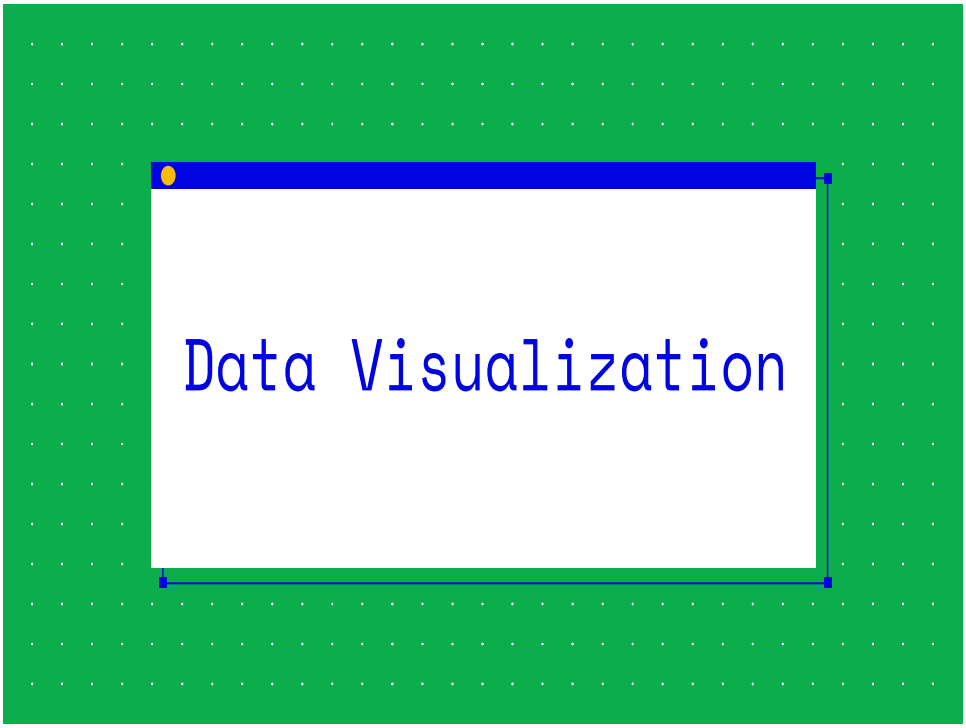
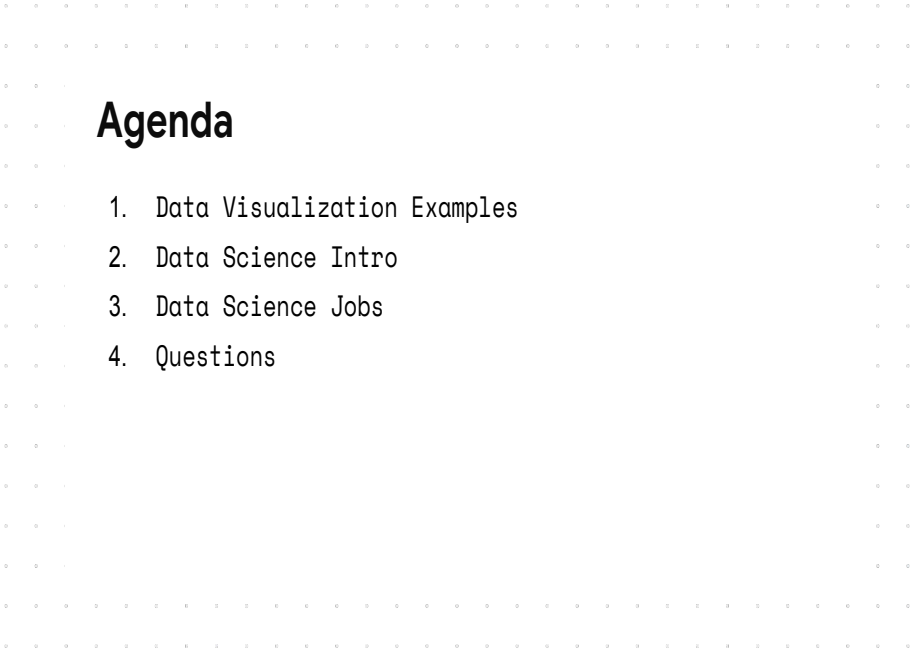
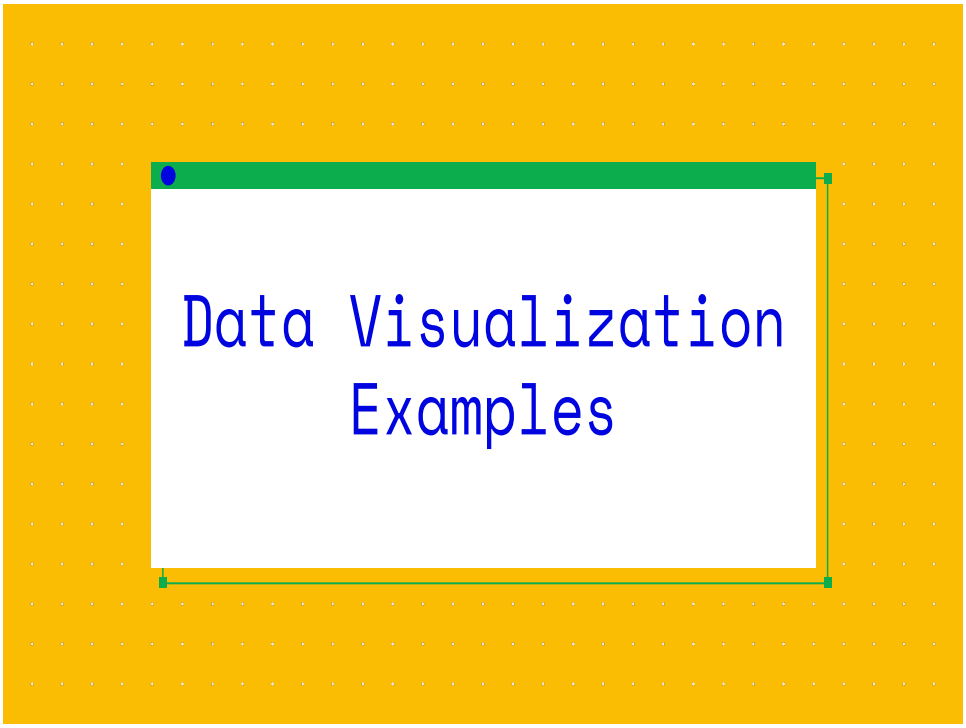

Data Visualization

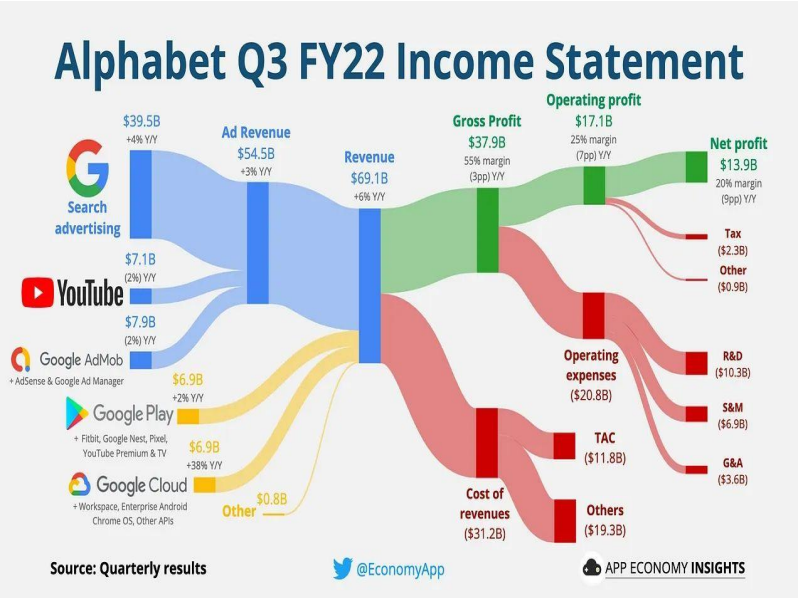


Here is a quick overview of what we'll cover today. First, we'll look at some compelling examples of data visualization, then we'll introduce the field of data science, discuss the different types of jobs within it, and finally, we'll open the floor for questions.



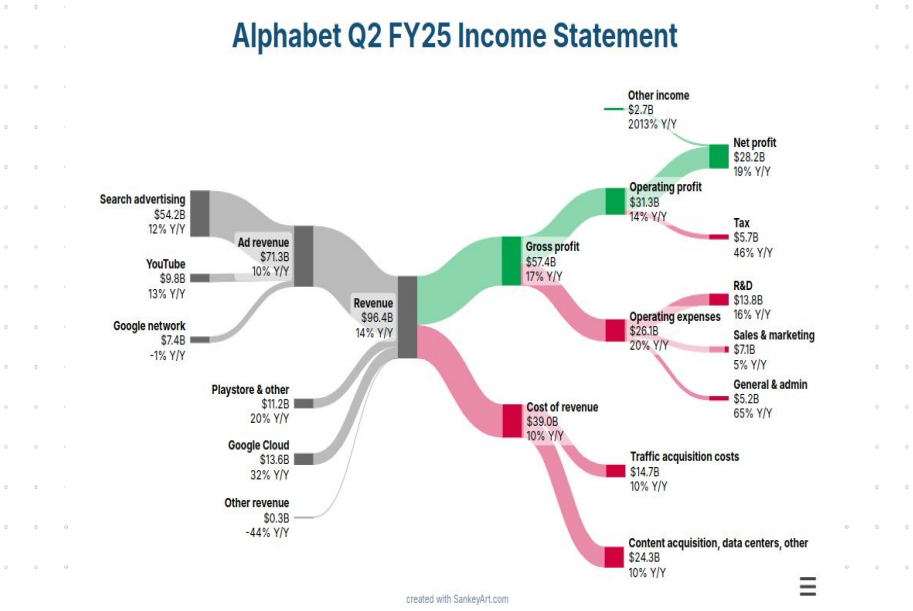
This slide transitions us into the first part of our agenda. We're going to dive right into some practical examples to show the power of data visualization.

Data Visualization Example 1: Sankey Diagram



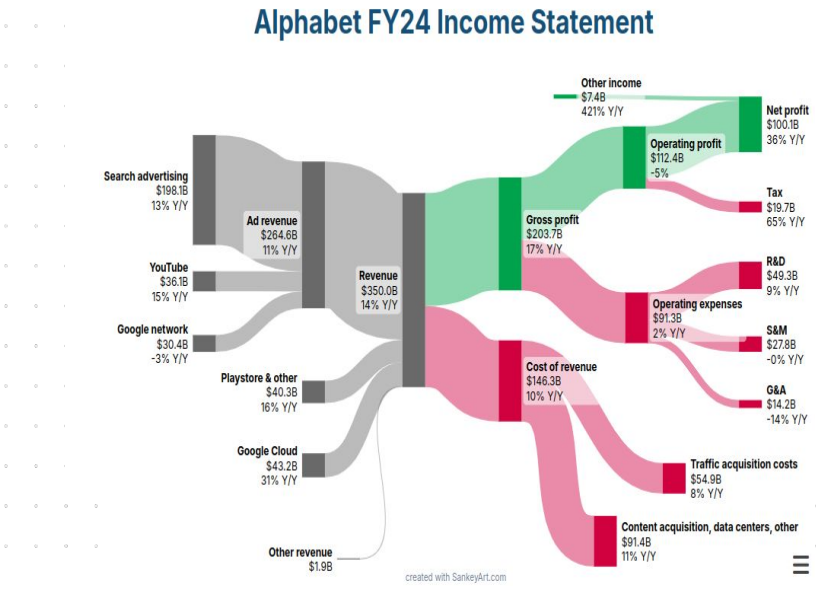
This is a great example of a Sankey diagram, which is a type of flow diagram. It's used here to visualize Alphabet's income statement. The chart effectively shows how revenue flows in from different sources and is then allocated to various expenses, profits, and taxes. You can quickly see the proportional contribution of each revenue stream, like Search and YouTube.

Data Visualization Example 2: Sankey Diagram



Please think about the differences between the current and previous graph, such as the background, the colors used, and their meaning.

Data Visualization Example 3: Sankey Diagram



Alphabet 2024 Income Sankey Diagram

https://www.sankeyart.com/sankeys/public/139100/

Data Visualization Example 4: Dashboard

This dashboard image was created via GenAI using the following prompt: Create an image that represents a dashboard created by Looker Studio.

Data Visualization Example 5: Dashboard

This dashboard image was created via GenAI using the following prompt: Create an image that represents a dashboard created by Looker Studio.

We're now moving into the second part of our agenda: an introduction to Data Science.

Data Science

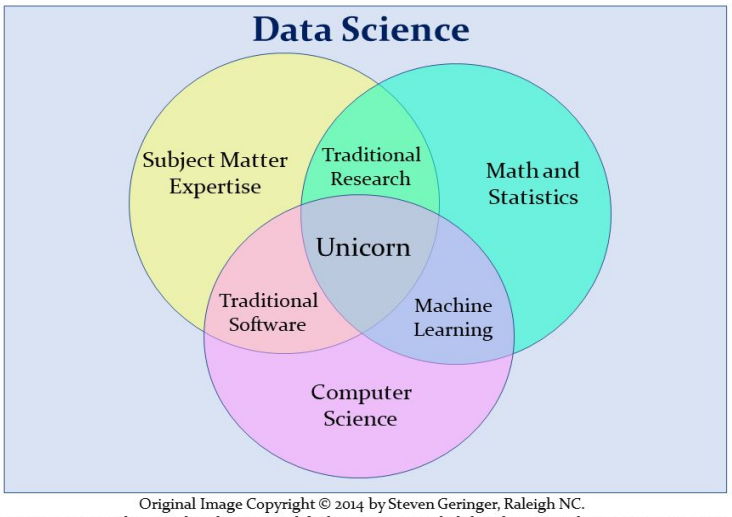
- combines several disciplines including:
  a. Math
  b. statistics
  c. computer science

- How to become a **Unicorn** in **Data Science**?

Data science is an interdisciplinary field that combines elements of statistics, computer science, and domain expertise to solve complex problems.

- https://www.1800flowers.com/blog/delivering-smiles/unicorn-facts/
- **Unicorn:** term used in the venture capital industry to describe a startup company with a value of over $1 billion.
  - https://www.inventiva.co.in/trends/top-10-unicorns-in-egypt/
  - https://www.google.com/search?q=Swvl++price

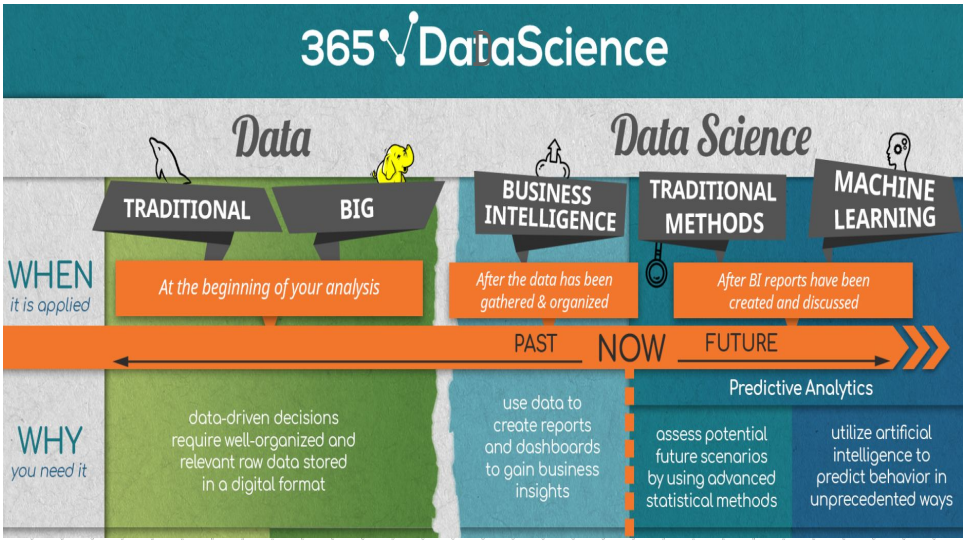
Data Science Venn Diagram 2.0



Visualize using Venn Diagram how to become a unicorn in data science

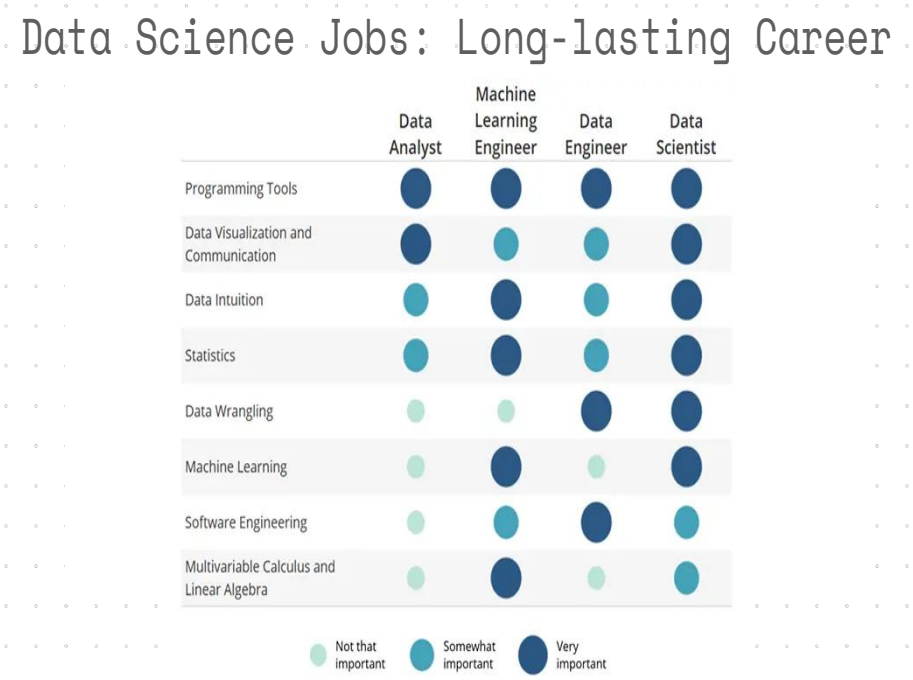
- This famous Venn diagram illustrates the core components of data science.
- It shows that data science is the intersection of three key areas:
    - **Hacking Skills**: The ability to code and work with data.
    - **Math & Statistics Knowledge**: The foundation for understanding data and building models.
    - **Substantive Expertise**: The domain knowledge needed to ask the right questions and interpret the results.

365 DataScience Infographic (When & Why)



Try to do your best to understand the above diagram.

We're now moving into the third part of our agenda: different Data Science jobs.

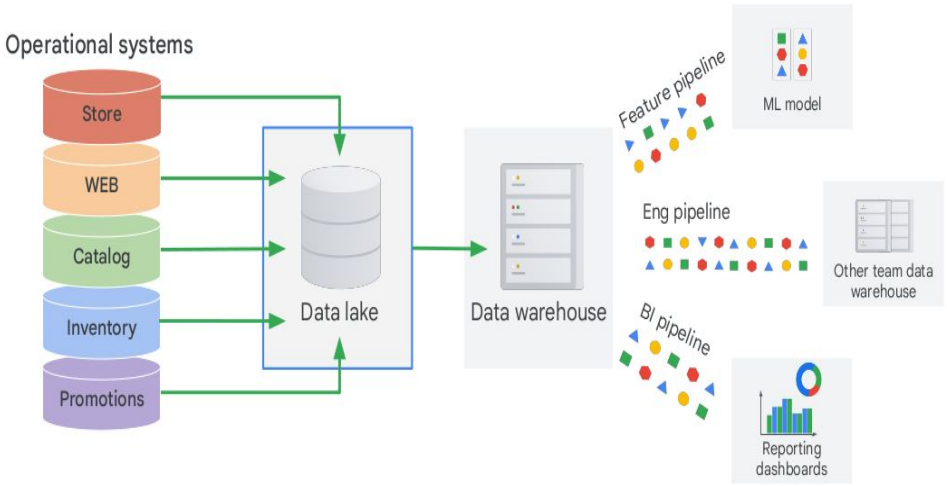


Now we'll focus on the different types of jobs within the data science field. The roles can vary significantly from one company to another.
https://www.udacity.com/blog/2018/01/4-types-data-science-jobs.html

Intuition حدس in Data Science

- is not about using your gut الغريزي feel.
- is the intuitive understanding of concepts (i.e.**how to apply the concepts**).
- A lot of people make the mistake thinking that to be a successful data scientist I have to learn all the math.

Data Science Jobs: Where is the role of each?



Try to find where each of the following will work in the above diagram:

1. Data Analyst
2. ML Engineer
3. Data Engineer
4. Data Scientist
5. BI Analyst

Data Science Jobs: 1) Data Analyst

- Job might consist of tasks like:
  - pulling data out of **SQL** databases
  - becoming an **Excel** or **Tableau** master
  - producing basic **Data Visualizations** and **Reporting Dashboards**
  - On occasion:
    - analyze the results of an **A/B Test**
    - take the lead on company's **Google Analytics account**

- Some companies: **Data Scientist** is synonymous to **Data Analyst**

A **Data Analyst** is typically the entry point for many in this field.

- Their primary role is to analyze data and present actionable insights.
- One of the key skills for this role is the ability to visualize data to tell a story.

Google Analytics

- is a web analytics service offered by Google that tracks and reports website traffic and also the mobile app traffic and events
- Google launched the service in November 2005 after acquiring **Urchin**, currently is a platform inside the **Google Marketing Platform** brand

Data Science Jobs: 2) ML Engineer

- Some companies: **data** or **data analysis platform** is the product
  - so, **Data Analysis** or **ML** can be pretty intense
  - so, there is a need for someone who:
    - has a formal mathematics or statistics background
    - is hoping to continue down a more academic path

- **ML Engineers**
  - often focus more on producing great **data-driven products**

An **ML Engineer**'s job is to build and deploy machine learning models in a production environment.

- This role requires strong a deep understanding of ML algorithms.

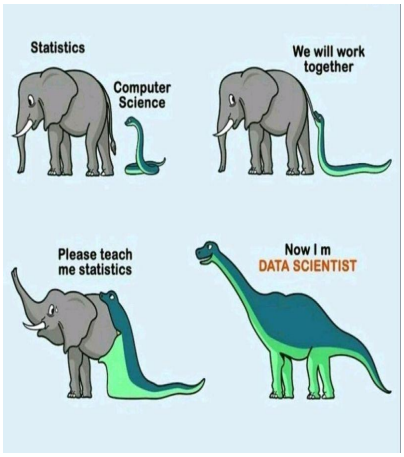
Data Science Jobs: 3) Data Engineer

- Some companies: have **lot of traffic** and **large amount of data**
  - so, there is a need for someone who:
    - set up lot of **data infrastructure** that the company will need
- Job postings listed under **Data Engineer**
  - strong software engineering skills are more important
  - heavy statistics and ML expertise are less important

Data Engineer's role is to build and maintain the data pipelines that data analysts rely on.

- They are responsible for making sure the data is clean, reliable, and accessible.
- This role is heavily focused on software engineering and data systems architecture.

Data Science Jobs: 4) Data Scientist



Often used as a blanket title

- to describe jobs that are drastically different

The term "**Data Scientist**" can be a bit of a blanket title.

- In some companies, a **Data Scientist** is a generalist who does a little bit of everything—from data cleaning to model building to visualization.

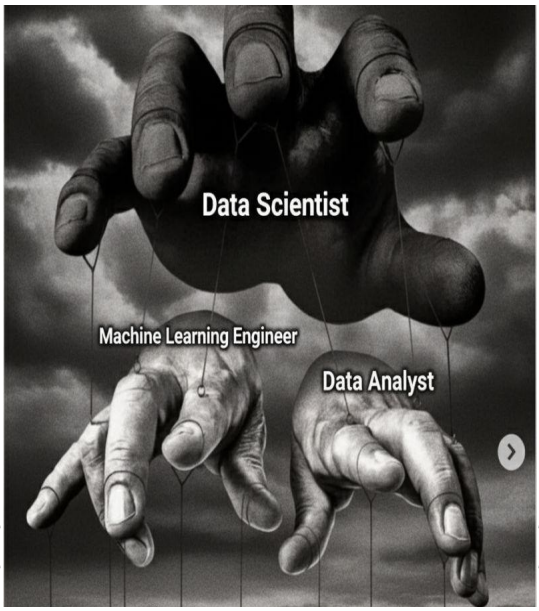
Data Science Jobs: 5) Data Generalist

- Some companies: look 4 **Generalist** to join data scientists team
  - these companies cares about data but probably isn't a data company
- **Data Generalist** most important skills:
  - familiarity with tools designed for **Big Data**
  - experience with messy, real-life datasets
- **Data Generalist** can:
  - perform analysis
  - touch production code
  - **Visualize Data**

A Data Generalist is a "jack of all trades" in the data science world.

- These are individuals who are comfortable with a wide range of tasks, including data analysis and data visualization.
- They are valuable in companies that are data-aware but may not have a fully specialized data team.

Data Analyst & ML Engineer vs Data Scientist



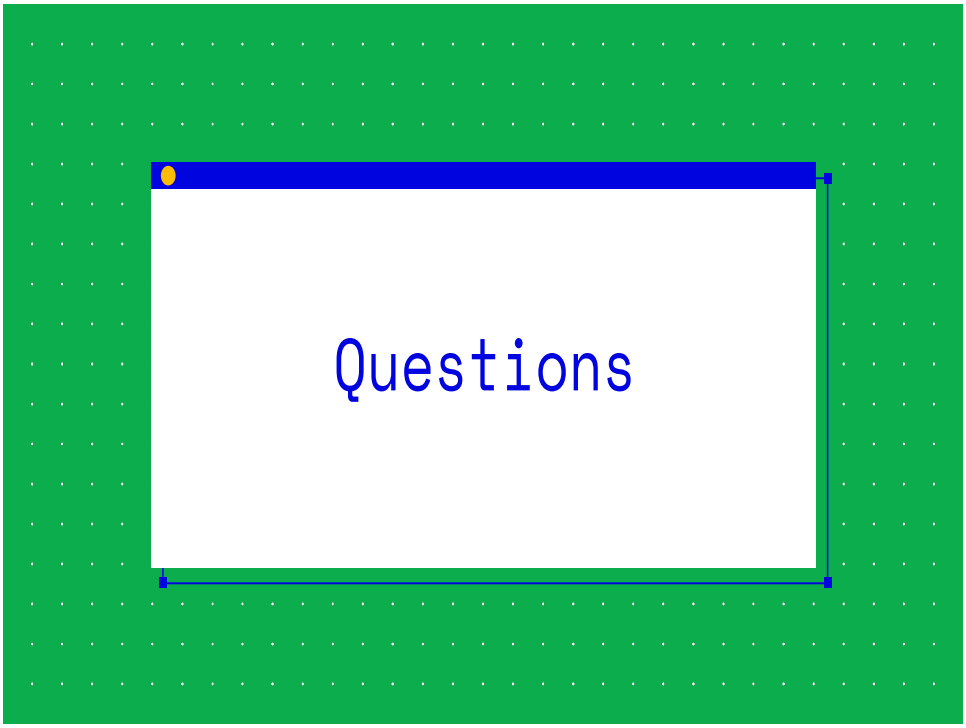
What is your opinion about the above image?

Job Confusion

HR Manager offers a job title different from the actual job

This slide addresses a common problem in the industry: job confusion.

- Many companies post job titles that don't accurately reflect the day-to-day responsibilities.
- For example,
    - a "**Data Scientist**" job might actually be a "**Data Analyst**" role, or
    - an "**ML Engineer**" role may require skills that are more aligned with a **data scientist**.
- It is always important to read the job description carefully.


Questions

Links

https://github.com/fcai-b/dv

This slide provides link to the GitHub repository related to this course, where you can find more resources and information.

- Please keep checking the repo link weekly, especially before the next lecture, to have the latest update.

References

1. https://www.udacity.com/blog/2018/01/4-types-data-science-jobs.html
    - 4 Types of Data Science Jobs
2. https://www.coursera.org/articles/data-analyst-vs-business-analyst
    - Data Analyst vs. Business Analyst
3. https://learn.365datascience.com/courses/intro-to-data-and-data-science
    - 365 Data Science - Introduction to Data and Data Science

And finally, here are the source links for the content of this presentation, including articles and blogs that dive deeper into the topics we've discussed today.