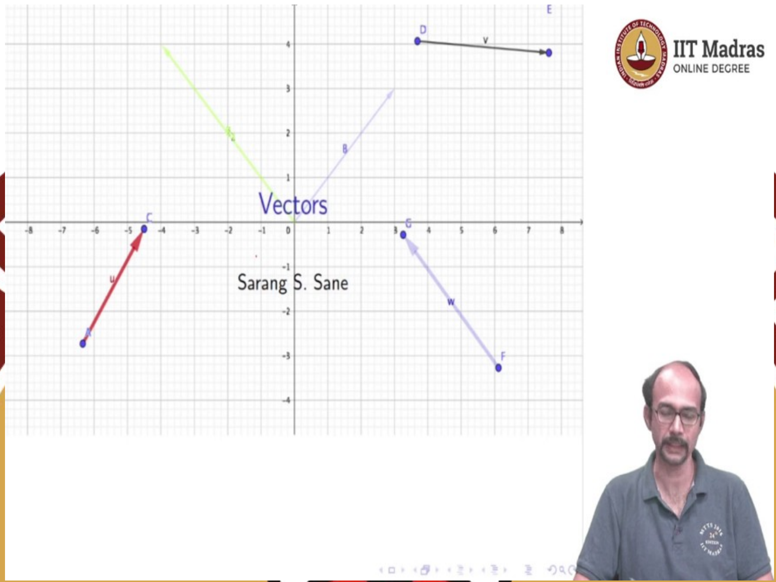
IIT Madras

ONLINE DEGREE

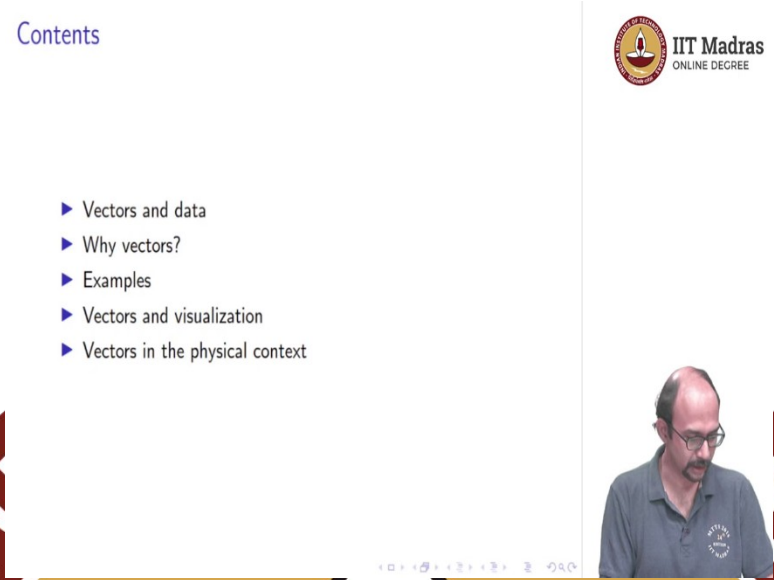
Mathematics for Data Science 2
Professor Sarang Sane
Department of Mathematics
Indian Institute of Technology, Madras
Lecture No. 01
Vectors

(Refer Slide Time: 0:14)



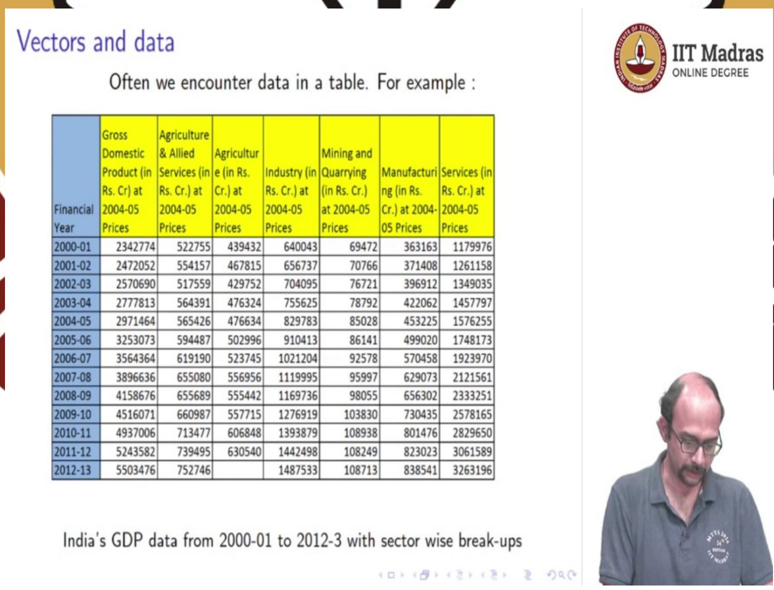
Hello and welcome to the maths 2 component of the online BSC program. My name is Sarang Sane, I will be instructing for this course and in the beginning of this course we will study a little bit of linear algebra. So, today is this first lecture is going to deal with vectors. So, what are vectors? In the screen appear we can see some vectors that you have probably studied in previous course. So, something with arrows and directions and so on. So, I am going to tell you a slightly different story about what vectors are.

(Refer Slide Time: 0:54)



So, let us see what are the contents of this video. So, in this video we are going to talk about vectors and data, we are going to talk about why vectors? We will do some examples. We will do vectors and visualization, this is sort of on the same ideas that you saw in the first slide. And then we look at vectors in the physical context. So, let us start with vectors and data.

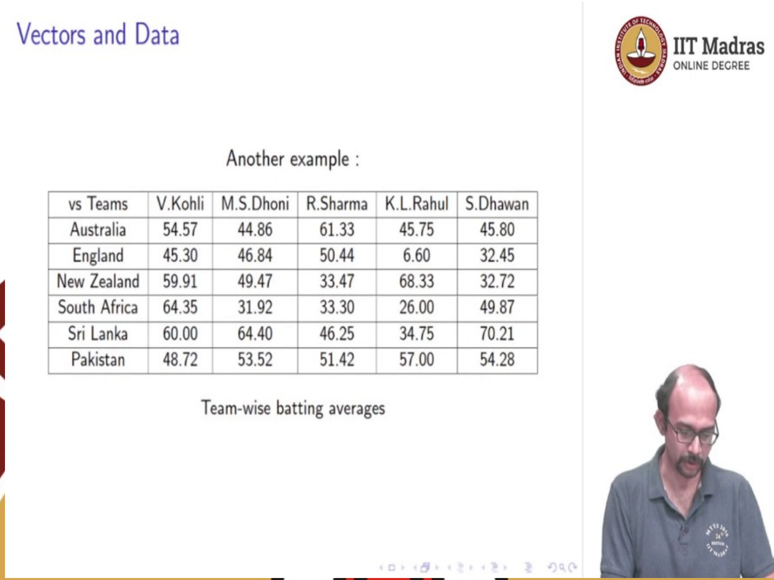
(Refer Slide Time: 1:23)



Financial Year	Gross Domestic Product (in Rs. Cr) at 2004-05 Prices	Agriculture & Allied Services (in Rs. Cr.) at 2004-05 Prices	Agriculture (in Rs. Cr.) at 2004-05 Prices	Industry (in Rs. Cr.) at 2004-05 Prices	Mining and Quarrying (in Rs. Cr.) at 2004-05 Prices	Manufacturing (in Rs. Cr.) at 2004-05 Prices	Services (in Rs. Cr.) at 2004-05 Prices
2000-01	2342774	522755	439432	640043	69472	363163	1179976
2001-02	2472052	554157	467815	656737	70766	371408	1261158
2002-03	2570690	517559	429752	704095	76721	396912	1349035
2003-04	2777813	564391	476324	755625	78792	422062	1457797
2004-05	2971464	565426	476634	829783	85028	453225	1576255
2005-06	3253073	594487	502996	910413	86141	499020	1748173
2006-07	3564364	619190	523745	1021204	92578	570458	1923970
2007-08	3896636	655080	556956	1119995	95997	629073	2121561
2008-09	4158676	655689	555442	1169736	98055	656302	2333251
2009-10	4516071	660987	557715	1276919	103830	730435	2578165
2010-11	4937006	713477	606848	1393879	108938	801476	2829650
2011-12	5243582	739495	630540	1442498	108249	823023	3061589
2012-13	5503476	752746		1487533	108713	838541	3263196

So, we often encounter data in a data. This entire course is supposed to be about data, so we better start with some data. So, here is an example of data. So, this is an example I took from the government website. So, it talks about the GDP, the Gross Domestic Product of the country and it breaks it sector wise. So, what was the GDP from agriculture and industry and mining and so on and across the years 2000-2001 to 2012-2013?

(Refer Slide Time: 2:03)

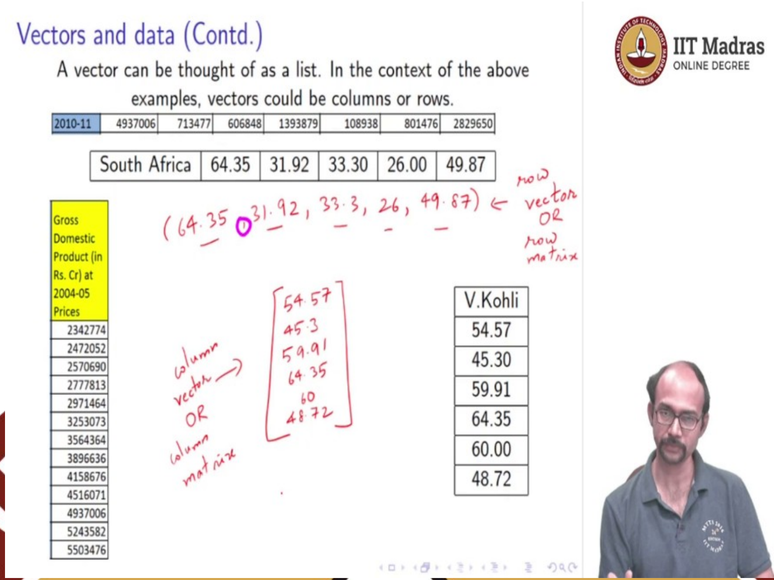


vs Teams	V.Kohli	M.S.Dhoni	R.Sharma	K.L.Rahul	S.Dhawan
Australia	54.57	44.86	61.33	45.75	45.80
England	45.30	46.84	50.44	6.60	32.45
New Zealand	59.91	49.47	33.47	68.33	32.72
South Africa	64.35	31.92	33.30	26.00	49.87
Sri Lanka	60.00	64.40	46.25	34.75	70.21
Pakistan	48.72	53.52	51.42	57.00	54.28

So, maybe here is another example from our national passion. So, this is the team-wise batting average for some players where one of them was a player when I made these slides, but unfortunately no longer is so I will have to say ex-player, ex team India player. Anyway, so this is Virat Kohli, Dhoni, Rohit Sharma, K.L. Rahul and Shikar Dhawan and the teams are Australia, England, New Zealand, South Africa, Sri Lanka and Pakistan and in the table we see for each team the average for each player.

So, these are two examples of data. As you can see the data is arranged in a table and this is typically how we get data and often the things we are interested in are about a row or a column or some rows or some columns. So, this is exactly where the idea of a vector stems from.

(Refer Slide Time: 3:05)



So, what is a vector? A vector can be thought of as a list. So, in the context of the above examples, vectors could be columns or they could be rows. So, here is an example from the GDP table. So, this is a row corresponding 2010-11. So, the total GDP and then the sector-wise GDPs, these are all in crores by the way. Or here is an example from the cricket table. So, this is the batting averages of the players that we saw, Kohli, Dhoni, Rohit Sharma, etcetera with respect to South Africa or we could pull out a column from let us say the GDP table.

So, this is the GDP, the full GDP so across the years 2000-2001 to I think it was 2012-13. And then here is a column maybe the averages for Virat Kohli with respect to the various teams. So, South Africa and then New Zealand and England and so on, Australia and so on. So, these are all examples of vectors. So, when I say vectors of course, you have to drop the heading in these rows or columns.

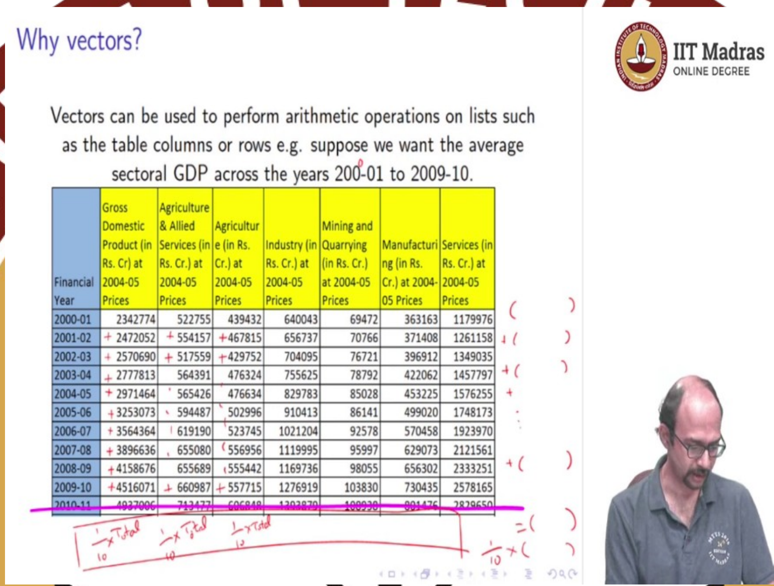
So, the vector corresponding to the South Africa row would be written as a 64.35, 31.92, 33.3, 26 and 49.87. So, this is a vector. So, this vector has how many components? It has 1, 2, 3, 4, 5 components. This vector has 5 components and I am putting commas in the middle just to represent that this number is ended and the next one has started, otherwise we really do not need commas.

Similarly, the vector corresponding to the column of Virat Kohli's averages, so this is written again as a column. So, 54.57, 45.3, 59.91, 64.35, 60 and 48.72. So, this is again a vector. So, this is a, this is what we will call a column vector and this is what we will call a row vector.

So, sometimes we are going to study very soon what are matrices, we can also call these as a column matrix and we can think of this as a row matrix.

So, depending on what we want to do, whatever is convenient, we think of this as a row vector or a row matrix and similarly as a column vector or a column matrix. So, the important point here is what is a vector? A vector is a list, it is a list of numbers that is what you have to remember.

(Refer Slide Time: 6:33)



Financial Year	Gross Domestic Product (in Rs. Cr) at 2004-05 Prices	Agriculture & Allied Services (in Rs. Cr.) at 2004-05 Prices	Agriculture (in Rs. Cr.) at 2004-05 Prices	Industry (in Rs. Cr.) at 2004-05 Prices	Mining and Quarrying (in Rs. Cr.) at 2004-05 Prices	Manufacturing (in Rs. Cr.) at 2004-05 Prices	Services (in Rs. Cr.) at 2004-05 Prices
2000-01	2342774	522755	439432	640043	69472	363163	1179976
2001-02	+ 2472052	+ 554157	+467815	656737	70766	371408	1261158
2002-03	+ 2570690	+ 517559	+429752	704095	76721	396912	1349035
2003-04	+ 2777813	564391	476324	755625	78792	422062	1457797
2004-05	+ 2971464	565426	476634	829783	85028	453225	1576255
2005-06	+3253073	594487	502996	910413	86141	499020	1748173
2006-07	+ 3564364	619190	523745	1021204	92578	570458	1923970
2007-08	+ 3896636	655080	556956	1119995	95997	629073	2121561
2008-09	+4158676	655689	555442	1169736	98055	656302	2333251
2009-10	+4516071	+ 660987	+ 557715	1276919	103830	730435	2578165
~~2010-11~~	~~4937006~~	~~713477~~	~~606848~~	~~1303879~~	~~108938~~	~~801476~~	~~2829650~~

So, let us go ahead and ask why vectors? So, we can use vectors to perform arithmetic operations on lists such as the table columns. So, for example if we want to average the sectoral GDP across the years 2000-2001 to 2009-10, I will take the relevant part of that table. So, we want 2000-2001 to 2009-10. I see a typo here, this should be 2000-2001 anyway. So, we do not want the, we want the average across the first 10 rows. So, let us first of all strike out that last row, we do need this and for these rows what do we want? We want the average.

So, how do I get the average? I add these numbers. So, if I want the average of the GDP, that is the total GDP, that is what we have in the first column. I add these numbers and then I whatever total I get I divide by 10. Similarly out here, what do I do? I add these numbers, and then I take the total and then I divide by 10. And what does that give me? That gives me the GDP contribution of Agriculture and Allied Services across the years 2000-2001 to 2009-10.

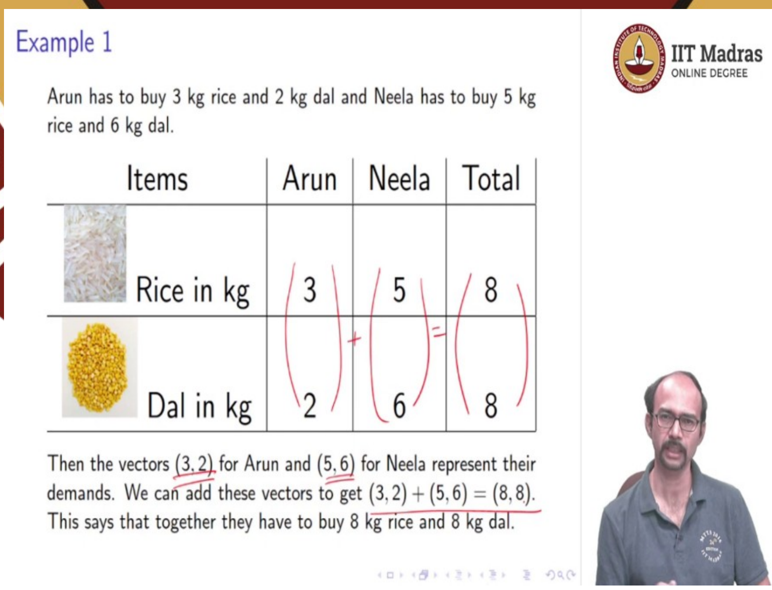
And then similarly I can for each of these I can add the corresponding entries and then I can divide by 10. Those will give me the averages. So, I can do this for each column and the idea is that well if you what is happening, you are repeating each operation across rows. So, instead of doing this for each column each entry, what I can say is I will take this entire row and I will add it to this entire row and I will add that to this entire row and I will add all these 10 rows.

So, what do I mean when I say add these rows? That means I add the corresponding entries, the corresponding components. So, the first entries are added, all the second entries are added, all the third entries are added, all the tenth, not tenth, seventh entry is added that is the last corresponding to the last column and then whatever I get that is the those that will give you me a vector of totals.

So, I will get the total GDP from 2000 to 2010, total GDP contributed by Agriculture and Allied areas from 2000 to 2010 and so on. So, I get all these totals. So, I have 7 numbers out here and then I divide each one by 10. So, I can do that by dividing the entire row by 10 or the entire row vector by 10.

So, 1 by 10×this thing that I got and that is exactly what I get out here. So, the point here is instead of doing it entry wise, I can think of it as if I am doing it for lists and that is what vectors are going to do of us. So, let us move ahead. We will see more examples and that might shed more light on what is happening.

(Refer Slide Time: 10:02)



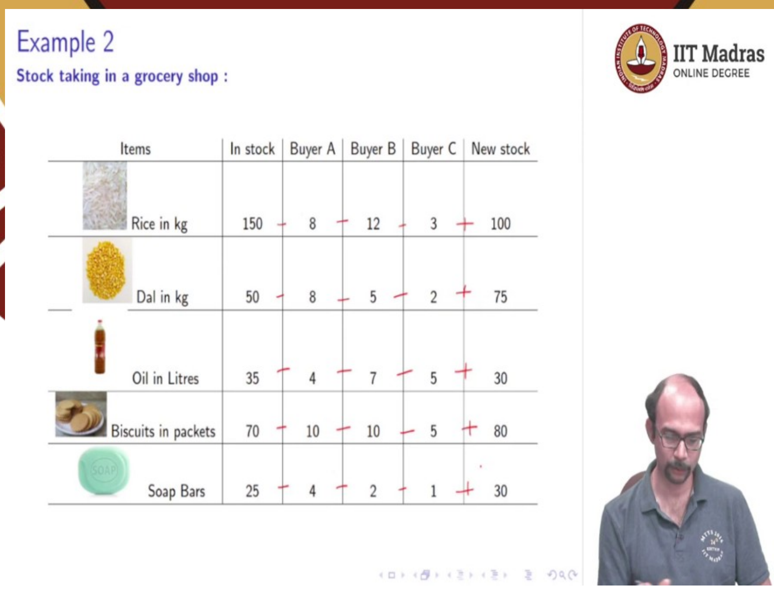
So, here is another example. So, Arun has to buy 3 kgs of rice and 2 kgs of dal and Neela has to buy 5 kgs of rice and 6 kgs of dal. So, let us put that into a table. So, this is rice in kgs and dal in kgs and the corresponding numbers for Arun and Neela. So, 3 kgs of rice for Arun, 5 for Neela and 2 kgs and 6 kgs of dal respectively. So, what are the total, what is the total amount of rice and dal that they want to get? They want to purchase 8 kgs of rice and 8 kgs of dal in total.

So, again we, how did we get this? We added these two and we got this, we added these two and we got this. So, I can think of this instead as we will think of 3,2 as the vector for Arun. So, now notice that here if I think of it like this, this is a column vector, but instead I have written it over here like a row vector. It is the same thing depending on what we want to do. So, here I have written 3,2 as the row vector, this is the demand for Arun.

So, the first component corresponds to the demand for rice, Arun's demand for rice. The second component corresponds to Arun's demand for dal. And then 5, 6 for Neela. So, again the first component is Neela's demand for rice, the second component is how much dal Neela wants to buy.

So, we can add these vectors. How do we add them? We add them component wise, exactly the way we have done over here. So, when I add these, I do 3 + 5 which is 8 and 2 + 6 which is 8 and that is how I get this column vector and instead we have just written these in terms of rows below. So, I have written these here in terms of rows. So, together they want to buy 8 kgs of rice and 8 kgs of dal. These are pretty simple example.

(Refer Slide Time: 12:02)



Items	In stock	Buyer A	Buyer B	Buyer C	New stock
Rice in kg	150	8	12	3	100
Dal in kg	50	8	5	2	75
Oil in Litres	35	4	7	5	30
Biscuits in packets	70	10	10	5	80
Soap Bars	25	4	2	1	30

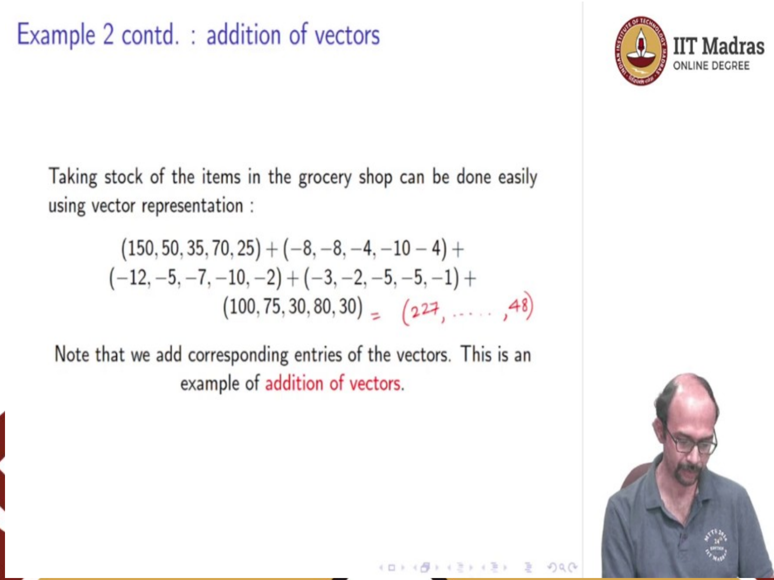
So, here is a slightly more a larger example I will not say it is more involved. So, this is examples about stock taking in a grocery shop. So, in this grocery shop we have 5 items. Let us say, so we have rice, dal, oil, biscuit packets and soap bars, and at the start of the day, the shopkeeper takes stock and finds that they have 150 kgs of rice, 50 kgs of dal, 35 litres of oil, 70 biscuit packets and 25 soap bars in stock. That is what they have in the beginning of the day.

So, maybe in the first hour, there are 3 customers, so, buyer A purchases whatever as per this column. So, this is the column corresponding to buyer A and then buyer B purchases according to this column. So, 12 kgs of rice, 5 kgs of dal, 7 litres of oil, 10 biscuit packets and 2 soap bars and then similarly buyer C purchases 3 kgs of rice, 2 kgs of dal, 5 litres of oil, 5 biscuit packets and 1 soap bar.

And after sometime, after that one hour has lapsed, this stock for that days arise, the new stock and that new stock 100 kgs of rice arrives, 75 kgs of dal, 30 litres of oil, 80 biscuit packets and 30 soap bars, this is what arrives. So, now the question is after that one hour, how much is in stock? What is the new thing in stock? So, how do we do this? So, the way to do this of course is we see how much stock we had at the start of the day.

So, let us look at rice. So, we had 150 kgs of rice, and then this is how much went out for buyer A, this is how much went out for B, this is how much went out for buyer C and then this what came in. So, $150-8-12-3+100$. So, this is how much rice is left at the end of one hour. This is how much is in stock. And similarly for dal, we will do the same thing and then oil, this is how much went out, this is how much came in, again this is how much, how many biscuit packets went out and 80 biscuit packets came in and then finally for soap bars, we had 25 to start with and then 4 and 2 and 1 were sold and then 30 new ones arrived after 1 hour.

(Refer Slide Time: 14:53)



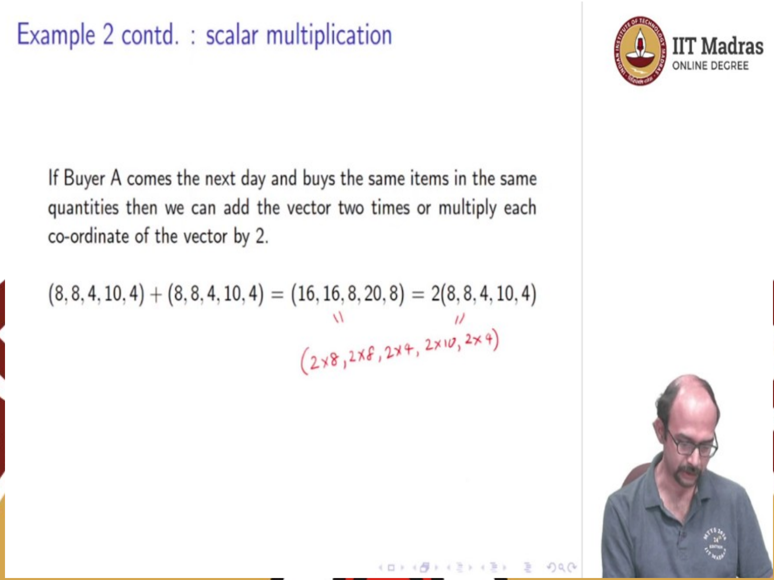
So, you can see that instead of doing it this way, we can think of this in terms of vectors. So, in vector representation how would we do this? You would write down the vectors for each of those columns. So, we have 150, 50, 35, 70, 25 which was in stock at the start of the day and then we have o. Notice that now we have put in a - sign which was not in the table so, that is because in vector notation now, we can keep track of what is being bought in, what is being sold.

So, the - corresponds to whatever is sold and the + which is down here corresponds to the new stock which has arrived. So, that is stock which the shopkeeper has brought from retailer. So, this is something that was not there in the table. This is something that we created. So, you can see that the vector has some value. So, this is - 8 - 8 - 4 - 10 - 4 corresponding to customer A and then the column for customer B has become this row, - 12 - 5 and so on.

And the column for customer C has become this row, - 3, this row vector, - 3 - 2 and so on and the new arrived stock, the fresh stock is this final vector + 100, 75, 30, 80, 30. And this is without a - sign as I said that is because this is being added to the stock, the existing stock. So, this is how you would take stock in a grocery store. So, so I will again reiterate that addition is in terms of how we in term is component wise.

So, we add corresponding entries of the vectors. So, over here if I wanted to add, then I would get a vector, row vector at the end and the first component of that row vector would be 150 - 8 - 12 - 3 + 100. So, that is a 150 - 23, so that is 127 + 100, so that would be 227 and you can do the rest. The last entry let us calculate that. So, you would have the last entry that would and 25 - 4 - 2 - 1, so 25 - 7 that is 18 and then + 30, so 48. So, I hope you understand how to add vectors. This is an example of addition of vectors.

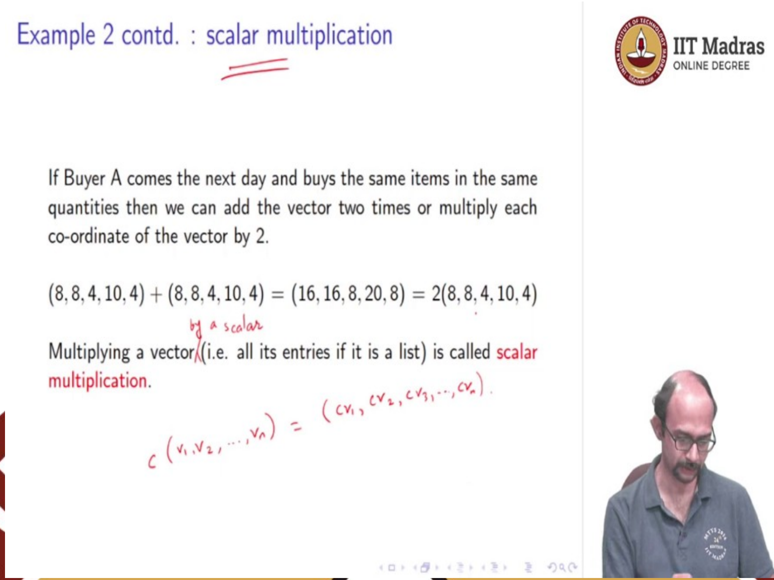
(Refer Slide Time: 17:55)



So, we can do one more thing. So, in the same example suppose buyer A comes the next day and they buy the same items that they bought the previous day. So, then we can add the vector two into or we can multiply each entry of the vector by 2. So, addition we learnt in the previous example, so if we want to see how much total items have been bought by the buyer across the two days, so then we would have 8, 8, 4, 10, 4 + 8, 8, 4, 10, 4. This is a vector corresponding to buyer A.

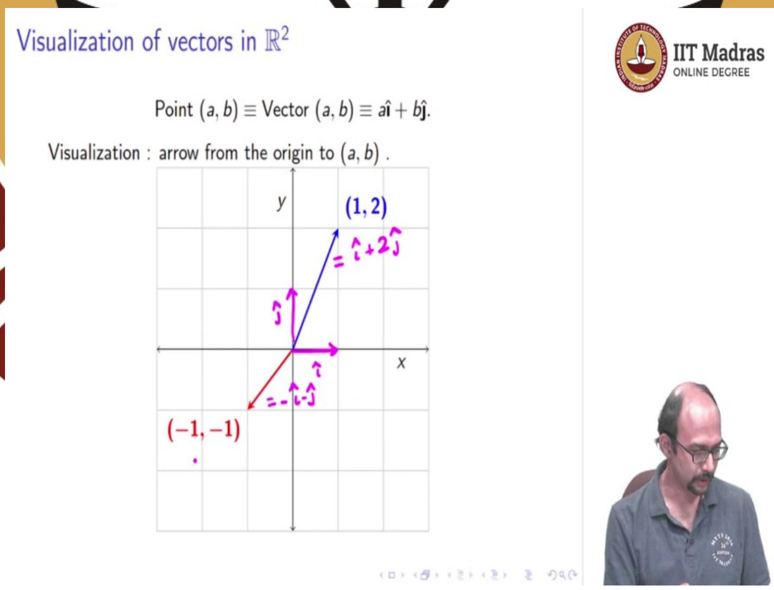
So, this was a column corresponding to buyer A which we have written as a row vector. And so we can say that so we add this addition as component wise, coordinate wise, so we get 16, 16, 8, 28, so, what this says is that buyer A bought 16 kgs of rice, 16 kgs of dal and whatever else. And the same thing can be written as, so we can think of this as 2×8, 2×8, 2×4, 2 ×10, and 2×8. So, we are adding this twice. So, each entry gets multiplied by 2. And then what we can do is we can take this 2 out. So, this is called scalar multiplication.

(Refer Slide Time: 19:29)



So, multiplying a vector is by a scalar so maybe here we should have by a scalar. So, that is what the slide says by a scalar is called scalar multiplication. So, that means if you have C into some vector let us say the entries of that vector are v_1 v_2 v_n., then scalar multiplication will be will mean that this is the same as the vectors $c\,v_2 cv_3 \ldots cv_n$; exactly as we had in this example upstairs.

(Refer Slide Time: 20:16)



So, let us talk about visualization of vectors. Maybe this is something some of you have possibly seen before. So, in R^2R^2, so let us start with R^2, so in R^2 we have points which come with bracket and two numbers with a, in the middle. So, the point a b, so we can instead think of this as the vector a b. A vector is just a list after all and the notation is that you put brackets and you have two numbers and a,in the middle if it is in, if it has two coordinates.

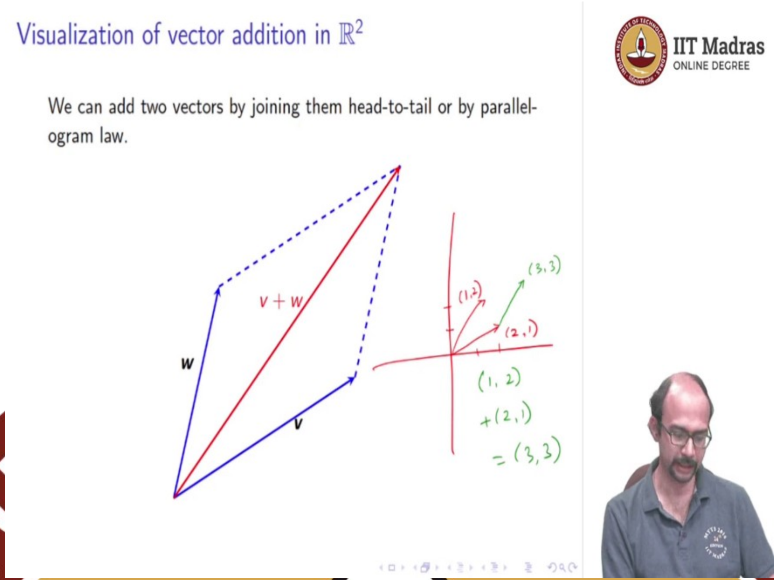
So, we can instead think of this as the vector a,b. So, this is the list and this is going to be identified with the vector $a\hat{i}+b\hat{j}$. So, if you have seen this notation before you will certainly know what this means. If you have not $i$and $j$correspond to the unit, so I maybe I should not use vectors, so it is an arrow on the x axis from 0,0 to 1,0 and j corresponds to an arrow on the x, y axis f, 0,0 to 0,1.

So, maybe here is a picture, so what is i and j over here? So, I is this, this is i that i over there and this is j. And $a\hat{i}+b\hat{j}$ means you scale i by a so a$\times i$ will mean you take a on the x axis and b×j will mean you take b on the, b×j on the y axis. So, 0,b and then when you add them you get we think of that as a arrow from the origin to the point a,b. So, here we have these 3 ways of thinking. We can think of this point as 1,2 or we can consider this arrow which is a line drawn from 0, 0, the origin to the point 1,2 with the arrow head at the tip of that arrow is at 1,2.

And we can also write this as$i+2\times j$. Similarly, here we have - 1,- 1, so that is your point and if you think of it as a vector in R^2 in the traditional sense, you draw a line from arrow from 0, 0 to - 1,- 1 with the tip at - 1,- 1. And we can also think of this as $-i-j$. Certainly, this is how you will do it for example in physics.

So, what do I want to point out over here, what I want to point out here is that the list that we have we are thinking of are nothing very different from vectors in the sense we may have seen before. So, that is in R^2.

(Refer Slide Time: 23:51)

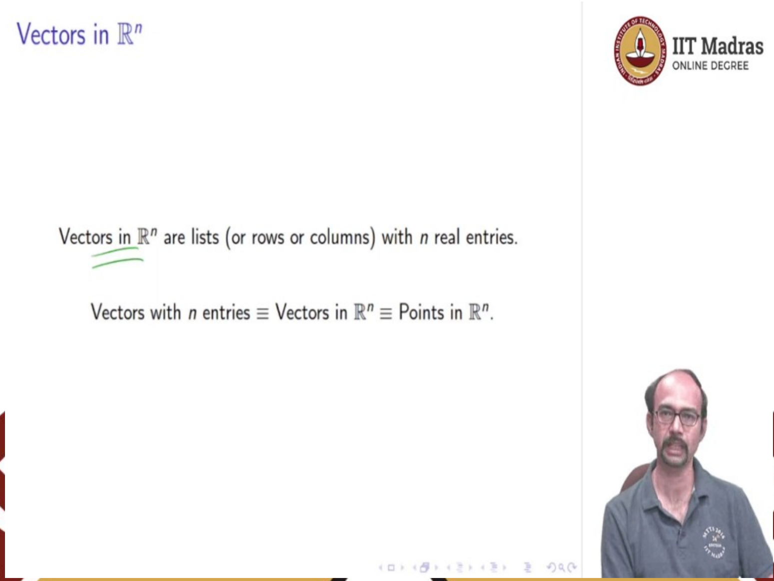


So, let us also quickly see something special about R^2 namely, how do we add two vectors in R^2. So, we can add two vectors by joining them head to tail or by the parallelogram law. So, what does that mean? So, here are two vectors v and w. So, we have drawn them with from the origin to some point and then if we want to add them, what you can do is you can either move w. So, you can move w from its starting point at 0, 0. Instead you start it at the tip of v. Or you can move v to start at the tip of w but you have to move it parallelly.

Remember that you cannot change the direction. So, you have to move it parallelly. When we say move, we mean move parallelly. And then what is v + w? v + w is exactly the vector that you obtain by drawing a arrow from the origin to the point that you obtain by completing this parallelogram. So, you could do it in either way. Maybe let us do an example here. So, we could think of the x y plane and maybe let us do, so let us say this is 1,1; this is my v, my path this is 2, 1 and this is 1, 2 and now when I add them what do I do?

I move the tip of the arrow or the starting point of the arrow 1,2 to over here. And then you can see how much, where it starts and where it ends and if you do that, you will see that this is the point 3,3 and this is exactly how we defined vector addition. So, 1,2 + 2,1 is 3,3. This was how we defined addition when we thought of them as lists and this indeed corresponds to how we add them when we do it in the traditional way we may have learnt in physics. So, this is using the parallelogram law.

(Refer Slide Time: 26:40)



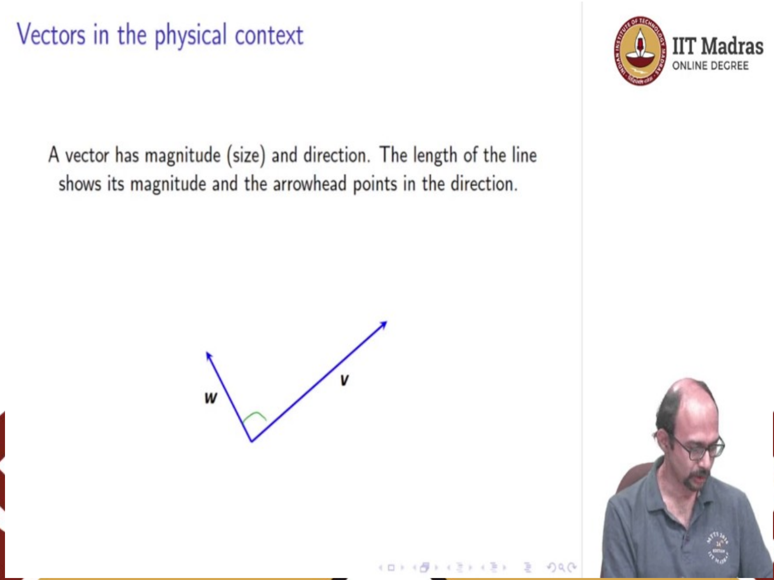
So, now let us talks about vectors in Rn and you will see why we insist on lists. So, vectors in R^n are lists that is rows or columns with n real entries. So, n entries from the real number, so n numbers. And now you cannot really think of this in terms of arrows and geometry. We have to think of this as abstract entities as just lists. And so vectors with n entries which is what we studied previously is the same as vectors in R^n.

So, what I mean is vectors in R^n in this sense and vectors in R^n as we know we can look at so you start from, you think of them as 0 and then you end at some point, then you can think of them as points in R^n. So, a vector in R^n is just a list and points in R^n are also you can think of them as just lists.

So, this is nothing very special happening. The difference is usually when we take points, we think of it as a set, so there is no concept of addition, whereas when you think of them as a vector, we think of addition or we think of scaling that vector and so on. We have no concept of scaling a point.

That is why we want to differentiate between points vectors. But just as entities, they are the same, they are just lists. So, in that sense there is nothing different in terms of representation.

(Refer Slide Time: 28:14)

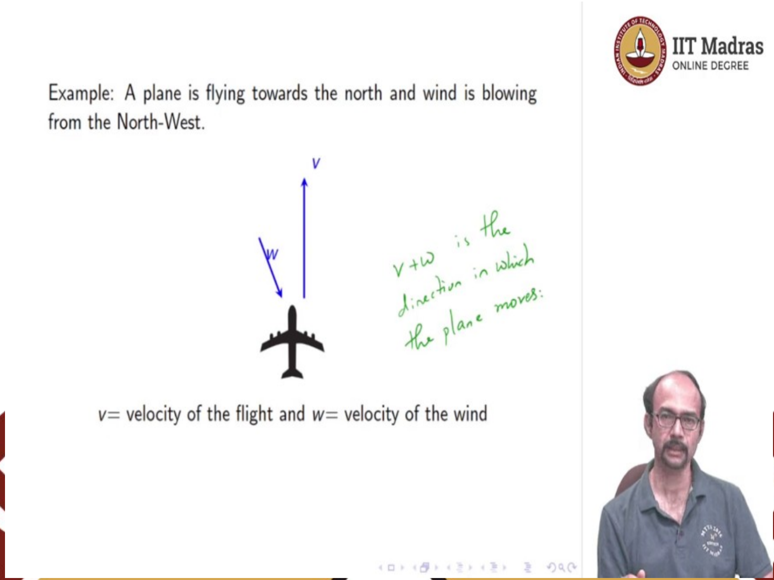


So, in the physical context, typically we see this statement in textbooks, so a vector has magnitude and direction. So, magnitude is supposed to represent size and the length of the line shows its magnitude and the arrow head points in the direction. So, for example here are two vectors v and w. So, note here that there is no axis given. So, if you assume that this is the axis and then this is a vector starting at 0,0 and ending at some point and we can say that this is, this vector is a relatively long vector.

So, it has more magnitude, the vector v has more magnitude than the vector w so it is longer and it points approximately in the north east direction, the vector w has relatively shorter magnitude and it is, it points in the north west direction. So, again this depends on where on where you draw the axis and on. So, we cannot really, directions and so on depends on the axis. So, we cannot really say much unless we know the axis, but there was somethings that we can say.

If we know the axis, then we can talk about direction and even if we do not know the axis, we can talk about the length that is something universal. And we can talk about direction in the sense that given two vectors we can talk about relatively what their, where they are pointing. So, in other words what is the angle between them? So, that is, but usually I mean when study this in physics for the first time, we do think of this in terms of axis.

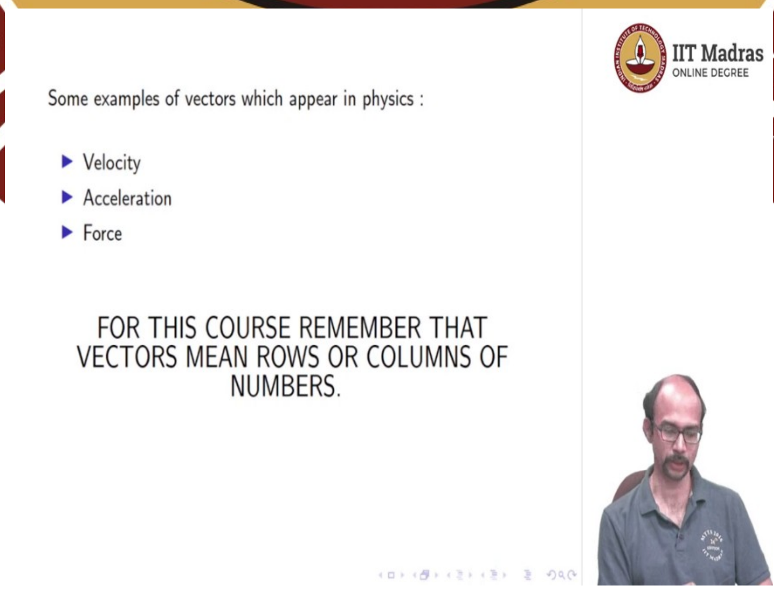
(Refer Slide Time: 30:06)



So, here is maybe an example maybe from real life. So, a plane is flying towards the north and wind is blowing from the north west. So, v is the velocity of the flight and w is the velocity of the wind and then you want to calculate what is the trajectory taken by the plane or in which direction is the net velocity. So, that will tell you the approximate direction of movement of the plane.

So, to do that you would have to do the sum. So, v + w is the direction in which the plane moves. So, as we can see vectors are quite useful in physics to determine things like this, so trajectories and so on.

(Refer Slide Time: 31:05)



So, here are some examples from physics. So, for example, velocity which we saw in the last slide or acceleration or force. So, what is important about these is that these are vector quantities meaning the direction is very important, it is not only whether the force is large or small or the acceleration is large or small, it also depends on the direction of the force or the direction of the acceleration. That will tell you the movement, something about movement. So, that is why these are vectors.

So, we have seen things like this maybe in physics or you may not have in which case what I want to say next is may not be of much use. So, what I want to say next is for this course do not keep this intuition in mind, this is not the intuition in mind. What do we want to do in this course? We want to study vectors in the context of data. Remember that is what we started with.

So, data is going to typically come out of tables or things like that and so your vectors are typically going to be lists. So, this intuition of geometry should be kept in your mind, but when you actually the algebra, think of them as lists and the addition and scalar multiplication is exactly done the way we have described, addition is coordinate wise or entry wise and similarly scalar multiplication is coordinate or entry wise.

So, these are brute force algebra algebraic operations and that is how we should remember them. So, in particular I would say the $\hat{i}, \hat{j}, \hat{k}$ is may be not something you want to think of at least for the linear algebra part of this course. So, do not think of that for vectors. Of course, when we do calculus, which we will do towards the end, we will naturally introduce coordinates and so on. So, there we may come across $\hat{i}, \hat{j}, k$ and so on where vectors will be useful. So, $\hat{i}, \hat{j}, \hat{k}$ etcetera are useful for vectors, but vectors we should think of as lists. So, with that I hope you have some feeling for vectors in this video. Thank you.