Business Tasks

Mittwoch, 2. Juli 2025 16:37

Business Tasks
+ real examples of business using data analytics
+ they all have issues to..
L> explore, questions to answer
L> or problems to solve
L> related to data

Business Task
L> question/problem data analysis answers for a business

Issue
L> topic or subject to investigate

Question
L> designed to discover information

Problem
L> obstacle/complication that needs to be worked out

Data-driven decision-making
L> using facts to guide business strategy

Datei auswählen Keine ausgewählt

Data and Fairness

Mittwoch, 2. Juli 2025 16:37

Fairness
L> ensuring that your analysis doesn't create/reinforce bias
L> example - gender differences not included in data
L> men succeed in company, women don't
L> but does not include surrounding conditions

Methods for fairness
L> teaming analysts with social scientists
L> provide insights on human bias and social context
L> self-reported data in separate system
to avoid potential for racial bias

Key takeaways

As a data professional, you will need to ensure you always consider fairness. This will allow you to avoid creating or reinforcing bias or accidentally drawing misleading conclusions. Using these best practices can help guide your analysis and make you a better data professional!

Best practice	Explanation	Example
Consider all of the available data	Part of your job as a data analyst is to determine what data is going to be useful for your analysis. Often there will be data that isn't relevant to what you're focusing on or doesn't seem to align with your expectations. But you can't just ignore it; it's critical to consider all of the available data so that your analysis reflects the truth and not just your own expectations.	A state's Department of Transportation is interested in measuring traffic patterns on holidays. At first, they only include metrics related to traffic volumes and the fact that the days are holidays. But the data team realizes they failed to consider how weather on these holidays might also affect traffic volumes. Considering this additional data helps them gain more complete insights.
Identify surrounding factors	As you'll learn throughout these courses, context is key for you and your stakeholders to understand the final conclusions of any analysis. Similar to considering all of the data, you also must understand surrounding factors that could influence the insights you're gaining.	A human resources department wants to better plan for employee vacation time in order to anticipate staffing needs. HR uses a list of national bank holidays as a key part of the data-gathering process. But they fail to consider important holidays that aren't on the bank calendar, which introduces bias against employees who celebrate them. It also gives HR less useful results because bank holidays may not necessarily apply to their actual employee population.
Include self-reported data	**Self-reporting** is a data collection technique where participants provide information about themselves. Self-reported data can be a great way to introduce fairness in your data collection process. People bring conscious and unconscious bias to their observations about the world, including about other people. Using self-reporting methods to collect data can help avoid these observer biases. Additionally, separating self-reported data from other data you collect provides important context to your conclusions!	A data analyst is working on a project for a brick-and-mortar retailer. Their goal is to learn more about their customer base. This data analyst knows they need to consider fairness when they collect data; they decide to create a survey so that customers can self-report information about themselves. By doing that, they avoid bias that might be introduced with other demographic data collection methods. For example, if they had sales associates report their observations about customers, they might introduce any unconscious bias the employees had to the data.
Use oversampling effectively	When collecting data about a population, it's important to be aware of the actual makeup of that population. Sometimes, oversampling can help you represent groups in that population that otherwise wouldn't be represented fairly. **Oversampling** is the process of increasing the sample size of nondominant groups in a population. This can help you better represent them and address imbalanced datasets.	A fitness company is releasing new digital content for users of their equipment. They are interested in designing content that appeals to different users, knowing that different people may interact with their equipment in different ways. For example, part of their user-base is age 70 or older. In order to represent these users, they oversample them in their data. That way, decisions they make about their fitness content will be more inclusive.
Think about fairness from beginning to end	To ensure that your analysis and final conclusions are fair, be sure to consider fairness from the earliest stages of a project to when you act on the data insights. This means that data collection, cleaning, processing, and analysis are all performed with fairness in mind.	A data team kicks off a project by including fairness measures in their data-collection process. These measures include oversampling their population and using self-reported data. However, they fail to inform stakeholders about these measures during the presentation. As a result, stakeholders leave with skewed understandings of the data. Learning from this experience, they add key information about fairness considerations to future stakeholder presentations.

Datei auswählen Keine ausgewählt

Explore your next job

Mittwoch, 2. Juli 2025 16:37

Introduction
+ every industry has specific data needs
+ even in one company there are different data analysts
+ finance, tech, telecom

Think about your interests first
+ when we can speak about our motivation
+ we can stand out in a great way

To name a few others that sound similar but may not be the same role:

- Business analyst—analyzes data to help businesses improve processes, products, or services
- Data analytics consultant—analyzes the systems and models for using data
- Data engineer—prepares and integrates data from different sources for analytical use
- Data scientist—uses expert skills in technology and social science to find trends through data analysis
- Data specialist—organizes or converts data for use in databases or software systems
- Operations analyst—analyzes data to assess the performance of business operations and workflows

Decoding the job description

	Data Analysts	Data Scientists	Data Specialists
Problem solving	Use existing tools and methods to solve problems with existing types of data	Invent new tools and models, ask open-ended questions, and collect new types of data	Use in-depth knowledge of databases as a tool to solve problems and manage data
Analysis	Analyze collected data to help stakeholders make better decisions	Analyze and interpret complex data to make business predictions	Organize large volumes of data for use in data analytics or business operations
Other relevant skills	• Database queries • Data visualization • Dashboards • Reports • Spreadsheets	• Advanced statistics • Machine learning • Deep learning • Data optimization • Programming	• Data manipulation • Information security • Data models • Scalability of data • Disaster recovery

Other industry-specific specialist positions that you might come across in your data analyst job search include:

- Marketing analyst—analyzes market conditions to assess the potential sales of products and services
- HR/payroll analyst—analyzes payroll data for inefficiencies and errors
- Financial analyst—analyzes financial status by collecting, monitoring, and reviewing data
- Risk analyst—analyzes financial documents, economic conditions, and client data to help companies determine the level of risk involved in making a particular business decision
- Healthcare analyst—analyzes medical data to improve the business aspect of hospitals and medical facilities

Datei auswählen Keine ausgewählt

Glossary

Donnerstag, 3. Juli 2025 16:33

Terms and definitions for Course 1, Module 4

Business task: The question or problem data analysis resolves for a business

Fairness: A quality of data analysis that does not create or reinforce bias

Oversampling: The process of increasing the sample size of nondominant groups in a population. This can help you better represent them and address imbalanced datasets

Self-reporting: A data collection technique where participants provide information about themselves

Datei auswählen Keine ausgewählt

>!Glossary!<

Donnerstag, 3. Juli 2025 16:35

Glossary

Data Analytics

Terms and definitions from Course 1

A

Analytical skills: Qualities and characteristics associated with using facts to solve problems

Analytical thinking: The process of identifying and defining a problem, then solving it by using data in an organized, step-by-step manner

Attribute: A characteristic or quality of data used to label a column in a table

B

Business task: The question or problem data analysis resolves for a business

C

Context: The condition in which something exists or happens

D

Data: A collection of facts

Data analysis: The collection, transformation, and organization of data in order to draw conclusions, make predictions, and drive informed decision-making

Data analyst: Someone who collects, transforms, and organizes data in order to draw conclusions, make predictions, and drive informed decision-making

Data analytics: The science of data

Data design: How information is organized

Data-driven decision-making: Using facts to guide business strategy

Data ecosystem: The various elements that interact with one another in order to produce, manage, store, organize, analyze, and share data

Data science: A field of study that uses raw data to create new ways of modeling and understanding the unknown

Data strategy: The management of the people, processes, and tools used in data analysis

Data visualization: The graphical representation of data

Database: A collection of data stored in a computer system

Dataset: A collection of data that can be manipulated or analyzed as one unit

E

F

Fairness: A quality of data analysis that does not create or reinforce bias

Formula: A set of instructions used to perform a calculation using the data in a spreadsheet

Function: A preset command that automatically performs a specified process or task using the data in a spreadsheet

G

Gap analysis: A method for examining and evaluating the current state of a process in order to identify opportunities for improvement in the future

H

I

J

Date

K

L

M

N

O

Oversampling: The process of increasing the sample size of nondominant groups in a population. This can help you better represent them and address imbalanced datasets

Observation: The attributes that describe a piece of data contained in a row of a table

P

Q

Query: A request for data or information from a database

Query language: A computer programming language used to communicate with a database

R

Root cause: The reason why a problem occurs

S

Self-reporting: A data collection technique where participants provide information about themselves

Stakeholders: People who invest time and resources into a project and are interested in its outcome

Structured Query Language: A computer programming language used to communicate with a database

Spreadsheet: A digital worksheet

SQL: (Refer to Structured Query Language)

T

Technical mindset: The ability to break things down into smaller steps or pieces and work with them in an orderly and logical way

U

V

Visualization: (Refer to data visualization)

W

X

Y

Z