Transcript for Session 004

Listen to the podcast session, see resources & links:
http://chandoo.org/session4/

Transcript:

Hi everyone. Welcome back to chandoo.org podcast session # 4. This show is all about making you awesome in data analysis, charting, dashboards and VBA, using Microsoft Excel.

I hope you are having a fantastic week. I enjoyed quite a bit this weekend - my family and I went to a very nice beach near my house (about 250 miles away from where we live). We drove down there and had a lot of fun for about three days visiting temples and spending time on the beach. The kids had a lot of fun playing in the hotel swimming pool and all-in-all it was quite a fun week.

Now I am back to record another podcast where we talk about whether you should make any pie charts in public or not. The title is "Can I pie-chart in public?" So let's talk about it.

But before we do that, I want to invite my kids who are at home today because they have a week's break from school, before they start their next academic year. You can probably hear them making some noise in the background. I thought I would invite them and they can say hello to you all. Nishanth is my son and Nakshatra is my daughter. If you are curious what their names mean - Nishanth means 'moon or an object that ends darkness' and Nakshatra means 'a star'. There is a story behind their names. Actually, my full name is Purna Chandra Rao and Chandoo is my web name (or short name). Chandoo really means 'moon'. And, my wife's name is Jyosthna which really means 'moonlight'. Since both of our names are objects in the sky, we wanted to give our kids similar names and that's why we chose Nishanth and Nakshatra when they were born. Anyway, here they are and they are just going to say hello to you and then will move into our session.

Nishanth: Hello, how are you?

Nakshatra: I am Nakshatra. Thanks for listening.

Nishanth and Nakshatra: Bye, enjoy the podcast.
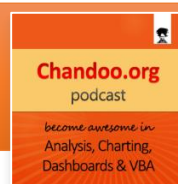
Chandoo: So let's talk about pie-charts in our podcast session # 4. As a reminder, you can download all the show notes, resources and links mentioned in the podcast at http://www.chandoo.org/session4. Please go there and you will be able to access all the material, links and any book recommendations that I make in this podcast.
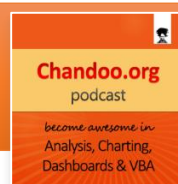
So moving on, our topic for today is - "Can I pie-chart in public?" Now when it comes to charting, if you are somebody who is just starting out with data analysis work, then you would consider pie-charts to be in the same league as every other chart in Excel. At least this is how I used to think when I started working as a Business Analyst. As part of my analysis work, even before I got a real job - I got a two-month summer internship, with a pharmaceutical company in India, in the summer of 2005. The job constituted going out and meeting a lot of doctors and conducting a survey to understand how various pharmaceutical companies are perceived by them in terms of the services rendered by the companies.

I was surveying about 50 different doctors in India and I had to make a report at the end of my internship for which I used (I distinctly remember) pivot tables to understand the data, and I made a lot of charts. I think I made a 50 or 60 slides presentation because that showed how much work I had done in the entire summer! So this entire report was really a massive PowerPoint presentation and a booklet-style report made in MS Word. But the backbone for this entire thing was Excel and most of the reports were really just charts that depicted the underlying information and presented the facts and findings. When I was doing that, nobody told me what kind of charts to use - I was just picking the charts randomly - I would select the data and go into the 'Insert-Chart' menu of Excel (back then I think I was using Excel 2003) and insert any chart that I fancied. I would format them in any way that looked okay. So there was no rule, rhyme or rhythm to what I was doing back then. It was just purely intuition - I used quite a few pie-charts, donut-charts, bar charts, column charts etc.

After I started blogging about Excel, charting and data visualization, I tried to learn the theory behind various charts and why a certain type of chart works or doesn't work etc. And then I came to realize some of the limitations and advantages of pie-charts. So today let's talk about these things and understand where pie charts work and where they don't. Before we talk about where they work and where they don't, let us quickly understand what a pie-chart really is.

A pie-chart is a circular diagram. And (it goes without saying) most of us would see pie-charts at least once a day. So what is a pie-chart? It is a circle that is divided into multiple slices, depending on how much each slice should indicate.

Let us take a simple example - let's say you are managing the budgets for a company which has only three departments. So the total budget is $100 million and the three departments that we have are

Operations, Marketing and HR. Let us say that the Operations department spends $50 million and the HR department spends $20 million and the Marketing department spends the balance $30 million. So, as a budget analyst, you are looking at these three numbers - 50, 30 and 20. Together these three numbers tell the story of how the overall budget is spent.

When you want to make a chart in such a case, a pie-chart is really one of the good fits. I am saying 'one of the good fits' because there are many ways to depict this kind of data or information and a pie-chart is just one of the many. So when you make a pie-chart from these three data points (i.e. 50, 30 and 20) all you are really doing is that you are drawing a big circle and you are dividing it into three parts - one part shows 50% (for the Operations department) and the other two parts would show 20% and 30% respectively (for the HR and Marketing departments).

So this is what a pie-chart does. It is a circular diagram that depicts how much portion each of the departments is consuming. So again pie-charts are something that are universal and we see them practically every day - if you open up a newspaper or if you are watching a TV report or if you are reading an article on the web that talks about information visualization or anything - chances are high that there would be a pie-chart in that particular element.

Pie-charts, as you can see, are really good for telling the story of how various parts contribute to the whole picture. - This is because, by nature, they are circular in shape and you are drawing slices. So when somebody looks at those three slices of 50, 20 and 30, they'll also understand that the three contribute to an overall of 100. That's a pie-chart.

They really work because they are universal. People are seeing pie-charts everyday - you open a newspaper and there is a pie-chart. You switch on a business news channel on television and you would see a pie-chart. You go to any leading newspaper website (or anything like that) and there would be a pie-chart somewhere on their site, as part of their news commentary. So pie-charts work because they are well-known. That's one of the main reasons why people use pie-charts. The second main reason is because they are really easy to make - you take any popular data analysis software, including Excel, and pie-charts would be available as one of the default charting options for data. Just select the data and making a pie-chart is as simple as eating (not making) a pie! So pie-charts are like that - they are well-known and they are easy to construct. And these two are the most important reasons why people make pie-charts. There is also a third reason which is that people want to see pie-charts. So if you have a boss, a manager or a customer who insists on following a reporting template, dashboard or scorecard format where there should be a pie-chart at a certain location - naturally they would insist that you use the pie-chart and, as an analyst, you would make that. So that's probably the third reason.

These are the three reasons 'why' pie-charts are used.
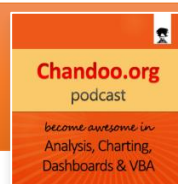
Now let us understand a little bit about why they should not be used. What is wrong with a pie-chart? Why are we even having a discussion about them? If they are so good then we might as well use them and not talk about them. But we are having an entire podcast dedicated to pie-charts, so there must be something wrong with them. Well, when it comes to the negatives of pie-charts, there are quite a few. The very first problem with a pie-chart is that (by means of evolution and training) our human eyes are trying to do a lot of things intuitively, out of which one is measuring length or height. Even a small child can measure these, and I remember when my kids were young and friends or relatives brought over candy, they would always pick the bigger candy - the candy that looked longer or taller, irrespective of whether they knew about lengths and heights or not. So this just indicates that we are naturally good at comparing heights or widths. If you look at two buildings you don't have to be an architect or a civil engineer to know which building is taller. - You see it and you know it. So this kind of training is quite natural and this is not something that we are learn in school - we just know it by means of evolution - our ancestors, their parents and grandparents were trying to do this because when humans were hunting or anything like that, measuring the bigger object was one of the primary goals.

So, anyway, coming back to the theory of how our eyes behave - measuring height or width is one of the good things that our eyes can do. Similarly, our eyes are also good at a lot of things - for example our eyes can spot colors better, we can understand when something is contrasted or not etc. But when it comes to measuring angles, i.e. when you look at an angle you cannot tell whether it is 30 degrees or 33 degrees or 27 degrees. You may be able to guess - that also depends on a lot of things - but you cannot tell one angle from another angle precisely. Well if it is a 90 degree angle you would probably be able to say with confidence that it is 90 degrees. The same goes for 180 degrees. But angles that are in between - like 45 or 60 or 30 or 75 degrees are somewhat dicey, as you cannot tell their angle or how much bigger they are as compared to another angle with precision or accuracy, simply by looking at them.

So our eyes are not very good at measuring angles. This is why we have special instruments for measuring angles. The second thing is that our eyes are also not very good at measuring areas. We are good at measuring height and width, but when it compares to measuring area and comparing one area with another area, we are not so good.

Again, in the podcast format, where I am just talking and you are listening, some of this might be slightly tricky to visualize, so as part of the show notes I am going to put 2 or 3 pictures that make this concept abundantly clear. You can see a picture of a 30 degree angle and a 33 degree angle and you won't be able to tell the difference with a lot of confidence. I'm going to put that in the post. Our eyes are not very good with both angles and areas. Maybe you are an exception, but most people (on average) are really poor at comparing angles or areas.
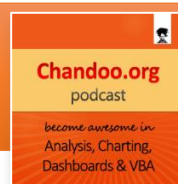
When it comes to a pie-chart, it essentially tells you that each of the slices is 20% or 30% or 50% - so it's either telling you the information in the angle or (if we look at a little more mathematically) it's telling you the information as an area. And we are not good at measuring either of these.

So a pie-chart is a medium where information is communicated in a way that is really challenging for our eyes - we have to look at it, but we won't understand it. This is also the reason why people add data labels for most pie-charts that we see in business situations. The pie-chart only makes sense when the label is present. Otherwise it is tricky to understand a pie-chart. So this is one of the limitations. The first limitation is that we are not good at measuring angles or areas.

The second limitation is that when you look at any kind of reporting or data analysis platform, like Excel for example, everything is a rectangle - if you look at a cell it is a small rectangle, if you look at a box, a report, a print out, your computer screen or your projector output – they all look like rectangles. So everything is a rectangle in the world of data analysis and visualization, whereas a pie-chart is a circle. So when we have a pie-chart in a report it kind of messes up all the alignment that you have there. Rectangles are naturally easy to align, as you can align one rectangle with another and they form a natural order of things. However a circle doesn't fit into this rectangle paradigm. It sticks out like a sore thumb and it doesn't fit into the layout, no matter how you shape, align or color it. It reduces the balance of your entire report or output, and that is another limitation.

Pie-charts also throw another tricky challenge for us. - Let's say we are talking about a pie-chart where there are more than 3 or 4 slices - the very first slice would start at the 0 or 90 degrees point and so, it is easy to see how much the first slice is - if you are looking at a circle from the top and it is divided into 20, 30 and 50 percent respectively - then the 20% slice from the top would be about 20% of the area or 20% of the angle. After that comes the 30% area, but that is not starting at a straight line position (i.e. either 0 or 90 degrees) - it's starting at an arbitrary position, and this makes it very difficult to compare one slice with another.

Again this concept will be very clear to you if you visit the show notes page on http://www.chandoo.org/session4 where I will add some pie-chart examples to illustrate how comparing one slice with another is made more difficult because of the way the angles are added up, as it creates a layout where it becomes tricky to compare one slice with another. You could probably compare the very first slice and the very last slice, but anything in between cannot be compared very well. Right?

Another negative with pie-charts is that when you add other Excel supported elements like 3-D or shadows, pie-charts can be very distracting. I wrote an article some time ago entitled 'Why 3-D pie-charts are evil' and I will link to that in the show notes. It explains why 3-D pie-charts can create a lot of
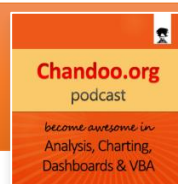
distortion and convince you that a small slice is actually bigger than it is. This is because of the way 3-D perception works - the angles or slices in the front tend to look bigger compared with the angles or slices in the background. So it is important to understand this limitation of pie-charts.

So these (in my opinion) are the four main reasons why pie-charts don't work.

To summarize:

1. We can't measure angles or areas very well, and that is what a pie-chart does.

2. Pie-charts are circular in shape whereas everything else in our reports, dashboards and analysis is usually rectangular in shape; and so pie-charts don't fit into the rectangular world and stick out in a very ugly way.

3. Pie-charts make comparison very difficult. They are very good with telling the story of how various parts contribute to the whole - but even there, as in our example of the HR, Marketing and Operations budget scenario - as a budget analyst or a CFO of the company as you look at the pie-chart you ask how much more Marketing is spending in comparison with HR, or how much Operations is spending in comparison with Marketing. These kinds of questions are also there in the back of your mind and pie-charts don't do very well to help you compare on those fronts.

4. And, finally, pie-charts distort data easily.

These are the reasons why pie-charts don't work. So now we have seen why pie-charts work and why they don't work. They work because they are well-known and easy to make. They don't work because they present data in a manner that is challenging for our eyes, i.e. with respect to measuring of angles or areas.

Before we talk a little about the alternatives to pie-charts and whether you should use them or not, let us quickly take a detour and understand the siblings and cousins of pie-charts.

The first is a *donut-chart* which is really a pie-chart with the center portion removed. - So it looks like a donut.
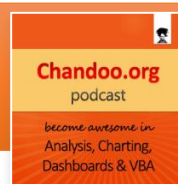
Then there is a *3-D pie-chart* which is really cumbersome and tricky for our eyes. Again, I really encourage you to read the article on 'Why 3-D pie-charts are evil' to understand why you should never use them!

There are *gauge charts* which are basically speedometer kind of charts that you see in cars, or any other kinds of dashboards or vehicles. You have a speed of 180 or 160 miles and the needle kind of moves around between 40, 60 and 70 miles.

Then there is an *area chart*, which can be considered to be a sister or cousin of a pie-chart because essentially both pie-charts and area charts are trying to convey information by means of total area given to that particular item. So $50 million for Operations department would take half the space. The same, when visualized as a bubble chart or an area chart, would also kind of indicate the same sort of thing. So these are all in the same space as our pie-charts.

More than area charts, I would say that bubble charts are similar to pie-charts. Now that we understand various kind of charts that are similar in the way they present information - that is donut charts, 3-D pie-charts, gauge charts, speedometer charts etc. - let us understand whether you should make a pie-chart or not and when you should make them.

In my opinion, pie-charts are not completely evil even though some people consider pie-charts to be 100% evil and say that you should always avoid them. For example, if you are reading a book named 'Information Dashboard Design' by Stephen Few (which is considered to be one the really good books for people who are making dashboards, reports or scorecards as Stephen Few lays out a lot of recommendations for structuring dashboards, presenting, highlighting and laying-out information etc. - it's a very good book that I highly recommend to you to read), on page 51 of Chapter 6 on 'Effective Dashboard Display and Mediums' (I am not sure if the page number would be the same as the edition that I have) which really refers to 'charts for your dashboards'.

In that Stephen Few talks about the charts that you should use, and he does not mention pie-charts at all. He mentions bar charts, column charts and others, but he does not talk about pie-charts. He has some final thoughts on missing elements where he gives some idea on why he is omitting a lot of charts that are common and there he also talks about pie-charts.

I am now going to read directly from his book (to ensure that there is no distortion of the information), "*Pie-charts were designed to display part-to-whole information such as the individual products that make up an entire product line. As we have already discovered, however, part-to-whole information can*
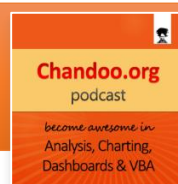
be communicated more clearly using a bar graph...............Viewers can process the information in the bar graph more quickly and easily than the pie-chart. Why? Whereas a bar graph uses pre-attentive visual attribute of line length, pie-charts encode values as two-dimensional areas of the slices and their angles as they extend from the center towards the circumference of the circle." Again, this is a very mathematical and jargon way of saying 'bar charts show lengths whereas pie-charts show angles'.

Stephen Few further writes, "*Our visual perception does a poor job of accurately and efficiently comparing 2-D areas and angles*." This is the same point that we highlighted earlier in the podcast stating that our eyes are good at measuring lengths and heights and comparing them, whereas they are really poor when it comes to measuring areas or angles. The only things that pie-charts have going for them is that when you look at one, you automatically know that you are looking at measures that are part of a whole. Because bar graphs can be used for other kinds of comparisons, when you use them to display part-to-whole data, you must label them in a manner that makes it clear. As long as this is done, bar graphs are far superior.

Now this is the stance that Stephen Few takes, and he is a well-known author and proponent of good dashboard design. You can visit his website - I'll add a link to it in the show notes, so that you can learn more from him as well as his perspective.

But, I tend to differ from this kind of purist thinking where it simply suggests to avoid pie-charts altogether. I subscribe to a theory that is a little more accommodating.

I say that you can make pie-charts for 3 or 4 reasons:

If you have only 2 to 3 slices, make a pie-chart. It is really simple, the metaphor of a pie is easy to understand, and everybody knows what it is. This is a good area where pie-charts can be used. I am not saying that only pie-charts should be used, but pie-charts are an alternative when you have only 2 or 3 pieces of information.

If your data is such that one value is disproportionately high (for example if we have data related to the budget of 10 departments and one department has most of the budget allocated to it - this is quite common - for example, a department like Operations or Sales or Marketing usually dominate the budget as they have a lot of people and they are the ones bringing in revenue and thus tend to have a lot of budget allocation), then a pie-chart is a good fit because it throws maximum light or emphasis in that area. Again a bar chart or column chart would also do justice, but a pie-chart would naturally fit into that flow. So, when one value is disproportionately high, you can use a pie-chart.
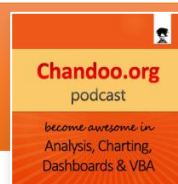
Pie-charts are also good for non-business situations. For example, if you are making a report that you would display on your school notice-board or in a library or any other informal situation like a fund collection drive for your local church, a pie-chart is a good informal means of displaying how much progress is made, how things are done and how various people are contributing etc. It's okay to have a pie-chart in such a scenario as people are used to seeing and understanding pie-charts. So you may as well make them and use them. It makes it a little more fun and interesting.

And the final reason why you can use a pie-chart is when 'you MUST use it'. This is quite common - actually it is more common than all the other three. You always have bosses, clients, customers or senior managers who insist that there should be a pie-chart in a slide or in a part of the report. I had a boss like that! Sometimes people are really keen to have a pie-chart in the report. As an analyst, your job is not just to present and analyze data, but also to please customers and do what they want to do. So, as a good and awesome analyst, you would use pie-charts if you must!

Just to recap - use a pie-chart if you only have 2 or 3 slices of data; or if one of the values is disproportionately large and it needs to be emphasized in the report; or for non-business situations like a school notice board, library or a local fundraising event; or when you must use them, i.e. when your boss insists that she wants a pie-chart. So these are the reasons why a pie-chart should be considered and created.

For those of you who are curious to know the alternatives to pie-charts, there are quite a few. As you can guess, a simple bar or column chart, is a naturally good way to replace pie-charts. They are as easy to make, and as universal as pie-charts. So that's the best alternative for pie-charts.

The second alternative is that if a pie-chart has too many slices (25 for example), you could technically consider a 'pie of a pie-chart'. This is also something that Excel natively supports. I will link to an article that shows you how this can be done and you can use it. What happens in such a scenario is that Excel will make two pie-charts - the first pie-chart will show only 3 or 4 slices, and one of the slices would be shown as contributing to all these other slices. I will link to a picture and an article that will explain how this can be constructed. Again this is strictly a scenario where you must make a pie-chart, the pie-chart has too many slices and you don't want to clutter it like that. In such cases you can consider a 'pie of a pie-chart'.

The third alternative is when your boss or somebody insists on a pie-chart, but you don't personally think that pie-chart is doing a good job. In such a case you can add data labels to the pie-chart. This is a

simple way of saying "This is a pie-chart and since I know you cannot read the data and understand what is going on from the pies alone, I am going to add labels here!"

And, a final alternative is just to use the numbers. If there are only 10 departments and their budgets, you can just make a nice beautiful table that shows the budget - people can read it and compare easily just by looking at the raw data itself.

So these are some of the alternatives. I have written an article that explains quite a few alternatives to pie-charts and I will link to that in the show notes. Please visit http://www.chandoo.org/session4 to access the resources and links that I am mentioning in the podcast.

So these are some of the alternatives to pie-charts. To summarize, I want you to remember that these are the mistakes that you should avoid:
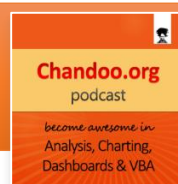
No 3-D pie-charts, regardless of what kind of pie-chart you are making. 3-D pie-charts are an absolutely poor way to depict information. They can distort the data a lot, and make it really difficult to understand what is going on. So, don't make 3-D pie-charts.

Don't make pie-charts when your data has one or two slices that are incredibly small. In other words, nobody likes a small slice of cake! - You want to have a big slice of cake when you go to a birthday party. I like a big slice of cake too. So, don't make a pie-chart when one of your slices is really small, because it is really tricky to figure out what the number is.

Excel also has a nasty option called 'explosion of pie-chart'. Instead of sticking all the pie slices together in a circle, this exploded pie-chart appears as if somebody lit a small bomb inside the circle because of which all the slices are moving away from the circle. Do not explode! Don't be tempted by this option of Excel. Do not explode your pies! - It isn't pretty.

Finally, color your pie-charts intelligently. In other words, avoid a rainbow. Since pie-charts can have a lot of slices, if you choose the default coloring option, you could create a really ugly coloring in Excel. And, this makes it very difficult to understand what each slice is and compare one with another. Pie-charts are as it is poor for comparison and if the colors are poorly chosen, there won't be enough contrast between the slices and it will make life very difficult for a person viewing the pie-chart and trying to understand what is going on.

So these are the mistakes you should avoid.

To summarize, no 3-D pie-charts; don't put too many small slices in your pie-chart (consider a bar or column chart or even a table as an alternative in such a scenario); don't explode your pie-chart and don't use rainbow colors. Make sure the colors are well contrasted and there are not too many colors, so that it is easy to look at the data and understand what is going on.

These are my thoughts on pie-charts.

To conclude, and answer our main question about 'should you pie-chart in public'; obviously you can make a pie-chart as long as you are using 3 or 4 data points, or if you must use a pie-chart. Otherwise consider one of the alternatives like bar charts. They work tremendously well and they create beautiful graphs that fit into the natural layout of your spreadsheet, report or presentation. Make sure that you don't use too many slices when you make pie-charts, especially if the slices are too small or too big, as that creates a lot of confusion and distortion of the data. Don't use any kind of 3-D pie-charts or exploded pie-charts.

So, in a nutshell, this is what I think about pie-charts. I am sure you also have strong opinions about pie-charts because whenever we talk about pie-charts on http://www.chandoo.org quite a few people comment and some people take the stance of 'no pie-charts ever'! On the other extreme, we have people who always say, "Nah...I don't mind pie-charts; I am going to use them." So pie-charts seem to be a divisive force - if you put a pie-chart in a room - you could be sure that there will be some people who love and you can also be sure that there will be some people who absolutely hate it!

I am sure you have an opinion about pie-chats too. So if you feel pie-charts should be used, should not be used, when they should be used, when they should not be used, and you want to chip in, please visit http://www.chandoo.org/sesssion4 and leave your comments there.

I hope you have enjoyed this podcast and let me take a moment to thank you for spending time with me, listening to this podcast and learning something about pie-charts.

Please visit http://www.chandoo.org/session4 to access all the resources, links and other information mentioned in this podcast. Before I end, I want to quickly remind you that you can visit http://www.chandoo.org/podcast to go to our podcast page. You can access all the previous episodes on this page. You can also subscribe to our podcast on iTunes, Stitcher or Zune windows phone and get the updates. Whenever there is a new episode you can access it on your phone or podcast device and enjoy it on the go.

Thank you so much and stay awesome. Bye.