Business Data Management
Professor. G. Venkatesh
Department of Humanities and Social Sciences
Indian Institute of Technology, Madras

Introduction

So, welcome to this course on Business Data Management. The course content is about managing businesses using data. This would probably be a better title for the course. This course is going to be taught by a number of people. In fact quite a few people.

The first four weeks are going to cover the business context, which comes mostly from the world of economics. So, along with me, my name is Venkatesh and I am just called GV. So, you will hear GV being referred to quite a lot. As we go along we have Sureshbabu, who is a professor of economics at IIT Madras and he will be sharing the load of teaching the first four weeks. We both will be discussing real concepts from economics, in the first four weeks.

Subsequent to that, we have a bunch of case studies that I will go through. I will explain to you what we are going to do, which will be for the remaining seven weeks. We have very interesting case studies from four different domains. To help me with those case studies, I have Dr. Milind Gandhe, who is at IIIT Bangalore. He is the head of the machine intelligence and robotics centre there.

He has many years of industrial experience and understands this world of data very well. So, Milind Gandhe and I will be taking the next seven weeks. Between myself, Suresh babu, and Milind Gandhe we will try and get you a reasonable coverage of what we want to do in this course.

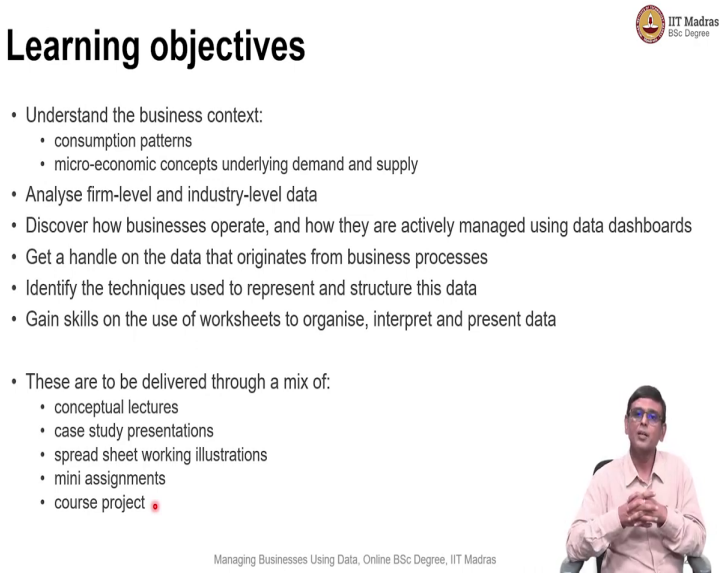
But the case studies itself we are going to have some guests, who are going to basically help us with the case study. For Flipkart, we have Omkar Karandikar who is heading the logistics division of Flipkart and he is going to basically help us with the kind of a fictitious case you have created. But based on data you can see in Flipkart.

So, this is called Fabmart, which is the case that we have for Flipkart. We are going to have a manufacturing case for which we have a person, Sivakumar, who has got many years of experience in the manufacturing industry and he has been consulting in manufacturing and so on. So, he will basically be helping us with the data as well as the discussions of that case.

And then we have Varsha from Mercer consulting which is a very prominent well known HR consulting firm. So, she is going to help us with a case in HR. Finally, we have Venkat from

PayPal, who will help us with the case on payments, Fintech. So, we have four different kinds of cases.

(Refer Slide Time: 02:55)



I will quickly explain to you what we want to do in this course. The learning objectives basically are just like, in some sense, the computational thinking course which got you inducted into the world of computing. We are assuming that you do not know anything in computing and we are kind of trying to introduce you to the world of computing.

So also, this course is a kind of introduction into the world of business. So, we are assuming that you probably do not know anything about business, and we are trying to introduce you into the world of business and trying to do that very fast, in some sense, because we have only 12 weeks to do it, and we are trying to do that very fast.

But again, just like in the computational thinking course, there is a slight variation. We are put on the content which is coming from the lens of a data scientist. So, just like in commercial thinking, we try to introduce computation from the view of a data scientist, through the lines of data. Here also we are trying to introduce you to the world of business through the lens of data.

So, that is how it is very different from what you might find in other places. So, the learning objectives of this course are to understand the business context, which means understand how consumers consume, what they consume, how they will do it, and what kind of data you look at to figure out consumption patterns. Underlying these consumption patterns, there are some

fundamental concepts that are lying around such as micro-economic concepts like the demand supply curve and things like that.

The elasticity of demand, how does demand vary or consumption vary with price or with income or something like that? So, how do we look at consumption patterns and how do you look at the relationship of consumption patterns to other variables, like income or price or some other determinants of consumption? And then we look at the supply side issues like production, the cost of production, how do you manage the cost of production? What do firms optimise? How do they optimise, under what conditions do they optimise these kinds of things? We will look at these in the first four weeks.

Beyond that, we start looking at the analysis of data that is coming from a firm. So, typical accounting data that the firm produces is in the nature of profit and loss statements, the balance sheets of the firm, the cash flow statement of the firm and other information, operational information that the firm provides. How do we take all this information and make some sense of the firm itself? How competent it is, whether it is growing, is it profitable? Is it managing its cash well, things like that will be discussed.

But also, we try to position the firm within its industry. Industry is a collection of firms who are competing for market share. How do we look at the competition that is going on in the industry and whether that firm itself or the industry itself is well located in the sense that is it getting too much pressure from suppliers or is it getting too much pressure from customers or is it under threat of being substituted away its products and so on.

These kinds of things that we will discuss in terms of analysing the firm level and industry level data and put the two together to get a composite picture of a firm. So, this constitutes the first four weeks and we will basically give you a chance to work on many of these things on your own through assignments. I will discuss a little bit more, as we go along.

We then go to the case studies and discuss how businesses operate. So, what exactly goes on inside a business? What are the business functions that are there? What is the role and responsibility of each of these functions and specifically, how do businesses manage their functions through data dashboards.

So, every business typically has some key performance areas that they have to manage and there are some measurements of these key performance called as KPIs and how do you need

to generate data and present this data in some meaningful way. How do businesses manage using data is what we will be discussing. To manage using data first you have to get the data.

So, where is the raw data stored inside an organisation? What form is it made available for analysis typically and so we will actually be looking at live data that comes from organisations, from various disciplines in the organisation like manufacturing, sales, accounting and so on and try and use some simple techniques to make sense of this data to process this data, organise this data and to present this data using some charts to make some meaning out of it.

And through this process, we hope that you will come away with some really good skills on the use of worksheets like Excel or Google sheets or any form of spreadsheet and how to use it well to organise your data. So, in this course, the data sets will be reasonably small such as 1000, 2000, 3000 entries, so that we can work with Excel.

In other courses, you will probably be working with much larger data sets which may require you to write software and process it using some statistical software and all. So, we do not have much statistics and we are not going to do any statistics in this course. It is basically looking at the raw data, organising the raw data in the form of a spreadsheet and then pulling and pushing, pulling the data into different pivot tables and then using these tables to organise the information in some way, to draw some charts from it, and make meaning out of it.

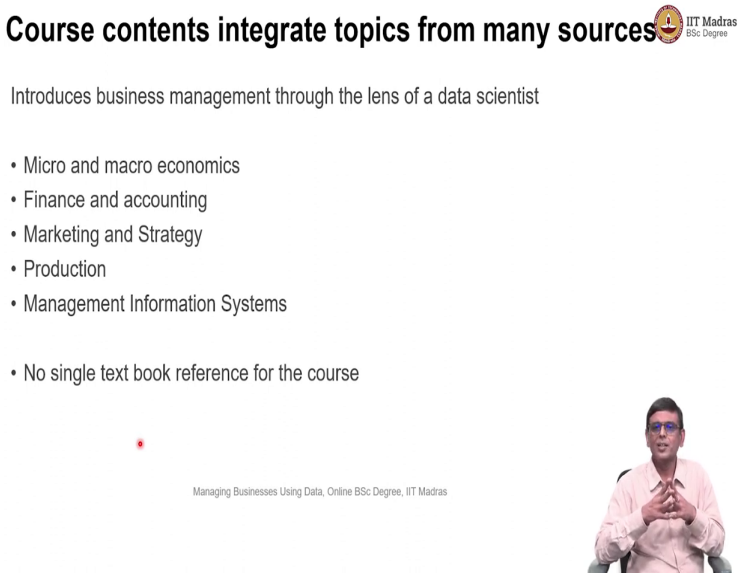
This is what we will be doing in this course and the way we will be delivering this course is through a mix of conceptual lectures, which you will see mainly in the first four weeks. These will be tested using quizzes and we will have presentations of case studies and we will basically be doing working illustrations. So, we will actually be working through those case studies with data and we will be giving you many assignments, quite a few of them.

So, that you have a chance to work on these datasets on your own and submit the results which we will evaluate. I will discuss with you a course project. So, for every student we expect you will work on a project of your own, where you will try and collect some field data from some relevant enterprise, small enterprise, and you will try and organise that data of their enterprise in the way in which we are discussing some of the issues in the case studies that will be presented to you.

So hopefully with all this, you will get a good understanding of how businesses are structured, how they are run, how what kind of data they generate, what is the thing they do with the data, what do they do with the data in terms of organising the data and then convert that organisation

into some form of presentations that they can make to their management and so on. So, this is what we hope you will come away with, a fair amount of skills on managing businesses.

(Refer Slide Time: 09:47)



Now, much of this content that we are doing in this course is coming from various sources and it is not from one book. It is not from one topic or one course or anything like that. So, we have material that is taken from the subject of micro and macroeconomics, material that is taken from finance and accounting courses, material from the marketing and strategy courses in a business management degree, material from the production operations, management, scheduling, and similar courses like that and finally, we have something coming from management information systems type of area.

So, these are very diverse different things, each of them probably is a full course and we have picked a little bit here and there and we have tried to put it together into a course. So, it might look a little bit ad-hoc, but please bear with us and you will start to appreciate the value of doing it this way.

There is, unfortunately, the problem with this method of doing things is that there is no single textbook reference that we can give you for the course. We are suggesting a couple of pointers which you can take a look at, but you will not find what we are teaching in this course in those books directly.

So, unless we write a book, which we plan to do, and share it with you, there is no single book which gives you this material directly. So, you have to make best use of all the material we are

distributing in the course and if you have any questions, feel free to write to us and then we can try and answer your questions.

(Refer Slide Time: 11:14)

With that, I want to talk a little bit deeper into what we are doing week-wise. In the first week and the second week, we will be discussing consumption and demand and the macroeconomic concepts underlying consumption demand. So, there will be a tutorial on spreadsheets, which somebody will give you. How to use spreadsheets and Google Sheets and so on, because throughout this course we are going to use it liberally. So, you need to get a good handle of how to use a spreadsheet yourself.

So, that we will be giving you an as a prelude. So, we expect that you know it before we get into this first course, but we are going to give you a small tutorial on that. In the first week, we will be discussing consumption. So, we are basically discussing how economics is organized, how economists take a look at the world in terms of different agents. Basically, there are the two agents that are very significant for us in economics, one is the producers, the firms, and the consumers, that is you and me, households.

So, on one side firms are producing the goods and services that the households consume. On the other side the households are providing the capital and labour that the firms consume. So, in some sense, there is a kind of a circular diagram that you can draw between firms on one side and consumers on the other side and basically there is a flow of goods and services from firms to consumers and there is a flow of labour and capital from the consumers to the firms.

And this flow basically depicts the world's economic model and we are trying to figure out how to organise this economic model in a way that is efficient. So, economists like to believe that price as a mechanism, shifts or makes allocation of resources in this model efficient. All these kinds of things will be discussed in week one.

What is consumption? What is production? What is this business of exchange that goes on between producers and consumers? What are the different kinds of consumption baskets that we can find like different households with different incomes consume different things? So, how does one take a look at the data that is generated from a consumption survey for example.

And there are sources of customer survey data that are available like the national sample survey, which is produced by the Ministry of Statistics in India. How do you look at that, how do you make sense out of that or other survey forms? For example, the Centre for Monitoring Indian Economy has something called consumer pyramid household survey CPHS, which is much more frequent. How do we look at that data and make sense out of it or some other kinds of survey data that are generated by consulting companies? These aspects will be looked at in week one.

Basically, the idea is to get your head around this whole idea of consumption and where do you look for data on consumption and how do you make sense out of that data. Week two is about the various underlying theoretical concepts that are there in understanding economy. So, the first theoretical principle is that of utility.

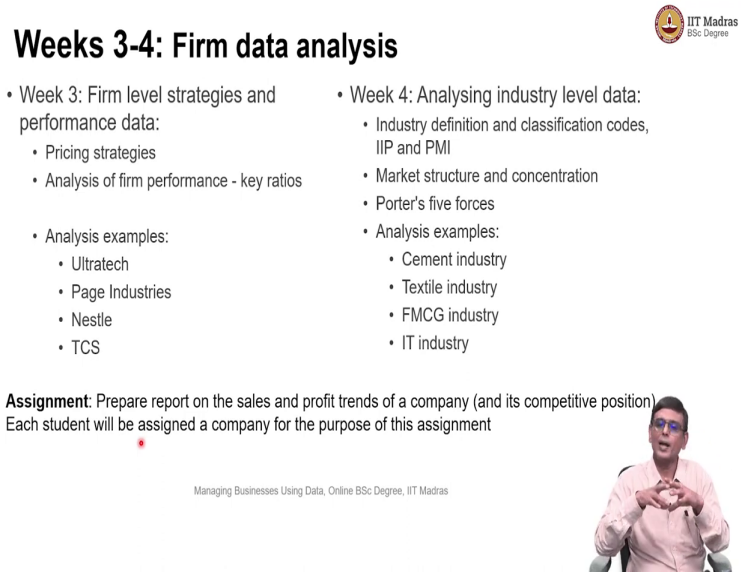
Customer satisfaction is when you buy and consume something which gives some satisfaction and that is expressed by economists using this thing called utility. Utility can be cardinal which means it can be measured or ordinal. It can be ranked. So, we look at both cardinal and ordinal utility and the idea of indifference curves.

Then we look at business of drawing a demand curve by aggregating the consumption pattern of several households, drawing the supply curve aggregating the production of various firms and how these two curves, that is the demand and supply curves, interact via the price mechanism and then we talk about how demand changes with various things like income and price and so on. Business of changing demand with various things is called elasticity of demand. The various factors of elasticity and what it means will be discussed in week two.

Simultaneously on the other side, just like we look at consumption, we will look at production and production cost curves and try to figure out how you can make decisions on how much to

produce and what quantity to produce? What kind of prices you should sell your goods at, pricing decisions and so on. How do you make those kinds of decisions by looking at some curves? So, this constitutes the material for weeks one and two.

(Refer Slide Time: 15:41)



In weeks three and four, we are going to get into analysis of data coming out of the firms. So, in week three, we will be looking at firm level performance data. In week four, we will be looking at industry level performance data. So first, we discuss pricing strategies of firms in week three and then we take cases for specific companies such as Ultratech from the cement industry, Page Industries which is a textile company, Nestle, which makes consumer FMCG products, and TCS as we all know is an IT company.

So, we have from four different industry segments cement, textile, FMCG and IT. So, first we will analyse the firms in week three and then we will look at the industry structure. Now, before we discuss the industry structure in week four, we will give you a framework within which you can understand industry structure.

So, how do you first of all classify industries, what are the ways to classify industries? National NIC codes give codes to every company and industry. We will see how this classification code comes about and the data that is generated by inspecting factories and governments who publish this data. The Index of Industrial Production gives you a high frequency data for every month which you can find out for every sector and what kind of production is happening in that sector across the country.

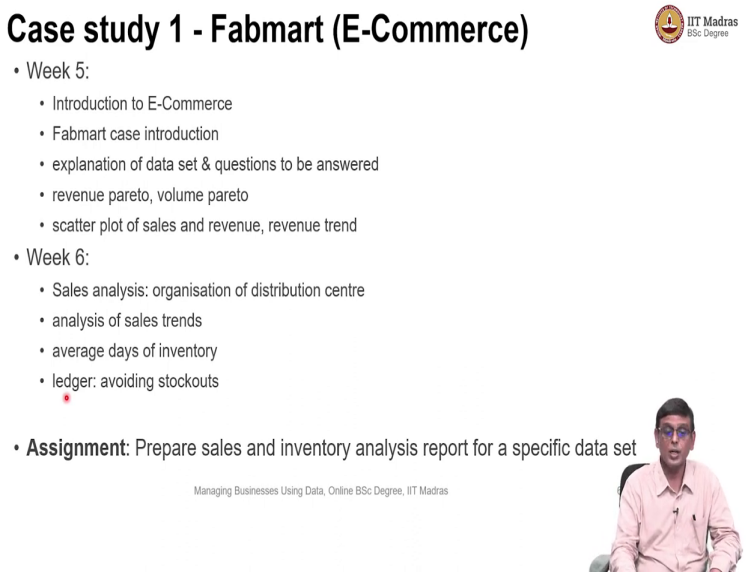
How do you look at PMI data, which is called Purchasing Managers Index. These kinds of indices tell you a fair amount of information about how an industry is performing but even with that you still have to do some analysis of where a firm or a specific industry is located. We survey other industries that are around it and we use the Porter's five forces to analyse that.

And then using those we will basically get into the analysis of the four different industries from where these four companies ,Ultratech Page Industries, Nestle and TCS, come from. We are hoping that while we do this, you will get a very good grip on how this analysis is done. We will make you then practice this analysis on your own firm.

We will give each of you a company to analyse and there is an assignment that you will have to do. In the assignment you will have to go and write a report on this company using the methods we are discussing in the class. So, using the same method that we have used in the class you will basically use it to do an analysis of your own on this company. We will tell you where to go, look for data on this company and so on and you have to write a report and that will be an assignment.

So, hopefully with this assignment, you get a very good understanding of the context within which companies operate, what are the terms that we use typically when we discuss performance of a company and then how do we analyse the firm's performance using various financial data that the firm provides, as well as industry data that is available.

(Refer Slide Time: 18:51)



From there, we move to case studies. So, the first case study that we have for you is called the Fabmart case study. It is based on the Flipkart data set. So clearly, it comes from the world of

e-commerce. It is an e-commerce case study. The guest lecture that we have is Omkar. He will introduce us to the world of e-commerce, what is e-commerce and then he will tell us a little bit about what this case is about and the data set that we have got in this case.

And then we, Milind and I, will take the data work on it. We will show you how to work with the data. We will take the real data in some sense and then both of us will take spreadsheets and work through the data to show you how to do a revenue Pareto, how to do a volume Pareto, how to do a scatter plot, how to do portfolio management, and few things like that.

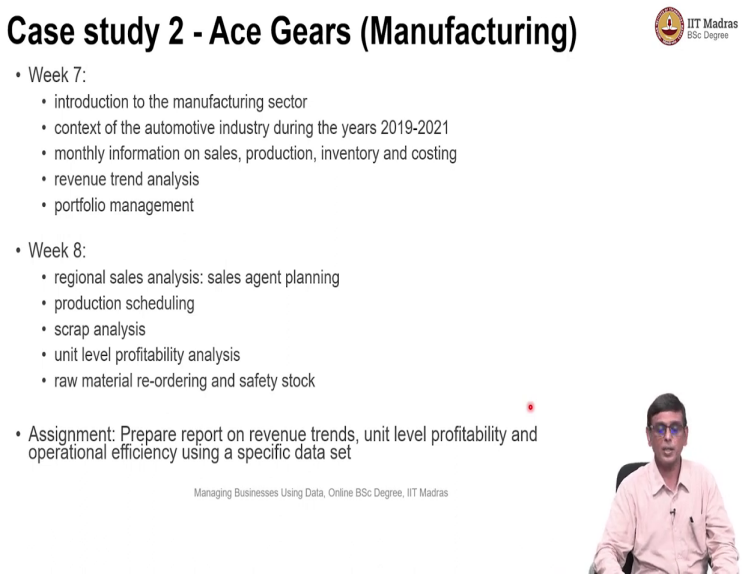
So that you get a grip on not only understanding the raw data, but also understanding how to process all this raw data and to put it together in the form of charts that we can present. Then in week 6, we will get down into analysis of the sales and organising the distribution centres Flipkart has, in this case, Fabmart.

In the Fabmart, they have distribution centres in three locations which are Hyderabad, Chennai and Kochi. How do they organise a distribution centre, so that you can efficiently deliver things to customers speedily? It turns out that you can do this by looking at the sales data and then look at some analysis of trends and sales trends. We will also look at things like inventory and how you manage the inventory well.

Then we will introduce to you this concept of a ledger and how do you look at this ledger and use this ledger to figure out how much talk to keep at each of these places. How do you avoid what is called stockout? Stockout basically happens when you do not have material to sell and the customer is looking for something and you do not have that thing to sell. This situation is called a stock out, which apparently is very bad for ecommerce companies. So, they keep worrying about having the stock out.

How do we avoid this by using this concept called the ledger? There will be an assignment for you which will be to prepare sales and inventory analysis for a specific data set. So, you will have a data set given to you which is picked up from the case study and you will have to take this data set and you have to do your own sales and inventory analysis report. So hopefully, with that, you will learn quite a lot about how to look at sales data, how to look at inventory data and make sense out of it.

(Refer Slide Time: 21:29)



The second case is for a manufacturing company. This is a company which makes gear assemblies. It is based on an actual manufacturing company's data. But we have kind of fudged it or massaged it to make sure the data is kept confidential, so that we do not get to the secrets of the company.

We have hidden some things, but still the data is meaningful as it has come from real life. We have taken this manufacturing data and Sivakumar will basically introduce us to the manufacturing sector. He will set the context of the automotive industry and what has been happening in the last two years, especially in the time of COVID. There are a lot of technology transformations that are happening and he will tell you basically that such transformations keep happening in the automotive industry and how they have to deal with it.

How do we use all the data that is there in terms of sales, production to do planning? Because one of the things that we learned in manufacturing is that you have to plan. So, how do you plan your production schedules? How do you plan your shift schedules? How do you plan your raw material purchases? How do you manage your inventory? All these are huge issues in manufacturing. So, we will try and do all that.

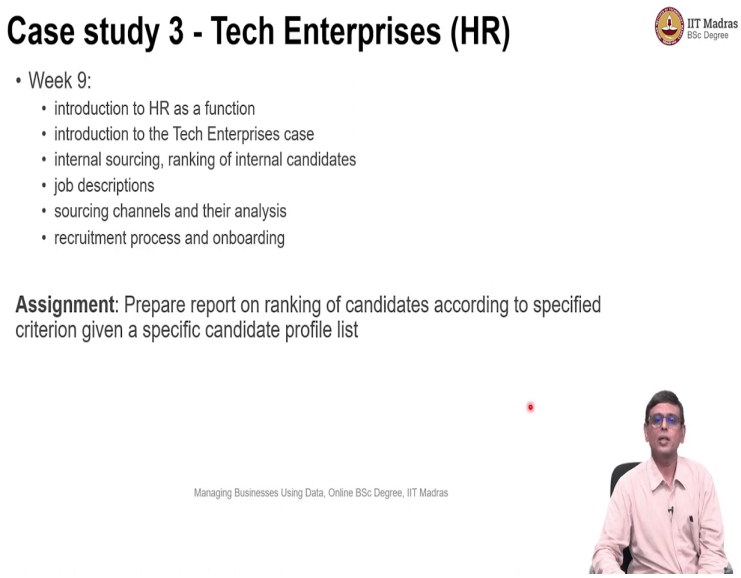
But before we do that, we will again try and do some revenue trend management and portfolio management using their sales data. We have in the case of fabmart data for 15 days. Whereas in the ace gears we have monthly data for two years, 24 months.

And then we will do some regional sales analysis and try to plan the number of sales agents. We will do some scheduling, some analysis of scrap and some unit level profitability analysis. So, that gives you an idea of the finance side of the world and how you do reordering of raw

materials using this concept called safety stock, reorder points and reorder quantities and things like that.

Again, here there will be an assignment that will help you to get to your own sense of everything. So, we will give you a data set and we will ask you to prepare a report on revenue trends, unit level profitability, operational efficiency and things like this using your own data set.

(Refer Slide Time: 23:44)

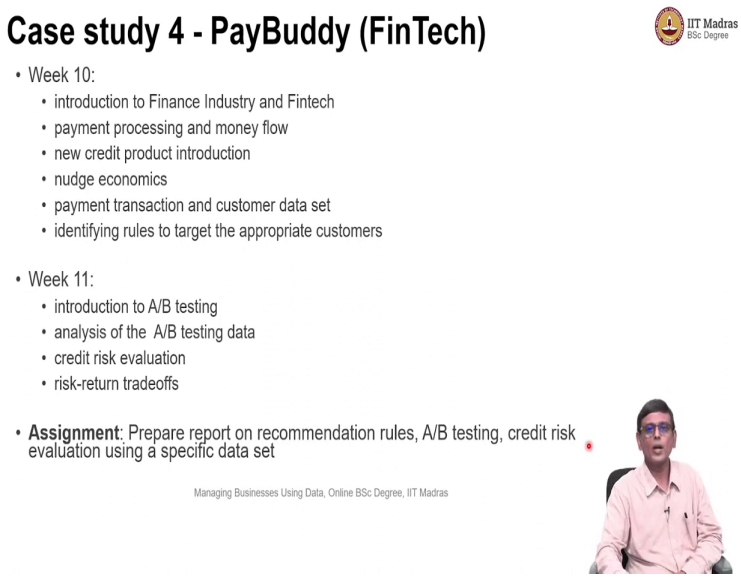


The third case comes from the world of HR. Here we have an IT company which is called tech enterprises which is a fictitious company and the data is based on the real world, but we have massaged it and put it into the form of tech enterprises. So, here we get an introduction to HR as a function and also an introduction to the data set.

Here the basic thing is a recruitment problem. So, we need to recruit some candidates. So we have two cases. The first case is about internal sourcing, which means we try to see whether there is a fitment from within the company for a particular position. In the second case let us look at recruiting from outside.

So, how do we perform management of this data, recruitment data, candidate profile data, and how do we take all of that, put it and then rank the candidates. You will see some examples of unstructured data here compared to the previous cases. This will be more useful for you to understand how unstructured data can be put together and how you can do ranking of candidates and so on using unstructured data.

(Refer Slide Time: 24:45)

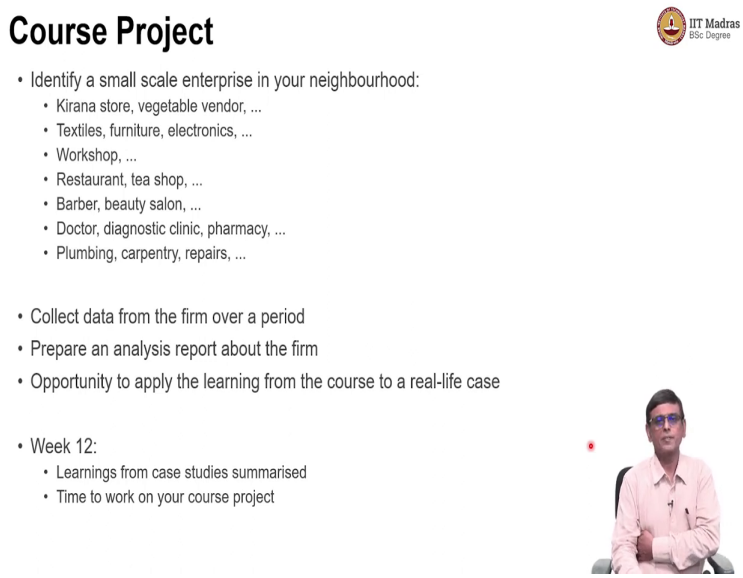


Finally, the fourth case that we have for you is from the FinTech world. This is from our guests who are from PayPal. So, the data is also from PayPal, but we massaged it again and so it is a fictitious case called paybuddy here. They will be an introduction to the finance industry, FinTech, in which we will understand payment processing and how money flows in payment processing.

And then we will introduce a new credit product and our guests will explain to us basically this whole idea of nudge economics, and how customer behaviour can be altered or modified by using this method called nudge. By giving default options, we will learn that it is quite interesting to learn and basically use nudge economics for a new product or trading product you want to introduce. You will want to see how well that credit product is doing and for that we use what is a method which is called AB testing.

How does AB testing work and how do we analyse AB testing data? And finally, how you evaluate credit risk, how credit risk is actually evaluated by firms like this and how we do risk-return trade-offs are some of the things we will do. Again, there will be an assignment which will let you practice everything and you will have to write a report where you have to come up with your own recommendation rules, do some AB testing, analysis and so on. You will get to learn how to do all these things yourself.

(Refer Slide Time: 26:03)



Finally in the course project, we expect every one of you to do a project, where you will have to identify a small-scale enterprise in your neighbourhood. If you have difficulty with identifying a small-scale enterprise, then we can help you with either identifying or getting this data source for you. So, we will make sure that you get your hands on some data that you can work on.

This will be your own data unlike all other assignments and so on which is packaged data that is already processed data which is cleaned. Here it will be uncleaned with all kinds of issues and noises which you have to deal with yourself. So, talk to your neighbourhood kirana store, vegetable vendor, a textile store, furniture store, electronic store, workshop, tea shop, barber, beauty salon, doctor or anybody who is doing something which is business.

And we expect you to collect some business data from this firm over a period and then prepare an analysis report that will tell you how to do it and use all the techniques that we learn in this course while you prepare this report. Then you will present this and we will have people who will help you with it.

We have a large number of tutors, who are going to be assigned for this course, who will basically help you with preparing, looking at this data, preparing the report and so on. So, up to eleventh week we will be doing case studies. Week twelve we want to use it for you to work on your course project work and we will also be presenting the learnings from the case studies that we have gone through in the course.

So, this is basically how our twelve weeks course is going to run. I hope you all enjoy this course just as much as we enjoyed putting all this material together and delivering it to you. In case you find any, face any problems while going through the material, feel free to write to us and then we will be as responsive as possible to give you clarifications and help you to get along with this content. Thank you.