# Crunch your numbers

Data Science for Team Leaders and Business Unit Managers

By Dmytro Karabash and Maxim Korotkov

## Introduction

You’re the boss now. You have a team or a business unit doing consultancy. Probably a bunch of Consultants helping clients, some Project Managers leading your projects in a nice Agile or whatever way. Everyone fills their timesheets (and you cut their bonuses if they don’t) – time is billed to clients and you have Fixed Fee projects also. Maybe smaller teams being parts of your team also – you can lead ten, hundred or thousand people (*hey, dude – you run a unit of 1000 employees? Don’t read this article – hire someone who did)* You probably even have an ERP or something and a CFO in a nice corner office. Do you have a good way to estimate how much profit every team member and project brings and ways to predict it with good precision? If you have a nice ERP dashboard which gives you all then – then you’re lucky and you don’t need this article either. But 70% chances are [a reference to research would be nice - trello] that you get a simple calculation end of year – like “let’s take all revenue finance recognized on your projects in a year and man day cost at $800 to get your unit profitability” ($800 can seem crazily high or unacceptably low, depending on business you’re in). So that’s the precision you have to the money your team brings to firm? Do we need to remind you that this is the only reason why your job exists?

What you probably have is – all the timesheets, projects with their budget and you can even get to approximate costs nicely (we’ll get there). Shitload of data for a year even for a team of 10. What I will tell you here is that you don’t need an ERP to do the rest – the notebook in Jupyter will do for now.

Let’s remember - management starts with measurement. How can you supervise your team and projects if you don’t have their data at your fingers?

## Managerial Accounting

Let me bore you for a minute. Timesheets and projects budgets are simple. You are probably frightened by IFRS revenue recognition and all that – that’s why your CFO has a nice corner office and you don’t, right. Fortunately, there is a simpler thing called Managerial Accounting [reference - trello].

So, profit is Revenue – Cost and we intend to measure by it and keep it simple.

**Revenue**: Let’s say that for a Time & Material project you generate as much revenue as you bill (on agreed rate) up to its budget. For a Fixed Fee you split all the revenue per month as per % of effort you spend / are going to spend in the month (changes can be complicated, but they are easy to be managed in the way you want – just go through introduction and wait for the notebook we promised)

**Cost**: So, let’s start with the fact that the average cost of $800 (or whatever) is bullshit. You might have people working in different countries and of totally different levels. Talk to your finance team and get some estimates from them (or make it together). We’d say if you were appointed to run a team or business unit – they owe you that much. We will assume below that you managed to get 3 grades (Associate, Base, Senior) per Country you have staff. Though costs per individual will go nicely also.

Let me summarize it here for you – in any decent firm, if you are a team leader or business unit manager you are involved in merit processes – you can get

* Dictionaries
  + Cost per region or default cost for “outside” contributors
  + Average revenue per hour for projects not “owned” by you
* HeadCount
  + Graded costs
  + Engagement model (employee or contractor)
* Projects
  + Budget, effort estimate, deadline
  + Project type (e.g. “On demand support” or “Subscription” will not have any actual deadline to control) - though it is ok if you don’t have it
  + Type of revenue recognition (Time and Material or Fixed Fee)
  + Note: we might need an internal project to “fill” the difference to full month (e.g. “INTERNAL” with revenue rate of 0 covering everything to full week)
* Timesheets (your team / your projects, note that other teams can contribute to your projects and your team might work in other projects)
  + Typically arrive monthly or weekly
* (optional) Plans
  + If there are no plans, below we talk about an option to “plan auto-fill”

But your ERP probably does not present you anything nicely in terms of parameters you’d want to control on your level (otherwise you won’t read this anyway). There is even maybe a project running so it will do so by 2025.

If you are steering this boat long enough – you might have ended up with some quick and dirty calculation, SQL queries to SAP (with field names still in German) or Excel file. Why? Because, well, your bosses don’t give a shit – they already hired you and a better ERP is coming in 2025 anyway – so they want to know how much money your team makes and why is it not so profitable (because it never is)

## Data Science

So, let’s get this done. First – yes, you need to know a little bit of python to get the files in. Basic level will do. (If you are a manager in 2020 and can’t write a simple script – hm, you probably missed something). The objective is to have an easy-to-use toolkit. We are definitely not developing any ERP here. What you will see is a very simple set of tools - you spend an hour or two to load your files into the notebook and then start playing with it - just make the analysis you want. You can end up checking a few numbers or building your own charts and dashboards. It is all [copyright model goes here, probably LGPL] free to use, even if you work for a corporation (that’s where you need it most) - just install Anaconda and download the notebook.

So, we first get and transform out input files (if you want to follow text AND code - read it on cohab, here we hide some code - even in 2020 loading few bad formatting excels will require some coding)

[get 3 files in – timesheets, costs, projects data]

The beauty is that this is it - basic calculations are available already

[make some basic calculations – effective rate, profitability]

effective cost rate (per project or per project per period) = SUM(cost)/SUM(days ? all or only billable)

effective revenue rate (per individual) = SUM(revenue)/SUM(working days in a period)

profit (per project) = revenue - costs

profit margin = (revenue - costs)/revenue

Now let’s do something a bit more complex - e.g. do a projection of how the budget will be spend

[profitability projections based on effort forecast, some simple approximation till budget ends with the linear drop at the end of the project]

??? The intelligent projection will require some fitting on completed projects data

If you want to use all power of machine learning - we can easily identify anomalies

[some more interesting staff (e.g. correlation of budget efficiency with PMs, clients, or regions)]

How can one quickly answer simple questions, e.g. which PM has a tendency to finish projects in over budget

And, maybe even more interesting thing - load to your model all fields you get from ERP, even the ones you don’t care about and mark projects which are worth your extra attention (e.g. monitoring) - you can then train the Decision Tree model to identify similar projects in the future (can be - junior PMs, getting close to budget expiration before end date, etc.)

## Conclusion

What we really want to highlight here is that the whole thing (including parsing of some crazy excel files) is <100 lines of code. So learn some python and a bit of data science and do your own analytical tools, stop being dependant of some PMO making reports for you!

## Objectives

How is the data we analyze helping managers.

1. Decision making - e.g. identify top loss-making projects
   1. Define possible non-billable spendings
   2. Identify projects which require management attention
2. Identify top performers (profit-wise)
   1. Identifying non-billable hours and internal projects
3. Identify suspicious behaviours [\* potentially some intelligent stuff \*]
   1. Employee enters fixed hours per project per day over a period
   2. Anomaly detection
4. Revenue and effort projections
   1. Capacity calculations: load all data but plans and generate automated projection (effort \* person \* month / revenue \* month)
5. Create consolidated analytic on demand for higher level of management (e.g. profitability forecast, revenue forecast, unallocated capacity) in case your ERP will not do anything like that
   1. **Example: crisis cost saving - if you decrease hours for some contractors, how much cost do you save vs how much revenue to you lose**

## Scope of Work

* Article
  + Text
  + “Simple” examples of operating with data
  + Addendum on how to load data
* Website
  + Generated by fastpage
  + From there you get to “open in cohab”