Guideline for project data input

In this document, the different elements (tabs, columns,...) of the corresponding *Project data input sheet* are explained. Moreover, the real-life construction project from that sheet will be used here to elucidate some concepts. The illustrations from the example project can be identified by an indent and a smaller grey font. It is also indicated whether a certain item (i.e. a certain column in the data sheet) is to be considered necessary input or not. The different possibilities are:

Necessary input	Data that always have to be filled in.
Additional input	Data that are very interesting to fill in, however not necessary (i.e. these fields can be left blank, but preferably not). Do note that, when in a certain tab one additional input field is filled in, the other additional input items become necessary input and thus also have to be completed. For example, if in the Resources tab a Name is filled in, the ID, Type, Availability, Cost/Use and Cost/Unit also have to be defined.
Output	Data that can be calculated from the input, or that were already inputted in another tab and repeated for clarification. These fields may always be left blank, and the corresponding concepts will therefore not be illustrated based on the example project.

Baseline Schedule

The baseline schedule presents the planned timing of activities and the precedence relations between those activities, or in other words, the network structure of the project. Moreover, estimated costs are assigned to the activities. The baseline schedule provides the basis for all further project assessments regarding risk analysis and project control (see later sections).

General

ID Necessary input

Activity ID. This is needed later for expressing the precedence relations between activities and the assignment of resources to activities. Note that these IDs do not necessarily have to be consecutive, as is the case in the example project.

> For example, 'hire contractor' can also be indicated as activity 3.

Name Necessary input (work packages = additional input)

Manually chosen activity name.

> In the example project, there are also work packages identified which contain a group of activities and correspond to a project phase (e.g. preparation, foundation, shell, etc.).

WBS Additional input

The work breakdown structure ID of the activity, indicating the project phase or work package the activity belongs to.

For example, 'placing sewerage' with WBS ID 1.2.4 is part of the 'foundation' phase with WBS ID 1.2.

Relations

Predecessors Necessary input

List of activities that immediately precede the current activity through precedence relations.

Successors Necessary input

List of activities that immediately succeed the current activity through precedence relations.

All relations in the example project are standard finish-start relations with a minimal time-lag of 0 time periods. For example, 2FS (in the predecessor field of activity 3) and FS3 (in the successor field of activity 2) represent a zero time-lag finish-start relation between activity 2 and 3, meaning that activity 3 can start as soon as activity 2 is finished.

These relations are indeed the most common. However, there are other, more general precedence relations possible:

- Finish-start (FS): An activity can only start x days after the finish of its predecessor activity, e.g. 1FS+xd (predecessor field) or FS2+xd (successor field).
- Finish-finish (FF): An activity can only finish x days after the finish of its predecessor activity, e.g. 1FF+xd (predecessor field) or FF2+xd (successor field).
- Start-start (SS): An activity can only start x days after the start of its predecessor activity, e.g. 1SS+xd (predecessor field) or SS2+xd (successor field).
- Start-finish (SF): An activity can only finish x days after the start of its predecessor activity, e.g. 1SF+xd (predecessor field) or SF2+xd (successor field).

Note that time-lags between activities can be positive as well as negative, e.g. 1FS-xd. Also, it is possible for an activity to not have any predecessors or successors, thus being an independent activity that influences no other activity.

For example, activity 4 is such an activity that is not related to any other activity.

Baseline

Baseline Start Necessary input

Planned starting date of the activity, according to the baseline schedule.

Baseline End Output

Expected end date of the activity, according to the baseline schedule.

Duration Necessary input

Expected duration of each activity, according to the baseline schedule.
Note that all durations in the data sheet are always expressed in WORKING days (and working hours)! More information on the definition of working days and hours will be provided in the section Agenda.

> In the example project, all days from Monday till Friday are defined as working days (see section Agenda). Activity 3 is expected to take 20 days, that is, 20 WORKING days. Indeed, when we add 20 working days to the baseline start of 10/07/06, we are at the baseline end of 04/08/06.

Resource Demand

Resource Demand Additional input

The list of resources (including the required units) that are needed to perform a certain activity.

> For example, activity 6 requires 4 of the 8 available workers from the subcontractor team.

Note that if no unit requirements are added, it means that only one unit of that resource is used for the considered activity.

> This is for example the case for activity 17, for which only one subcontractor worker is employed.

Resource Cost Output

The expected cost of the listed resources that are needed for a certain activity (see section Resources for a more elaborate discussion).

Baseline Costs

Fixed Cost Necessary input

The planned fixed cost of an activity; this is a fixed amount of money which is independent of the duration of the activity.

Cost/Hour Output

The expected hourly cost of an activity which will define the variable cost of that activity (see next item).

Variable Cost Necessary input

The expected variable cost of an activity, calculated by multiplying the expected activity duration (in hours) with the expected hourly cost (see previous item). This is thus a variable amount of money which is dependent on the activity duration.

In the example project, resources are explicitly entered. In such a case, the above fixed and variable costs do not contain any cost of resources; it are costs that are set for an activity and not for a resource (i.e. it are activity costs and not resource costs). A typical example of a variable activity cost is activity overhead, which is indeed not related to resources. The cost of the resources are then calculated separately (i.e. the resource cost that was mentioned earlier and will be discussed further in the Resources section).

However, the data provider might not dispose of explicit resource data (Which resources were used for the project? How many units were available? How much did they cost per hour? Etc.). This would imply that the Resources tab in the data sheet could not be filled and thus remains blank. However, the data provider might have access to the general fixed cost and variable cost data of the activities, which would also have to include the cost of resources used for those activities. In that case, resources and resource costs are thus not explicitly entered (the explicit resource cost mentioned earlier would appear to be 0), but they can be counted for through the implicit incorporation of the resource costs in the fixed and variable activity costs.

Total Cost Output
The total of fixed costs, variable costs, and resource costs.

Resources

The definition of the available resources in a project and the resource requirements of the different activities is in fact a component of the baseline schedule. Here, we placed the description of the resources in a separate section (i.e. tab) for clarity reasons.

General

ID Additional input
Resource ID.

Name Additional input
Manually chosen resource name.

Type Additional input
There are two types of resources:
- Renewable: are available on a period-by-period basis, i.e. the available amount is renewed from period to period (i.e. per hour). Only the total resource use at every time instant is constrained. Typical examples are manpower, machines, tools, equipment, space, etc.
- Consumable (non-renewable): are not constrained on a periodic basis and have an unlimited consumption availability for the entire project. Typical examples are money, raw materials, energy, etc.

> In the example project, all resources are renewable (moreover, all related to manpower), which is also the most common category.

Availability Additional input
The availability of the resource, expressed in units. In principle, this availability might differ in time, but usually it remains constant. Resource availability is of course only relevant for renewable resources as consumable resources have an unlimited availability.

> For example, there are 8 workers available in the subcontractor team.

Resource Cost

Cost/Use Additional input

The one-time cost that is incurred every time that the resource is used by an activity. The per use cost is thus unrelated to the activity duration (in fact, this per use cost can be seen as the fixed portion of the resource cost).

For renewable resources, the total per use cost is obtained by multiplying the per use cost by the required number of resource units for that activity.

> For example, if the per use cost of a crane is € 200 and 3 cranes are used for a certain activity, then the total per use cost of the resource for that activity is € 200 x 3 = € 600.

For consumable resources, the per use cost is only applied once from the moment the resource is used, which is logically, as there is an unlimited availability of those resources.

Cost/Unit Additional input

This is a cost that depends on the amount of resources required by the activity and on the activity duration (thus, this per unit cost can be seen as the variable portion of the resource cost). The per unit cost rates are calculated per time unit (normally hours) and per resource unit (e.g. per worker).

> In the example project, where there are only renewable resources with per unit cost (the most common case), the per unit cost of the resource 'team subcontractor' expresses that one worker (i.e. one unit) of the subcontractor team cost € 38.56 per hour.

Resource Demand

Assigned To Output

The list of activities (identified by their ID) for which a certain resource is used, including the number of resource units needed for that activity. Logically, this number has to be lower than the resource availability.

Total Cost Output

The total expected cost of a certain resource over all activities for which it is used. This cost thus includes per use costs and per unit costs based on the expected baseline durations.

Risk Analysis

Risk analysis depends upon the definition of distribution profiles for the activity durations. Based on this input, sensitivity information of activities and simulation-based predictions of project duration and cost can be obtained.

General Output

Already explained earlier.

Baseline

Duration Output

Expected duration of the activity, but now expressed in hours since hours are used to define the activity duration distribution profiles here.

Activity Duration Distribution Profiles

Description Additional input

The description of the nature of the risk distribution profile of a certain activity's duration. First of all, we assume that all profiles are triangular.

There are two main options for defining the profiles: you can choose one of the four standard (predefined) profiles or you can manually enter a more specific risk profile for an activity's duration.

The four standard profiles are:

- No risk: the activity entails no risk and the duration is a single point estimate.
- Symmetric: the activity is subject to risk within a certain range, with worst case and best case scenario symmetric above and below the average.
- Skewed (to the) right: the activity is subject to risk within a certain range, where activity delays are more likely than early activity durations.
- Skewed (to the) left: the activity is subject to risk within a certain range, where early activity durations are more likely than activity delays.

It is important to mention that, if the data provider decides not to select the most appropriate standard profile for an activity or - even better - manually enter an activity-specific one, the standard symmetric profile is assumed for that activity.

All the predefined profiles are expressed in terms of relative durations (logically, as it should be possible to apply them to any activity), more specifically, as percentages of the baseline duration of the activity.

All profiles can be described through three duration estimates:

- Optimistic: the lowest possible duration of the activity, the activity could never be completed faster (i.e. best case scenario). This estimate corresponds to the (left) start point of the risk triangle.
- Most probable: the most likely duration of the activity, the duration that we expect to have the greatest probability of occurring. This estimate corresponds to the top of the risk triangle.
- Optimistic: the maximal possible duration of the activity, the activity could never take longer (i.e. worst case scenario). This estimate corresponds to the (right) end point of the risk triangle.

The standard profiles are described by the duration estimates in the table below. For these standard profiles, the estimates are thus fixed and do not have to be entered manually. Furthermore, an example activity using one of the standard risk profiles is presented in the table.

	Optimistic	Most probable	Pessimistic	Example
No risk	99	100	101	activity 16
Symmetric	80	100	120	activity 4
Skewed right	80	90	120	/
Skewed left	80	110	120	activity 67

As already mentioned, all these estimates are percentages of the activity's baseline duration, thus they indeed express relative distributions.
Manually inputted profiles can also be represented as relative distributions, but here there is another possibility, namely to characterize the distribution by absolute durations (in hours or days) instead of percentages of the baseline duration.

> This is done for several activities in the example project. Here, hours are used as a time measure. For example, consider activity 2. The optimistic estimate is 402 hours, expressing that the activity will never be completed faster than that. The most likely duration of the activity is set on 480 hours. And the pessimistic estimate is 812 hours, meaning that we do not expect the activity ever to exceed this duration.

Project Control

Project control comprises the monitoring of the actual progress of a project, concerning both time and cost. Performing project control thus consists of periodically keeping track of the schedule and cost performance of all activities. This periodic monitoring approach thus yields schedule and cost data for multiple time instances during the project, called tracking periods (not. TP).
It is important to note that one might not dispose of such periodic monitoring data, but only of the final actual progress data obtained after project completion. This would come down to the existence of only one TP, situated at the end of the project. These data can also be of interest, however, periodic tracking data is of course strongly preferred.

The definitions below apply for every TP tab in the corresponding data sheet. Obviously, if only the final actual progress data of the project (i.e. obtained after project completion) were provided, there will only be one TP (tab).

TP Status Date — Necessary input

The status date is the end date of a TP, more specifically, the date at which the status (or progress, or performance) of the project is evaluated.

> For example, the status date of TP3 is 16/03/07 (end of the day). This means that the tracking data that are shown in this tab reflect the progress that the project had already made by the end of 16/03/07.

TP Name — Additional input

The chosen name of the TP. For example, this name can indicate that tracking was performed when a certain project phase was completed.

> This was done in the example project. E.g. TP4 is called 'Shell' because it indicates that the tracking has been performed after the completion of the shell construction project phase.

General Output
Baseline Output
Resource Demand Output
Baseline Costs Output

All already explained, but also included in this tab as a potential basis for comparison with the tracking results.

Tracking

Note that tracking periods do not necessarily need to have the same length, however, that is the most common approach (e.g. monthly monitoring of the project).

> Another option is displayed by the example project, where the status dates coincide with the completion of certain project phases (e.g. preparation, foundation, shell, etc.). One can indeed observe that the interval between two status dates is not constant here.

Actual Start Necessary input

The actual starting date of the activity. Obviously, this has to be a date before (or on) the current status date.

> For example, the status date of TP3 is 16/03/07 (end of the day), and indeed, all activities that are started (or have already finished) have a starting date earlier than the TP status date (16/03/07 is the latest starting date, namely that of activity 23).

The actual starting date might of course differ from the expected starting date according to the baseline schedule.

> For example, activity 17 should have started on 11/12/06 according to the baseline schedule but is instead only started on 09/02/07.

Actual Duration Necessary input

The actual duration of the activity, or more specifically, the time actually spent on the activity (beginning on its actual start date) up to the current status date. For an activity that has not started yet, of course, the actual duration is 0.

> Consider TP3. Activity 19 has not started yet, so indeed, its actual duration is 0.

For an activity that has already started but has not yet been completed on the current status date, the actual duration would be the number of (working) days between the actual activity start date and the status date.

> Consider TP3. Activity 18 has already started on 28/02/07 but is still in progress at the status date of 16/03/07. Therefore, the actual duration of this activity is 13 days, being the time span (in working days) between the actual start date and the current status date.

For an activity that has finished before the current status date, the actual duration is obviously the number of (working) days between that activity's actual start date and actual end date.

> Consider TP3. Activity 17 started on 09/02/07 and ended on 22/02/07. Indeed, there are 10 (working) days lying between these two dates, explaining the actual duration. Also notice that the end date (22/02/07) lies before the current status date (16/03/07), thus indeed indicating that the activity is finished.

Again, the actual duration of activities that have already been finished on the current status date can differ from its baseline duration.

> For example, activity 17 was expected to take 14 days but instead was completed in only 10 days.

Remember that the actual duration, just like the baseline duration, is expressed in WORKING days!

PAC Output

The planned actual cost is the cost that was planned to occur following the presumption that fixed costs remain the same as planned and variable costs (both activity and resource costs, see earlier) evolve linearly with the actual activity durations.

PRC Output

The planned remaining cost (same principle as the PAC) is the cost that is anticipated to occur in the future following the presumption that fixed costs remain the same as planned and variable costs (both activity and resource costs, see earlier) evolve linearly with the (anticipated) remaining activity durations (see next item).

Note that fixed costs are always incurred in their totality at the start of an activity, so they are never part of the PRC if an activity is already started (because then they are already added to the PAC).

Remaining Duration Output

The anticipated time still needed - given the current schedule performance of the activity - to finish the activity. Thus, this is in fact a forecast.

Also remark that when all tracking data is received at once after the project has ended, one could in fact, for every status date and for every activity that is still in progress at that time, know exactly what the remaining duration for that activity will be. However, this information was not available at the actual moment of that status date and, therefore, using it would inflict a bias. Therefore, it is advised not to adapt the remaining duration forecast based on the current schedule performance, unless management had, at the time of the considered status date, indeed made a concrete prediction (e.g. based on their own prospect or those of workers) for the remaining duration of the activity.

PAC Dev Necessary input

The deviation between the PAC and the actual cost (i.e. did the activity cost more/less than anticipated under the PAC presumptions?).

> In the example project, the PAC Dev is always 0, indicating that the PAC presumptions always seem to be correct here (perhaps because of the very thorough definition of the resources) and PAC thus equals the actual cost.

However, for many other projects this will not be the case and the actual costs incurred for an activity on the current status date will differ from the PAC anticipations (e.g. because of resources working less efficient than expected and thus requiring overtime, because of fixed cost appearing to be higher than expected, etc.).

PRC Dev Additional input

The deviation between the PRC and the remaining cost (i.e. is it foreseen that the activity will cost more/less in the future than expected from PRC calculations?).

> In the example project, the PRC Dev column is left blank. In that case, a PRC Dev of 0 is adopted for every activity, which expresses that the future costs of the activities are assumed to be equal to the PRC.

Actual Cost Necessary input

The actual cost that has already been incurred by the activity on the current status date. For activities that are already finished on the current status date, the actual cost of that activity will of course be the final actual cost of that activity. For activities that are still in progress on the current status date, the actual cost will be some fraction of the final actual cost of that activity.

> An example. At the status date of TP3, activity 18 is still in progress (PC = 65%) and has already cost € 110,554.45. The activity is completed during TP4, however it took an extra € 8,328.96 of expenses to finish it, leading to a final actual cost of € 118,883.41 for activity 18.

Remaining Cost Output

The foreseen remaining cost of the activity. Note that, whereas the actual cost is a known number at the current status date, the remaining cost is in fact - just as the remaining duration - a forecast, as it concerns costs that will be made in the future and are therefore not yet known.

Also remark that when all tracking data is received at once after the project has ended, one could in fact, for every status date and for every activity that is still in progress at that time, know exactly what the remaining cost for that activity will be. However, this information was not available at the actual moment of that status date and, therefore, using it would inflict a bias. Therefore, it is advised to leave PRC Dev at 0 (so assume that the remaining cost is equal to PRC, as was always done in the example project), unless management had, at the time of the considered status date, indeed made a concrete prediction (e.g. based on their own prospect or those of workers) for the remaining cost of the activity.

Percentage Completed Necessary input

The percentage complete (PC) is the portion (percentage) of an activity that is estimated to be completed at the current status date. The PC thus relates to the estimated physical progress of an activity. This estimation can be made by the people performing the activity when quantifying the physical progress is not straightforward (e.g. writing a software program; it is difficult for an outsider to determine the PC of such a task, but the programmer himself should be able to provide an estimation).

> E.g. for activity 57 of the example project this means that 80% of the floor surface was laid by the end of TP6. Here, the physical progress is of course easy to be quantified as it is very tangible. Indeed, if 1,600 m^2 of the total floor area of 2,000 m^2 was already laid, the PC = 1,600 / 2,000 = 80%.

Tracking Output

Indicates whether an activity has not started yet (PC = 0%), has already started but has not yet finished (0% < PC < 100%), or has already finished (PC = 100%) at the current status date.

Earned Value (EV) Output

The value that has actually been earned at the current status date.

Planned Value (PV) Output

The value that should have been earned (according to the baseline schedule) at the current status date.

More information on EV, PV and related Earned Value Management (EVM) and Earned Schedule (ES) concepts can be found in the books *Measuring Time* (2009) and *Project Management with Dynamic Scheduling* (2012) by Mario Vanhoucke.

Agenda

In this tab, the data provider can specify the working hours and working days for the considered project, as well as the holidays.

Working Hours Necessary input

The hours that work is performed on a normal working day.

For the example project, there are standard eight-hour working days from 8 am to 5 pm, with a lunch break from 12 pm to 1 pm.

Working Days Necessary input

The days that work is performed in a normal working week (i.e. a week without holidays).

For the example project, there is a standard working week from Monday till Friday.

Holidays Additional input

A list of days that would normally be working days, but on which no work is performed due to e.g. construction leave, Christmas holidays, etc.

For the example project, no holidays were inserted. However, for illustration purposes, 18/06/2015 and 19/06/2015 (by which the project had long been completed) were entered as holidays.