Transcriptions Taken from The *The Importance of Fair Business Decisions* Course Videos

Video One - The Power of Data in Business

As a data analyst, you'll be tackling business tasks that help companies use data, Coming Up, we'll talk more about what a business task actually is, and some examples of what they might look like in real analyst jobs. Let's take a second and think back on the real examples of businesses using data analytics and their operation we've seen before. You might have noticed a common these across every example. They all have issues to explore questions to answer, or problems to solve. It's easy for these things to get mixed up. Here's a way to keep them straight when we talk about them in data analytics.

- **An issue is a topic or subject to investigate**
- **A question is designed to discover information**
- **And a problem is an obstacle or complication that needs to be worked out.**

Coca-cola had a question about new products. Data analysis game them insights into new flavors customers already like. The City Zoo and Aquarium had a problem with staffing. Data helped them figure out the best staffing strategy. These questions and problems become the foundation for all kinds of business tasks that you'll help solve as a data analyst. A business task is the question or problem data analysis answers for business. This is where you focus a lot of your efforts in the work you'll do for future employers. Let's stick with our zoo example and see if we can imagine what a business task for a zoo might look like. We know the problem, *unpredictable weather* was making it hard for the zoo to anticipate staffing needs, Maybe the business task could be something like, *analyze weather data from the last decade to identify predictable patterns*. The data analysts could then plan out the best way to *gather, analyze, and present* the data needed to solve this task for the zoos goals. Then, using data, the zoo would be able to make informed decisions about their daily staffing. We talked a little about data-driven decision making in previous videos. But just in case you need a refresher, here it is. *Data-driven decision-making is when facts that have been discovered are used to guide a business strategy.* The simplest way to think about decision-making is that it's a choice between consequences. good, bad, or a combination of both. In our zoo example, the zoo had the data they needed to make an informed decision that solved their problem. But what if they had made this decision without data? Let say they just relied on observation and memory to track the weather and make staffing schedules. Well, we already know that wouldn't solve their problem long-term. Data analytics gave them the information they needed to find the best possible solution to their problem. That's the power of data. Observation and intuition are powerful tool in decision-making, but they can only take us so far when we make decisions based on just observation and gut-feelings, we're only seeing part of the picture. Data helps us see the whole thing. With data, we have a complete picture of the problem and it's causes, which let's us find new and surprising solutions we never would have been able to see before. Data analytics helps businesses make better decisions. It all starts with a business task and the question it's trying to answer. With the skills you'll learn throughout this program you'll be able to ask the right questions, plan out the best way to gather and analyze the data, and then present it visually to arm your team so they can make an informed, data-driven decision. That makes you critical to the success of any business you work for. Data is a powerful tool. With great power comes great, well, you know the rest. Up next, we'll take about your responsibility as a

data analyst to make sure you're gathering, analyzing, and presenting data in a way that's fair to the being being represented by that data.

Video Two - Rachel: Data Detective

Hi, my name is Rachel, and I'm the *Business Systems and Analytics Lead* at *Verily*. There are a lot of different types of problems that a data analyst can solve. I've been lucky enough over my career to have see a lot of them and to take in a lot of very different ty[pes of data and help turn that into meaningful answers. I think one of the most important things to remember about data analytics us that data is data. I'm a finance data analyst and so my role at *Verily* is to take all of our financial information. All of the information of the money we're spending and the money we're making, and turn that into reports and insights so that our business leads can understand what we're doing. One of the most important things I've done at *Verily* recently was help create what's called a *Profit and Loss Statement* for each of our business units. That means that in real time, our teams can see what their budget is and how they're spending against that budget. What that does, is help our our teams keep to their budget by either increasing their revenue streams so that they'll have more money to play with, or by pulling back their spending so that they can keep themselves on track within that budget. All of that really helps keep us on track as a company in making sure that we're hitting our goals. I found that data acts like a living, breathing thing. When you have a ton of data points, it can be a bit overwhelming when you first sit down to try and make sense of it. You have tons of columns, tons of records, tons of different types of data, and finding a way to make sense of that is really hard and that's where the expertise of a data analyst comes in. It has been some of the most frustrating moments of my career, but also some of the most rewarding work I've ever done when it finally comes together. The best advice I have fort any data analyst just starting out is to keep at it. If the angle you're taking doesn't work, try another one. Try to come at the problem in a different way, try to ask a different question. Eventually the data will yield and you'll get the insights you're looking for.

Video Three - Understanding Data and Fairness

So far, we've covered the different roles data analysts play in business environments and the kinds of tasks that come with those roles. But data analysts have another important responsibility: making sure that their analyses are fair. Now, I know what you're probably thinking, *data is based on collected facts, how can it be unfair?* WEll that's a good question! Let's learn what *fairness* means when we talk about data analysis and why it's important for you as an analyst to keep in mind. Fairness means *ensuring that your analysis doesn't create or reinforce bias*. In other words, *As a data analyst, you want to help create systems that are fair and inclusive to everyone*. Sounds simple enough? Well, here's the tough part about fairness in data analytics. There isn't one standard definition of it, but hopefully the way we've just described it can give you one way to think about fairness for right now, but it's about to get a bit trickier. Sometimes conclusions based on data can be true and unfair. What can you do about then? Well, let's find out with an example. Let's say we have a company that's kind of notorious for being a "boys club". There isn't much representation of other genders. This company wants to see which employees are doing well, so they start gathering data on employee performance and their own company culture. The data shows that men are the only people succeeding at this company. Their conclusion? That they should hire more men, because after-all, they are doing really well there. But that's not a fair conclusion for a couple of reasons. First, it doesn't even consider all of the available data on company culture, so it paints an incomplete picture. Second, it doesn't think about the other

surrounding factors that impact the data, or in other words, the conclusion doesn't consider the difficulties that people of different gender identities have trying to navigate a toxic work environment. If the company only looks at this conclusion, they wont acknowledge and address how harmful their culture is and they wont understand why certain people are setup to fail within it. That's why it's important to keep fairness in mind when analyzing dat. The conclusion that only men are succeeding at this company is true, but it ignores other systemic factors that are contributing to this problem. But don't worry, there's a way to make a fair conclusion here. An ethical data analyst can look at the data gathered and conclude that the company culture is preventing some employees from succeeding, and the company needs to address those problems to boost performance. See how this conclusion paints a much more complete and fair picture. It recognizes the fact that some people aren't doing as well in this company and factors in why that could be instead of discriminating against a huge number of applicants in the future. As a data analyst its more responsibility to make sure your analysis is fair and factors in the complicated social context that could create bias based on your conclusions. It's important to think about fairness from the moment you start collecting data for a business task to the time you present your conclusions to your stakeholders. We'll learn more about bias in the data analysis process later on in another course. For now, let's check out an example of a data analysis that does a good job of considering fairness in its conclusion. a team of Harvard data scientists were developing a mobile platform to track patients at rick of cardiovascular disease in an area of the United States known as the *Stroke belt*. It's important to call out that there were a variety of reasons people in this area might be more at risk. With that in mind, these data scientists recognized that fairness needed to be a priority for this project, so they built fairness in,to their models. The team took several fairness measures to make sure they were being as far as possible when examining sensitive and potentially biased data. First, they teamed analysts with social scientists who could provide insights on human bias and the social context that created them. They also collected self reported data in a separate system to avoid the potential for racial bias, which might skew the results of their study and unfairly represent patients. To make sure this sample population was representative, they oversampled non-dominant groups to ensure the model was including them. It's clear that the team made fairness a top priority every step of the way. This helped them collect data and create conclusions that didn't negatively impact the communities they were studying . Hopefully these examples have given you a better idea of what fairness means in data analysis. But, we're going to keep building on our understanding of fairness throughout this program and you'll get to practice with some activities.

Video Four - Alex: Fair and Ethical Data Decisions

Hi, I'm Alex. I'm a research scientist at Google. My team is called the *Ethical AL Team*, we're a group of folks that are really concerned not only about how AI technology operates, but how it interacts with society and how it might help or hard marginalized communities. When we talk about data ethics, we think about what is the good and right way if using data, When it comes to data ethics, it's just about minimizing hard but it's actually the concept of beneficence. How do we actually improve the lives of people by using data? *When we think about data ethics we're thinking about, who's collecting the data, Why are they collecting the data, why are they collecting it, and How are they collecting it and for what purpose.* Because of the way that organizations have imperatives to make money or to report to somebody or provide some analysis, we also have keep strongly in mind how this is actually going to benefit people at that end of the day. Are the people represented in this data going to be benefited by this? I think that's the thing you never want to lose sight of as a data scientist or data analyst. I think aspiring data analysts need to keep in mind that a lot of the data you're going to encounter is data that comes from people so at the end of the day, data are people. You want to have a responsibility to those people that are represented in those data sets. Second, is thinking about how

to keep aspects of their data protected and private. we don't want to go through our practice thinking about data instances as something we can just throw on the web. No, there needs to be considerations about how to keep that information and likenesses like their images, or their voices, or their text. how do we keep that private? We also need to think about how we can have mechanisms of giving users and giving consumers more control over their data. It's not going to be sufficient to say: *We collect this data, just trust us with it*. But we need to ensure that there's always actionable ways in which people can consent to giving those data, and ways that they can as for it to be revoked or removed. Data's growing and at the same time, we need to empower people to have control over their own data. The future is that data is always growing, we haven't seen any evidence that data is actually shrinking. With the knowledge that data's growing, these issues become more and more piqued, and more and more important to think about.