OvalEdge Academy Transcript:
Data Discovery

Video: Navigating Lineage

Hi, I'm Becca from OvalEdge. Today we'll discuss how you can navigate lineage.

At a high level, lineage is data movement from one source to the next, repeated many times. A typical data environment includes source systems, ETL or ELT tools, to extract, perform transformations, and load the data into another place. This would typically be a data lake, warehouse, or store, and the data would ultimately end up on a dashboard or report.

Lineage has many benefits that increases confidence, speed to market, quality remediation, and tech debt retirement, making it a worthy endeavor for organizations to invest that quickly has a return on its investment.

The data is moved from place to place, which could also include some transformations. Additionally, it is used in formulas for metrics and used as filters. Lineage can provide the data flow from its inception to its consumption.

Data governance promises consistency and clarity. Lineage is a powerful tool that shows you the source and movement of data. Have you ever looked at a report and asked yourself, "Where did this data come from?" Well, now we can show you!

Data Lineage for Business Users
Business users are often fed analytics through a report and are responsible for making a business decision based on this information. However, sometimes, they don't know what the data provided means, they don't trust it, they don't know who wrote it, and they don't know where the data came from or how current it is.

So, if I click on a report, let's focus on this 'SuperStore Summary,' I can see the report's lineage. Users can follow where an object originated and how it traveled to its reporting destination. Now, you can understand where your information is going and where it's coming from.

You can also look at an individual column's lineage. In the 'Column Details' tab, I can go to the 'Profit' column and select 'Column Lineage.' From here, the 'Profit' column is traced back to its source in the 'Superstore DM' report. This is helpful because I now know the source from which my profit value calculation originated.

Impact Analysis
Data lineage also allows organizations to run upstream and downstream impact analyses to address changes in the environment and keep data assets intact. It is often seen as a game

changer for all data management and pipeline participants. For more information on impact analysis, reference the Impact Analysis video.

Thanks for watching!