# Excel Fashion Trends Analysis

**Cleaning the data**

1. Removing duplicates: Data>Remove Duplicates and then confirming with ‘Ok’.

A screenshot of a computer

AI-generated content may be incorrect.

1. Checking for Missing or Blank values.

This was done using the count blank function.

=COUNTBLANK(fashion\_data\_2018\_2022!A1:T661)

This returned zero, indicating that there were no blanks.

1. Ensure consistent text entries.

Although it is likely this data was already validated, consistent spelling and casing was tested for using Power Query.

Data > From Table / Range > Get & Transform Data.

A screen shot of a computer

AI-generated content may be incorrect.

Apply desired cleaning:

* Trim and clean
* Consistent case
* Fix spelling variants.

Close and Load to being cleaned data back to excel.

This enables future data that is imported to be automatically cleaned by these rules every time it is refreshed, handling bulk replacements and separating raw data from cleaned data.

At this point, after inspecting the data I realised that some of it was likely incorrect, for example:



Chino Shorts for Males being categorised as Shorts, but for female being categorised as Blouse. There were similar anomalies throughout this, such as having skirts for ‘Male’, despite a quick search identifying Zara doesn’t sell Male skirts. For now, this was disregarded due to the data likely being generated and simply for practise purposes rather than driving real insight.

As this dataset contained data solely from Zara, the brand column was removed.

**Sales performance analysis**

Added a column ‘revenue’ which is calculated to be (price x sales\_count). This was then sorted in descending order to quickly identify top performers.

A pivot table was then constructured, allowing comparison between material, gender, age and category of clothing with revenue.

A screenshot of a computer

AI-generated content may be incorrect.