1) Uploading data files

In this video I'm going to guide you through the process of uploading files to Gooroo Planner using our website. This creates a dataset which is the precursor to producing reports – they're covered in the next video.

I'll assume that you've already created all the necessary data files, and if you haven't then the place to start is with the SQL code that we publish on our website. Go to the Publications section, and you'll find the Getting Started SQL code there as well as the documentation.

I should also say that, once you've settled into using Gooroo Planner, you can automate the whole process of uploading your data and running reports by getting a Planner API licence, so do drop us an email if you'd like to take advantage of that.

Anyway let's get started.

To start uploading data we need to go to Gooroo Planner, so we click My Dashboard and Visit Planner, and here we are in Planner and we click the green Upload button to start uploading a dataset. This launches the dataset wizard.

Let's give our new dataset a name. It's up to you, but personally I like to use a name that is fairly descriptive and put a number at the end so that it's easy to have further versions. I'll just call this one "Test Dataset 1". The description is optional.

Now we select the first data file that we want to upload. Assuming you have followed the recommended naming conventions for these files, the first file will begin with the number 01. So select that one.

While it's uploading, you can adjust any other settings that you need to. Again, if you have followed the recommended naming conventions the instructions will be at the end of the filename – in this case the word "Statistical" tells us to select "Statistical data" in the dropdown.

Now we can click Next.

And here there are some more settings. The first two dates are the dates of the past period, and it's important to get these right from the outset, because the system is going to measure things like activity over this date range and we would end up with inconsistent data if we tried to change these past period dates later on.

And here I need to make a little digression and tell you about the three date ranges that are available to you here. These are also discussed in the Planner documentation which is in the Publications section of our website.

The first kind of date range is the Past Period, and this is the period over which demand is measured. The guiding principle is that it should be recent enough to be relevant, and long enough to be representative, but other than that it can run from any date to any date. If in doubt, just use the last 13 months or 54 weeks so that you have a full year of seasonal fluctuations in the data.

The second period is the whole period covered by the additions and activity data, and this is used to measure seasonality.

The third period is called the Performance Period, and this is the period over which statistics like average length of stay and the proportion of urgent patients are measured. By default the whole data period is used as we can see at the bottom of the screen.

At this point we don't actually have any choice about which Past Period to choose because it was set in stone when we generated the data files. If you aren't sure what the Past Period is then there are two ways of finding out.

Either: look in the DateSettings file if there is one, because it's there under Past Period; or you can look in either of your waiting list snapshot data files, because two of the snapshot census dates must be within a week of the start and end of the past period.

So having looked up the dates I'll put them in now.

Then we have the future period dates, and here you have a bit more flexibility because you can change both these dates later on when you create a report. The only exception is if your waiting list snapshot file includes a third snapshot marking the start of the future period, and in that case the future period needs to start within a week of that census date and should not be changed later unless you want to assume that waiting list size for the new date too.

Then we have some fields that describe how clinical priorities are being described, and by default it assumes that we book cancers within 2 weeks and other urgents within 4 weeks. You can change this if you have described urgency any differently from that.

And we can leave the defaults for the rest of the fields, and click Next. It takes a little while to upload the data, and then it gives us a report of what happened.

You can see that it has found 125 different service lines in the data, which is what we were expecting. And it has logged 40 errors and we can click this icon to download a list of those. You can also download this later from the dataset manager later if you like.

We want to look in the Data Log which is this one here, and the issue is that some fields have NULLs which have been interpreted as null, so in this instance the errors are not a problem. We might have been more worried if for instance none of the fields had uploaded correctly. So we'll carry on by clicking Next.

And we've landed back in the dataset manager. Now all the remaining files are going to be thrown into the same dataset, using the Update dataset icon which is the 5th one along.

So we start by clicking Select and we're on file number 2. Personally I like to keep the folder open beside the browser and highlight the file I'm currently working on, so that I don't lose track. If you do forget where you're up to then you can always download the dataset from the dataset manager, and look in the History file which will tell you the last file that you uploaded.

While that's uploading we can adjust the settings. The filename tells us to select "Allow overwriting of existing data?" and select "Patient level data – additions" in the dropdown.

If you are uploading new data for the first time, one tip can be that it's a helpful diagnostic to tick "Allow creation of new records?" for every file, because then if there are any typos or inconsistencies in the names you have given to fields like specialty or subspecialty then they'll be obvious because you'll see similar-looking duplicates scattered through the resulting dataset.

Now we can click Next, and we got all the remaining settings right the last time around so we just click Next again. And wait while the file uploads.

When that's done we will get the usual report, and the numbers about numbers of services and errors are cumulative so as long as they're the same as the ones we got from the first file, there will be nothing to worry about here. And there they are, they're the same numbers, so we click Next.

Back in the dataset manager we click the 5th icon along again to update with the next file, It's file number 3. It's a waiting list snapshot with nothing else to select, so click Next, Next... and there were no new errors so Next...

Update again... Select file number 4. Now until now we have been uploading outpatient data, and this is the first of the admitted patient files. So (as the filename suggests) we need to click "Allow creation of new records?" to allow it to create all the inpatient and daycase and non-elective services in the dataset. And it's an activity type file.

Next... and Next... and we can see there are lots of new services now, and we've got another 10 items in the error log which we should take a look at, as we did before.

We just keep going like this until all the files have been uploaded

And when it's done it's a good idea to download the dataset and take a look at the statistics that have been measured to check they look reasonable.

So we've downloaded it, and we're going to open it up, and open the file that doesn't have a suffix which is the dataset itself. Personally I like to use 'freeze frames' so that I can always see the headers, and sort by HeadType and specialty so that similar services are grouped together. And then scan the data for completeness.

The first thing to check is that we have all the service lines we were expecting. If you've been loading data for the first time and ticked "Allow creation of new records?" for every file, then you'll see duplicates for any services where you weren't consistent with your naming conventions. For instance, you might see services called Trauma and Orthopaedics, and some called T&O, and some called Orthopaedics, with some of the data in one and some in the other, which means you'll need to revisit your queries to make sure the naming conventions are consistent.

For non-waiting list services such as non-electives, we are only expecting to see non-waiting list activity, and a full set of demand and activity seasonality profiles (at least for services with lots of activity).

For waiting list services the main thing is that we want to see a full set of waiting list movement data. That means: waiting list activity, additions to the waiting list, and waiting list snapshots at both the start and the end of the past period, and removals. If you're missing any of these, then you'll need to check the files you're uploading to find the reason why.

If one or more of the waiting list snapshots is missing then a likely explanation is that the start or end of the past period is not within a few days of a snapshot census date in the file you uploaded.

You will also want to see clinical priority data for waiting list services, bed data for bed using services like average length of stay for the inpatient services, and clinic and theatre data in the right places. And for waiting list services you'll want to see a complete set of demand and activity seasonal profiles, certainly for the larger services.

Uploading these files over the web interface is not the most exciting thing in the world, but it's good idea to do it this way when you're starting out with Gooroo Planner because you'll find any problems with your data files a lot quicker this way. But once you're comfortable that everything is uploading correctly, you should consider getting a Planner API licence so that the whole process can be fully automated.