Ask Questions to Make Data-Driven Decisions

Saturday, January 28, 2023 22:21

Intro to problem-solving and effective questioning

Wednesday, September 14, 2022 11:15

Structured thinking The process of recognizing the current problem or situation, organizing available information, revealing gaps and opportunities, and identifying the options

Case study

Business Anywhere Gaming Repair wants to advertize to get more customers. But which medium of advertisement should AGR choose? print, billboards, TV commercials, public transportation, podcasts, radio... ? >> must think about: target audience (specific people AGR is trying to reach), what is the budget and how much does each ad medium cost? The data analyst started by defining the problem that needed to be solved. >> looked at the whole situation in context, collaborated with stakeholders (owner, VP of coms, director of marketing and finance) >> Problem: Determine what advertising method is best for reaching AGR's target audience. >> collected data on different advertising methods and people with gaming consoles (target audience). >> cleaned data (remove outliers, transform data, create more complete information). >> analyze: Who's most likely to own a video gaming system? where are these people most likely to see advertisement? >> people ages 18-34, tv commercials and podcast >> recommendation: given budget and high cost of tv commercials, advertise on podcasts. >> summarize & share result using clear visuals

Solve problems with data

Wednesday, September 14, 2022 20:39

Problems are opportunities to put your skills to work and find creative and insightful solutions.

Common problem types:

1 making predictions
2 Categorizing things
3 Spotting something unusual
4 Identifying themes
5 discovering connections
6 finding patterns

making predictions: using data to make an informed decision about how things may be in the future

Categorizing things: assigning information to different groups or clusters based on common features

spotting something unusual: Identifying data that is different from the norm

identifying themes: grouping categorized information into broader concepts

discovering connections: finding similar challenges faced by different entities and combining data and insights to addres

finding patterns: using historical data to understand what happened in the past and is therefore likely to happen again

Craft effective questions

Thursday, September 15, 2022 21:26

Effective questions follow **SMART**:

Specific question are simple, significant and focus on a single topic or a few closely related ideas

Measurable questions can be quantified and assessed

Action-oriented encourage change

Relevant questions matter, are important, and have significance to the problem you're trying to solve

Time-bound questions specify the time to be studied

- Avoid leading questions, closed-ended (y/n) questions, questions that are vague and have little context.
- **Fairness** also means crafting questions that make sense to everyone · to be clear and have straightforward wording that does not imply/suggest a response
- Questions should be **open-ended**

Understand the power of data

Saturday, September 17, 2022 12:37

Data-inspired decision-making: Explores different data sources to find out what they have in common

Algorithm: A process or set of rules to be followed for a specific task

Today, we create so much data that scientists estimate that 90% of the world's data has been created in just the last few years. But just having tons of data isn't enough; whe have to do something meaningful with it. Individual datapoints become more useful when collected and structured but we need to interpret data to turn it into information. There are limitations to data analytics: sometimes we don't have access to the right data, data is measured differently across programs.

Some wrong conclusions from data: New coke, Mars orbiter loss

Quantitative data: Specific and objective measures of numerical facts · what? how many? how often? Charts, Graphs

Qualitative data: Subjective or explanatory measures of qualities and characteristics · why? Why are the numbers the way they are?
→ adds context

Quantitative data tools: Structured interviews
Surveys
Polls

Qualitative data tools: Focus groups
Social media text analysis
In-person interviews

Follow the evidence

Saturday, September 17, 2022 13:54

Report: Static collection of data given to stakeholders periodically.

Dashboard: monitors live, incoming data. Strategic, Operational, or analytical

Reports	⊕	⊖
	high-level, historical data → finance firm's monthly sales can be designed and sent out periodically (weekly or monthly), organized and easy to reference information, quick to design and easy to use, reflect cleaned and sorted data	need continual maintenance, less visually appealing, static

Dashboards:	⊕	⊖
	dynamic, automatic and interactive (can play with filters) have long-term value because they're dynamic, more stakeholder access, can be a time saver instead of pulling reports repeatedly, low maintenance visually appealing	Labor-intensive design, can be confusing, needs a lot of maintenance to fix if it breaks, potentially uncleaned data

Pivot table: A data summarization tool that is used in data processing. Pivot tables are used to summarize, sort, reorganize, group, count, total or average data stored in a data base. Allows users to transform colums into rows and rows into columns.

Metric: Single, quantifiable type of data that can be used for measurement. Data starts as a collection of raw facts, until we organize them into individual metrics that represent a single type of data. ie revenue by sales person, ROI (profit & cost of investment), customer retention rate

Metric Goal: a measurable goal set by a company and evaluated using metrics

Mathematical thinking

Saturday, September 17, 2022 15:14

Mathematical thinking : looking at a problem and logically breaking it down step-by-step to see the relationships of patterns in the data, and using that to analyze the problem

Small data : Specific
short time period
day-to-day decisions
> spreadsheets

Big data : large and less specific
long time period
big decisions
> SQL
4 V's > Volume – amount of data
Variety – different kinds of data
Velocity – how fast can the data be processed
Veracity – quality and reliability of the data

Working with spreadsheets

Saturday, September 17, 2022 15:51

Some common math functions : Sum
average
Count
Min
max

Spreadsheet tasks:

- organize your data
  - pivot table
    - sort and filter
- Calculate your data
  - formulas
  - functions

Operator: A symbol that names the type of operation or calculation to be performed.

+ addition
− subtraction
* multiplication
/ division

Cell reference: A cell or a range of cells in a worksheet that can be used in a formula
letter of column and number of the row where the data is

Absolute referencing: marked by a "$" sign, A10 has absolute referencing for both the column and row value, switch between absolute and relative referencing using the F4 key

use the **fill handle** in the bottom right of a cell to auto-fill the sequence

Combining with functions: COUNTIF() is a formula and a function, the function runs based on criteria set by the formula
COUNT is the formula ; it will be executed IF the conditions

you create are true e.g. COUNTIF(A1:A16, "7") to count the cells that contained the number 7

Errors and fixes

#DIV/0! A formula is trying to divide a value in a cell by 0 or by an empty cell

fix: IFERROR(B4/A4, "Not applicable")

#ERROR (google sheets only) A formula can't be interpreted as input (aka a parsing error)

fix: most likely need to fix syntax

#N/A Data in formula can't be found by spreadsheet often happens with VLOOKUP()

fix: most likely a typo

#NAME? A formula or function name isn't understood

fix: also typo

#NUM! A formula or function calculation can't be performed as specified fix: could be data entry error

#VALUE! A general error that could indicate a problem with a formula or referenced cells

#REF! A formula is referencing a cell that is no longer valid or has been deleted

Functions and structured thinking

Monday, September 19, 2022 20:39

Function: A preset command that automatically performs a specific process or task using the data

- SUM(A1:A10)
- MIN()
- AVERAGE()
- MAX()

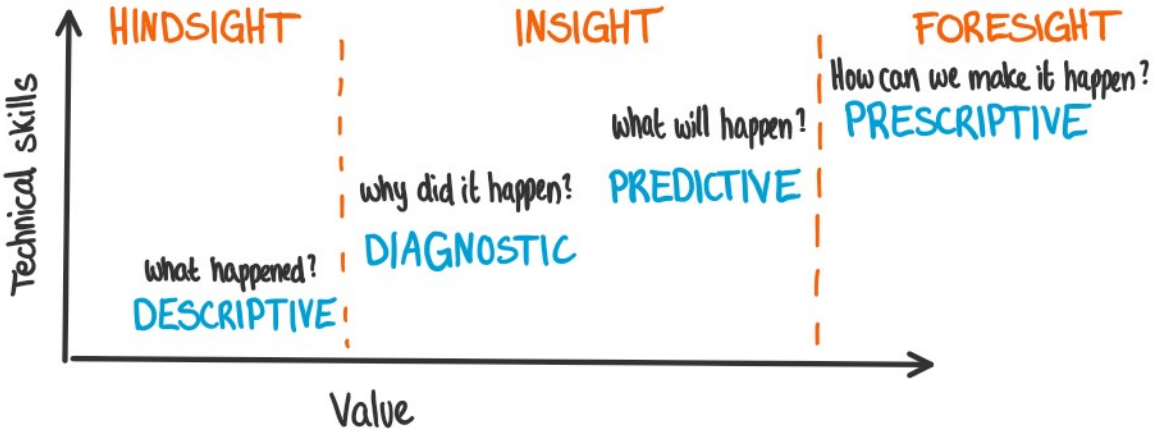
Problem domain: The specific area of analysis that encompasses every activity affecting or affected by the problem

Structured thinking: The process of recognizing the current problem or situation, organizing available information, revealing gaps and opportunities, and identifying the options

Scope of work (SOW): An agreed-upon outline of the work you're going to perform on a project

- deliverables
- milestones
- timeline
- reports

Data needs **context** & staying **objective**



To really understand what the data is about, think through

who, what, where, how, why.

Who collected the data? What is it about? When was the data collected? Where and how was the data collected? Sometimes data is made up or collected to serve an agenda.

Balance team and stakeholder needs

Tuesday, September 20, 2022 20:18

3 common stakeholder groups:

- **Executive team**: they set goals, develop strategy, and make sure that strategy is executed effectively, VP, chief marketing officer, looking for headline news about the project
- **Customer-facing team**: anyone in an organization who has some level of interaction with customers and potential customers, they compile information, set expectations, and communicate customer feedback to other parts of the internal organization
- **Data Science team**: other data analysts, data scientists, data engineers, collaborating to find new angles of the data to explore

Working effectively with stakeholders:

- discuss goals
- feel empowered to say "no"
- plan for the unexpected: make a list of potential roadblocks
- know your project: keep track of your discussions, understand why you are doing an analysis
- start with words and visuals to get on the same page more quickly
- Communicate often: share notes about milestones, setbacks, and changes, use a change-log

Focus on what matters

1. Who are the primary and secondary stakeholders?
2. Who is managing the data?
3. Where can you go for help?

Communication is key

Tuesday, September 20, 2022 20:57

Before you communicate, think about

1. Who your audience is
2. What they already know
3. What they need to know
4. How can you communicate that effectively to them

Tips for effective communication

- Learn as you go and ask questions
- Practice good writing habits
- Read your emails out loud
- Answer in a timely manner

Balancing expectations

- set a reasonable and realistic timeline
- flag problems early for stakeholders
- set realistic expectations at every stage of the project

There is a tradeoff between **speed and accuracy**

before jumping to conclusions on a problem to get a quick result, consider:

- reframe question
- problems
- challenges
- solutions
- timelines

Limitations of data:

- **incomplete (or nonexistent) data**

  may not be able to reach a conclusion, may have to pivot to a different problem, might be able to find alternate source of data

may not be able to reach a conclusion, may have to pivot to a different problem, might be able to find alternate source of data

- misaligned data

  different datasets might use different business rules, ie. measure and define things in different ways, the same metric could be calculated in different ways

  > establish early on how to measure, standardize, don't merge those datasets

- dirty data

  contains errors, requires good data cleaning effort

- Tell a clear story
  - compare the same types of data
  - visualize with care
  - leave out needless graphs
  - test for statistical significance
  - pay attention to sample size
- Be the judge, make sound judgements, know the limitations of your data

Amazing teamwork

Wednesday, September 21, 2022 17:37

Meetings best practices

Do

- come prepared
  - bring what you need
  - read the meeting agenda
  - prepare notes and presentations
  - be ready to answer questions
- be on time
- pay attention, take notes, follow up
- ask questions
- the meeting should focus on making a clear decision and should include the person who ultimately makes that decision
- schedule meetings as soon as a decision needs to be made - don't leave it until next week
- keep # of people at the meeting under 10
- respect your teammembers' time
- keep everyone involved
- send agendas ahead of time

From conflict to collaboration

- there can be many reasons for conflict
- mismatched expectations and miscommunications are some of the most common
- try to be objective and stay focused on the team's goals

Possible to shift a situation from problematic to productive:

- reframe the problem

- start a conversation
- take time cool off
- understand the context