Audit Your Data for Missing Values and Duplicates - WindSurf

What the app should do

- A small web app called Excel Data Auditor

- Runs fully **client-side** (no backend)

- Built using **Next.js only** (no extra UI libraries)

File Upload Experience

- A single page with:

  - "Choose File" button.

  - Accepts **.xslx** files

- Shows the **raw table** immediately after upload

Cleaning Features (toggle each on or off)

- Trim extra whitespace

- Case-Insensitive comparison

- Treat empty space/missing values as missing

- Show only rows with issues (Missing values (empty or whitespace-only cells))

- Make differences visually clear

Mini Project: Excel Data Auditor

A professional mini project built to improve data quality and make spreadsheet review easier.

Key Capabilities

- Identifies and highlights **missing values** across all columns

- Detects **duplicate rows** using either selected key columns or full-row matching

- Allows configurable rules such as trimming whitespace, case-insensitive comparison, and treating blank/space-only cells as missing

- Displays each sheet in a clean, readable table view

- Provides a compact summary with total rows, missing-value percentages, and duplicate counts

- Offers filters to view **only rows containing issues**

- Supports exporting a cleaned workbook or issue-only CSV

Tech Stack Used

- **Next.js** for the client-side UI and fast, lightweight rendering

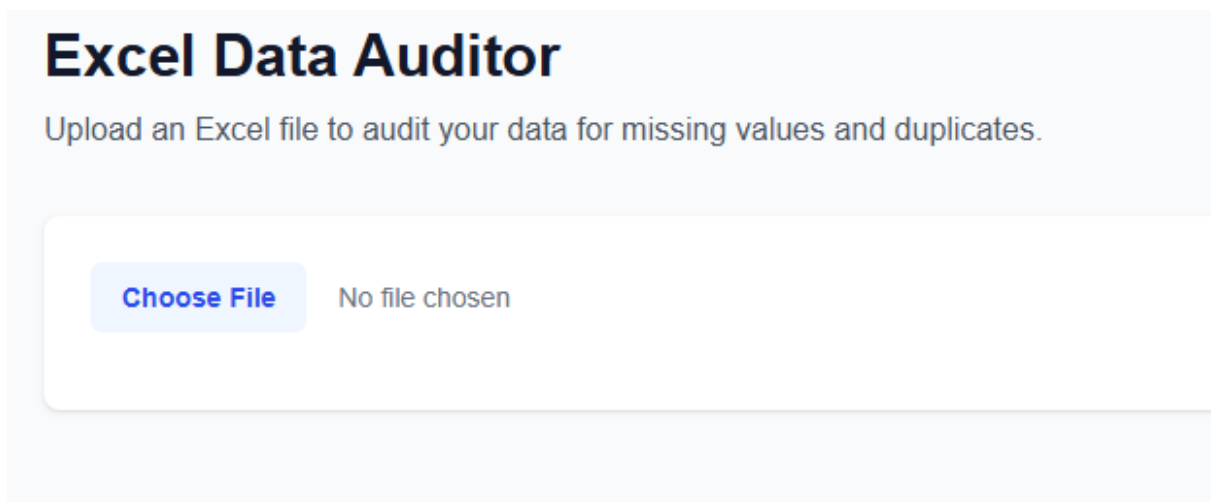- **Sheets (xlsx)** to read and process Excel files directly in the browser

Prompt Text

1. Build a small web app called Excel Data Auditor that runs entirely client-side.

2. I should be able to upload an .xlsx file and preview each sheet as a table.

3. Highlight missing values and duplicate rows. Let me choose a key column set for duplicate detection (e.g., ID+Email) and also offer a simple 'entire row' duplicate mode.

4. Provide toggles for: trim whitespace, case-insensitive comparisons, and treat empty strings/space-only cells as missing.

5. Show a compact summary per sheet with total rows, missing cells per column (count and %), and duplicate count.

6. Allow me to filter to 'only rows with issues'.

7. Include buttons: Download cleaned Excel, Download issues-only CSV, and Reset.

8. Keep the UI minimal with basic CSS and accessible keyboard navigation.

9. Use only Next.js + SheetJS (xlsx)—no other UI libraries.

10. Code should be easy to read so I can adjust rules during review.

Prompt Output - Windsurf Project Setup

1. **Past the URL in the local browser**

2. Upload the excel file

3. Review the result

Excel Data Auditor

Upload an Excel file to audit your data for missing values and duplicates.

Choose File   csv_cleaner_test.xlsx

MessyData

Audit Options

☑ Trim whitespace   ☑ Case-insensitive comparison   ☑ Treat empty/whitespace as missing   ☐ Show only rows with issues

Key Columns for Duplicate Detection

Name   Email   Department   Score   Comments

Audit Summary

Total Rows	Duplicate Rows	Missing Values
9	0	4

Missing Values by Column:

Score: 2 (22.2%)   Comments: 2 (22.2%)

Download Cleaned Excel   **Download Issues CSV**   **Reset**

#	NAME	EMAIL	DEPARTMENT	SCORE	COMMENTS
1	IAN	ian402@example.com	finance	75	excellent
2	IAN	ian402@example.com	finance	75	excellent
3	CHARLIE	charlie200@example.com	it	75	needs improvement
4	CHARLIE	charlie200@example.com	it	75	needs improvement
5	evelyn	evelynexample.com	it	(empty)	(empty)
6	FRANK	frank429@example.com	finance	75	good performer
7	julia	julia455@example.com	Sales	(empty)	(empty)
8	CHARLIE	charlie45@example.com	it	42	good performer
9	CHARLIE	charlie45@example.com	it	42	good performer

4.For Duplicate Detection - Select Name and WindSurf Highlights the rows with duplicate names

Key Columns for Duplicate Detection

Name Email Department Score Comments

Audit Summary

Total Rows	Duplicate Rows	Missing Values
9	6	4

Missing Values by Column:

Score: 2 (22.2%) Comments: 2 (22.2%)

Download Cleaned Excel Download Issues CSV Reset

#	NAME	EMAIL	DEPARTMENT	SCORE	COMMENTS
1	IAN	ian402@example.com	finance	75	excellent
2	IAN	ian402@example.com	finance	75	excellent
3	CHARLIE	charlie200@example.com	it	75	needs improvement
4	CHARLIE	charlie200@example.com	it	75	needs improvement
5	evelyn	evelynexample.com	it	(empty)	(empty)
6	FRANK	frank429@example.com	finance	75	good performer
7	julia	julia455@example.com	Sales	(empty)	(empty)
8	CHARLIE	charlie45@example.com	it	42	good performer
9	CHARLIE	charlie45@example.com	it	42	good performer

5.When You click on Download Cleaned Excel, it removes duplicates and replaces empty rows with N/A and Title Case for Name

Name	Email	Department	Score	Comments
Ian	ian402@example.com	finance	75	excellent
Charlie	charlie200@example.com	it	75	needs improvement
Evelyn	evelynexample.com	it	N/A	N/A
Frank	frank429@example.com	finance	75	good performer
Julia	julia455@example.com	Sales	N/A	N/A
Charlie	charlie45@example.com	it	42	good performer