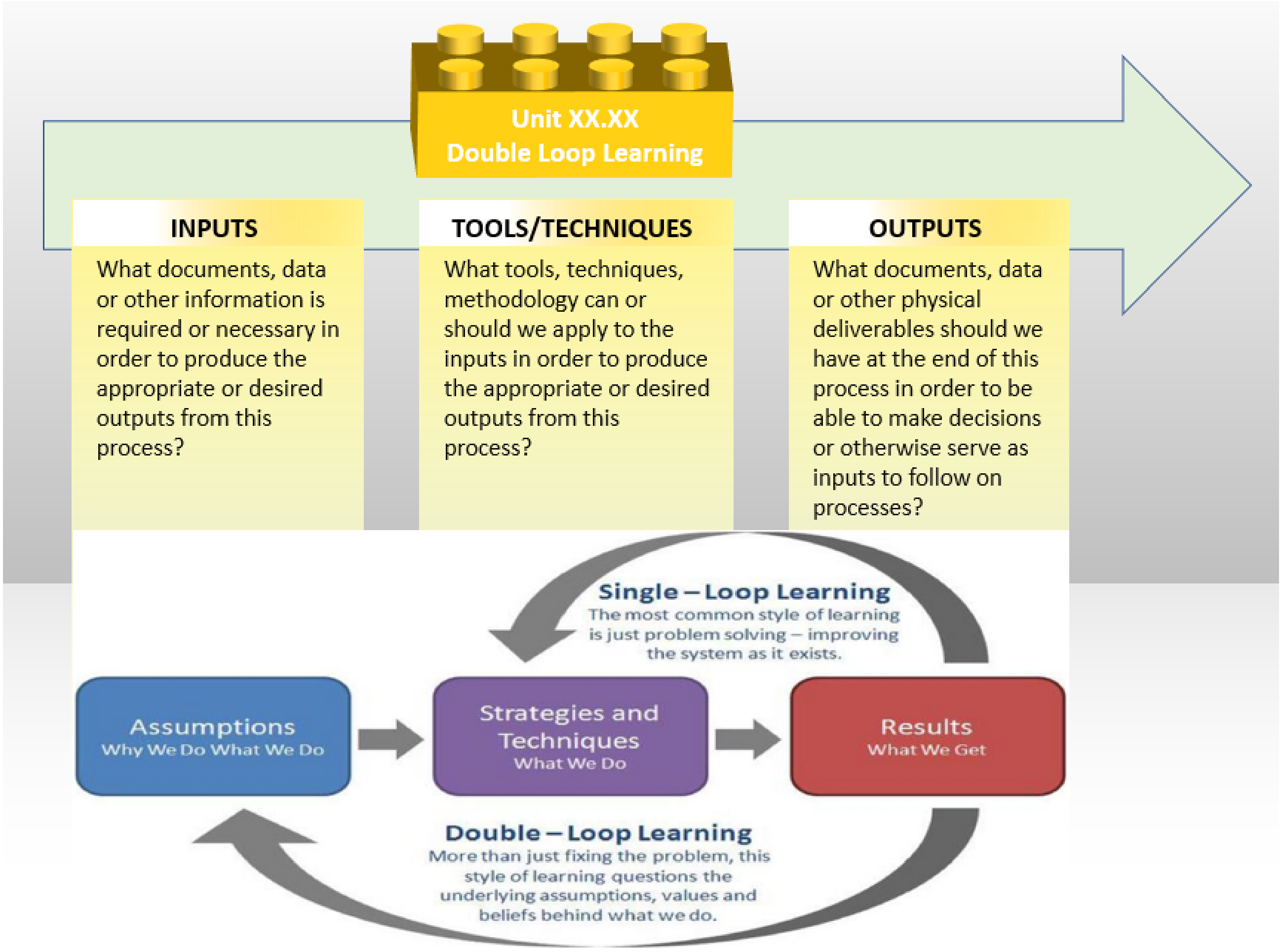


Figure 21– Process Mapping from 100 Meters Showing Continuous Process Improvements

Figure 21 shows that while PMI and AACE (and many other professional societies) have adopted Shewhart's "Plan-Do-Study-Act" Cycle PDSA / and Deming's "Plan-Do-Check-Act" PDCA Cycle, PTMC has long believed that Argyris & Schon's Double Loop Learning is much more relevant and appropriate for use in a PROCESS based model and that the Argyris/Schon approach is more SCALABLE, making it easier to use for both OWNER's and CONTRACTOR's alike. Therefore, we can only recommend you try both approaches, and then YOU decide which one works best for YOU.

Consistent with the PTMC Team's belief in the importance and relevance of Argyris and Schon's "Double Loop Learning," there is another important database we, as project controllers, should create and maintain, and that is a repository where we can catalog, file and be able to locate the many supporting pieces of information which may or may not lend themselves to entry as data points but which are important as supporting or supplemental references. This includes journal articles, photos, videos, frag nets, case studies, or legal briefs- any and all documents which contain valuable and/or useful information but do not lend themselves to being entered as data into the database.

To make this relevant to database management, we continue to make the same mistakes repeatedly, so how can we capture "lessons learned" to quickly share them with others, so they don't make the same mistakes?
There are three ways we can approach this from a database management perspective:

1. One is to scan the documents as Acrobat files (.pdf) or upload them as audio (i.e. .MP3, WMA or . WAV) or video files (i.e. . MOV, AVI, FLV, MP4, and MXF) and then EMBED them in the design object (assuming an object-oriented or hybrid database) With Building Information Modelling, this is how documents such as installation instructions, operating and maintenance manuals and HAZOPS reports and hazardous materials sheets are being handled. Given a choice, this is probably what the future will look like, so this would be the "better" or "best recommended" practice whenever possible. Using this approach, those in the field who are using mobile technology have all these documents at their fingertips in real-time. Ideally, this would include comments and recommendations from those who had previously installed this same piece of equipment or performed the same task so they can be aware of any tricks they should be aware of.
2. The second way is to scan these documents, converting them to Acrobat (.pdf) files or upload them as audio (i.e. .MP3, WMA or . WAV) or video files (i.e. . MOV, AVI, FLV, MP4, and MXF) and then using a relational or flat file database, archive them, creating a "keyword" field so that others in the organization can find these files. While this method too can be accessed using mobile technology, by not linking the documents to an object, but requiring a keyword search slows down the process and is subject to important information being missed if the keywords don't match up.
3. Lastly, there is the old-fashioned way of storing the documents in a filing cabinet, and while this has worked well enough for at least 100 years, it is no longer an appropriate method given the technology we have today and the technological trends of the future.

Another example supporting the trend AWAY from paper-based systems favoring digitization is the number of companies in the business to digitize architectural and engineering drawings: Smithsonian Institute- the University of Florida- Archive Journal (2012) CentriPlan.

To summarize, the era of archiving documents in file cabinets (or shoe boxes) has ended, and the professional project control practitioner of the future knows how to turn these documents into a format that can be uploaded as part of a database, accessible in real-time to those who need to know. Implicit in this is the data is accessible electronically and that the people who need to access this information know how to do it.

✓ **Source of Legal Databases**

For our Forensic specialists, below is a list of Legal Databases:

- New York Law School-
- University of Oxford, Bodleian Law Library-
- Duke University Law Library-

- Stanford University Law Library-
- University of Sydney Law Library-

For more on how these can or should be used, refer to **Unit 14- Managing Forensics**.

However, suffice it to say that the project controls department should have access to these libraries even if there is a fee to do so.

✓ **Additional Cost & Productivity Databases**

For no other reason than R.S Means is probably the oldest (100+ years) and arguably has the largest or most complete databases, we have been using R.S. Means for our examples. (With their permission, of course)

However, many other organizations offer both general and specialized cost databases:

- SPONS-
- Hutchins-
- Griffiths-
- Richardson's-
- Compass-
- Marshal & Swift-
- Building News International-

From the perspective of practicality, instead of "reinventing the wheel," it is often preferable to purchase one of these commercial databases just for the structure and coding and then modify the cost, crew productivity, and other numbers to fit your area of operations than it is to try to create your own from scratch.

OUTPUTS

✓ A Cost Estimating And Productivity Database Which Provides Accurate, Reliable And Precise Cost And Duration Estimates, Appropriately "Fit For Purpose."

UNIT 13.4- UPDATING AND USING THE PROJECT DATABASE(S)

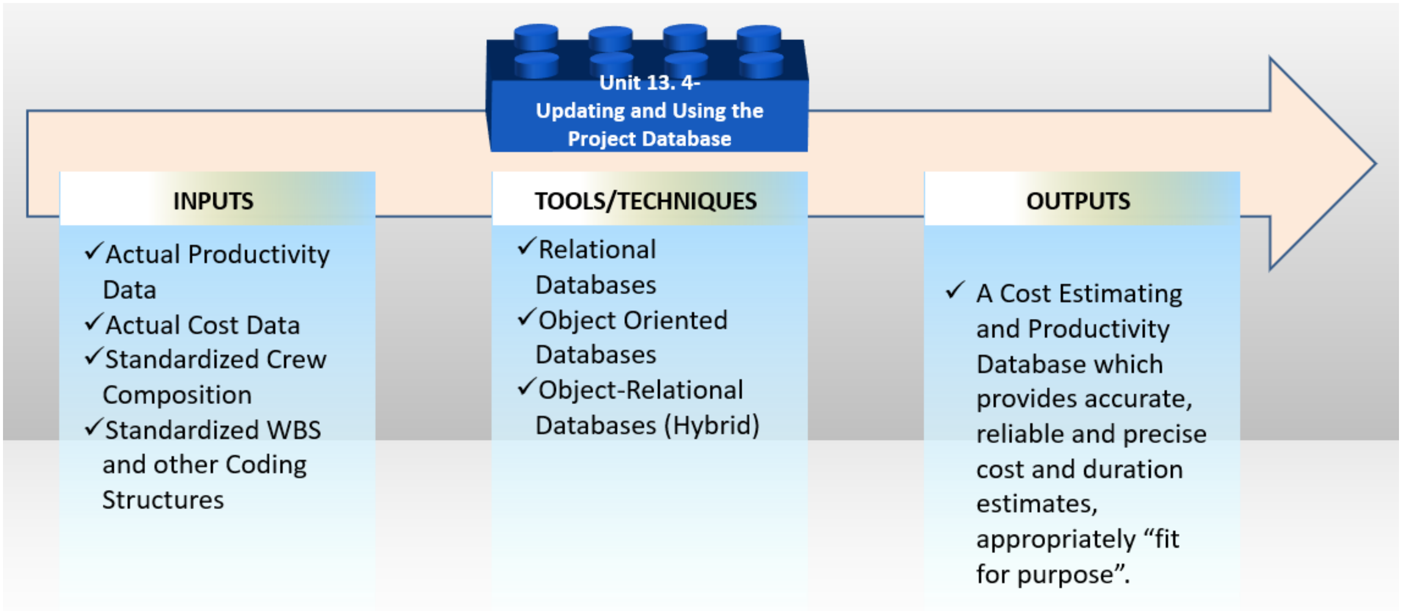


Figure 22- The Updating and Using the Project Databases Process Map from 100 Meters

Source: PTMC Team

INTRODUCTION

In the first sub-unit, we explored how to CREATE the database, and now we need to look closely at how to USE it.

✓ **Sources of the Cost and Productivity Databases**

For no other reason that R.S Means is probably the oldest (100+ years) and arguably has the largest or most complete databases and based on the fact that for over 50 years, we have been using R.S. Means for our company, is the reason we chose to also use RS Means for our examples. (With their permission, of course) However, many other organizations offer both general and specialized cost databases:

- SPONS-
- Hutchins-
- Griffiths-
- Richardson's-
- Compass-
- Marshal & Swift-
- Building News International-

From the perspective of practicality, instead of "reinventing the wheel," it is often preferable to purchase one of these "commercial off the shelf" (COTS) databases just for the structure and coding and then modify the relevant data to fit your area of operations than it is to try to create your own from scratch. Speaking candidly, we have been using the R.S. Means Database for over 50 years now, and that is what we do.

We've used the R.S. Means database as a TEMPLATE, and then when we get to a new country, we make these MODIFICATIONS:

- ✓ Crew Composition
- ✓ Crew Labor Rates
- ✓ Material Costs
- ✓ Equipment Costs
- ✓ Convert Units of Measure to Metric

It takes time but starting with a TEMPLATE makes it much easier and faster than starting from scratch.

INPUTS

- ✓ Actual (Current) Productivity Data
- ✓ Actual (Current) Cost Data
- ✓ "Lessons Learned"

TOOLS & TECHNIQUES

- ✓ **What Fields to Update?**

03 Concrete										
03 30 Cast in Place Concrete										
03 30 53 Miscellaneous Cast in Place Concrete										
03 30 53.40 Concrete in Place		1	2	3		4	5	6	7	8
0.0010 0.0020 0.0050	Including Forms (4 uses), reinforcing steel, concrete placement and finishing, unless otherwise indicated.	Crew Type	Daily Output per Unit	Labor Hours per Unit	Unit of Measure	Material Costs	Labor Costs	Equip-ment Costs	Total Costs per Unit	Total Price/Unit Including OH&P
0.0300	Beams- 5 kip per lineal foot, 10' long spans	C14-A	15.62	12.8	Cubic Yard (CY)	$315.00	$490.00	$48.50	$853.50	$1,225.00
0.0350	Beams- 5 kip per lineal foot, 25' long spans	"	18.55	10.78	CY	$325.00	$415.00	$40.50	$780.50	$1.100.00

Figure 23- Modifying the US Park Cost Estimating Database for Use in Scheduling Databases
Source: R.S. Means 208 Facility Cost Estimating Database

Once you select the commercial off-the-shelf database OR create your own, you need to populate it with "real" numbers appropriate to your country or region. This means whether you are an owner or contractor, you need to validate the following data fields; however, it is essential that your tracking and reporting, where you capture the input data from the field (Unit 11- Managing Progress), is at the same level of detail, which means that your data capture must be using "Activity-Based Costing (ABC) and Activity-Based Management (ABM) at the level of detail appropriate to your need or application.

(Normally, for an OWNER, it would be Level 3 or Level 4, while for a CONTRACTOR, it would be Level 4, Level 5, or even Level 6:

- ✓ **(23.1) Crew Type**- If the crew TYPE changes or if your crew COMPOSITION changes, then you need to update this field
- ✓ **(23.2) Daily Output Per Unit**- Based on real-time productivity, you need to continuously capture the daily output and, using the appropriate Statistical Process Control (SPC) tools/techniques, calculate the mean or average value and the probability of any single productivity figure being met or exceeded. See the learning curve and statistical process control charts below for more on how to analyze this data before inclusion in the database.
- ✓ **(23.3) Labor Hours Per Unit-** This will almost certainly exhibit variability and depend upon the crew sizes, how well they work together, and a host of other variables. As with Output Per Unit in place, the key is for the project controls professional to apply statistical process control chart analysis to the data, throwing out any outliers (those which fall outside of +/- 3 sigma above or below the mean) as well as looking at other patterns which develop in the data which may indicate problems with the process itself.
- ✓ **(23.4) Material Costs** are self-evident. This can either be validated by simply applying "purchasing power parity" to comparing a "market basket" of materials between two locations. See more on how to use purchasing power parity below.
- ✓ **(23.5) Labor Costs**, too, should be self-evident. These can easily be validated by contacting any one of many government agencies in nearly all countries who post the various wages for different trades, or you can purchase any one of a number of commercial off the shelf databases that contain labor rates for different countries and/or indices to enable you to compare labor rates and productivity between different countries or even different regions within the same country.
- ✓ **(23.6)** Same with **Equipment Costs**- With the proliferation of the internet, even in the most remote sites, it is possible to find out local equipment rental costs and the condition of the equipment, and the relative productivity.
- ✓ **(23.7) Total Unit Costs** are what is important, whether owner or contractor and while there is no single "silver bullet" source, the professional project control practitioner should be able to use his/her network combined with Google searches to locate the current information, analyze it and use it to keep the values in the database updated and current.
- ✓ **(23.8) Total Unit Prices**, which are the costs marked up to cover contractor's project overhead, home office overhead, contingency, and profit margin, become the OWNERS costs. To what the contractor submits to the owner, they also have to add in their project overhead costs, home office overhead, funding costs, and owner contingency to arrive at the "fair market value." This again is something that both the contractor's and owner's project control people have to find from within the organization. However, "fair market value" can also be found using networking and search engines.

- ✓ **"Real" or "Constant" Currency Using Purchasing Power Parity (Big Mac Index & Gold Equivalency)**

The key to consistently being able to produce accurate, reliable, and precise cost estimates, which are "fit for purpose," comes from being able to enter accurate numbers into the cost database in the first place, then keep those numbers updated using "real" or "constant" money. Real or constant money is defined as the "Purchasing power of a currency expressed in relation to its purchasing power in a specified year or period. So, for example, in inflationary times, wages are adjusted for the effects of inflation (are 'deflated') by using an index (such as a consumer price index or CPI) to find their worth in constant currency ('in real term's). See also the current dollar [or any other currency].

The easiest way to get started is to purchase a database that contains cost and/or productivity data and then update it to reflect local conditions.

However, in today's world of unstable currencies, professional project controllers are looking to use PURCHASING POWER PARITY as a way to NORMALIZE costs. And there are two approaches.

- One is to use the relative costs of a market basket of goods in one region or time period and then compare the same market basket of goods in another locality and/or point in time. The classic example of this is the Economist's "Big Mac" Index, which started out in 1986 to be a light-hearted story; as the index gained credibility, it is now used as a reasonably valid indicator of purchasing power parity between any two locations (provided of course they sell Big Macs there).

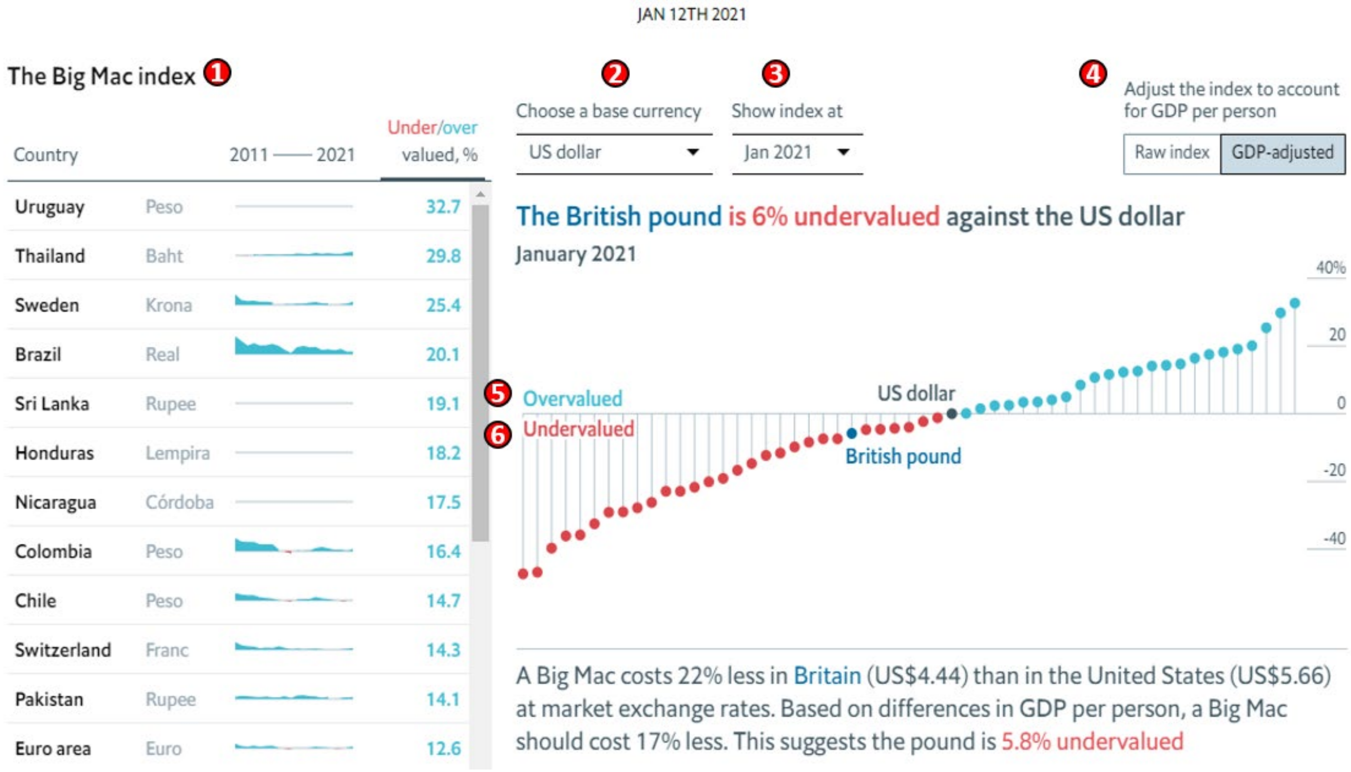


Figure 24- Big Mac Index

One skill set the project control practitioner of tomorrow needs is more or better financial or business analysis competencies. The two most common "Tools & Techniques" we use are the Big Mac Index or "Gold Equivalency."

- **(24.1) Big Mac Index has** been around since 1986, and while it started out somewhat as a joke, it soon gained a fair amount of credibility and is often used in retail settings as a Big Mac has all the elements of nearly all products. For those interested in following up, we recommend you start with this article in Investopedia by Troy Segal (2021), "What is the Big Mac Index?"

- **(24.2)** You can choose the US Dollar, British Pound, Euro, Chinese Yuan, or the Japanese Yen as the base currency and compare it against any other currency. Our R.S Means database is in US Dollars, we choose that as our base.
- **(24.3)** You can choose either January or July data going back to 2011 to make your comparison.
- **(24.4)** You also have a choice between using the Raw Index or the index modified by Gross Domestic Product. In our experience, we have found using the GDP-adjusted values works out to be the most "Accurate, Precise and Reliable."
- **(24.5)** and **(24.6)** As you can see from the graph, in terms of their PURCHASING POWER, some currencies are UNDERVALUED while others are OVERVALUED when compared against the US Dollar. Using their example, a Big Mac in the UK was selling for the equivalent of US $4.44 while the same Big Mac was selling in the USA for the US $5.66. Simple division shows us that $4.44/$5.66 = 0.78445. Going the other way, $5.66/$4.44 = 1.27477. Knowing these ratios, if we have our costs in US Dollars and want to make them EQUAL to prices in the UK, we would have to take our US Dollar value and DIVIDE it by 1.27477. Likewise, if we had UK Pounds and wanted to convert it to the equivalent Purchasing Power, we would have to take the COST in £ to $ we would have to DIVIDE by 0.78445 to get the equivalent cost in the UK. Far from perfect but certainly adequate for a Class 1, 2, or 3 Level of Estimate.

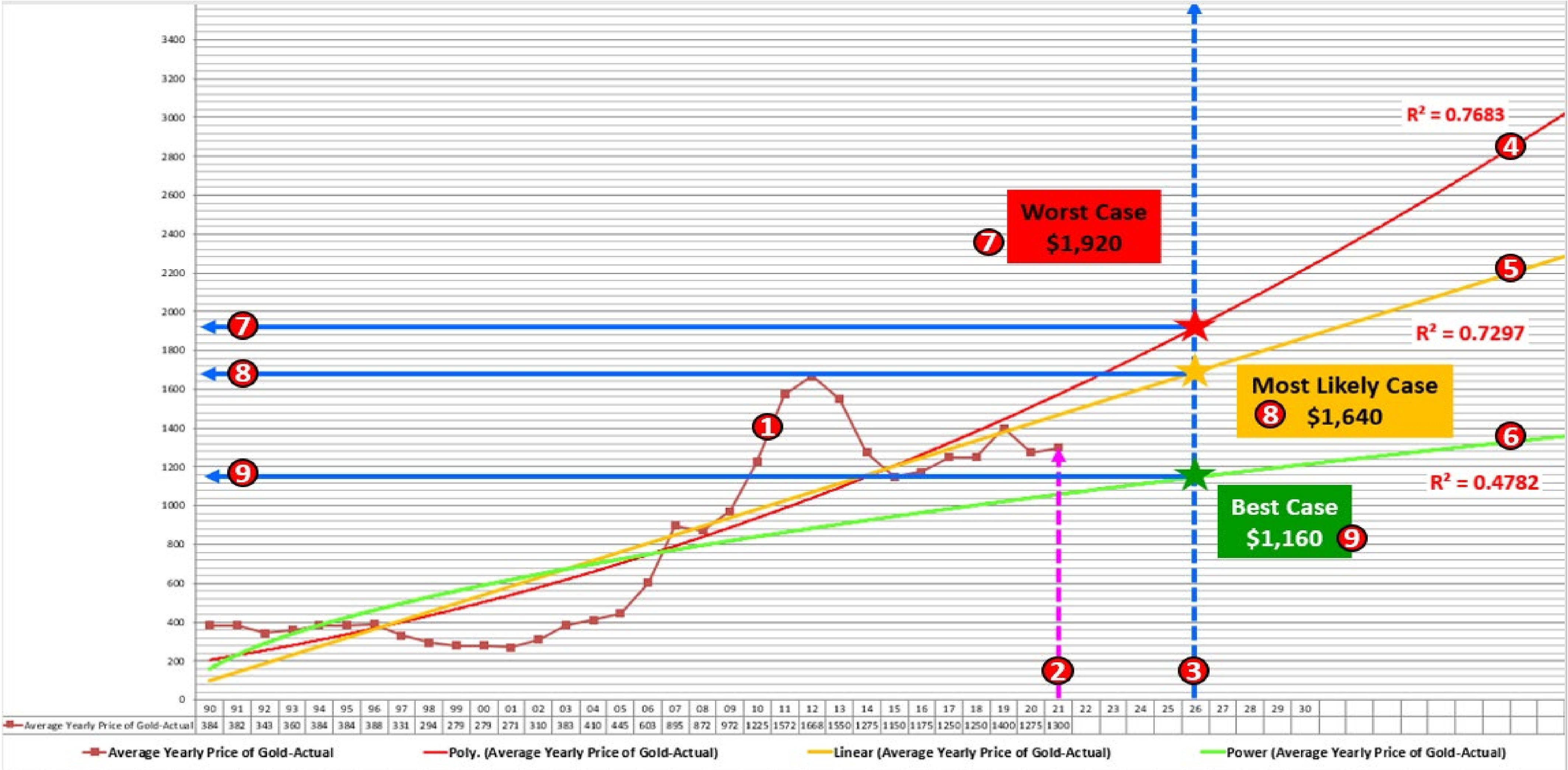
We do NOT recommend using the Big Mac Index on large industrial or commercial construction projects. "Common Sense" should tell you that the labor costs of a welder or heavy equipment operators are not the same as those of someone flipping burgers. Also, the materials are not even close to the materials used in construction. Nor do we recommend using this method as the basis for contractors to bid on projects. This is a "Top Down" tool appropriate for OWNERS to use, not contractors.

Another valid way to measure purchasing power parity which is quickly gaining adherents in today's world is gold equivalency. This is because the purchasing power of gold has remained remarkably stable over several hundred years. For example, in the 1800's it took approximately 1 ounce of gold to purchase a good quality man's suit. And today, it costs just about the same- an ounce of gold to purchase a good quality man's suit.

- **(25.1) MARKET PRICE OF GOLD-** As Gold is sold in just about every country globally, it is very simple to find the current and the historic price of gold in just about any currency in the world. And while we can see from this curve that the PRICE of gold is fairly volatile, what is NOT volatile is what a gram or an ounce of gold will PURCHASE at any given point in time. And because the mining process for gold is heavily dependent upon large capital investments, and relatively high salaries, it is better to use for construction and heavy industrial projects than the Big Mac.
- **(25.2) TIME NOW or DATA DATE,** which is Q2 2021. IF we were using data from an old project, we would have gone back and started mid-point between the time the project started and was completed.
- **(25.3) PROJECTION DATE-** This should be the MID-POINT of the project that is being ESTIMATED.
- **(25.4) FORECAST "WORST CASE" SCENARIO**- Using the "Best Fit" curve feature in Excel, we have forecast the past 20 years of gold prices five years into the future, until Q2 of 2026. In this case, we used

2nd Order Polynomial distribution, which yielded an R^2 (R-Squared) value of 0.7863. As generally speaking, when gold prices increase, it is a sign of inflation, political unrest, or other bad news; we have labeled it a “Worst Case” scenario.

- **(25.5) FORECAST “MOST LIKELY CASE” SCENARIO**- Using the “Best Fit” curve feature in Excel, we have forecast the past 20 years of gold prices five years into the future until Q2 of 2026. In this case, we used a Linear distribution, which yielded an R^2 value of 0.7297.
- **(25.6) FORECAST “BEST CASE” SCENARIO-** Using the “Best Fit” curve feature in Excel, we have forecast the past 20 years of gold prices five years into the future, until Q2 of 2026. In this case, we used a Power distribution, and it yielded an R^ value of 0.4782. As we know, the R^2 (R-Squared) value is a statistical measure of how close the data are to the fitted regression line. It is also known as the coefficient of determination or the coefficient of multiple determination for multiple regression. The definition of R-squared is fairly straightforward; it is the percentage of the response variable variation that a linear model explains. Or:
    - R-squared = Explained variation / Total variation
    - R-squared is always between 0 and 100%:
        - 0% indicates that the model explains none of the variability of the response data around its mean.
        - 100% indicates that the model explains all the variability of the response data around its mean.
    - In general, the higher the R-squared, the better the model fits your data.
- **(25.7) WORST CASE VALUE-** By drawing a horizontal line from the point where the “Worst Case” extended forecast line intersects the date we are forecasting to back to the X-Axis values, we can see that the “Worst Case” forecast price of gold in Q2 2026 is $1,920 per ounce.
- **(25.8) MOST LIKELY CASE VALUE-** By drawing a horizontal line from the point where the “Most Likely Case” extended forecast line intersects the date we are forecasting to back to the X-Axis values, we can see that the “Most Likely Case” forecast price of gold in Q2 2026 is $1,640/ounce.
- **(25.9) BEST CASE VALUE-** By drawing a horizontal line from the point where the “Worst Case” extended forecast line intersects the date we are forecasting to back to the X-Axis values, we can see that the “Worst Case” forecast price of gold in Q2 2026 is $1,160/ounce.



1281 **Figure 25- 30 Years of Gold History Used to Forecast AHEAD 5 Years (from Q2 2021 to Q2 2026)**

✓ **Here is a Case Study to play with:**

You constructed a Pump Station in Anchorage, Alaska, USA, between 2005 and 2007 at a final (depreciable) cost of $45,000,000. You will build an IDENTICAL Pump Station in Jakarta, Indonesia, starting in 2025 and ending in 2027. Using the forecast cost of gold in 2026 (the project's midpoint), what should you budget for the new project in both US dollars and Indonesian Rupiah? The exchange rate TODAY is 14,300 IDR to 1 USD. What ADDITIONAL adjustment might you want to make that is NOT shown here? How RISKY are these numbers? And how do you know how risky they are or are not? (Go back and revisit **Unit 6- Managing Risk & Opportunity**)

Steps	Activity	Values	Formula
1	Actual Cost of Project In Anchorage, AK in 2006	$45,000,000	Input Value
2	Cost of an Ounce of Gold in Alaska in 2006	$600	Input Value
3	Ounces of Gold Equivalency	75000	=$C$2/$C$3
4	Proposed Project in Jakarta, Indonesia 2026	???????	Choose from P40 to P90 in $$$ or IDR
5	Ounce of Gold in 2026 (Forecast)		
6	Worst Case Scenario	$1,920	From Forecast Values
7	Most Likley Scenario	$1,640	From Forecast Values
8	Best Case Scenario	$1,160	From Forecast Values
9	PERT MEAN	$1,606.67	=($C$7+$C$9+(4*$C$8))/6
10	Standard Deviation	$126.67	=($C$7-$C$9)/6
11	Variance	$16,044.44	=$C$11^2
12	Gold Price in 2026 P50 (= Mean)	$1,606.67	=$C$10
13	Gold Price in 2026 P40 (= Mean - 0.25 Sigma)	$1,575.00	=$C$10+($C$11*-0.25)
14	Gold Price in 2026 P85 (= Mean + 1.04 Sigma)	$1,738.40	=$D1$0+($D$11*1.04)
15	Gold Price in 2026 P90 (= Mean + 1.29 Sigma)	$1,770.07	=$D$10+($D$11*1.29)
16	P50 Cost of the Project in Jakarta in 2026 in $USD	$120,500,000	=$C$4*D13
17	P40 Cost of the Project in Jakarta in 2026 in $USD	$118,125,000	=$C$4*D14
18	P85 Cost of the Project in Jakarta in 2026 in $USD	$130,380,000	=$C$4*D15
19	P90 Cost of the Project in Jakarta in 2026 in $USD	$132,755,000	=$C$4*D16
20	P50 Cost of the Project in Jakarta in 2026 in IDR (14,300/$1)	IDR 1,723,150,000,000	=$D$17*14300
21	P40 Cost of the Project in Jakarta in 2026 in IDR (14,300/$1)	IDR 1,689,187,500,000	=$D$18*14300
22	P85 Cost of the Project in Jakarta in 2026 in IDR (14,300/$1)	IDR 1,864,434,000,000	=$D$19*14300
23	P90 Cost of the Project in Jakarta in 2026 in IDR (14,300/$1)	IDR 1,898,396,500,000	=$D$20*14300

Figure 26- Here is a "Gold Equivalency" Case Study Showing the Solution and the Formulae

Remember this is a TOP-DOWN cost estimating tool used by OWNERS to create a Level 1, Level 2, or Level 3 BUDGET estimate. Contractors would not be likely to use this at all unless it was a Design-Build or EPCC or any of the Open Book contracting options.

This is the kind of information that management and other key stakeholders have the right to expect from a PMO or Project Controls Department that adds value to the organization.

Without getting into any further details in this Unit of how to do this, here are two published articles that have attempted to validate the use of gold equivalency as the basis to project costs into the future: Kumar, Hari S (2012) Exploring Gold as Alternative Currency for Future Cost Estimation in Telecommunication

Projects and Asmoro, Trian (2013) Exploring Gold Equivalency for Forecasting Steel Prices on Pipeline Projects.

These papers provide a detailed step-by-step approach to projecting costs into the future, given the unstable global financial situation. Especially for those of you doing long-range planning or estimating megaprojects of 3+ years duration, this might prove to be a conservative approach.

- ✓ **"Cone of Uncertainty" and "Reference Class Forecasting"**

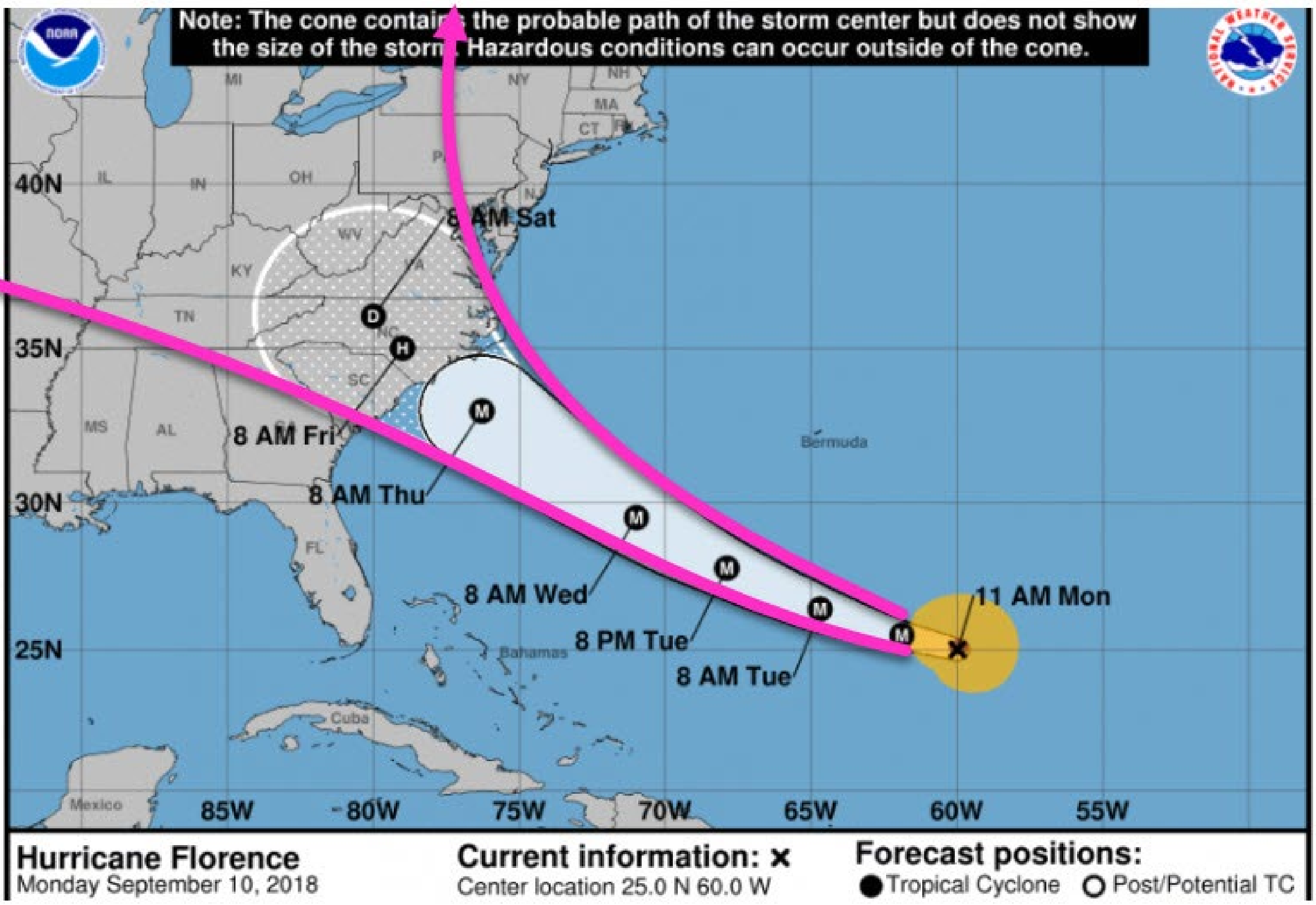
The case study above highlights another important concept that we need to embrace in project management in general and project controls specifically is the concept of the "Cone of Uncertainty, meaning that the further we look into the future, the greater the spread of possible outcomes we can expect. This is illustrated and reinforced by looking above at **(25.4), (25.5)** and **(25.6).**

This not only applies in forecasting Estimates to Complete (ETC) and Estimates at Completion (EAC) for both time and costs, as well as SPI and CPI (efficiencies). It also applies to Rolling Wave Planning.

The best way to illustrate this is using weather, understanding it applies equally to cost, time or any other value that we are predicting cost or direction or duration at some point in the future.

In the example shown in Figure 59, we can see that forecasting results in a RANGE OF POSSIBLE OUTCOMES and the further we look into the future, the wider that range becomes. This is also known as "Reference Class Forecasting," "Range Estimating," or "Comparison Class Forecasting" and is a method of predicting the future by looking at comparable historical situations and their outcomes. Reference Class Forecasting is so named as it predicts the outcome of a planned action based on the range of possible outcomes and their probabilities based on past historical results of comparable systems to that being forecast. Daniel Kahneman and Amos Tversky developed the theories behind reference class forecasting. This theoretical work helped Kahneman win the Nobel Prize in Economics.

For more on the topic of forecasting costs and durations into the future, taking into account the "cone of uncertainty" by producing a "range of estimates" including case studies from a real program, see this certification paper from one of our top Guild of Project Controls Expert Level certifications students, Stephen Paterson, ExxonMobil Singapore. "A Comparison Between 8 Common Cost Forecasting Methods"



https://www.newsweek...th-east-1114983

Figure 27- "Cone of Uncertainty" Illustrated

✓ **Construction Cost Indices**

Another tested and proven way to adjust or modify cost data from year to year and/or place to place is COST INDICES.

As cost indices are very much location-specific, even between one city and another in the same country, if they are not available, there is an alternative approach using various forms of "Purchasing Power Parity." Using Purchasing Power Parity (PPP), you take a "market basket" of goods and services and compare the prices for that same market basket in other cities.

The World Bank and other NGOs, as well as commercial companies, publish this data, but by far the easiest and some would argue the most reliable and realistic method is to use the Economist's "Big Mac Index." While this started out over 15 years ago as a satire, it quickly gained respect and trust as a relatively reliable, accurate, and precise way to compare "real-time" costs between any two locations. To use it as an index, IF we know that in Australia, the price of a Big Mac is $2.44 while the price of a Big Mac in America is $3.15.

- Thus to estimate the cost of the same or similar project done in Australia to be done in the USA, we would have to increase the cost of the project in the USA by $3.51/$2.44 = 143.9%

We could also take the same approach if the project in Australia were done five years ago.

- To do that, we would have to find out the price of a Big Mac in Australia and the price of a Big Mac in the USA today and performing the same calculation as above; we could adjust for both TIME and LOCATION.

This can also be done between any two countries.

- Taking the same example as above, we did a project in Australia where the Big Mac costs $2.44, and we want to do the same or similar project in Switzerland, where a Big Mac costs $4.93. To adjust the cost of the same project we did in Australia to be constructed in Switzerland, we would have to increase the price by $4.93/$2.44 or 202%.

It is very important, especially for owners to know and understand how to keep their cost databases updated, either based on "real-time" bids coming in from their contractors (the most accurate, reliable, and precise method or if that information is not accessible then using published indices such as those published by Engineering News-Record and if that information is not available then using the Big Mac Index.

Historical Cost Indexes

The following tables are the revised Historical Cost Indexes based on a 30-city national average with a base of 100 on January 1, 1993.

The indexes may be used to:

1. Estimate and compare construction costs for different years in the same city.
2. Estimate and compare construction costs in different cities for the same year.
3. Estimate and compare construction costs in different cities for different years.
4. Compare construction trends in any city with the national average.

EXAMPLES

1. Estimate and compare construction costs for different years in the same city.

A. To estimate the construction cost of a building in Lexington, KY in 1970, knowing that it cost $900,000 in 1996.

Index Lexington, KY in 1970 = 26.9
Index Lexington, KY in 1996 = 93.7

$$\frac{\text{Index 1970}}{\text{Index 1996}} \times \text{Cost 1996} = \text{Cost 1970}$$

$$\frac{26.9}{93.7} \times \$900{,}000 = \$258{,}400$$

Construction Cost in Lexington in 1970 = $258,400

B. To estimate the current construction cost of a building in Boston, MA that was built in 1960 for $300,000.

Index Boston, MA in 1960 = 20.5
Index Boston, MA in 1996 = 129.0

$$\frac{\text{Index 1996}}{\text{Index 1960}} \times \text{Cost 1960} = \text{Cost 1996}$$

$$\frac{129.0}{20.5} \times \$300{,}000 = \$1{,}900{,}000$$

Construction Cost in Boston in 1996 = $1,900,000

2. Estimate and compare construction costs in different cities for the same year.

To compare the construction cost of a building in Topeka, KS in 1990 with the known cost of $600,000 in Baltimore, MD in 1990.

Index Topeka, KS in 1990 = 83.2
Index Baltimore, MD in 1990 = 85.6

$$\frac{\text{Index Topeka}}{\text{Index Baltimore}} \times \text{Cost Baltimore} = \text{Cost Topeka}$$

$$\frac{83.2}{85.6} \times \$600{,}000 = \$583{,}200$$

Construction Cost in Topeka in 1990 = $583,200

3. Estimate and compare construction costs in different cities for different years.

To compare the construction cost of a building in Detroit, MI in 1978 with the known construction cost of $4,000,000 for the same building in San Francisco, CA in 1973.

Index Detroit, MI in 1978 = 52.6
Index San Francisco, CA in 1973 = 40.7

$$\frac{\text{Index Detroit 1978}}{\text{Index San Francisco 1973}} \times \text{Cost San Francisco 1973} = \text{Cost Detroit 1978}$$

$$\frac{52.6}{40.7} \times \$4{,}000{,}000 = \$5{,}169{,}500$$

Construction Cost in Detroit in 1978 = $5,169,500

4. Compare construction trends in any city with the national average.

To compare the construction cost in Reno, NV from 1965 to 1979 with the increase in the National Average during the same time period.

Index Reno, NV for 1965 = 21.6 For 1979 = 56.7
Index 30 City Average for 1965 = 21.5 For 1979 = 54.9

A. National Average Increase From 1965 to 1979

$$= \frac{\text{Index} - \text{30 City 1979}}{\text{Index} - \text{30 City 1965}} = \frac{54.9}{21.5}$$

National Average Increase From 1965 to 1979 = 2.55 or 255%

B. Increase for Reno, NV From 1965 to 1979

$$= \frac{\text{Index Reno, NV 1979}}{\text{Index Reno, NV 1965}} = \frac{56.7}{21.6}$$

Reno Increase 1965 – 1979 = 2.63 or 263%

Conclusion: Construction costs in Reno are higher than National Average and increased at a greater rate from 1965 to 1979 than the National Average.

Figure 28- Showing how to use Cost Indices

Source: Adapted from R.S. Means 2018 Facility Cost Estimating Database

Another approach is gaining traction in today's unstable economy, known as the Gold Equivalency Method. Because the purchasing power of gold has remained fairly stable for over 200 years (a good quality man's suit cost what an ounce of gold was back 200 years ago and to buy a good quality man's suit still costs the same as what an ounce of gold costs today) because it is so stable in terms of purchasing power, it makes an ideal tool to use as an index.

- To use the gold equivalency method, you take your project costs in Country A and divide it by the selling price of gold in Country A.
- This will tell you how many ounces of gold your project is worth in Country A.
- Then find out what an ounce of gold sells for in Country B.
- Knowing that you multiply the price of an ounce of gold in Country B x how many ounces of gold your project costs in Country A, and you have adjusted for location.

As with previous examples, you can also adjust for a time as well, and if you apply regression analysis, you can use any of these indices to project into the future.

Here are two articles, one coming from telecommunications project and another from oil and gas, which show you step by step how this is done:

- ✓ Sellapan, Hari Kumar, (2012) Exploring Gold as Alternative Currency for Future Cost Estimation in Telecommunication Projects
- ✓ Asmoro, Trian Hendro (2013) Exploring Gold Equivalency for Forecasting Steel Prices on Pipeline Projects
- ✓ Cost Indexes are published by many organizations, including:
  - Engineering News-Record (ENR)
  - R.S. Means
  - European Union Statistics (Eurostat)
  - Royal Institute of Chartered Surveyors (RICS)-
  - EC Harris- (Now Arcadis)

- ✓ **Statistical Process Control Charts**

Another important tool/technique that is often overlooked in analyzing cost or productivity data into a cost estimating database is identifying and eliminating the impact of outliers.

As we know from using Statistical Process Control Charts explained in **Unit 5- Managing QA-QC**, that any process has a normal variation of +/- 3 sigma or 3 standard deviations, and any data points which fall outside of +/- 3 sigma are not a normal part of the process but are caused by forces outside of the normal distribution. These are called special or identifiable causes. Looking at our Business Dictionary definition, we find that a "special cause" is a Quality control term for that cause of variation, which is not an inherent part of a process but arises from intermittent, unpredictable, and unstable factors. These extraordinary causes are indicated by data points that fall outside of the upper or lower limits of a control chart. Also called assignable cause, external or identifiable cause.

Applying this to our cost and productivity data, we need to plot our cost and productivity data and then throw out those readings that fall outside +/- 3 sigma. Failing to do that will result in our data having a high variation. High variation will result in our cost or productivity data being UNRELIABLE as only a few outliers can dramatically skew the values. Thus we need to eliminate them from inclusion in our database.

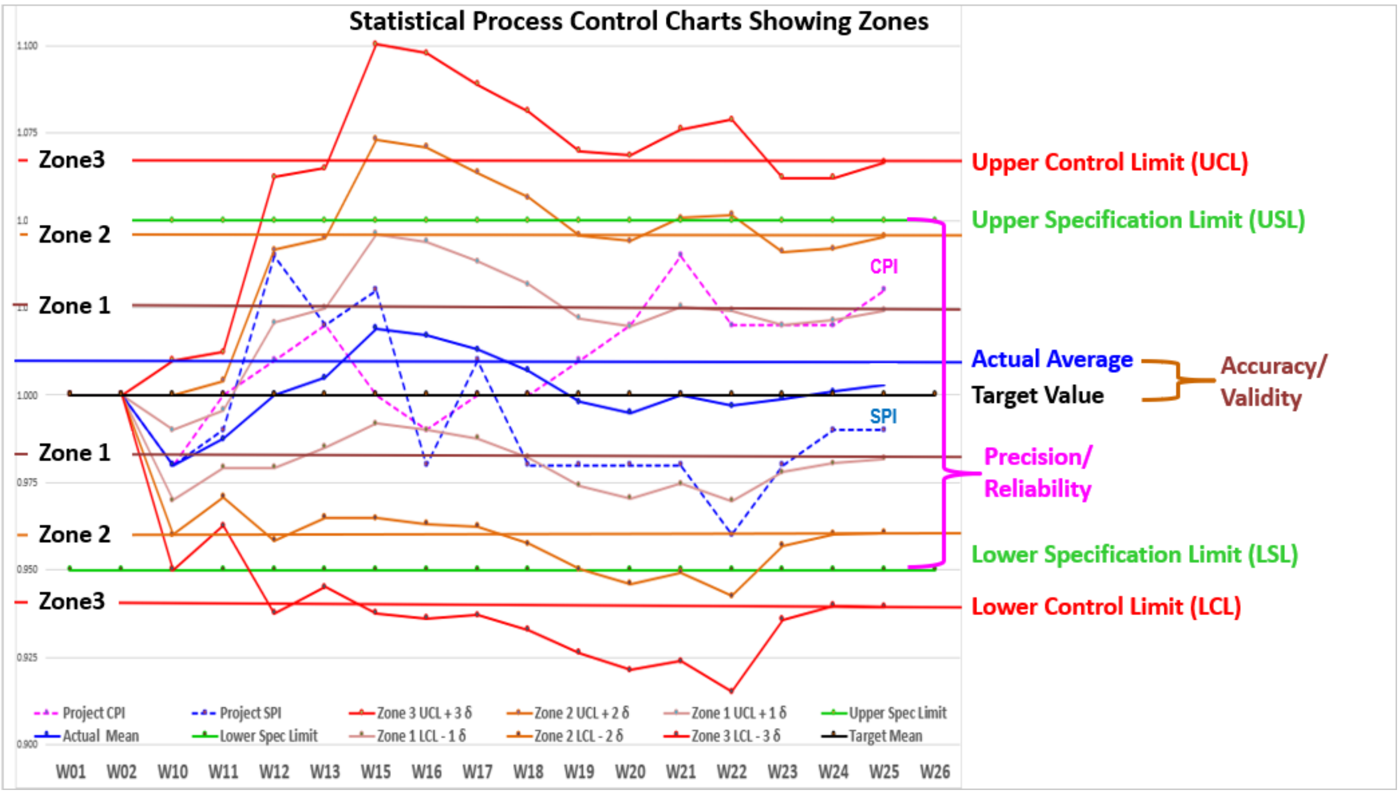


Figure 29– Simple Statistical Process Control Charts Applied to Analyze SPI and CPI.[1]

[1] Mahar, Arif 2020 https://arguniaace2020.wordpress.com/2020/10/17/w15_-efficient-project-monitoring-using-spi-and-cpi-with-statistical-process-control/ and https://arguniaace2020.wordpress.com/2020/11/08/w18_outliers-data-and-the-impact-to-forecasting/

As we can see from the SPC Chart (Figure 25), while the process is "In Control" because the actual SPI and CPI readings fall within or between the Upper and Lower Control Limits (UCL and LCL) as they also fall +/- 3

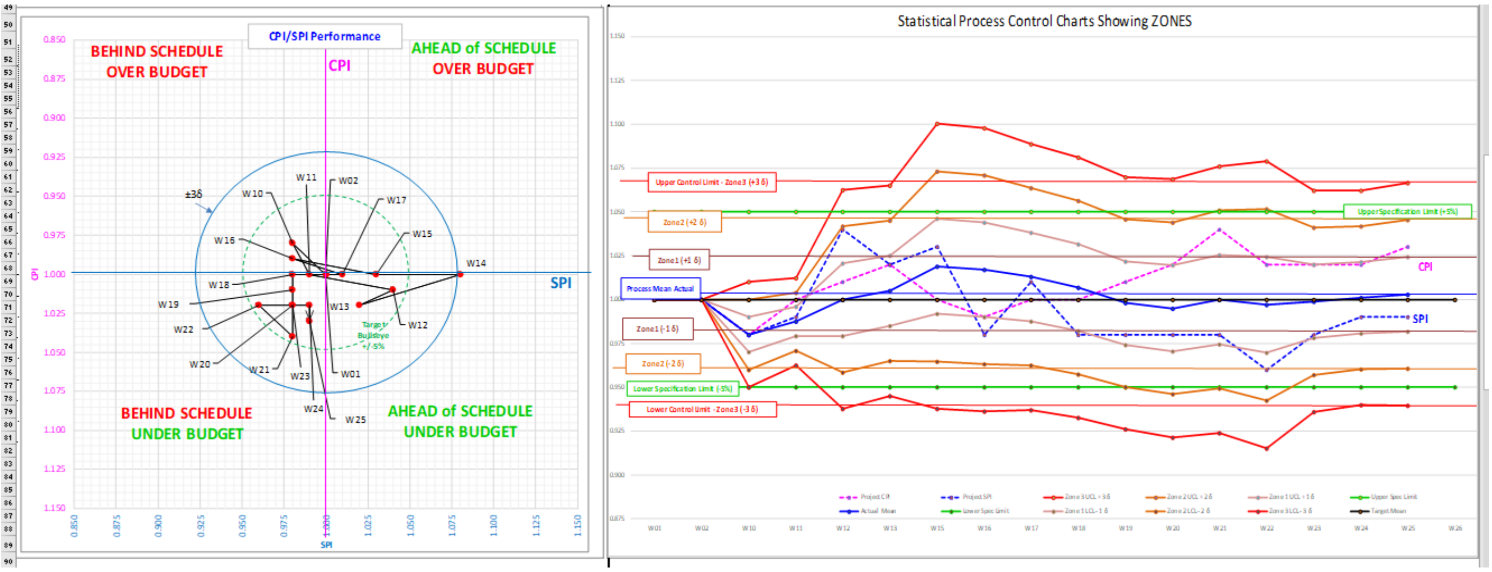


Figure 30- A CPI vs. SPI and SPC Charts from a Real Program.[2]

Sigma above and below the Upper and Lower **Specifications** Limits (USL and LSL) that the process as it is currently configured cannot produce deliverables that consistently meet the Specifications. This means an unacceptably high rejection rate. The advice to the team was to tighten up their processes to be more consistent. (They were not following the SOP the way they should have been, producing too much variability)

Keep in mind that statistical process control (SPC) analysis can be applied to any of the cost or schedule data illustrated above in our ideal database.

The other adjustments we have to make are for PRECISION which is measured by the number of standard deviations from the mean the data falls, and for accuracy, which is how close or far away our actual cost or durations are from our original cost or duration estimates. (Adjusted, of course, for approved change orders).

✓ **Quality Metrics that Apply to Data Analysis and Normalization**

THREE quality metrics that also apply to ANY set of numeric data and are thus useful for analyzing Risk or Opportunity outputs are the Accuracy, Precision, and Reliability of the data: (In this example, we used SPI and CPI, which are EFFICIENCY factors you should have learned more about in **Unit 11- Managing Progress**.)

[2] Mahar, Arif 2020 https://arguniaace2020.wordpress.com/2020/10/17/w15_-efficient-project-monitoring-using-spi-and-cpi-with-statistical-process-control/ and https://arguniaace2020.wordpress.com/2020/11/08/w18_outliers-data-and-the-impact-to-forecasting/

- **ACCURACY- (31.4)** Measures how close the actual "as-built" values came to the estimated values in terms of both time and cost.
- **PRECISION- (31.1)** and **(31.2)** What was the SPREAD of the actual data against the Specifications and the natural variability? In this case, the Upper and Lower SPECIFICATIONS limits were 5%. **(31.5)** In Figure **(31.1)**, all were within the Upper Specification Limits (USL) and Lower Specification Limits (LSL) circle. **(31.5)** In Figure **(31.3),** the actual data points not only fell outside the USL and LSL of 5% **(31.5)** but ALSO even exceeded the Upper Control Limits (UCL) and Lower Control Limits (LCL) of +/-3 δ. **(31.6)**
- **RELIABILITY- (31.3)** This is how USEABLE the data set is. In this example, while most of the data points were within the USL and LSL of +/-5%, some outliers fell outside the UCL and LCL of +/-3 δ. This explains why, when we have outliers beyond the UCL and LCL, we delete those readings once we have identified WHY they happened.

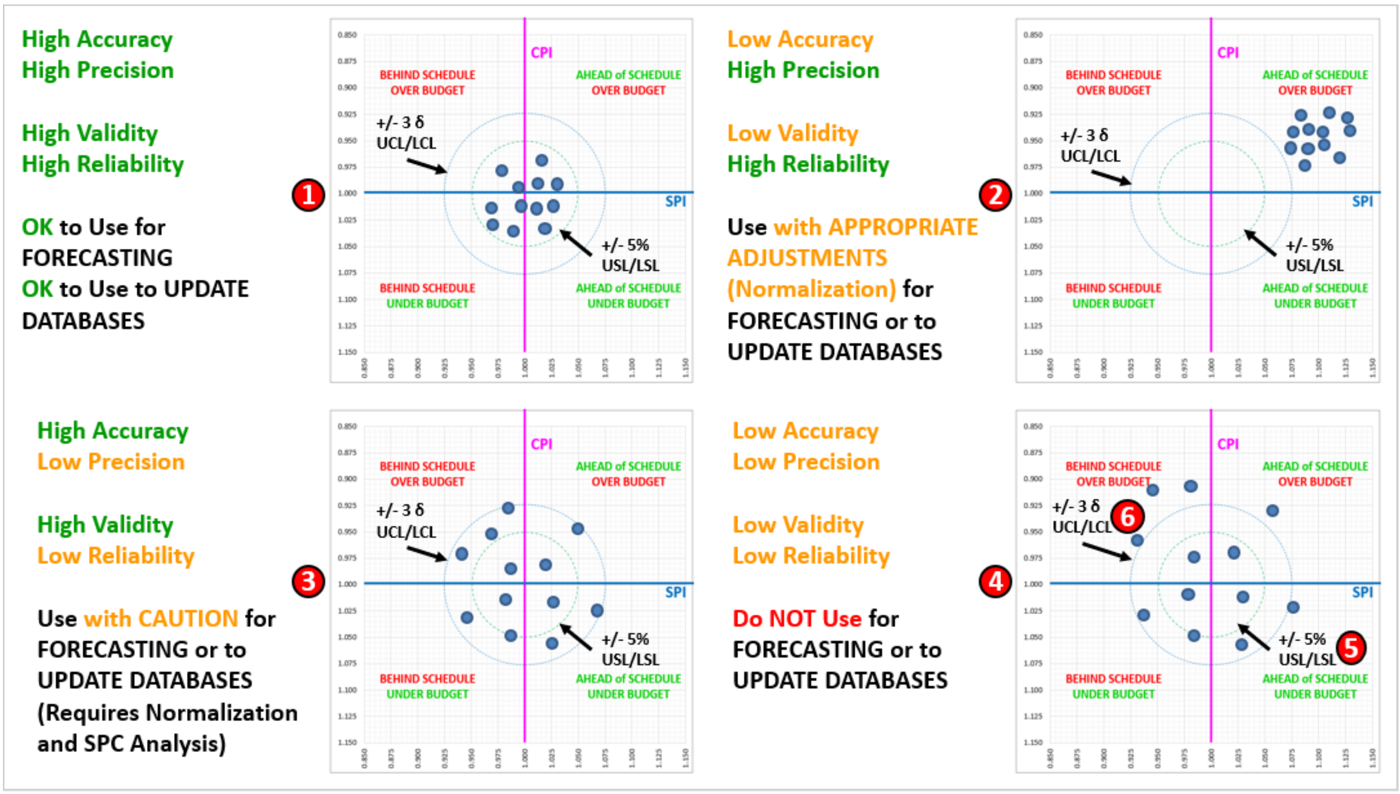


Figure 31- Illustrating Data Quality Metrics Precision, Reliability and Accuracy Source: Giammalvo, Paul D (2015) Course Materials. Adapted from Rizo, Chris (1999) "Precision, Accuracy, and Reliability Illustrated and Contributed Under [Creative Commons License BY v 4.0]

If any shooters are reading this, you can appreciate how you can apply your shooting (target acquisition and adjustment) knowledge to help you decide what adjustments you have to make to bring your SPI and CPI values back to the Bullseye.

Like the use of Statistical Process Control Charts, this analysis not only could but MUST be applied to any of the productivity or cost data from the database examples above. Failure to analyze the data and identify those outliers will serve to render your databases unreliable, and once they are deemed unreliable, it will be almost impossible to convince anyone to trust them or use them.

- ✓ **Learning Curves**

Knowing and understanding how learning curves impact durations is another "tool and technique" that planners/schedulers or cost estimators/project controllers can utilize to produce more realistic and achievable schedules and budgets. It can be used for scheduling (durations) as well as for costs.

It is applicable whenever there is a single activity or series of activities ("fragnets") that repeat on a project. Examples of this are repetitive floor layouts in a hotel or high-rise office building, installing pipelines, hanging doors, installing electrical lighting, or any other activity or series of repeat activities.

Learning curves help us know how long the first activities are scheduled to take; we can then apply a sound mathematical formula to justify what the subsequent durations are likely to be.

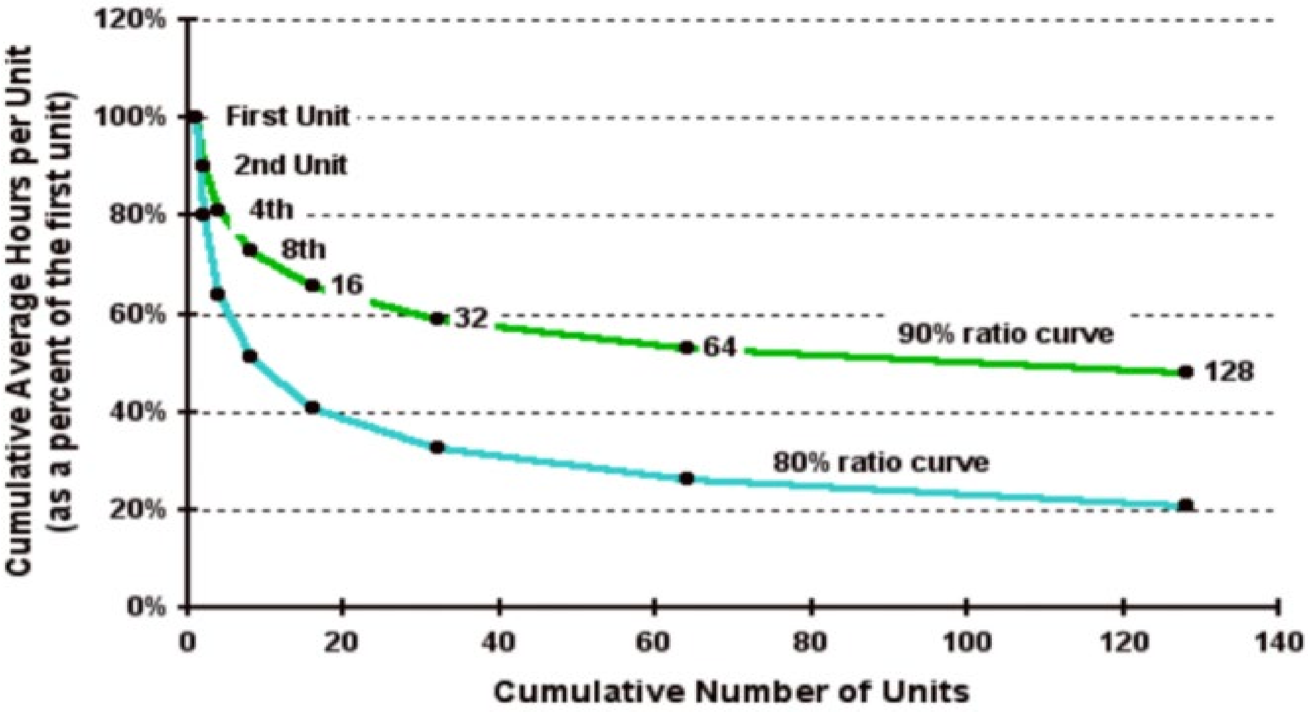


Figure 9: Illustration of learning curves

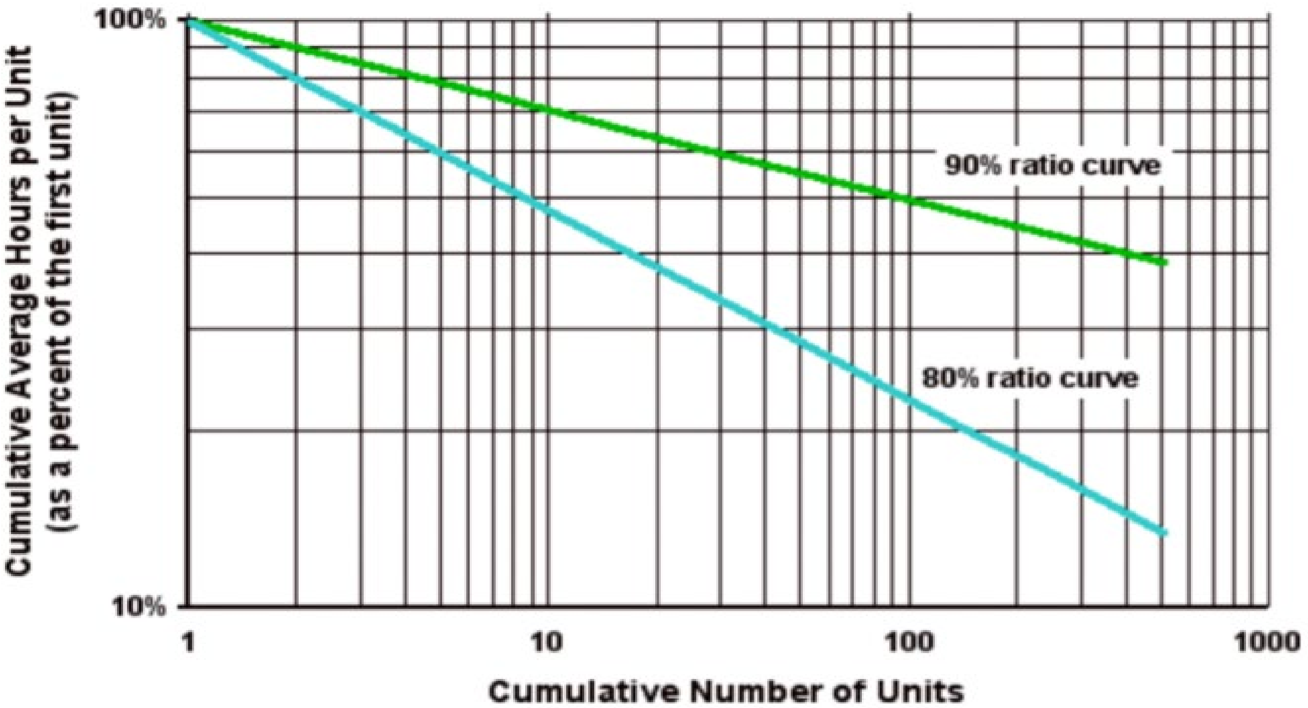


Figure 32- Learning Curves
Source: Wideman, Max (n.d.) Learning Curve Theory

While many planners/schedulers introduce a slower production rate (say 40 or 50% productivity) for the initial periods or instances for any repetitive operations, the professionally justifiable approach should include a formal analysis should be conducted and applied to the initial periods or instances of the repetitive operations, including a follow-up analysis to see if the original assumptions were valid.

The theory behind this is that a learning curve is a geometric distribution with the general form "Y = aXb" where:

- Y = cumulative average time per unit or batch.
- a = time taken to produce the initial quantity.
- X = the cumulative units of production or, if in batches, the cumulative number of batches.
- b = the learning index or coefficient calculated as log learning curve percentage ÷ log 2. So b for an 80 percent curve would be log 0.8 ÷ log 2 = – 0.322.

As we can see from Figure 5 above, we can plot the curve easily using Excel, or we can plot it manually using log-log paper to generate a straight line. Explained very simply:

- The first time we execute the activity takes us so many minutes, hours, or days.
- The second time we execute the activity, it only takes us between 80% to 90% of the time it took us to do it the first time.
- The 4th time we do the activity, it only takes us between 80% to 90% of the time it took us to execute the activity the 2nd time and so on.
- Each time we double the number of times we execute the activity, the time it takes (the number of periods required) is reduced anywhere between 10% (90% Learning Curve) to 20%. (80% Learning Curve)

As noted, there are two approaches, using units of production or batches and even though the formula is identical. The planner/scheduler can experiment to see which method yields the most accurate results for any specific application. As this tool & technique applies to both time and cost, it is an important one for all project control professionals to master; here are recommended supplemental references:

  - Robert Agar, (2020) "Reducing your DBA's Learning Curve"
  - JULIA KAGAN and ERIC ESTEVEZ (2020) "Learning Curve"
  - What is Learning Curve? (n.d)

✓ **Productivity and Cost Adjustment Factors**

Other adjustments which need to be taken into consideration when entering new data or updating existing data come to us from published research by the World Academy of Science, Engineering and Technology International Journal of Civil, Environmental, Structural, Construction, and Architectural Engineering Vol:8, No:10, 2014 "Labor Productivity in the Construction Industry - Factors Influencing the Spanish Construction Labor Productivity" by G. Robles, A. Stifi, José L. Ponz-Tienda, S. Gentes.

OVERALL RII AND RANKING OF ALL FACTORS SURVEYED

Rank	Code	Factor	Overall RII (%)
1	F25	Shortage or late supply of materials	87.40
2	F3	Clarity of the drawings and project documents	86.41
3	F12	Clear and daily task assignment	85.53
4	F27	Tools or equipment shortages	85.20
5	F5	Level of Skill and experience	83.16
6	F17	Delays in payments to workers	82.47
7	F22	Coordination between crews	82.00
8	F14	Improper coordination of subcontractors	81.59
9	F13	Insufficient supervision of subcontractors	81.03
10	F20	Communication problems	80.88
11	F6	Ability to adapt to changes and new environments	80.84
12	F15	Inadequate planning	78.10
13	F7	Labour motivation	77.47
14	F18	Delays in payments to suppliers	76.99
15	F26	Unsuitability of materials storage location	75.36
16	F19	Unrealistic scheduling	75.07
17	F10	Worker´s integrity	75.00
18	F16	High congestion	73.94
19	F24	Rework	73.19
20	F30	Motion´s limitation in the jobsite	72.08
21	F21	Reallocation of laborers	70.80
22	F23	Lack or delay in supervision	70.22
23	F11	Incentive policies	69.65
24	F32	High/low temperatures	69.53
25	F28	Performing work at night	67.93
26	F2	Complexity of the design	66.86
27	F1	Construction method	65.49
28	F33	Rain	64.39
29	F29	Influence of working at height	64.36
30	F4	Project scale	64.17
31	F34	High winds	63.25
32	F9	Number of breaks and their duration	62.67
33	F8	Working overtime	59.82
34	F35	Distance between construction sites and cities	54.23
35	F31	Air humidity	53.56

Figure 33- Productivity Factors

Source: World Academy of Science, Engineering and Technology International Journal of Civil, Environmental, Structural, Construction, and Architectural Engineering Vol:8, No:10, 2014 “Labor Productivity in the Construction Industry - Factors Influencing the Spanish Construction Labor Productivity” by G. Robles, A. Stifi, José L. Ponz-Tienda, S. Gentes.

Applying Pareto’s “80:20” rule, we need to at least consider whether or not any adjustments to the data can or should be made not only for the top 11 (Pareto’s 80%) but all 35 factors.

Another very important consideration we need to keep in mind is the number of days per week scheduled for work. Hours Worked per day/week.

Especially for those owners and contractors working in different countries, the labor laws are not the same, and while there is nothing wrong with exceeding the requirements, you may be subject to significant fines if you break these local laws. For example, in the Middle East, you cannot work your field people if the temperature exceeds 42 degrees Celsius. This means you need to change your work calendars during the hot months and/or schedule in two shifts per day rather than three.

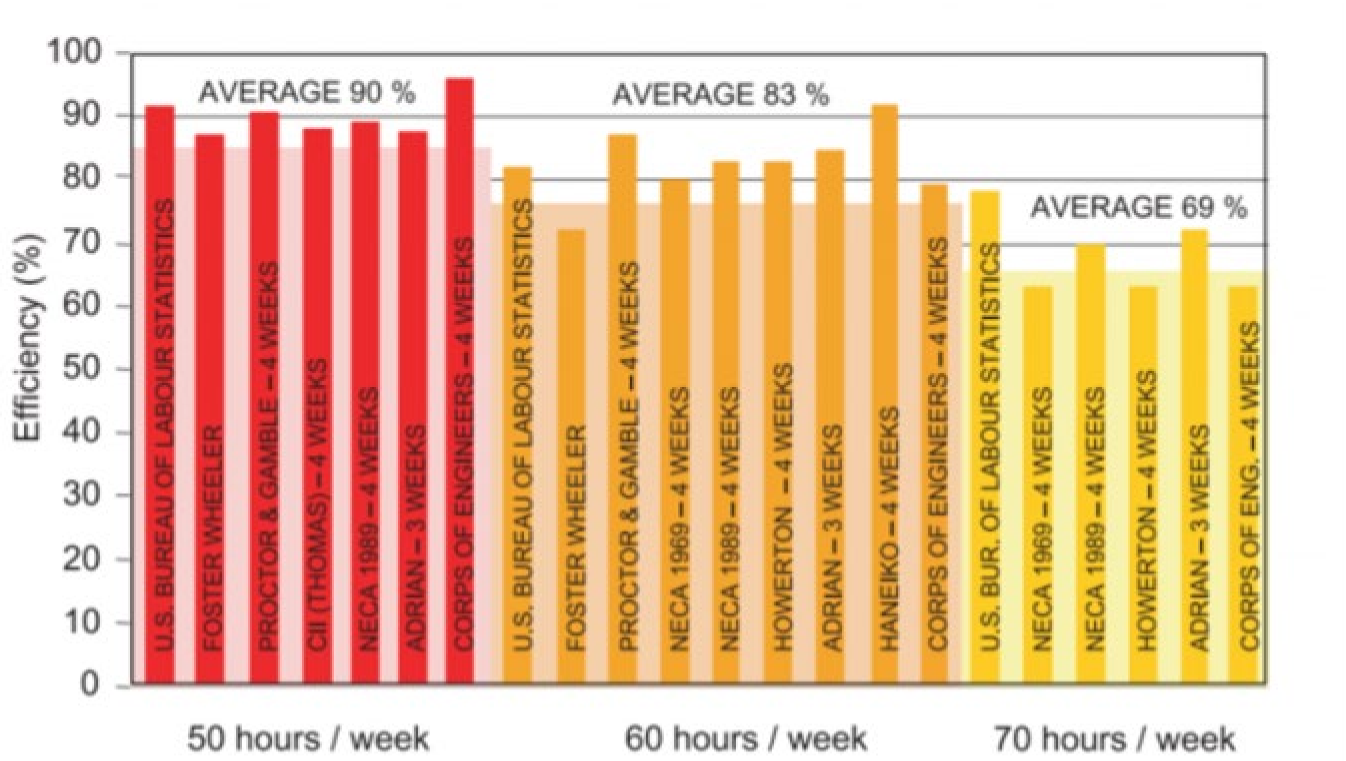


Figure 34- Revay Report

Source: Source: Revay & Associates(n.d.) The Revay Report

Probably the most complete and comprehensive analysis of the impact over time has on the project productivity comes to us from the November 2001 issue of The Revay Report “Calculating Loss of Productivity Due to Overtime Using Published Charts – Fact or Fiction” by Regula Brunies, FPMI, CCC, CQS, and Zey Emir, P.Eng, MBA Revay and Associates Limited

Their research concludes that:

- going from a 40-hour workweek to a 50-hour workweek, we lose on average 10% productivity;
- going to a 60-hour workweek, we drop 17% to only 83% from the base productivity and

- going to a 70-hour workweek, we lose 31% productivity, dropping down to only 69% of what we can expect working a standard 40-hour workweek.

OUTPUTS

A Cost Estimating And Productivity Database Which Provides Accurate, Reliable And Precise Cost And Duration Estimates, Appropriately "Fit For Purpose."

ARTIFICIAL INTELLIGENCE/ MACHINE LEARNING/DEEP LEARNING

With Artificial Intelligence (AI) and Machine Learning (ML) permeating everything we see, read, and hear, now is a good time to include a section on these topics in our Unit on Database Management, understanding that what we see here applies now or will in the future, apply to pretty near all the other Units.

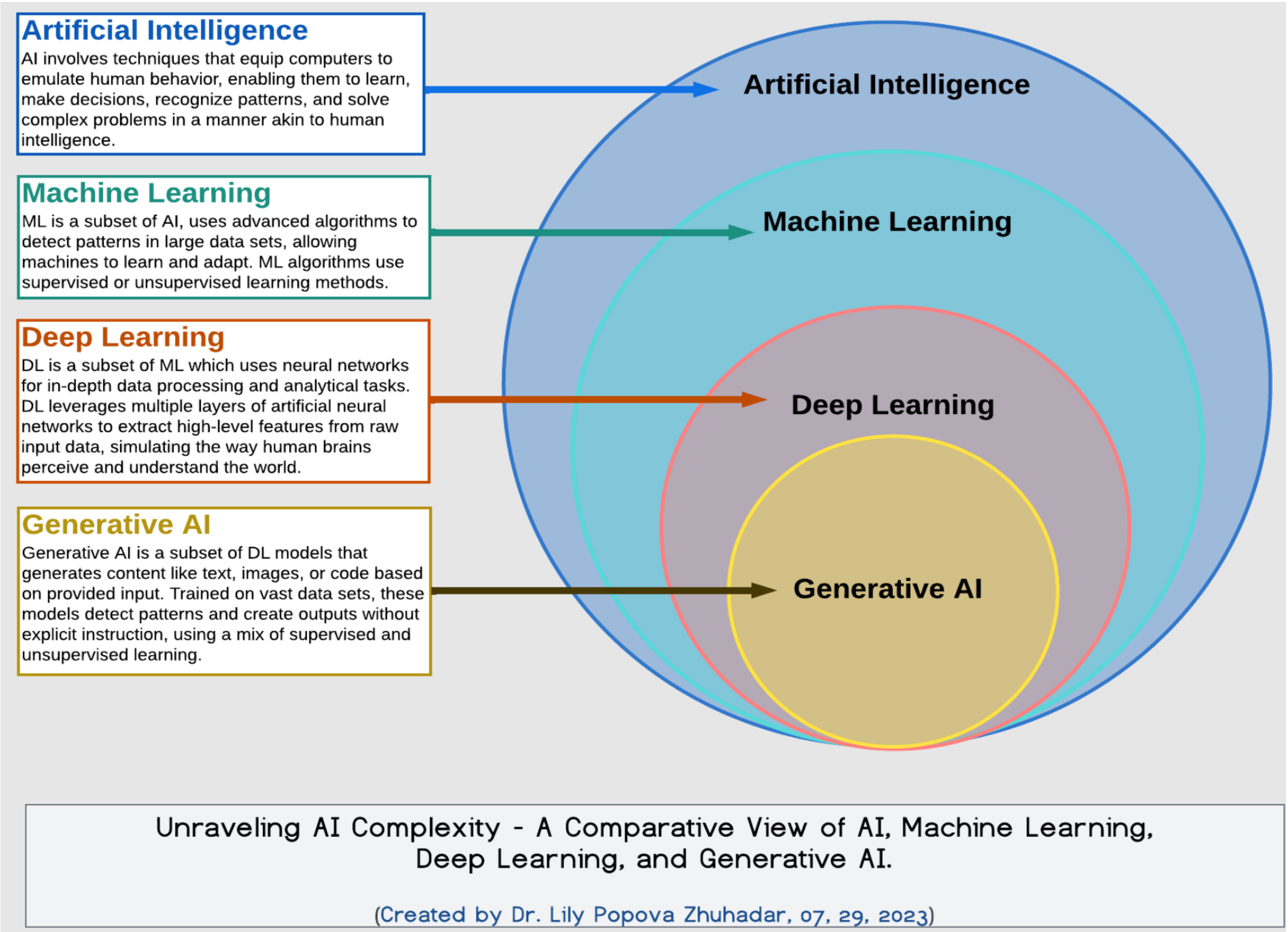


Figure 35- World of AI, ML, DL [3]

[3] Dr, Lily Popovada, Wikimedia 2023 https://commons.wikimedia.org/wiki/File:Unraveling_AI_Complexity_-_A_Comparative_View_of_AI,_Machine_Learning,_Deep_Learning,_and_Generative_AI.png#Licensing

- ✓ Artificial Intelligence Defined- Nearly all the major dictionaries and credible references define "artificial intelligence" as being "the ability of a digital computer or computer-controlled robot to perform tasks commonly associated with intelligent beings."[4]
- ✓ "Stuart Russell and Peter Norvig in their leading textbook, Artificial Intelligence: A Modern Approach, they delve into four potential goals or definitions of AI, which differentiates computer systems based on rationality and thinking vs. acting:
  - Human approach:
    - Systems that think like humans
    - Systems that act like humans
  - Ideal approach:
    - Systems that think rationally
    - Systems that act rationally."[5]

Given that we are NOT successfully managing projects in the world of project management raises some thorny questions about why we should be programming "systems that think or act like humans" when what we are doing is not working? This is a major concern and one we MUST address before moving forward with AI or ML "solutions."

- ✓ "Weak" AI vs. "Strong" AI
  "Weak AI—also called Narrow AI or Artificial Narrow Intelligence (ANI)—is AI trained and focused on performing a specific task or tasks. Weak AI drives most of the AI that surrounds us today. 'Narrow' might be a more accurate descriptor for this type of AI as it is anything but weak; it enables some very robust applications, such as Apple's Siri, Amazon's Alexa, IBM Watson, and autonomous vehicles."[6]
  Strong AI comprises Artificial General Intelligence (AGI) and Artificial Super Intelligence (ASI). Artificial general intelligence (AGI), or general AI, is a theoretical form of AI where a machine would have an intelligence equal to humans; it would have a self-aware consciousness that can solve problems, learn, and plan for the future. Artificial Super Intelligence (ASI)—also known as superintelligence—would surpass the intelligence and ability of the human brain. While strong AI is still entirely theoretical with no practical examples in use today, that doesn't mean AI researchers aren't also exploring its development. In the meantime, the best examples of ASI might be from science fiction, such as HAL, the superhuman, rogue computer assistant in 2001: A Space Odyssey[7]*.*

  "Deep Learning"

---

[4] B.J. Cope;and (2021) "Britannica" https://www.britannica.com/technology/artificial-intelligence
[5] Stuart Russel and Peter Norvig (2021) "Artificial Intelligence- A Modern Approach, 4th Edition" http://aima.cs.berkeley.edu/
[6] IBM (2020) "What is Artificial Intelligence?" https://www.ibm.com/cloud/learn/what-is-artificial-intelligence#toc-types-of-a-q56lfpGa
[7] IBM (2020) "What is Artificial Intelligence?" https://www.ibm.com/cloud/learn/what-is-artificial-intelligence#toc-types-of-a-q56lfpGa

Deep Learning analyzes the neural networks between all the Inputs>Tools & Techniques>Outputs. Explained another way, to teach a machine how to do something equal to and preferably better than a human being, we need to not only identify ALL the Inputs, Tools & Techniques and Outputs but, more importantly, how they interact with one another.

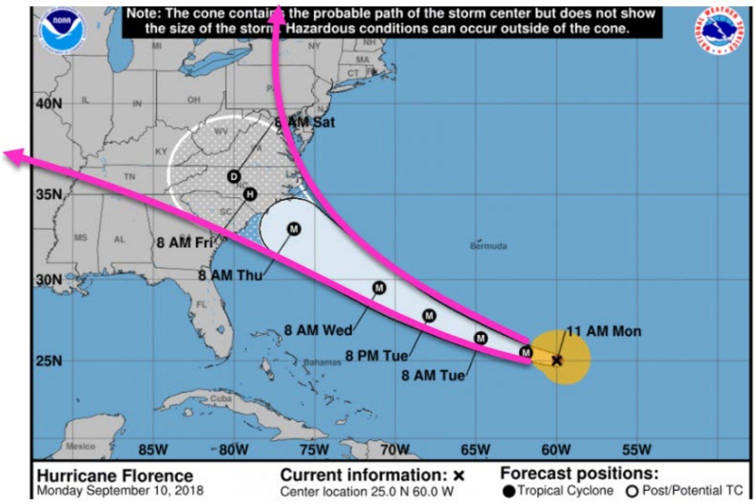


Figure 36- Weakness of Neural Networks

We can see examples of this type of "Neural Network Analysis" being used in Systems Dynamics Software with some of the more sophisticated models coming to us from the various Weather Agencies to predict not only hurricanes, typhoons, and tornadoes, but also the climate scientists. To fully understand where we are in the evolutionary development of these tools and techniques, just look at the "Cone of Uncertainty" that applies to the prediction of where a hurricane is going to be.

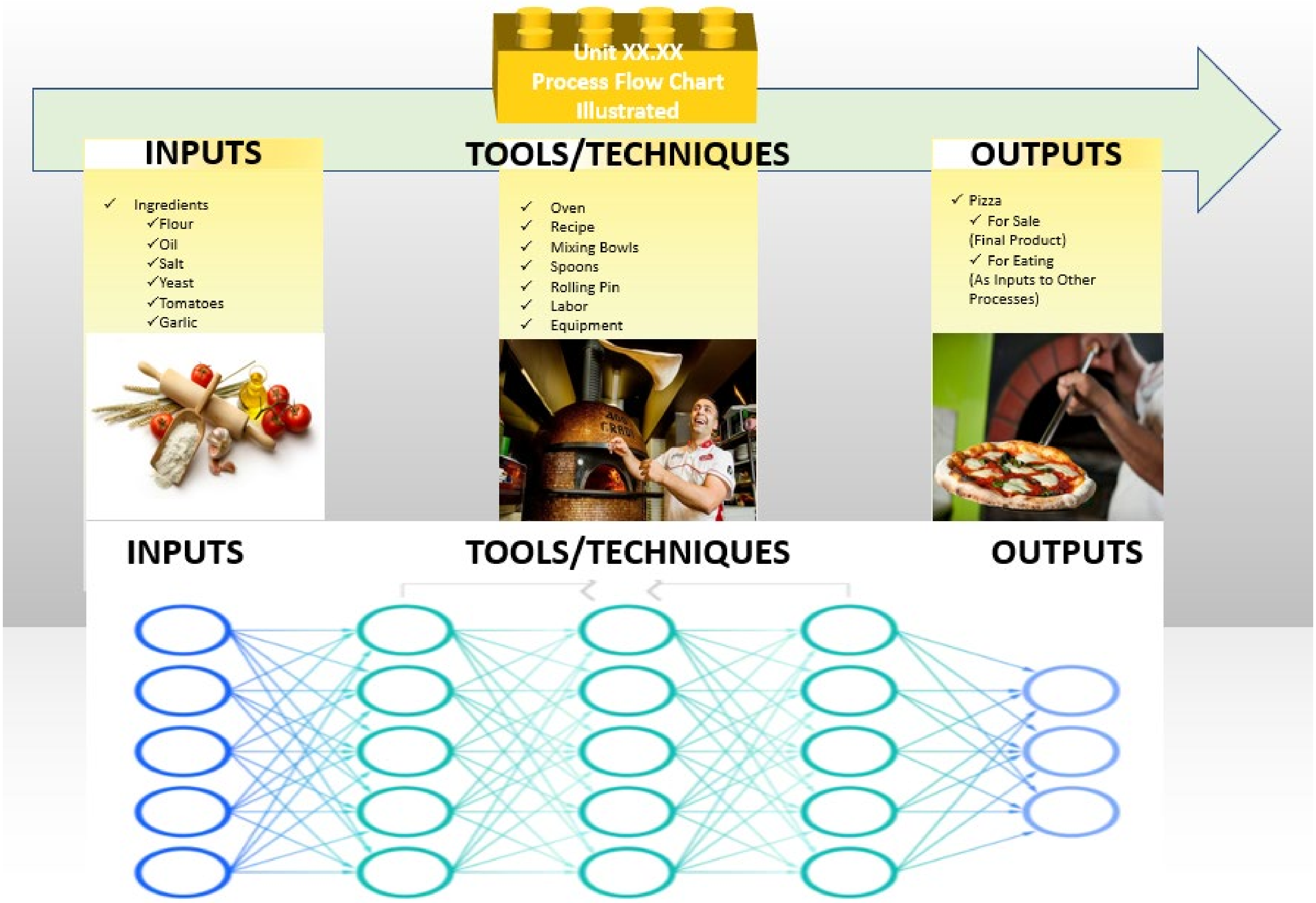


Figure 37- "As Is" Conceptual Process Flow Chart to the "Deep Learning" Model ("To Be")

As we can appreciate just looking at the differences between these "as is" vs. "to be" models, and knowing that even using the existing "tools, techniques and method" we cannot consistently deliver projects on time and within budget, much less solving the problem of exploiting the opportunity the project was initiated to accomplish it helps us to grasp the huge task we have before us.

To see more examples of the use of "Machine Learning" and "Deep Learning" specifically related to project management, we need to research the work done by the Systems Dynamics Society. As our background is mostly construction, we are familiar with the use of Systems Dynamics for construction project management, but the challenge is for anyone reading this who is coming from other sectors, we have provided you with references to help you get started researching how Systems Dynamics has been applied to other project environments.

Worth noting is one of the reasons we have switched from using Shewhart's PDCA and Deming's PDSA cycles as the basis for Continuous Improvement to Argyris and Schon's "Double Loop Learning" approach is because capturing and incorporating "Lessons Learned" which is essential to build a "Deep Learning" model has not worked well in the 50+ years PMI, AACE and IPMA have been in existence.

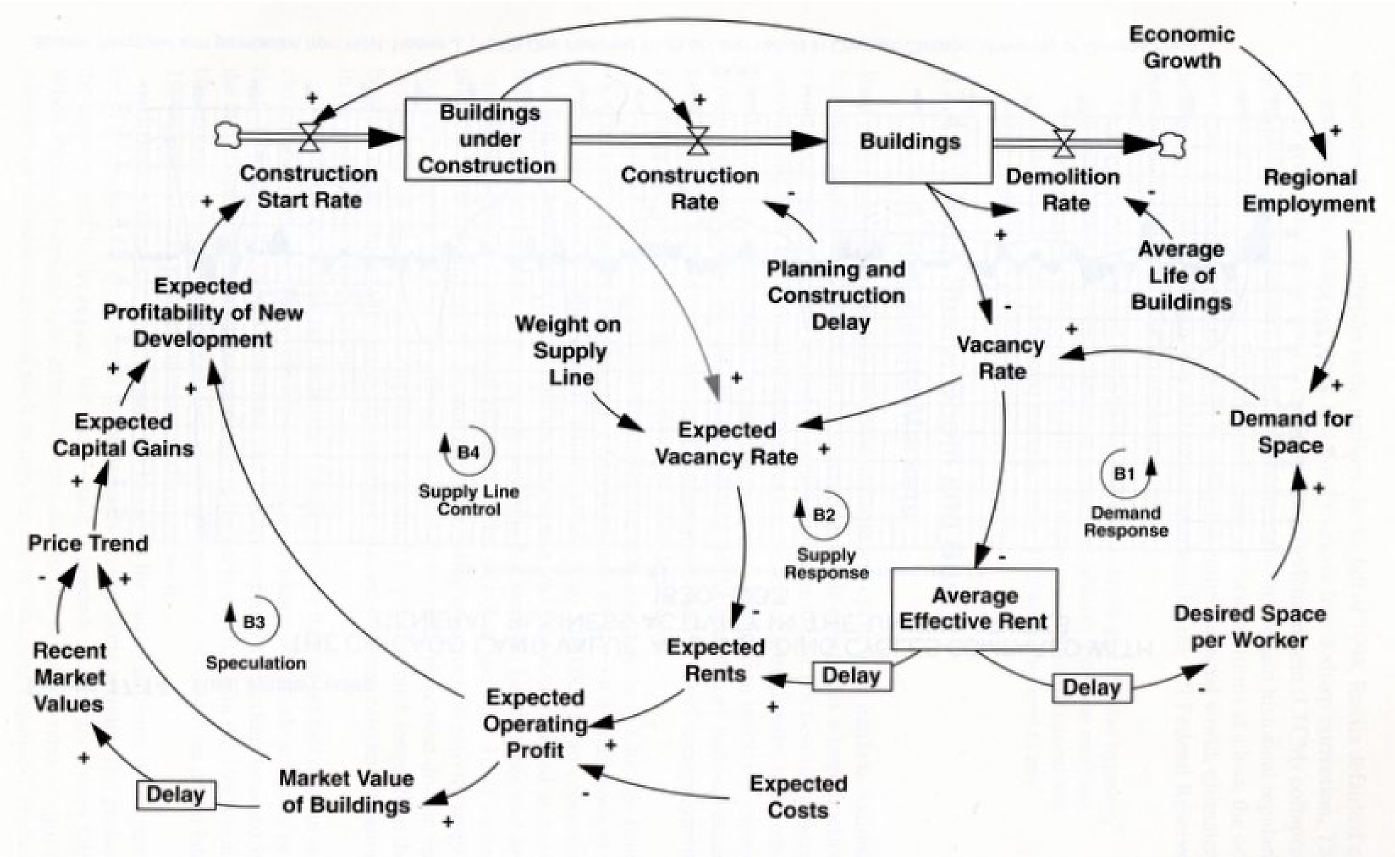


Figure 38- Systems Dynamics Model ("Theory of Change") for housing[8]

Figure 38 is a good example of a Machine Learning model illustrating how we can program the variables and their impacts based on historical data, but only if we have it to start with. To take it to the "Deep Learning" level, as we apply this model, the machine itself will start to learn if our original assumptions were valid and, if they were not, what changes to the model are necessary to bring what the model shows following what really happens.

In researching for this update, the work of Terrence Shin, Data Scientist @ KOHO | Data and Marketing Advisor | Top 1000 Writer on Medium | MSc, MBA | https://www.linkedin.com/in/terenceshin/ stood out in the context of making sense to those of us who are trying to discern the future of AI and ML applied to the project management processes.[9]

Shin identified 5 different types of Machine Learning Models:

- ✓ Ensemble learning algorithms
- ✓ Explanatory Algorithms
- ✓ Clustering Algorithms

[8] John D. Sterman (2000) "Business Dynamics- Systems Thinking and Modeling in a Complex World" page 700, Figure 17-15
[9] Terrence Shin (2020) "All Machine Learning Algorithms you Need to Know for 1922" https://towardsdatascience.com/all-machine-learning-algorithms-you-should-know-in-2022-db5b4ccdf32f

- ✓ Dimensionality Reduction Algorithms
- ✓ Similarity Algorithms

We will try to make this as real as possible in the context of IF and HOW we are using the precursors of these algorithms as project managers today, intending to help you get started developing and using these tools for the future.

- ✓ ***Ensemble Learning Algorithms (Random Forests, XGBoost, LightGBM, CatBoost)***[10]

What are ensemble learning algorithms?

To understand ensemble learning algorithms, you first need to know what ensemble learning is. Ensemble learning is a method where multiple models are used simultaneously to achieve better performance than a single model itself.

Conceptually, consider the following analogy:

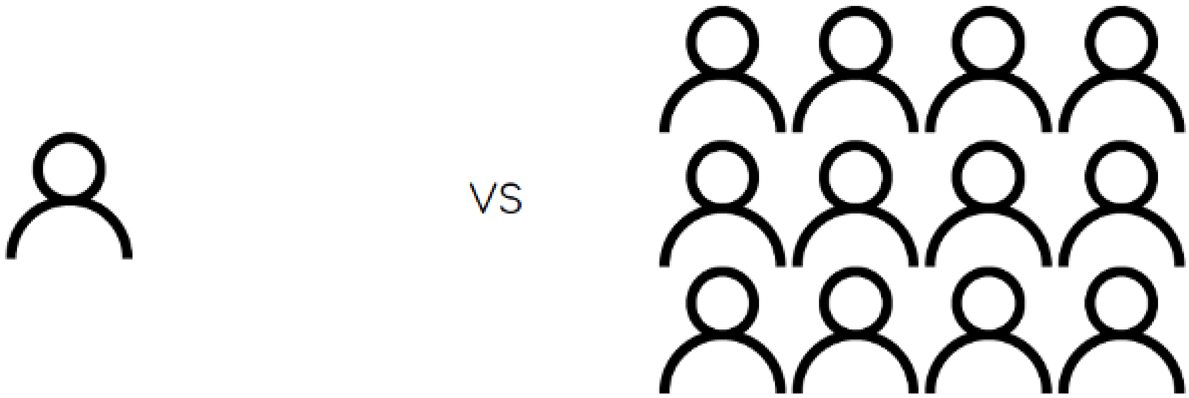


Figure 39- Ensemble Learning[11]

Imagine if one student had to solve a math problem versus an entire classroom. As a class, students can collaboratively solve the problem by checking each other's answers and unanimously deciding on a single answer. On the other hand, the individual doesn't have this privilege — nobody else is there to validate his/her answer if it's wrong.

And so, the classroom with several students is similar to an ensemble learning algorithm with several smaller algorithms working together to formulate a final response.

If you want to learn more about ensemble learning, check out this article:
[Ensemble Learning, Bagging, and Boosting Explained in 3 Minutes]

[10] Terrence Shin (2020) "All Machine Learning Algorithms you Need to Know for 1922" https://towardsdatascience.com/all-machine-learning-algorithms-you-should-know-in-2022-db5b4ccdf32f
[11] Terrence Shin (2020) "All Machine Learning Algorithms you Need to Know for 1922" https://towardsdatascience.com/all-machine-learning-algorithms-you-should-know-in-2022-db5b4ccdf32f

Intuitive explanations and demystifying fundamental concepts
towardsdatascience.com

When is Ensemble Learning useful?
Ensemble learning algorithms are most useful for regression and classification problems or supervised learning problems. Due to their inherent nature, they outclass all traditional machine learning algorithms like Naïve Bayes, support vector machines, and decision trees.

Algorithms
Random Forests
XGBoost
LightGBM
CatBoost

- ✓ ***Explanatory Algorithms (Linear Regression, Logistic Regression, SHAP, LIME)***[12]

What are explanatory algorithms?

Explanatory algorithms allow us to identify and understand variables' statistically significant relationships with the outcome. So rather than creating a model to predict values of the response variable, we can create explanatory models to understand the relationships between the variables in the model.

From a regression standpoint, there's a lot of emphasis on statistically significant variables. Why? You'll always be working with a sample of data, which is a subset of the entire population. To make any conclusions about a population given a sample, it's important to ensure enough significance to make a confident assumption.

In the example below, we illustrate one of many examples from Unit 9 from real projects where we use Regression Analysis to forecast ACWP **(40.7)** and BCWP **(40.6)** into the future to help us identify potential problems BEFORE they become critical. We can also use it to forecast when our current productivity will result in our scheduling having "Negative Float." **(40.9)**

We also use this same method in estimating costs into the future using the "Gold Equivalency" method as shown in Unit 10- Managing Cost Estimating and Budgeting, and we can see examples in this Guild Expert Level certification papers from Stephen Paterson:

1) "Comparison of 8 Common Cost Forecasting Methods" (2018) - https://pmworldlibrary.net/wp-content/uploads/2018/01/pmwj66-Jan2018-Paterson-comparison-of-8-common-forecasting-methods-featured-paper.pdf and

[12] Terrence Shin (2020) "All Machine Learning Algorithms you Need to Know for 1922" https://towardsdatascience.com/all-machine-learning-algorithms-you-should-know-in-2022-db5b4ccdf32f

2) “Incentivizing Early Completion of Oil and Gas Projects” (2017) https://pmworldlibrary.net/wp-content/uploads/2017/11/pmwj64-Nov2017-Paterson-incentivizing-early-completion-of-oil-and-gas-projects.pdf

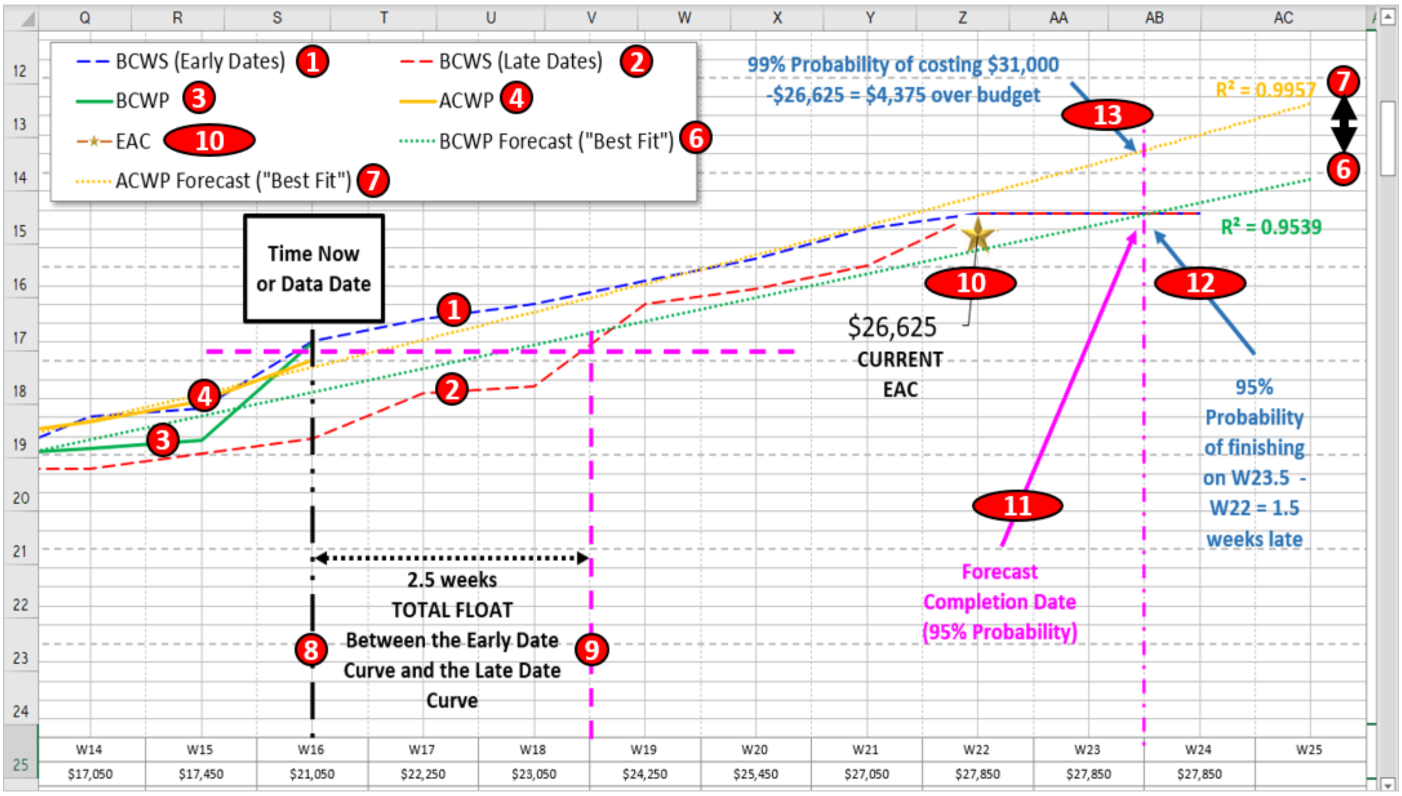


Figure 40- Actual Example Showing the Use of REGRESSION ANALYSIS (Excel “Best Fit” Curve Function)

So what is this telling us? We are already using these tools and techniques, but we have not AUTOMATED them yet. We are still setting them up MANUALLY. And if we want to move into “Deep Learning,” we have to teach the programs how to refine the algorithms using Neural Networking Analysis.

Recently, there’s also been the emergence of two popular techniques, SHAP and LIME, used to interpret machine learning models.

When are they useful?

Explanatory models are useful when you want to understand “why” a decision was made or when you want to understand “how” two or more variables are related to each other.

In practice, the ability to explain what your machine learning model does is just as important as the machine learning model's performance. If you can’t explain how a model works, no one will trust it, and no one will use it.

Algorithms

Traditional explanatory models based on hypothesis testing:

Linear Regression

Logistic Regression

Algorithms to explain machine learning models:

[SHAP](#)

[LIME](#)

✓ ***Clustering Algorithms (k-Means, Hierarchical Clustering)***

What are clustering algorithms?

Clustering algorithms are used to conduct clustering analyses, an unsupervised learning task that involves grouping data into clusters. Unlike supervised learning, where the target variable is known, there is no target variable in clustering analyses.

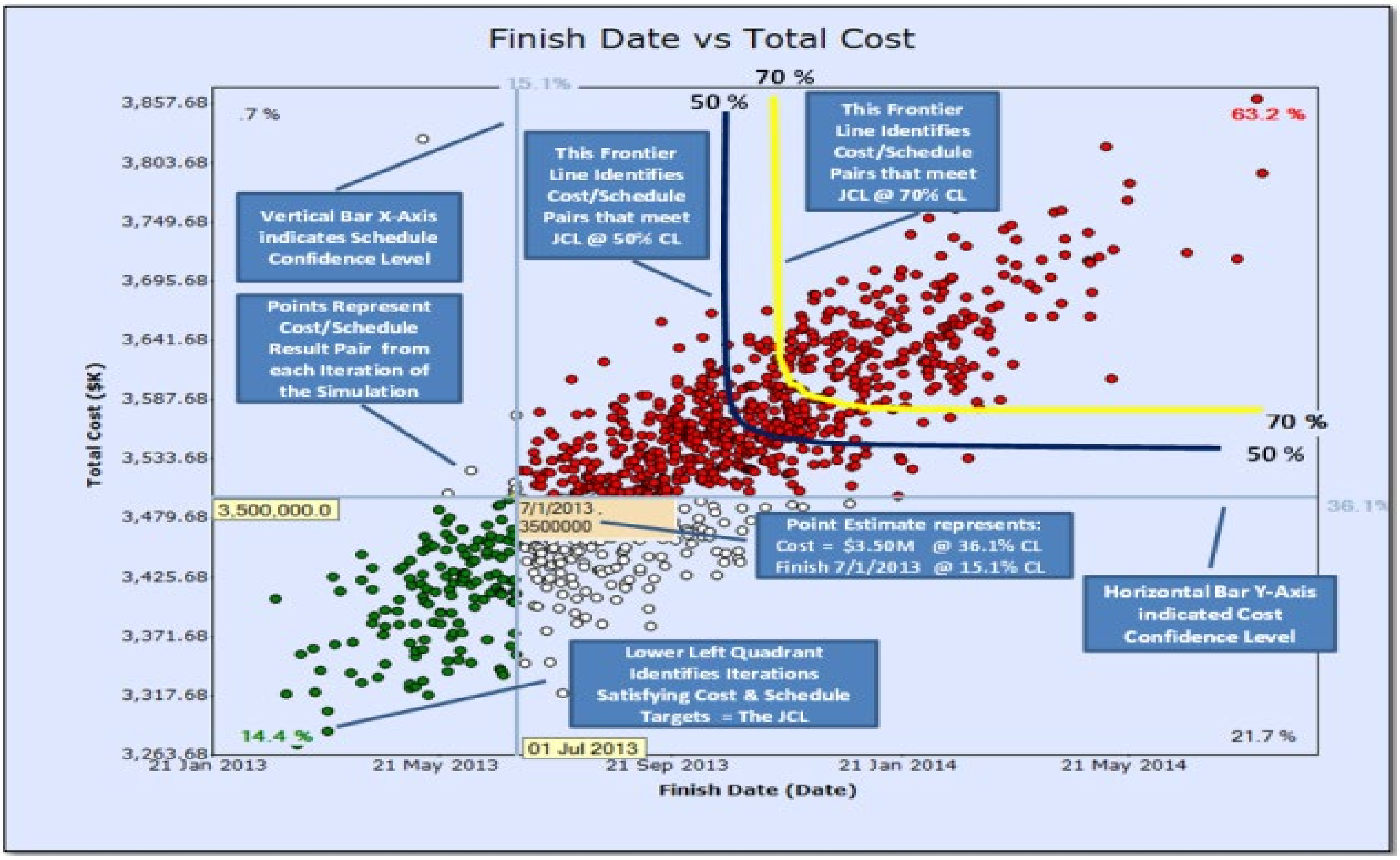


Figure 41- Example of "Stratification" from NASA's Cost Estimating Handbook[13]

[13] NASA's Cost Estimating Handbook (2015)Figure 13, Page 39. https://www.nasa.gov/pdf/263676main_2008-NASA-Cost-Handbook-FINAL_v6.pdf

A perfect example of using" Clustering" is what NASA calls "Stratification" in their Cost Estimating Handbook as shown in Figure 41.

Another example is the SPI vs. CPI graph we use for all our projects, shown in Figure 42.

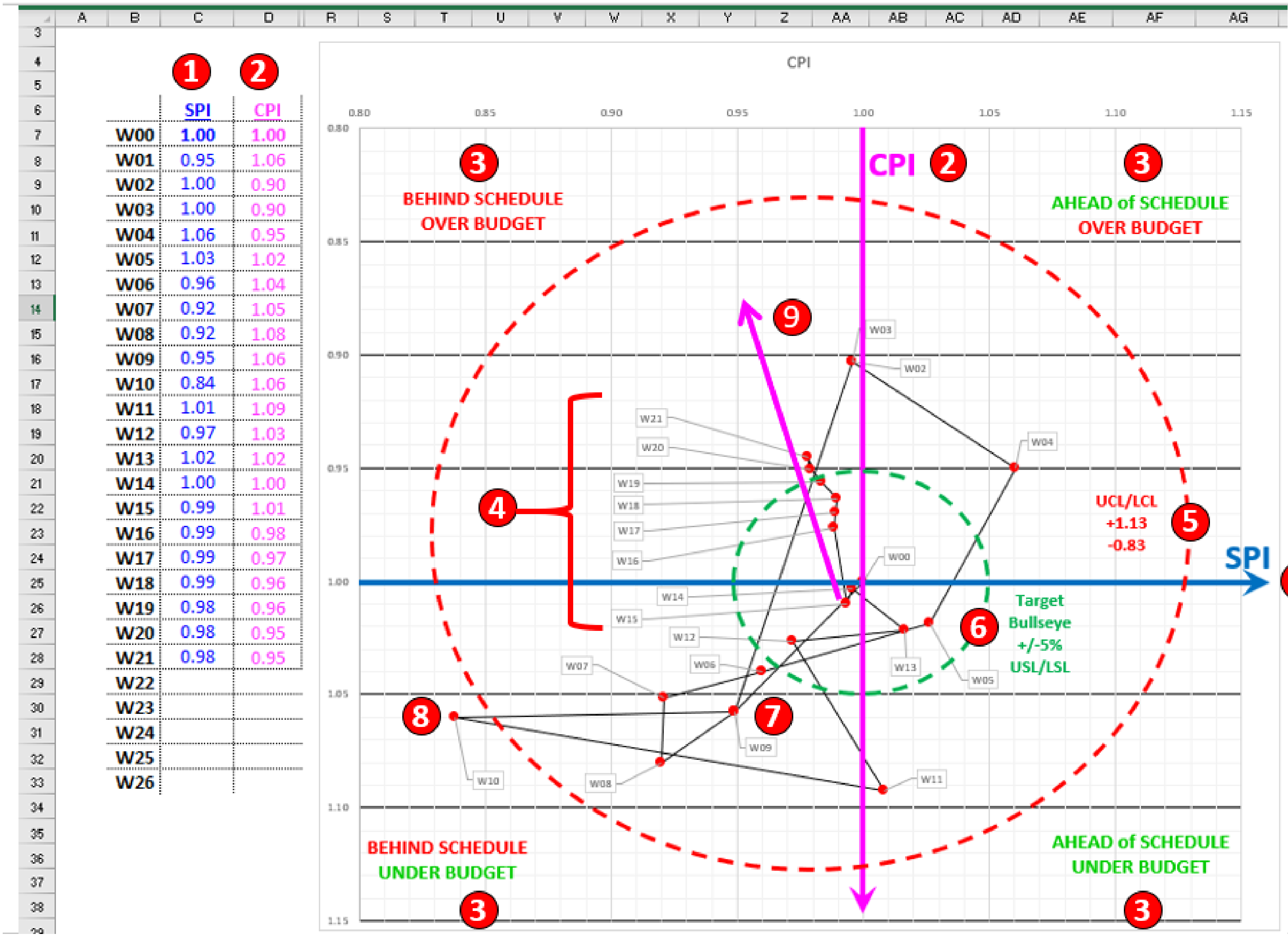


Figure 42- Scatter Diagram Showing SPI and CPI

So here are two "real-life" examples where we already are applying the fundamental tools and techniques associated with Artificial Intelligence. Still, you have to ask if we advocate these as "best tested and PROVEN" practices AND you are not implementing them on your projects yet, how close to being ready for Machine Learning or Deep Learning do you think you or your organization is? Remember that automating a broken system is not likely to add any value.

When are they useful?

Clustering is particularly useful for discovering natural patterns and trends in your data. It's very common for clustering analyses to be conducted in the EDA phase to uncover more insights about the data.

Similarly, clustering allows you to identify different segments within a data set based on different variables. One of the most common types of clustering segmentation is the classification of users/customers.

Algorithms

The two most common clustering algorithms are k-means clustering and hierarchical clustering, although many more exist:

K-means clustering

Hierarchical clustering

✓ ***Dimensionality Reduction Algorithms (PCA, LDA)***

Given the proliferation of Building Information Modeling (BIM) and the use of Virtual, Augmented and Mixed Reality that is automating the entire construction asset life span from concept through demolition, IF we expect the various 1D, 2D, 3D, 4D, 5D and 6 D apps to be able to exchange data, there are two requirements:

1) WBS and CBS coding structures must be STANDARDIZED across entire sectors AND
2) They cannot be "flat file" or hierarchically organized but must be structured as OBJECT-ORIENTED or RELATIONAL databases meaning they can be filtered, sorted and displayed in any way that makes the most sense to different stakeholders.

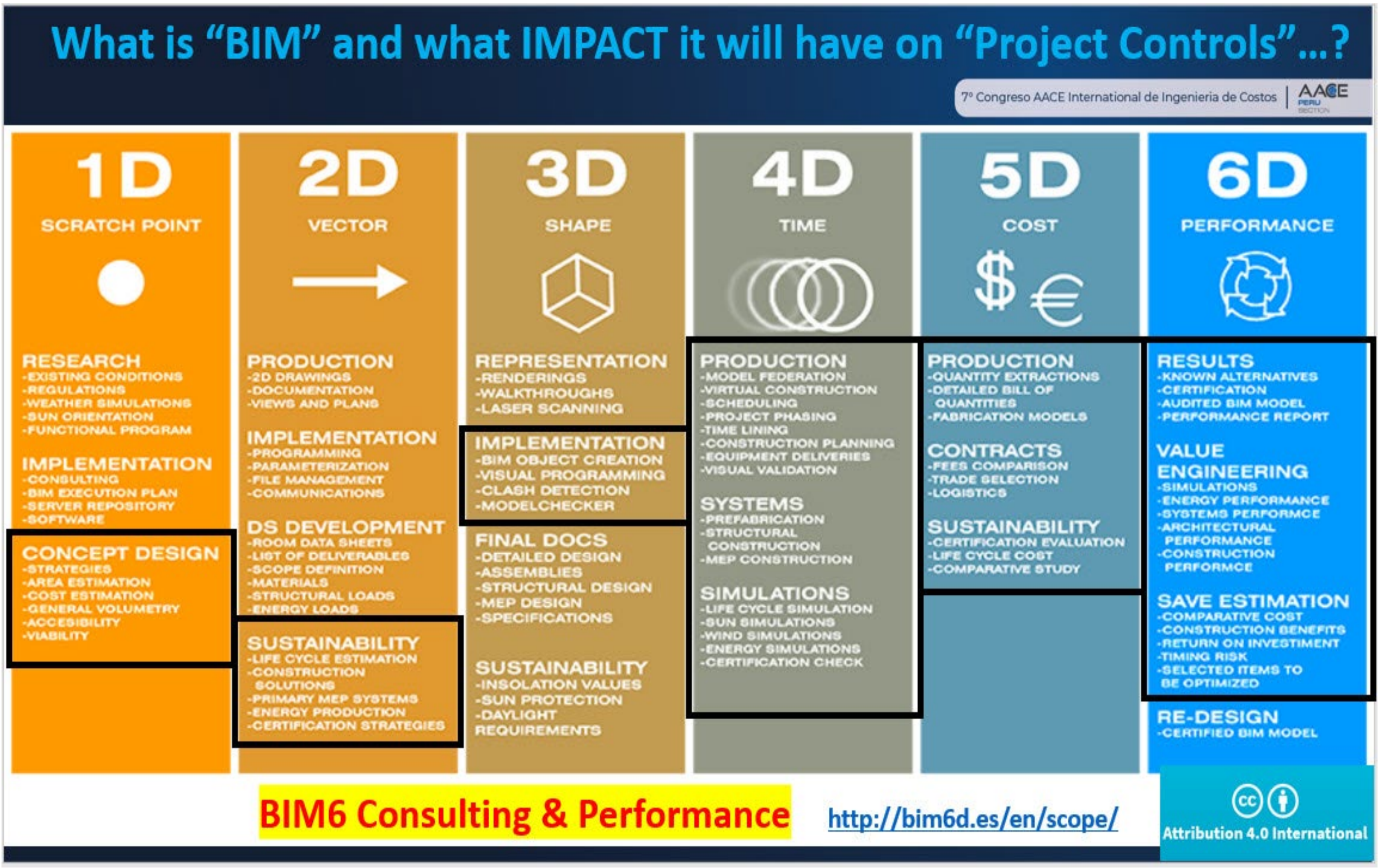


Figure 43- The Future is Here NOW- Building Information Modelling (BIM)[14]

In construction, Building Information Modeling (BIM) already exists and is growing by the day, with many countries (USA, UK, Singapore et al.) mandating the use of BIM for all government-funded projects. This genie is out of the bottle and already is having an impact on construction project management in general and in "project controls" or "Project Management Offices" (PMO's) specifically.

Given that construction, medicine, entertainment and new product development have been around for 6000+ years now, don't you agree that in 6000 years, we should have learned how to "initiate, plan, execute, control and close" projects successfully? And is it unreasonable to expect or urge our colleagues from less mature users of project management processes as the delivery system to "create, acquire, expand, update, repair, maintain and eventually dispose of ORGANIZATIONAL ASSETS?

[14] AACE Presentation in Milan Italy and Lima Peru

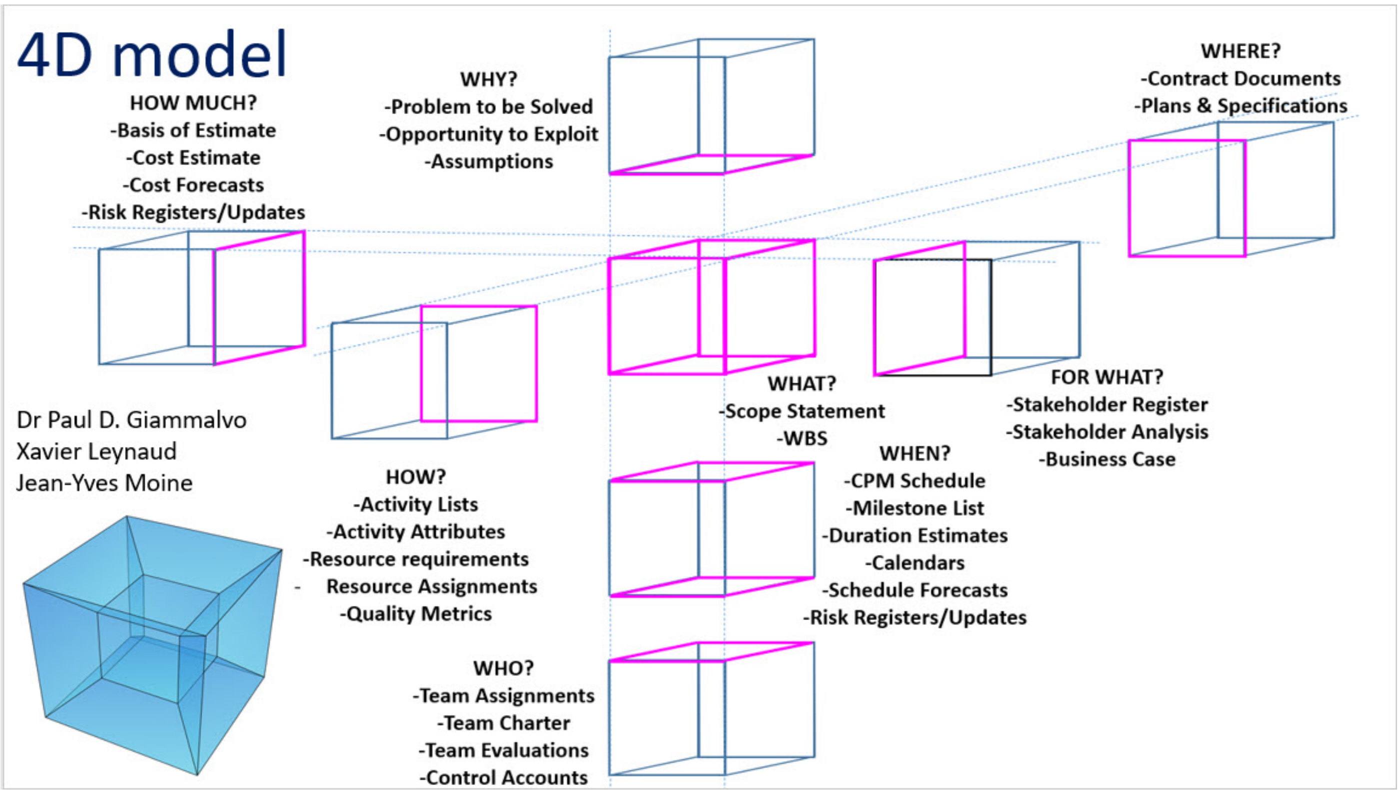


Figure 44- Multi-dimensional, multi-stakeholder WBS/CBS Coding Structures[15]

To get an idea of what the WBS/CBS structure of the future is likely to look like using Virtual, Augmented, and Mixed Reality, take 2 minutes to watch this video on Tesseracts or "Hypercubes" and then take 2 minutes to watch this video of a ship being designed and built using 3D BIM software. This is what is being done RIGHT NOW, and we see little or no evidence that PMI nor AACE, much less IPMA, APM/APMG or any of the other "professional societies," are being proactive in preparing tomorrow's practitioners for this world. (AACE is doing more than most but not as much as the Guild of Project Controls)

After watching these two videos, IF you do not recognize what we see happening right now in the early phases of AI, ML and DL starting, and that if you fail to prepare for it NOW, in a very short period, you are going to become UNEMPLOYABLE, just as happened to the buggy whip manufacturers in the early 1900s when the automobile replaced the horse. These are the times we live in right now. While project management will survive and thrive in this new world, just as it has for the past 6000+ years, there is no indication PMI, AACE, or any other "professional societies can recognize what is happening, much less are being proactive in preparing for it.

---

[15] **Xavier Leynaud, Paul D. Giammalvo Ph.D., Jean-Yves Moine (2019) "Multi-Dimensional Project Breakdown Structures – The Secret to Successful Building Information Modeling (BIM) Integration" https://www.amazon.com/Multi-Dimensional-Project-Breakdown-Structures-Information/dp/1948149125**

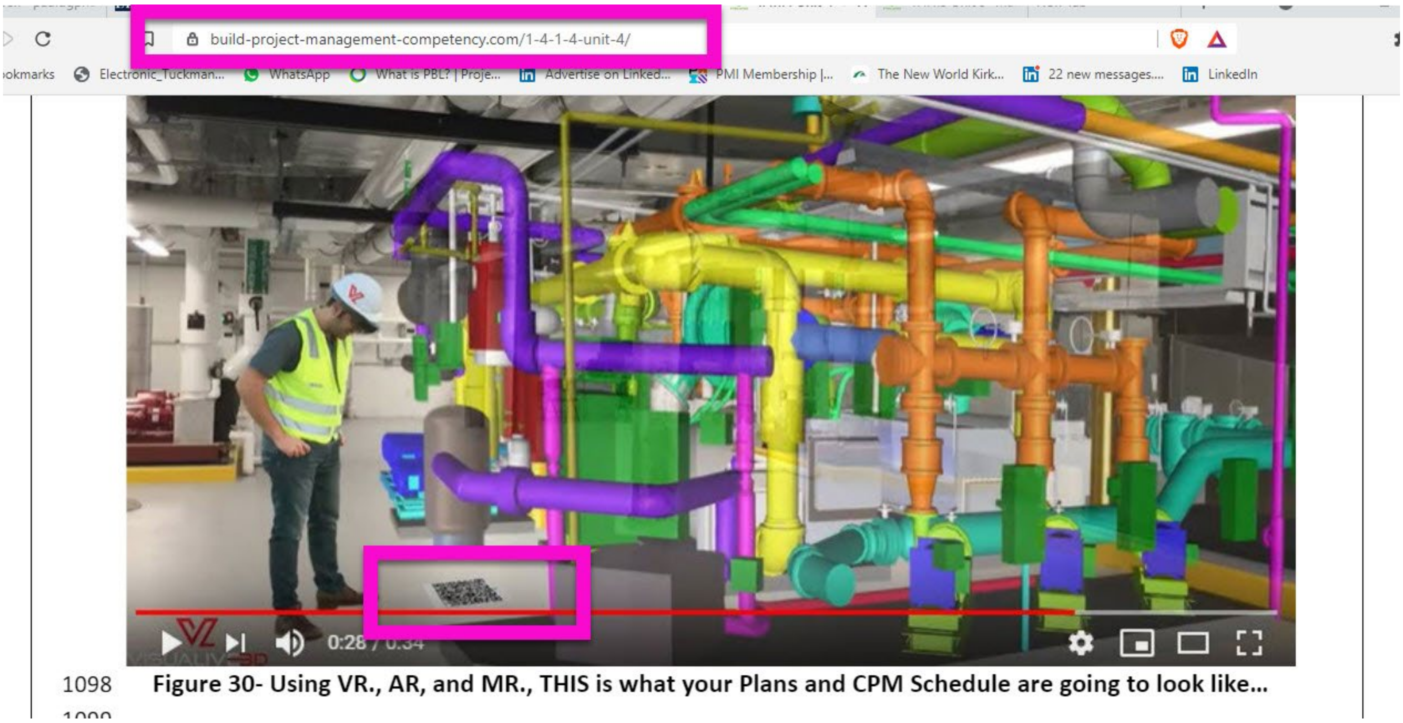


Figure 45- What Project Scheduling, Plans and Specifications Software Is Starting to Look Like.

For those of you who are currently working as "Quantity Surveyors, Cost Estimators, Planners and Schedulers, we hate to be the bearers of bad news, but the evidence is clearly pointing to the use of Virtual, Augmented and Mixed Reality that is going to pretty much replace what you are currently doing. Just as digitizers replace the old-fashioned manual calculation of quantities, how much longer will it be before planning/scheduling and cost estimating are AUTOMATED using 4D and 5D BIM software? Five years? Ten years?

FWIW, Clemson University just started a Drone Pilots course. We urge all our QS, Cost Estimators, Planners and Schedulers to get their Commercial Drone Pilot's licenses and start preparing NOW for the inevitable future.

Given the complexity of the data required to generate VR, AR and MR models or to collect and manage data from drone flights, it is highly likely that we will need to understand dimensionality reduction algorithms.

What are dimensionality reduction algorithms?

Dimensionality reduction algorithms refer to techniques that reduce the number of input variables (or feature variables) in a dataset. Dimensionality reduction is essentially used to address the curse of dimensionality, a phenomenon that states, "as dimensionality (the number of input variables) increases, the volume of space grows exponentially resulting in sparse data.

When are they useful?

Dimensionality reduction techniques are useful in many cases:

They are extremely useful when you have hundreds, or even thousands, of features in a dataset and you need to select a handful.

They are useful when your ML models are overfitting the data, implying that you need to reduce the number of input features.

Algorithms

Below are the two most common dimensionality reduction algorithms:

Principal Component Analysis (PCA)

Linear Discriminant Analysis (LDA)

✓ ***Similarity Algorithms (KNN, Euclidean Distance, Cosine, Levenshtein, Jaro-Winkler, SVD, etc.…)***

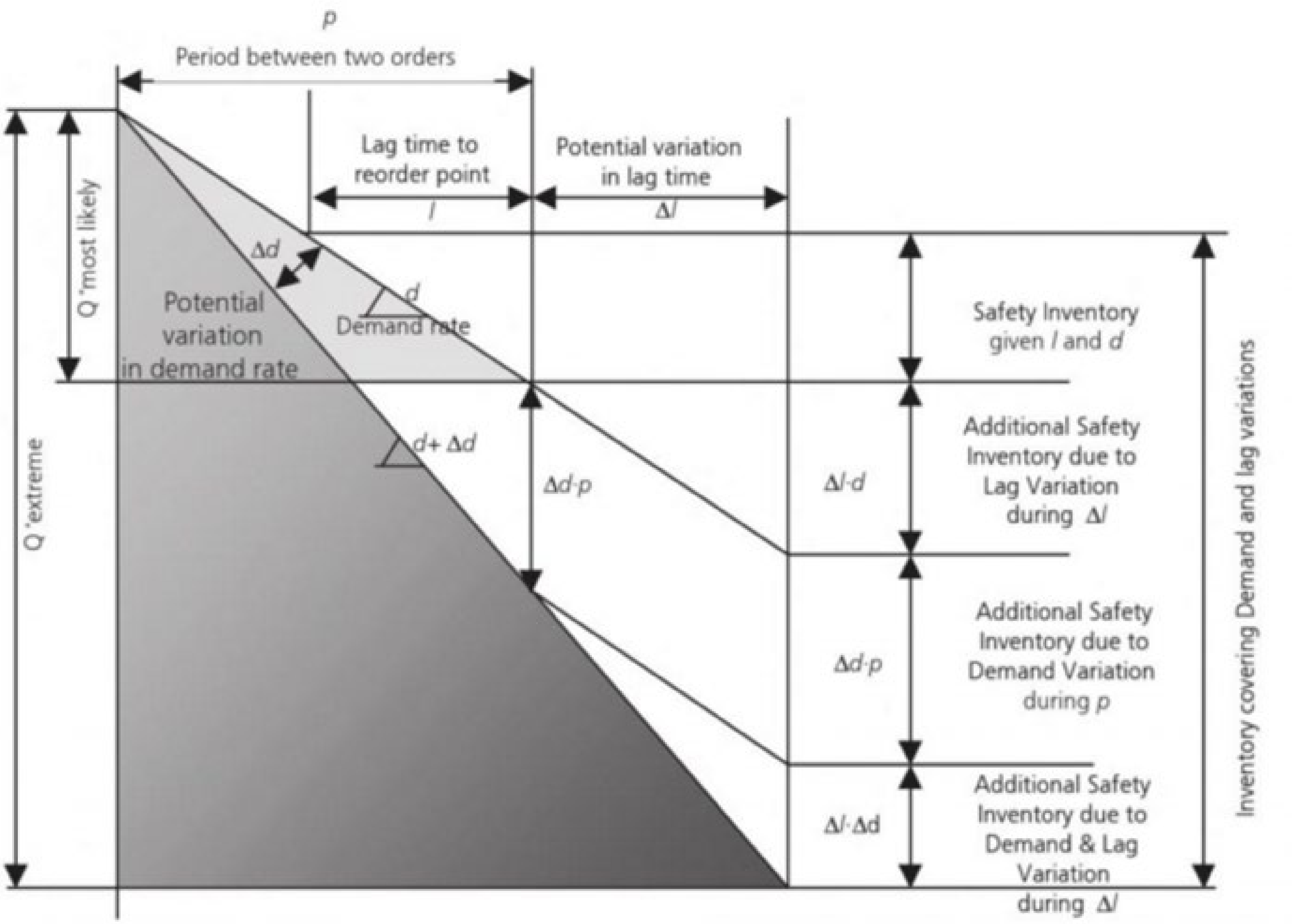
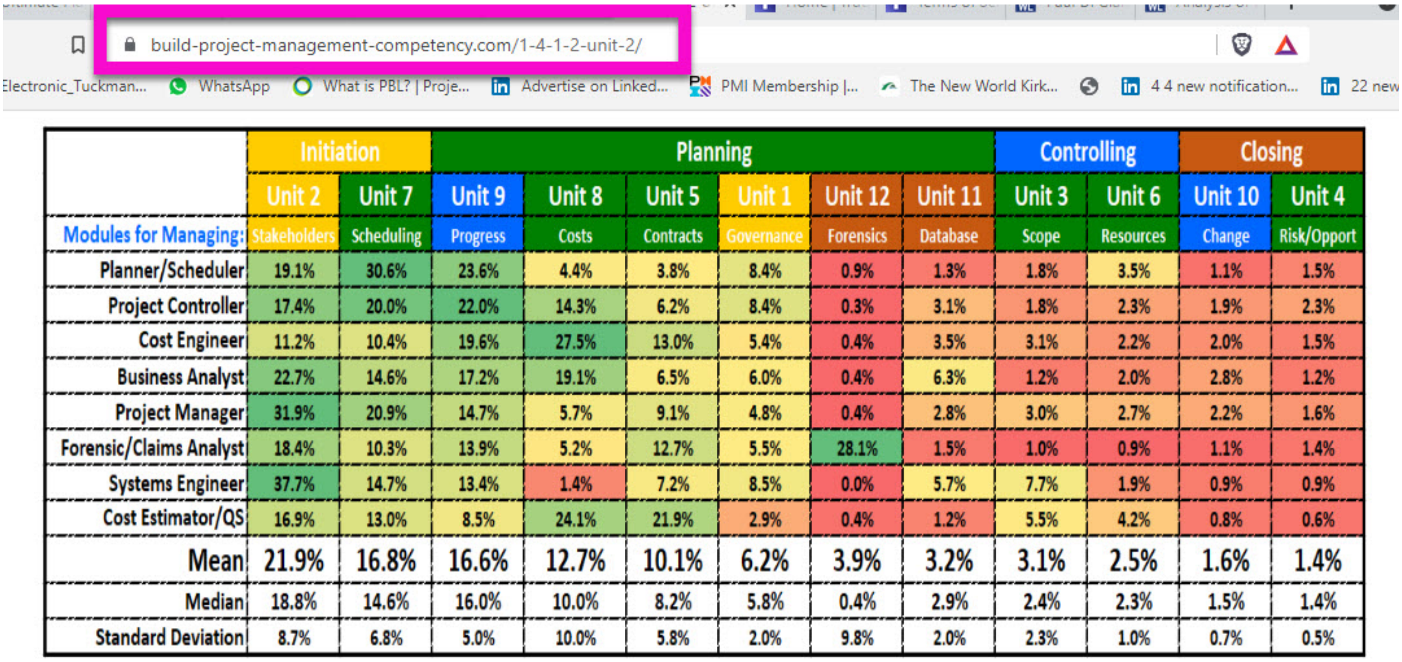


Figure 46- Similarity Algorithm for Procurement and Inventory Management

While the most obvious example showing where we are already using Similarity Algorithms lies in the areas of procurement and inventory management, here is an example where we used keyword analysis software on the global help-wanted advertisements to help us define what skill sets the marketplace were seeking for the following 8 common project management job titles:

1) Planner Scheduler
2) Cost Estimator/Quantity Surveyor
3) Forensic Claims Analyst
4) Project Controller
5) Business Analyst
6) Cost Engineer
7) Project Manager
8) Systems Engineer

And based on the data generated from that keyword analysis, here are the results:



	Initiation		Planning						Controlling		Closing	
	Unit 2	Unit 7	Unit 9	Unit 8	Unit 5	Unit 1	Unit 12	Unit 11	Unit 3	Unit 6	Unit 10	Unit 4
Modules for Managing:	Stakeholders	Scheduling	Progress	Costs	Contracts	Governance	Forensics	Database	Scope	Resources	Change	Risk/Opport
Planner/Scheduler	19.1%	30.6%	23.6%	4.4%	3.8%	8.4%	0.9%	1.3%	1.8%	3.5%	1.1%	1.5%
Project Controller	17.4%	20.0%	22.0%	14.3%	6.2%	8.4%	0.3%	3.1%	1.8%	2.3%	1.9%	2.3%
Cost Engineer	11.2%	10.4%	19.6%	27.5%	13.0%	5.4%	0.4%	3.5%	3.1%	2.2%	2.0%	1.5%
Business Analyst	22.7%	14.6%	17.2%	19.1%	6.5%	6.0%	0.4%	6.3%	1.2%	2.0%	2.8%	1.2%
Project Manager	31.9%	20.9%	14.7%	5.7%	9.1%	4.8%	0.4%	2.8%	3.0%	2.7%	2.2%	1.6%
Forensic/Claims Analyst	18.4%	10.3%	13.9%	5.2%	12.7%	5.5%	28.1%	1.5%	1.0%	0.9%	1.1%	1.4%
Systems Engineer	37.7%	14.7%	13.4%	1.4%	7.2%	8.5%	0.0%	5.7%	7.7%	1.9%	0.9%	0.9%
Cost Estimator/QS	16.9%	13.0%	8.5%	24.1%	21.9%	2.9%	0.4%	1.2%	5.5%	4.2%	0.8%	0.6%
Mean	21.9%	16.8%	16.6%	12.7%	10.1%	6.2%	3.9%	3.2%	3.1%	2.5%	1.6%	1.4%
Median	18.8%	14.6%	16.0%	10.0%	8.2%	5.8%	0.4%	2.9%	2.4%	2.3%	1.5%	1.4%
Standard Deviation	8.7%	6.8%	5.0%	10.0%	5.8%	2.0%	9.8%	2.0%	2.3%	1.0%	0.7%	0.5%

852 **Figure 30- Skill Sets Rank Ordered by Unit**

Figure 47- Skill Sets Rank Ordered by Unit

So what did this analysis show us? That essentially, from the perspective of those writing job descriptions that the marketplace doesn't see a whole lot of difference between job titles and what they need or want us to do. What is even more interesting (??Disturbing???) is that much of what we THINK we are doing to add value (i.e., Risk/Opportunity Management and Change Management) is not reflected in what employers include in their help-wanted advertisements. This tells us that more research needs to be done using more sophisticated software and refinement of the keywords, perhaps?

What are similarity algorithms?

Similarity algorithms compute the similarity of pairs of records/nodes/data points/text. Similar algorithms compare the distance between two data points, like Euclidean distance, and similar algorithms compute text similarity, like the Levenshtein Algorithm.

When are they useful?

Similarity algorithms can be used in various applications, but they are particularly useful for making recommendations.

What articles should Medium recommend to you based on what you previously read?

What ingredients can you use as a replacement for blueberries?

What song should Spotify recommend based on what songs you've liked already?

What products should Amazon recommend based on your order history?

These are just a few of the many examples where similarity algorithms and recommendations are used in our everyday lives.

Algorithms

Below is a non-exhaustive list of some similarity algorithms. If you want to read about more distance algorithms, check out this article. Likewise, if you want to read about more string similarity algorithms, check out this article.

K nearest neighbors

Euclidean Distance

Cosine Similarity

Levenshtein Algorithm

Jaro-Winkler Algorithm

Singular Value Decomposition (SVD) (not exactly a similarity algorithm, but indirectly relates to similarity)

What can we conclude from this research and analysis?

1) At least SOME of the "best tested and proven practice" tools and techniques associated with an asset, portfolio, program, and project management are already used on an "ad hoc" basis.
2) That most if not all of these "best tested and proven" tools and techniques are being run manually- that few if any have truly been automated, beyond Excel or Access databases.
3) That given so many projects continue to "fail" (late and over budget) that we are not yet ready to automate anything until we first fix what is broken.

4) Whether we like it or not, AI, ML and DL are coming to project management, and it represents considerable opportunities for those who can identify and exploit them.

REFERENCES TO UNIT 13- MANAGING DATABASES

- Larry P. English is the father of data and information quality management. His thoughts are still available here:
- Thomas C. Redman, aka the Data Doc, writes about data quality and data in general on Harvard Business Review. His articles are found here:
- David Loshin has made a book with the title The Practitioners' Guide to Data Quality Improvement
- Gartner, the analyst firm, has a glossary with definitions of data quality terms here:
- Massachusetts Institute of Technology (MIT) has a Total Data Management Program (TDQM)
- Knowledgent, a part of Accenture, provides a white paper on Data Quality Management here:
- Deloitte has published a case study called data quality-driven, customer insights enabled:
- An article on bi-survey examines why data quality is essential in Business Intelligence
- The University of Leipzig has a page on data matching in big data environments (they call it dedoop)
- A Toolbox article by Steve Jones goes through How to Achieve Quality Data in a Big Data context
- An Information Week article points to 8 Ways To Ensure Data Quality
- Data Quality Pro is a site, managed by Dylan Jones, with a lot of information about data quality:
- Obsessive-Compulsive Data Quality (OCDQ) by Jim Harris is an inspiring blog about data quality and its related disciplines
- Nicola Askham runs a blog about data governance: One of the posts in this blog is about what to include in a data quality issue log:
- Henrik Liliendahl has a long-time running blog with over 1,000 blog posts about data quality and Master Data Management:
- A blog called Victor Davis Data Craftsmanship provides some useful insights on data management:
- Talend (2021) "Definitive Guide to Data Quality"
- Cost Estimating Databases
  - Spons-
  - Hutchins-
  - Griffiths-
  - Compass International-
  - Marshal & Swift-
  - R.S.Means Cost Estimating Data-
- NASA Configuration Management Handbook

SUPPORTING TEMPLATES TO UNIT 13- MANAGING DATABASES

Owner-Contractor Change Order Templates

- FIDIC documents-
- AIA documents-
- EJCDC documents-
- AGC documents-

- CONSENSUS docs-

Internal Change Order Templates

- NASA Systems Engineering Handbook
- Housing and Urban (HUD) Change Order
- CalTrans (California DoT)
- US Federal Highway Administration Forms

PRINCIPLES, PHILOSOPHIES, BELIEFS OR TENETS TO UNIT 13- MANAGING DATABASES

- Brainy Quotes (n.d.) "Top Ten Change Quotes."
- Brainy Quotes (n.d) "Top Ten Change Management Quotes."

ARTIFICIAL INTELLIGENCE (AI)/MACHINE LEARNING FOR UNIT 13- MANAGING DATABASES

- Arun Singh (2020) "Significance of AI in Databases Management"
- Isabell van Rees Datamize (n.d.) "Artificial Intelligence Database Explained."
- Marina Chatterjee (2020) "Data Science vs. Machine Learning and Artificial Intelligence"
- Sanity Solutions (2020) "20 Data Management Trends for 2021"
- Database Trends and Applications (n.d.) "Definitive Guide to the Machine Learning Life Cycle."

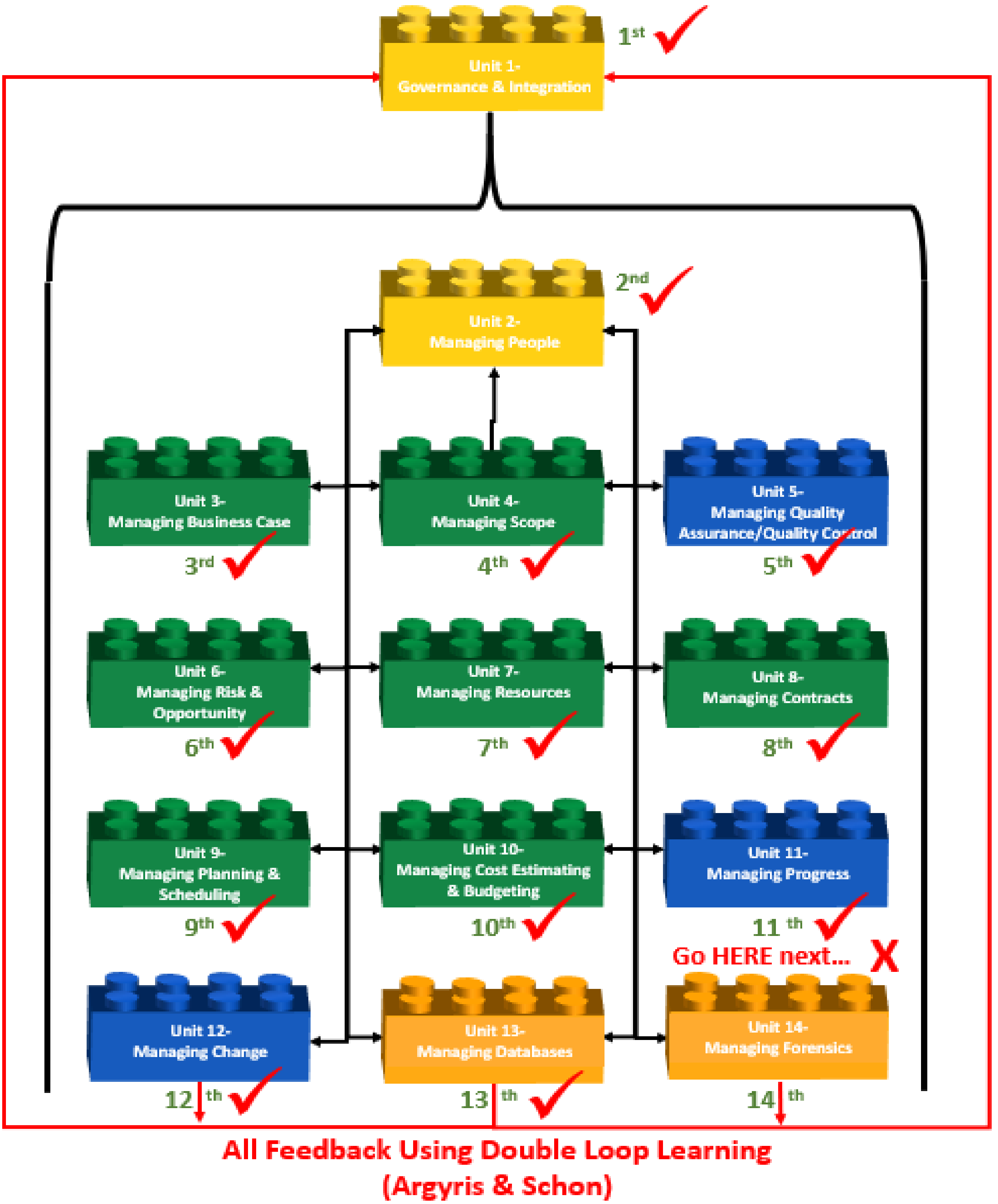


Figure 48- What's Next?