Course 2:

Ask Questions to Make Data-Driven Decisions

Week 1: Introduction to problem-solving and effective questioning

08-01-2022

Introduction

In the ask step, we define the problem we're solving and make sure that we fully understand stakeholder expectations. This helps in keeping focused on the actual problem, which leads to more successful outcomes.

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

Data in Action

Ask:

- Define the problem you're trying to solve
- Make sure you fully understand to stakeholder's expectations
- Focus on the actual problem and avoid any distractions
- Collaborate with stakeholders and keep an open line of communication
- Take a step back and see the whole situation in context

Prepare:

- What metrics to measure
- Locate data in your database
- Create security measure to protect the data

Process:

- Using spreadsheet functions to find incorrectly entered data
- Using SQL functions to check for extra spaces
- Removing repeated entries
- Checking as much as possible for bias in the data

Analyze:

- Perform calculations
- Combine data from multiple sources
- Create tables with your results

Share:

- Make better decisions
- Make more informed decisions
- Lead to stronger outcomes
- Successfully communicate your findings

Act:

- Providing stakeholders with recommendations based on our findings so they can make data-driven decisions

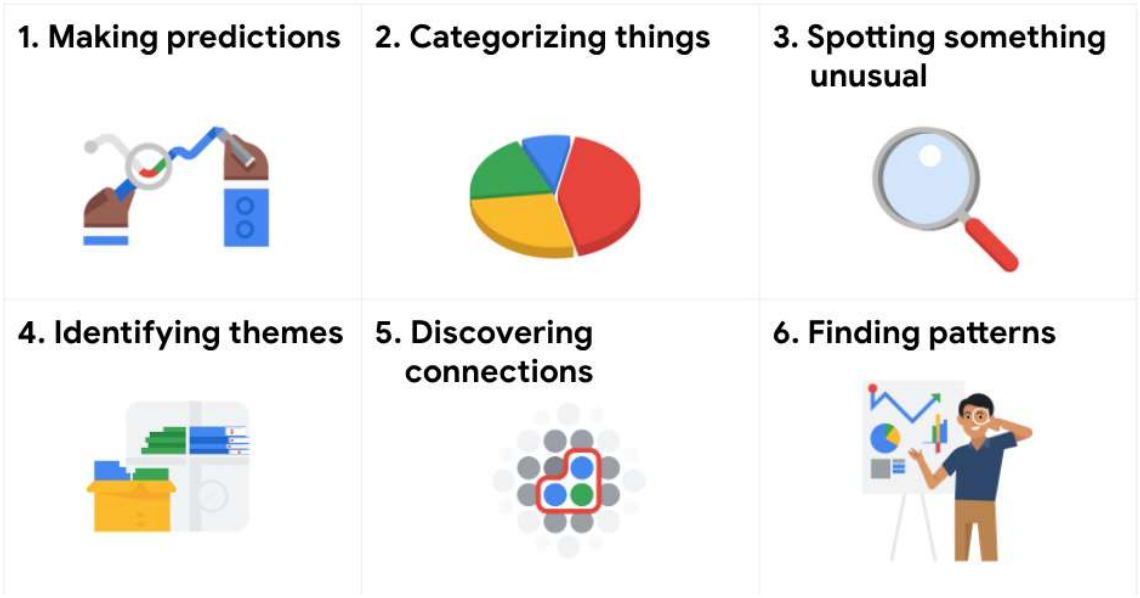
These steps can help to break the data analysis process into smaller, manageable parts, which is called **structured thinking.** This process involves four basic activities:

1. Recognizing the current problem or situation
2. Organizing available information
3. Revealing gaps and opportunities
4. Identifying your options

Common problem types

- Making predictions
  - Involves 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
  - Identify data that is different from the norm
- Identifying themes
  - Takes categorization a step further by grouping information into broader concepts
- Discovering connections
  - To find similar challenges faced by different entities, and then combine data and insights to address them
- Finding patterns

- Find patterns by using historical data to understand what happened in the past and is therefore likely to happen again



SMART Questions

Things to avoid when asking questions

Leading question: have a particular response.

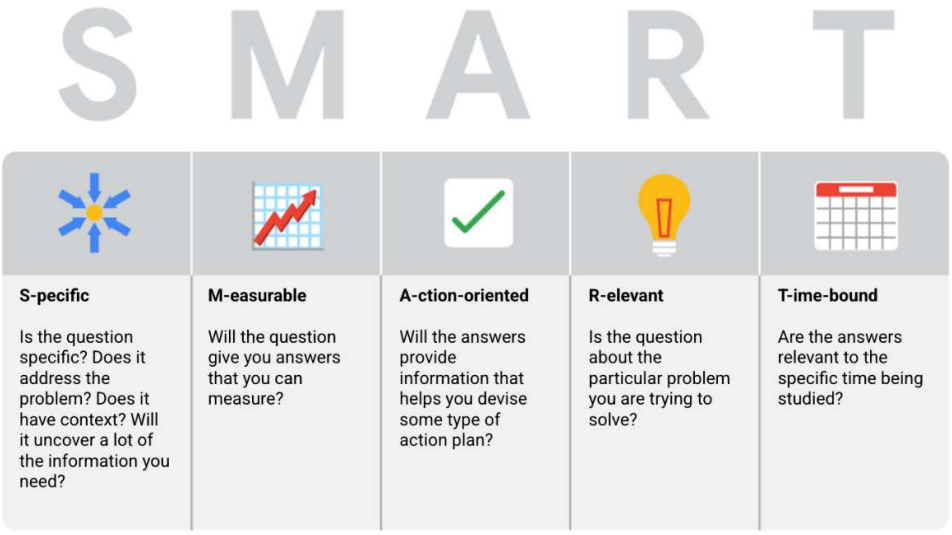
Close-ended question: ask for a one-word or brief response only.

Vague question: aren't specific or don't provide context.

Specific
Measurable
Action-oriented
Relevant
Time-bound

- **Specific** questions are simple, significant, and focused on a single topic or a few closely related ideas.
- **Measurable** questions can be quantified and assessed.
- **Action-oriented** questions 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.

There’s something that’s very important to keep in mind when crafting questions, fairness. **Fairness** means ensuring that your questions don’t create or reinforce bias.



Week 2: Data and Decisions

13-01-2022

How data empowers decisions

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.

Qualitative and Quantitative data

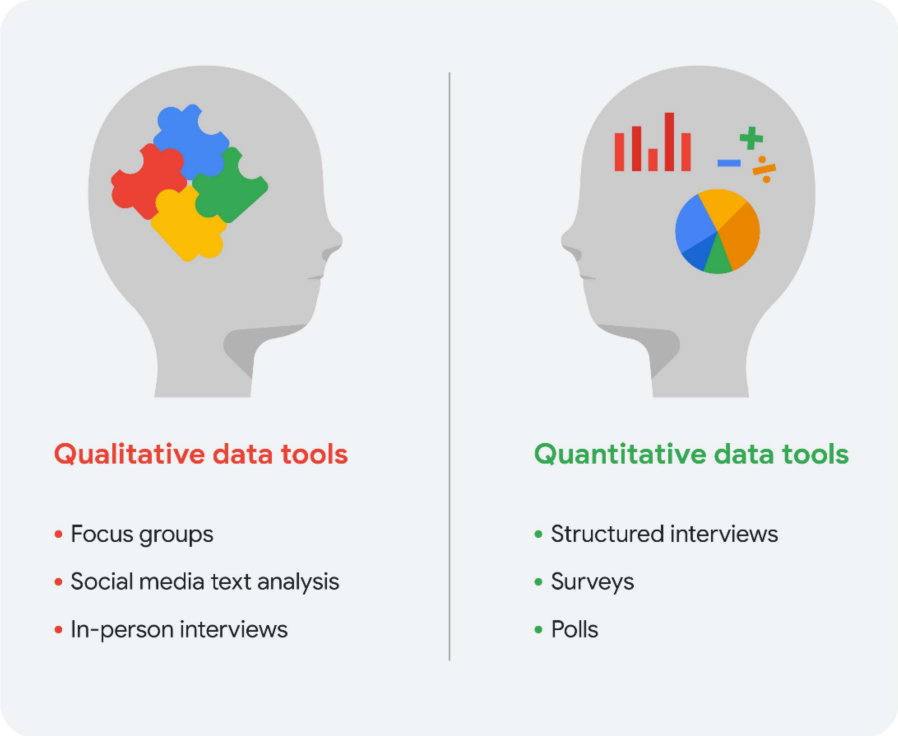
Quantitative data: specific and objective measures of numerical facts.

The what? How many? And How Often? About a problem.

Qualitative data: Subjective or explanatory measures of qualities and characteristics.

The Why?

Helps to add context to a problem.



14-01-2022

Sharing your findings

- Reports
    - A static collection of data given to stakeholders periodically
    - Pros
        - High-level historical data
        - Easy to design
        - Pre-cleaned and sorted data
    - Cons
        - Continual maintenance
        - Less visually appealing
        - static
- Dashboards
    - Monitors live, incoming data
    - Pros
        - Dynamic, automatic and interactive
        - More stakeholders access
        - Low maintenance
    - Cons
        - Labor-intensive design

- Can be confusing
- 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 database.

Data versus metrics

Metric: Single, quantifiable type of data that can be used for measurement.

Mathematical thinking

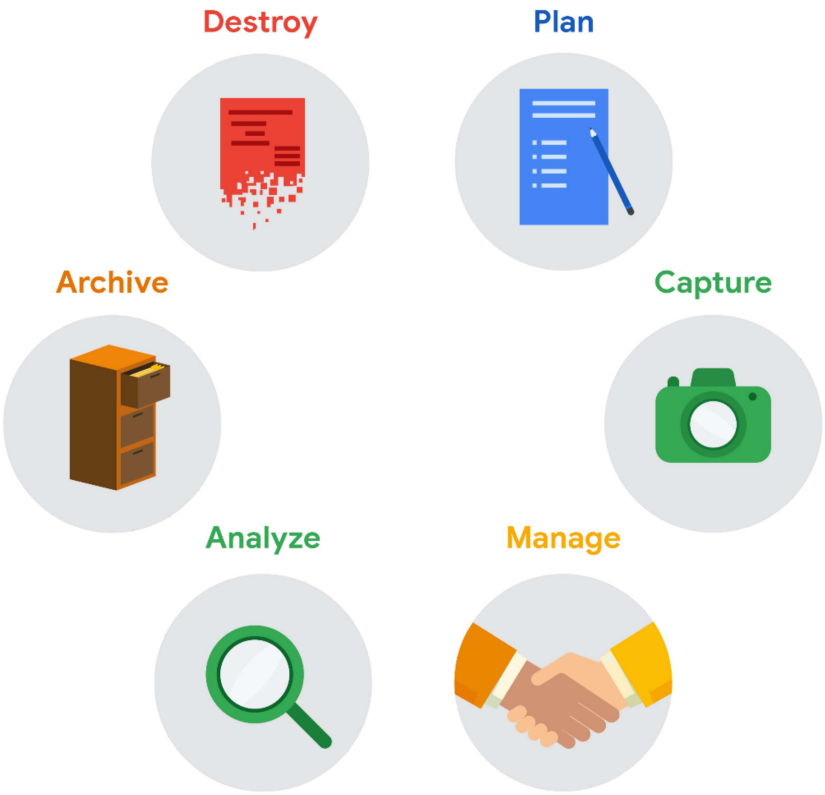
It means looking at a problem and logically breaking it down step-by-step, so its easier to see the relationship of patterns in our data, and use that to analyze our problem.

Week 3: More Spreadsheet Basics

Spreadsheet tasks:

- Organize your data
  - Pivot table
    - Sort and filter
- Calculate your data
  - Formulas
  - Functions

Spreadsheets and the data life cycle



- **Plan**: for the users who will work within a spreadsheet by developing organizational standards. By setting standards we improve communication, ensure consistency, and help people be more efficient with their time.
- **Capture**: data by the source by connecting spreadsheets to other data sources, such as an online survey application or a database. This will be updated in the spreadsheet thus keeping the information always as current and accurate as possible.

- **Manage**: different kinds of data with a spreadsheet. This involves storing, organizing, filtering, and updating information. Spreadsheets also let us decide who can access the data, how the information is shared, and how to keep our data safe and secure.
- **Analyze**: data in spreadsheet to help make better decisions. Some of the most common analysis tools include formulas and pivot tables.
- **Archive**: any spreadsheet not used often, but might be needed for future reference. Useful to store historical data before it gets updated.
- **Destroy**: when we are certain that it will never be needed again, or we have better backup copies, or for legal or security reasons.

https://support.google.com/docs/answer/181110

https://support.microsoft.com/en-us/office/keyboard-shortcuts-in-excel-1798d9d5-842a-42b8-9c99-9b7213f0040f

18-01-2022

Step-by-Step in Spreadsheets

Title: Short, Clear, and Relevant according to data it holds

Storage: Keep spreadsheets in relevant folders at relevant position

Formulas for success

Formulas: a set of instructions that perform a specific calculation.

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

Formulas start with an equal sign in spreadsheets.

Cell reference: a cell or range of cells in a worksheet that can be used in a formula. (A1 , B4, ..)

Range: A collection of two or more cells.

Absolute reference: do not change when a formula is copy and pasted to a different cell. (A1 - both column and row are absolute, $B3 - only column is absolute, row no will change,..)

Spreadsheet errors and fixes

- IFERROR(value, [message])

#ERROR! a formula can’t be interpreted as input (also known as parsing error) (in Google Sheets only)

#N/A data in a formula can’t be found by the spreadsheet

#NAME? a formula of function name isn’t understood

#NUM! a formula of function calculation can’t be performed as specified

#VALUE! a general error that 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 101

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

A ‘:’ between cell references implies that you are using a range.

Save Time with Structured Thinking

Problem Domain: The specific area of analysis that encompases 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
- timeline
- milestones
- reports

Staying Objective

Context:The condition in which something exists or happens

To really understand what the data is about, you have to think through:

- who
- what
- where
- when
- how
- why

Context is important as it helps to sift through huge amounts of disorganized data and turn it into something meaningful. Context can turn raw data into meaningful information.

Week 4: Communicating with your Team

19-01-2022

Balancing needs and expectations across your team

Stakeholders: People that have invested time, interest, and resources into the projects you'll be working on as a data analyst.

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

Clear Communication is key

Before you communicate, think about

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

Tips for effective communication

1. Learn as you go and ask questions
2. Practice good writing practices
3. Answer in timely manner

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

Amazing Teamwork

Do:

- come prepared
- be on time
- pay attention
- ask questions

Don’t:

- show up unprepared
- arrive late
- be distracted
- dominate the conversation
- talk over others
- distract people with unfocused discussion