Executive summary

Hayes Baked Goods

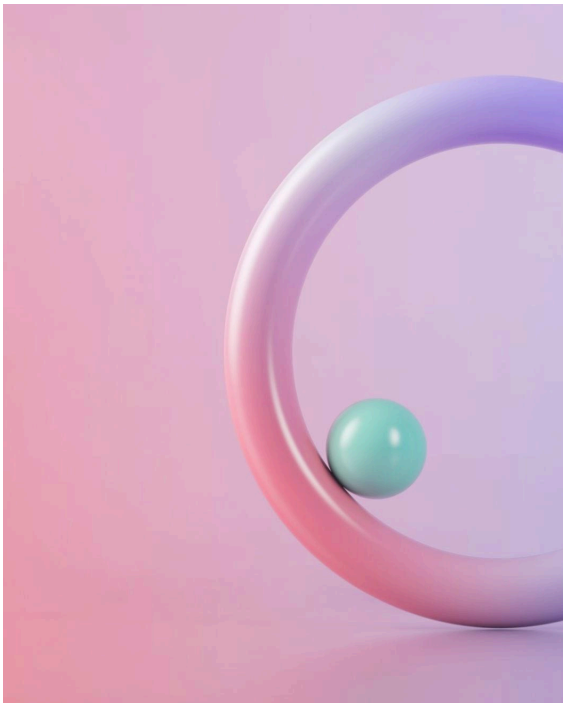
Data Audit & Roadmap

01. Overview

Client Profile and Goals

Hayes Baked Goods is a small, home-based bakery run by one full-time owner/manager, supported by a part-time baker and a contract bookkeeper. The business specializes in freshly baked cookies offered for local pickup and limited delivery.

Hayes Baked Goods wants to move from intuition-based decision-making to data-backed insights around pricing, product performance, profitability, and ingredient planning. This audit will review current systems and outline a simple, unified reporting approach.



02. Key findings

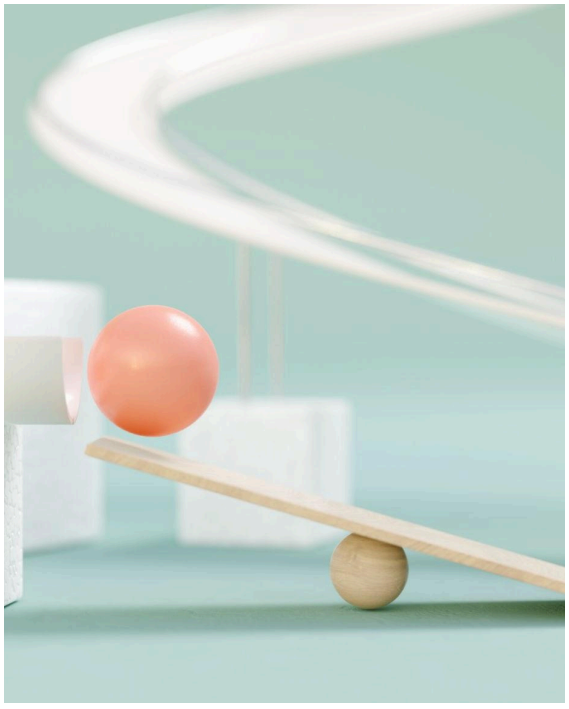
Identify the project's most noteworthy findings

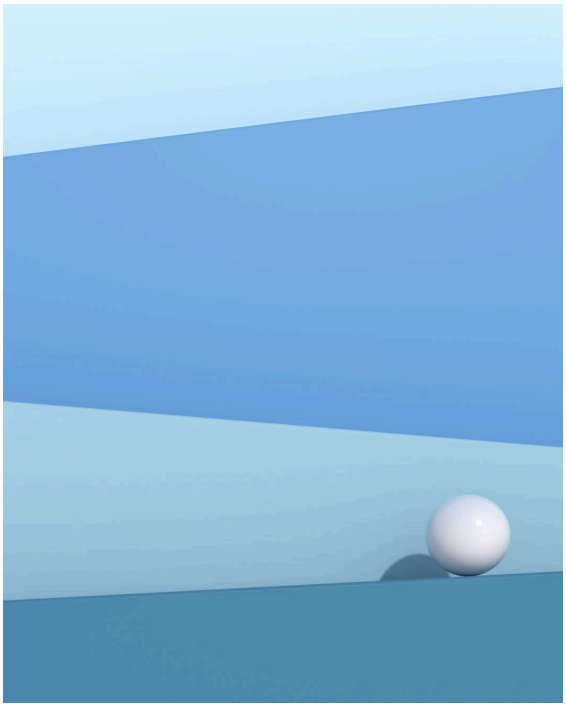
a. Key finding 1
b. Key finding 2
c. Key finding 3

03. Solutions

Document solutions that were developed

a. Solution 1
b. Solution 2
c. Solution 3





04. Impact

Summarize the project's greatest areas of impact

a. Impact 1
b. Impact 2
c. Impact 3

Introduction & recommendations

Background

01. Operations

Current offerings & menu

Hayes Baked Goods offers three kinds of cookies, each available by the dozen: Chocolate Chip Cookies, Emmy's Sugar Cookies, and Nutella Marbled Chocolate Chip Cookies. All three are priced at $28/dozen. Customers can add on Nutella stuffing or crunchy peanut butter stuffing for $4 per dozen, and/or sea salt topping for free.

02. Current data setup

Platforms and systems in use

a. **Shopify**: Reliable sales data but no visibility into SKU-level profit.
b. **Google Sheets**: Ingredient and recipe costs tracked manually; not tied to sales.
c. **Instagram**: Good marketing engagement, but no link to sales performance.

03. Business Goals

- Set data-informed pricing that reflects ingredient costs
- Understand best-selling and most profitable cookies
- Identify low-margin or underperforming SKUs
- Forecast weekly/monthly ingredient needs
- Plan better for holiday boxes and seasonal flavors

- Track customer behavior (AOV, repeat buyers)

Priority Data Questions

- **Sales**: Which cookies and pack sizes sell best? How does demand shift seasonally?
- **Profitability**: What does each SKU cost to make? Which items have the strongest/weakest margins?
- **Customers**: Who returns? Do deluxe items drive higher-value orders?
- **Operations**: How much of each ingredient is needed? When are peak order times?

Recommendations

Our strategy for answering questions, providing clarity, and setting goals

1. **Consolidate Data** from Shopify, Sheets, and QuickBooks; clean SKU and ingredient lists.
2. **Build a Simple Data Model** linking recipes → ingredients → SKUs → sales.
3. **Create Core Reports** for SKU profitability, sales trends, customer insights, and ingredient forecasting.
4. **Recommend KPIs** to define and track business success.
5. **Deliver Pricing & Product Recommendations** based on current costs and margins.

Data Dictionary

Data Dictionary

Source Tables

Orders (Shopify)

Transactional sales data representing customer orders placed through Shopify. Data is refreshed regularly via CSV export.

Grain: One row per order line (SKU per order)

Relationships:

Orders.sku → Products.cookie_sku (required)
Orders.add_on_sku → Products.cookie_sku (optional)

Column	Storage Type	Semantic Type	Notes
order_id	**int64**	Primary key (Order)	Unique per order; sourced from Shopify
date	**object**	Order paid date	ISO YYYY-MM-DD string
sku	**object**	Product SKU	Alphanumeric product identifier; consistent across systems
quantity	**int64**	Quantity ordered	Positive integer
add_on_sku	**object**	Optional add-on SKU	Nullable; present only when add-on is purchased
total_price	**float64**	Monetary (USD)	Order line total in USD; non-null; includes add-ons if applicable; stored as float; decimals reflect source system precision

Products (Shopify)

Master list of sellable products maintained in Shopify. Updated manually as product offerings change.

Grain: One row per sellable product SKU

Relationships:

Referenced by Orders.sku
Referenced by Orders.add_on_sku
Referenced by Recipes.sku

Column	Storage Type	Semantic Type	Notes
cookie_sku	**object**	Product SKU	Alphanumeric product identifier; consistent across systems
product_name	**object**	Title of the product	Descriptive language
category	**object**	Product category (enumeration)	Allowed values: “Dozen Cookies”, “Add-on”
price	**float64**	Monetary (USD)	Retail price in USD; manually maintained in Shopify; stored as float; decimals reflect source system precision

Recipes (Google Sheets)

Bill-of-materials style table defining ingredient requirements per product SKU.

Grain: One row per ingredient per product SKU

Relationships:

Recipes.sku → Products.cookie_sku (required)
Recipes.ingredient → Ingredients.ingredient (required)

Column	Storage Type	Semantic Type	Notes
sku	**object**	Product SKU	Alphanumeric product identifier; consistent across systems
ingredient	**object**	Name of ingredient	Matches Ingredients.ingredient exactly
quantity_unit	**object**	Measurement unit	Text unit used consistently with Ingredients.unit
quantity	**float64**	Quantity of unit needed	Quantity required per product unit (e.g., per one dozen cookies); non-null

Ingredients (Google Sheets)

Ingredient cost, supplier, and unit-conversion reference table used for cost and inventory calculations.

Grain: One row per ingredient

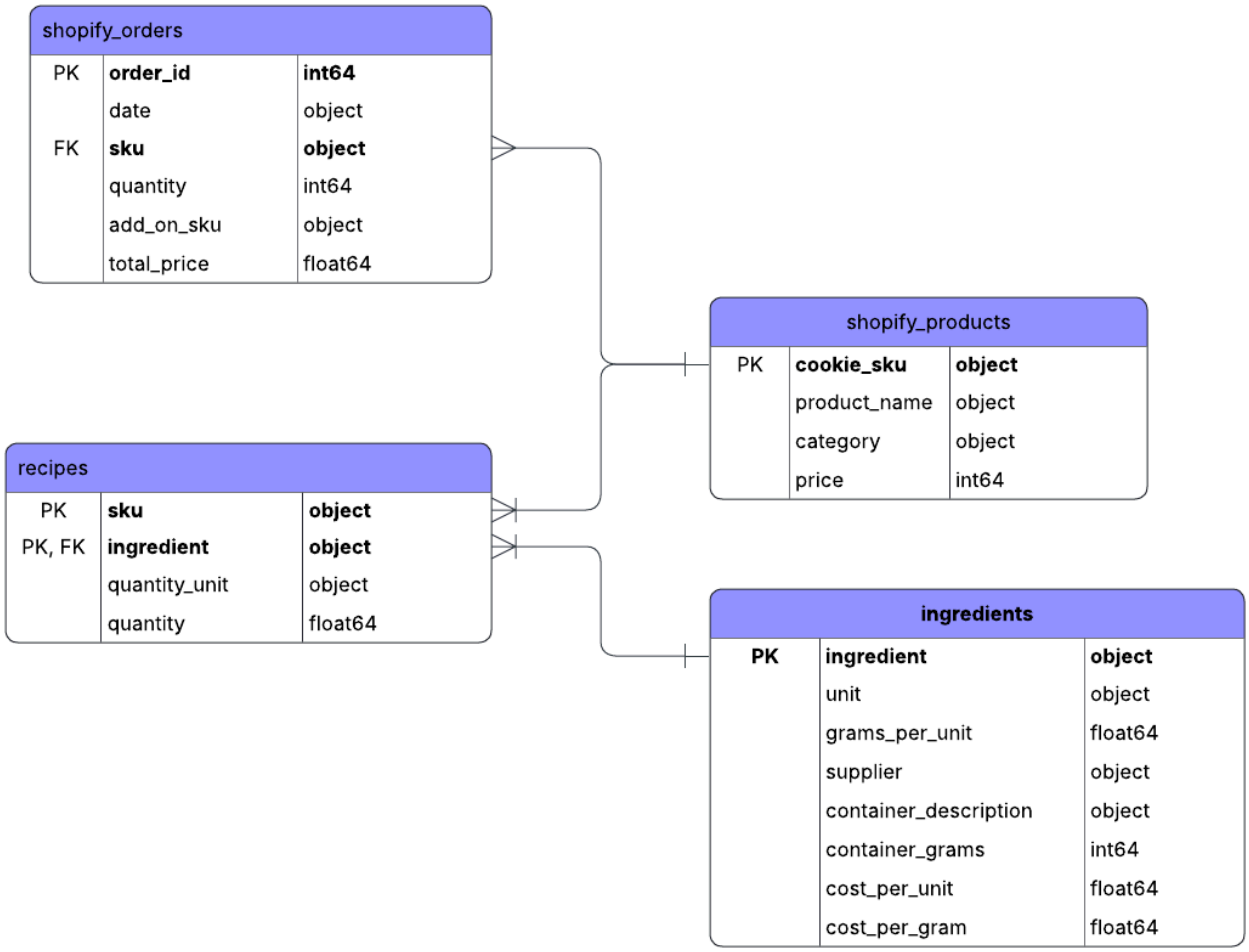
Relationships:

Referenced by Recipes.ingredient

Column	Storage Type	Semantic Type	Notes
ingredient	**object**	Name of ingredient	Descriptive language
unit	**object**	Measurement unit	Canonical unit for recipe conversions
grams_per_unit	**float64**	Number of grams per unit used in recipes	Positive number, constant per ingredient

Column	Storage Type	Semantic Type	Notes
supplier	**object**	Supplier identifier	Single-token supplier name, no spaces
container_description	**object**	Packaging description	Human-readable container size (e.g., “1 lb bag”); not used for calculations
container_grams	**int64**	Grams per purchasable container	Positive integer
cost_per_unit	**float64**	Cost per unit used in recipes	Stored as float; decimals reflect source system precision
cost_per_gram	**float64**	Cost per gram of ingredient	Derived from container_grams and container cost; stored as float; decimals reflect source system precision

Relationships & Constraints



Analytical Model

Fact Table - fact_orders

Business Purpose: Represents individual product-level sales transactions used to analyze revenue, volume, and cost.

Grain: One row per order line per product SKU per order date.

Source Tables

- Orders (Shopify)
- Products (Shopify)

- Recipes (Google Sheets)
- Ingredients (Google Sheets)

Measures

- quantity (sum)
- total_price (sum)
- ingredient_cost (derived)
- gross_margin (derived)

Foreign Keys

- product_key → dim_product
- date_key → dim_date
- add_on_product_key → dim_product (nullable)

Notes

Add-on SKUs are modeled as a separate product relationship.
Monetary fields are stored in USD.
Cost and margin calculations are based on recipe ingredient quantities.

Dimension Table: dim_product

Business Purpose: Provides descriptive and categorical context for sellable products.

Grain: One row per product SKU.

Source Tables

- Products (Shopify)

Attributes

- product_key (surrogate key)
- product_sku
- product_name
- category
- price

Notes

Acts as a conformed dimension across analytical tables.
Product pricing reflects the current configuration in Shopify.

Dimension Table: dim_date

Business Purpose: Enables time-based analysis of sales and cost metrics.

Grain: One row per calendar date.

Source Tables

- Derived from Orders.date

Attributes

- date_key (YYYYMMDD)
- date
- day_of_week
- month
- year

Notes

Date values reflect the order paid date as recorded in Shopify.
Used consistently across all fact tables.

Dimension Table: dim_ingredient

Business Purpose: Provides descriptive and cost-related context for ingredients used in product recipes.

Grain: One row per ingredient.

Source Tables

- Ingredients (Google Sheets)

Attributes

- ingredient_key (surrogate key)
- ingredient
- unit
- supplier
- grams_per_unit
- cost_per_unit
- cost_per_gram

Notes

Used primarily for cost analysis and ingredient-level rollups.
Supports auditing of ingredient cost calculations.

Bridge Table: bridge_product_ingredient

Business Purpose: Maps products to their constituent ingredients and required quantities to support cost calculations.

Grain: One row per ingredient per product SKU.

Source Tables

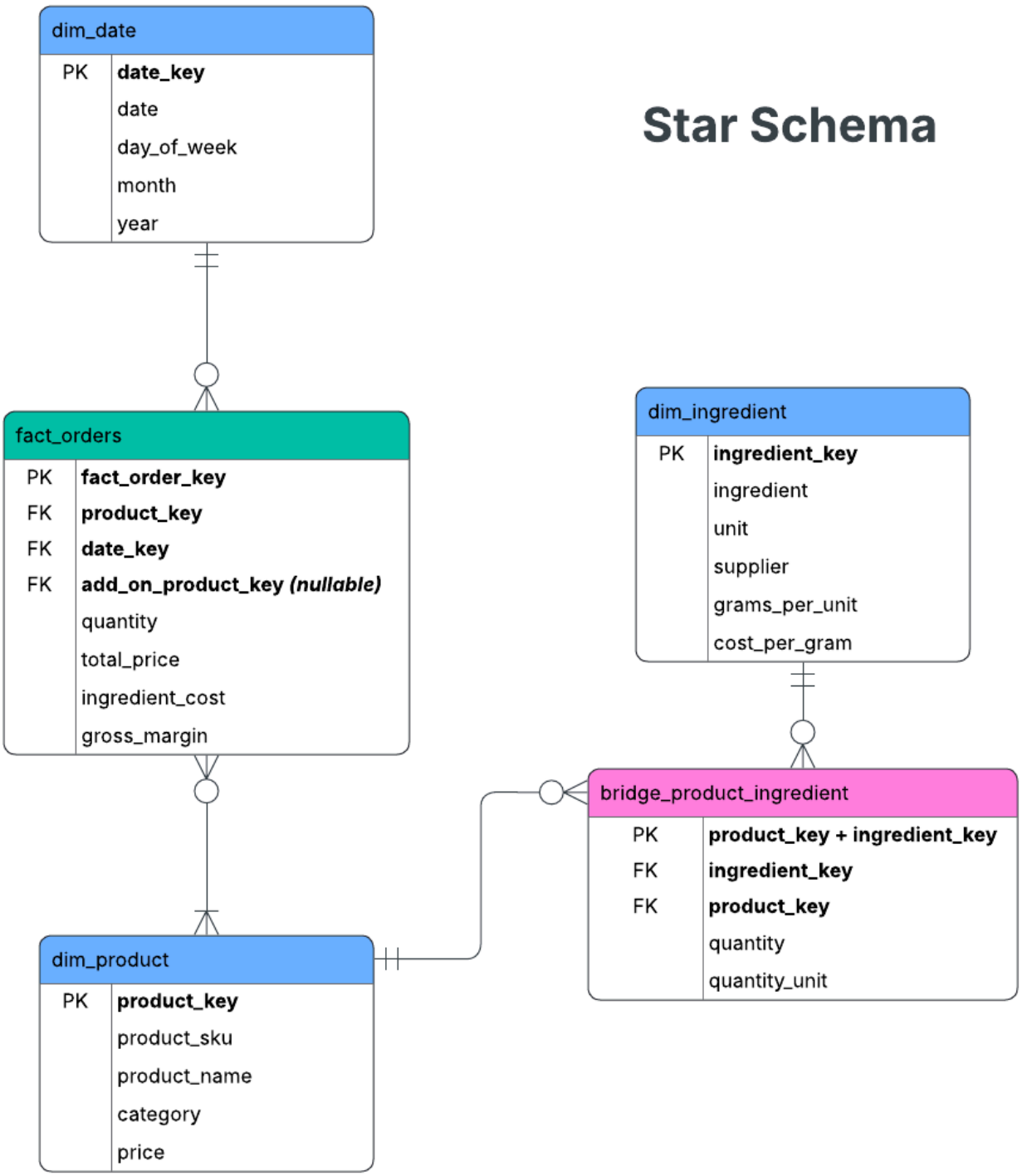
- Recipes (Google Sheets)
- Ingredients (Google Sheets)

Attributes

- product_key
- ingredient_key
- quantity
- quantity_unit

Notes

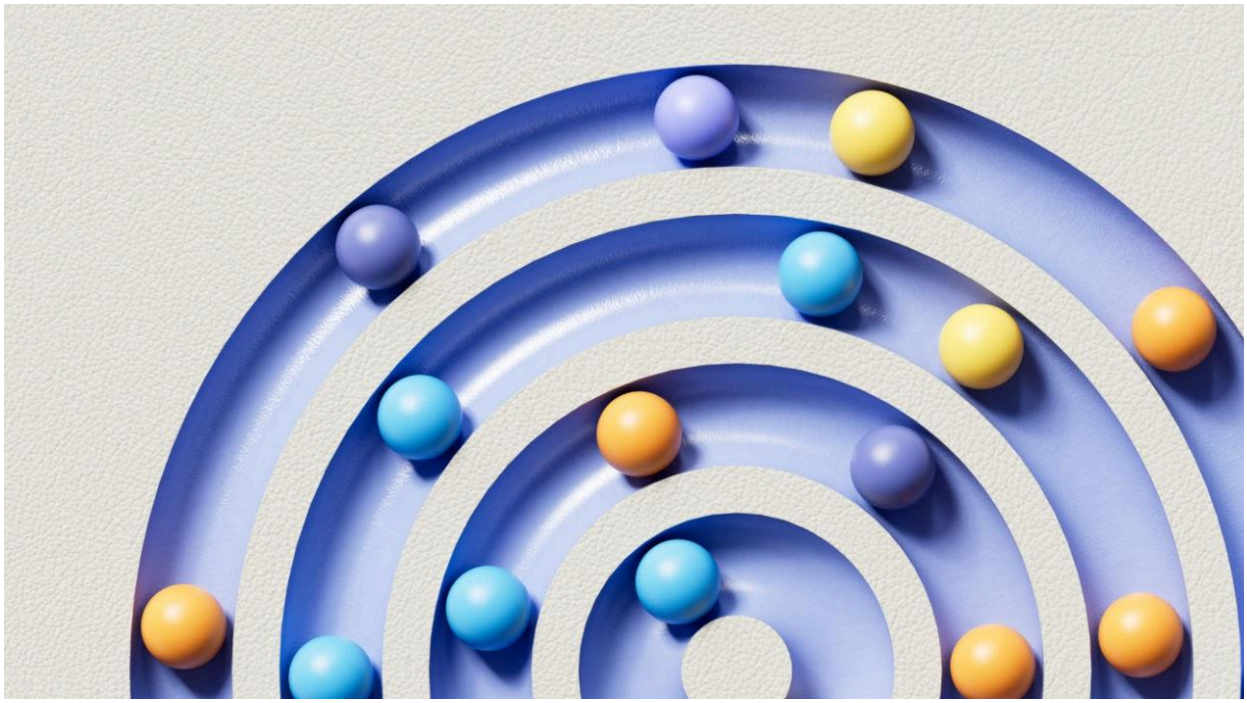
Serves as a bridge between dim_product and dim_ingredient.
Enables allocation of ingredient costs to sold products.
Not treated as a fact or dimension table.



// to do next: ETL

Conclusion

Conclusion



01. Summary of findings

Summarize the main findings or conclusions of the case study

This should be a more comprehensive summary, taking the whole project into consideration and incorporating key findings from the entire document. Use multiple paragraphs if necessary.

02. Implications of the study

Describe the impact of the completed study

Focus on real-world impact, concrete outcomes and how they addressed the project's goals. Evaluate outcomes using both objective data and subjective feedback. Draw well-supported

conclusions based on analysis and interpretation of relevant information.

> **“Feature a quote from the case study. It should be a key takeaway or learning that you want audiences to remember.”**

If your implications explanation is lengthy, you can break it up with a quote to help communicate a key takeaway.

References

References

Check what reference style your organization typically uses. Below is an example of the format for citing references called APA. Other common formats include MLA and Chicago Style.

Examples:

Lastname, A. (Year). *Book Title*. Publisher. DOI

Lastname, A., Lastname, B., & Lastname, C. (Year). Article title. *Journal Title, Volume#*(Issue#), Page(s)#. DOI

Lastname, A. (Year, Month Day). *Title*. Site Name. URL

Supplementary materials

FILE	NOTES
File	Files might include surveys, questionnaires, or data tables
File	Add file notes
File	Add file notes