### This is the details I have gotten from the Exploratory Data Analysis of the Fashion Inventory Dataset

I decided to use MySql for the EDA. i began by creating a new database, a table int it then importing the cleaned dataset into my workbench.

These are the key summary variables I explored and their results:

| Most Common Brands   * Women - Anouk * Men - Deyann * Boys - Aj Dezines * Girls - Biba * Unisex - Ddecor * Overall - Anouk |
| --- |
| Most Common Product Types   * Women - Straight Kurta * Men - Straight Kurta * Boys - Kurta With Pyjamas * Girls - Lehenga & Blouse with Dupatta * Unisex - N/A * Overall - Straight Kurta |
| Percentage of Clothing   * Women - 61.4% * Men - 26.3% * Boys - 3.9% * Girls - 4.7% * Unisex - 3.7% |
| Most Discounted Product   * Chanderi Blend Kurta with an average discount of 75% |
| Most stocked product |

* Straight Kurta with 1777 In Stock inventory

I also used MySQL to answer my business questions which are:

1. Which Products or Brands are often out of stock?

Anouk, Biba, and Shree are the brands with the most out of stock items in the dataset. Straight Kurta and A-line Kurta are the most out of stock products in the dataset.

1. Are discounted Items being ignored by buyers?

more discounted products are sold than non-discounted products. The ratio of out of stock discounted items to in stock discounted items is less than 1 meaning these Items are selling fine.

1. Do features like size, color, and "ideal for" affect product availability?

size- adult sized products are more available than children sized. Color- common colours like blue,black,beige, white are more available than others. Blue is the most dominant color in the dataset. Ideal for- women and girls have less products in stock than out of stock meaning they sell faster than boys, men and unisex clothing.

|  |
| --- |

| Moving on from SQL I returned to Excel for Inferential Statistics.  I measured the correlation between Variant Price and Variant Compare At Price. The result was 0.785.  I also performed chi-square test comparing the in-stock, out-of-stock and unknown values of the top 10 products in the dataset. The result was 2.928E-95  Then I calculated the 95% confidence interval for the price ofWomen and Men Clothing. This was the result:   | Variables | Women | Men | | --- | --- | --- | | Mean | 13,634.47 | 13,737.82 | | Confidence Interval Upper Bound | 13,864.86 | 14,083.07 | | Confidence Interval Lower Bound | 13,404.09 | 13,392.57 | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |