

Konecta

Task (5) Documentation

Web Scraping + Dashboarding

Malak Ahmed Saber

1. Web Scraping

- Main Site (English Version) :
 - <https://dream2000.com/en>
- Scrapped 5 Categories :
 - <https://dream2000.com/en/mobiles.html>
 - <https://dream2000.com/en/tablets.html>
 - <https://dream2000.com/en/laptop-notebook/laptop.html>
 - <https://dream2000.com/en/accessories/mobile-accessories/power-bank.html>
 - <https://dream2000.com/en/accessories/audio/speakers.html>

2. Data Cleaning with Power Query

- Renamed all queries with their category name
- Dropping Irrelevant Columns:
 - “Add to cart”, “Wishlist”, “Compare”
 - Also 3 columns filled with null values in tablets query
- Removed the “After Discount Price” column from all queries and will calculate it as a measure from “Original Price” – “Saved”
- For “Original Price” column I converted the values like EGP2000 to 2000, using replace EGP -> empty space

- “Save x EGP” column fixed with Transform -> column from examples
- Checked no nulls or ubnormal values in that new column
- Dropped old column

The screenshot shows a data processing interface with a table on the left and a panel on the right.

Table:

	A ^B _C Column1	A ^B _C Column2	A ^B _C Column3
1	Save 1100 EGP	Realme	Realme Note 60X 3GB Ram, 64GB - Marble Black
2	Save 1100 EGP	Realme	Realme Note 60X 3GB Ram, 64GB - Wilderness Green
3	Save 1500 EGP	Xiaomi	Xiaomi Redmi A3x 4GB Ram, 128GB - Midnight Black
4	Save 1500 EGP	Xiaomi	Xiaomi Redmi A3x 4GB Ram, 128GB - Aurora Green
5	Save 1500 EGP	Xiaomi	Xiaomi Redmi A3x 4GB Ram, 128GB - Moonlight White
6	Save 5509 EGP	Samsung	Samsung Galaxy M55 5G, 8GB Ram, 128GB - Light Green
7	Save 5509 EGP	Samsung	Samsung Galaxy M55 5G, 8GB Ram, 128GB - Black
8	Save 2600 EGP	Samsung	Samsung Galaxy A26 5G, 6GB Ram, 128GB - Black
9	Save 4700 EGP	Samsung	Samsung Galaxy A56 5G, 8GB Ram, 128GB - Awesome Pink
10	Save 4700 EGP	Samsung	Samsung Galaxy A56 5G, 8GB Ram, 128GB - Awesome Lightgray
11	Save 5100 EGP	Samsung	Samsung Galaxy A56 5G, 8GB Ram, 256GB - Awesome Lightgray
12	<		>

Text Between Delimiters:

1100
1100
1500
1500
1500
5509
5509
2600
4700
4700
5100

Save 1100 EGP.

- Renamed all columns to have the following for all 5 queries:
 - “Brand”
 - “ProductName”
 - “OriginalPrice”
 - “Saved”
- Changed the “OriginalPrice” and “Saved” columns data types to Whole Number instead of strings
- To finish up :
 - Added “Category” column in each query
 - Combined the 5 queries into 1 table “Dream2000Data” to use for dashboarding
 - Disabled Loading for the original 5 queries : right click -> disable load
 - Close and apply so only “Dream2000Data” is loaded

3. DAX

New column “Sale Price” :

- SalePrice = Dream2000Data[OriginalPrice] - Dream2000Data[Saved]

New Measure “Total Products” :

- Total Products = COUNTROWS(Dream2000Data)

New Measure “Average Sale Price” & Average Original Price:

- Average Original Price = AVERAGE(Dream2000Data[OriginalPrice])
- Average Sale Price = AVERAGE(Dream2000Data[SalePrice])

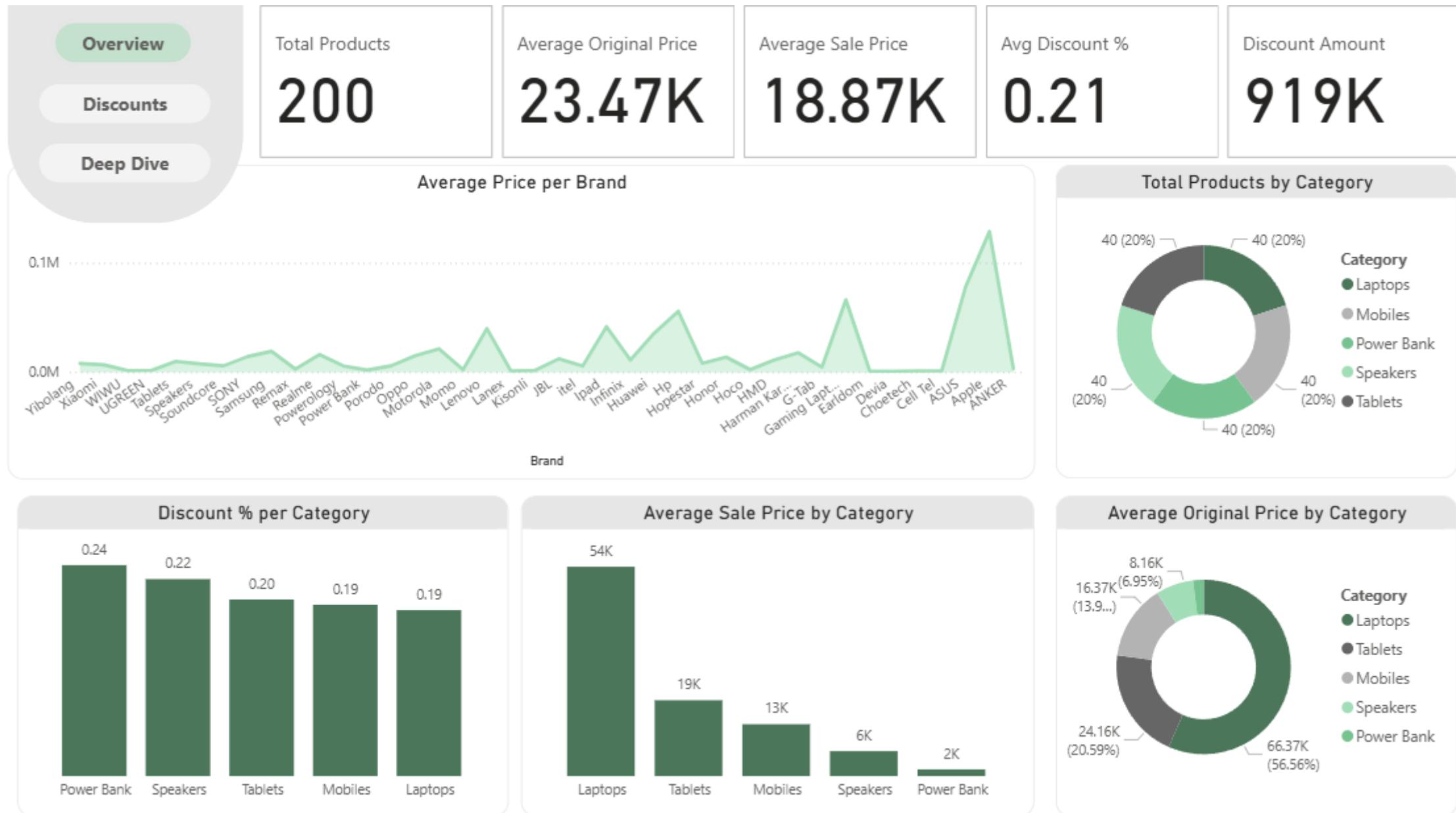
New Measure “Discount Amount” and “Average Discount %”:

- Discount Amount = SUMX(Dream2000Data, Dream2000Data[OriginalPrice] - Dream2000Data[SalePrice])
- Avg Discount % = AVERAGEX(
 • FILTER(Dream2000Data, Dream2000Data[OriginalPrice] > 0 && Dream2000Data[SalePrice] > 0),
 • DIVIDE(Dream2000Data[OriginalPrice] - Dream2000Data[SalePrice], Dream2000Data[OriginalPrice])
 •)

New column “Price Band” to be used in a visual:

- Price Band =
- SWITCH(
 • TRUE(),
 • Dream2000Data[SalePrice] < 2000, "Below 2000",
 • Dream2000Data[SalePrice] >= 2000 && Dream2000Data[SalePrice] < 5000, "2000 - 4999",
 • Dream2000Data[SalePrice] >= 5000 && Dream2000Data[SalePrice] < 10000, "5000 - 9999",
 • Dream2000Data[SalePrice] >= 10000 && Dream2000Data[SalePrice] < 20000, "10000 - 19999",
 • Dream2000Data[SalePrice] >= 20000, "20000+",
 • "Unknown"
 •)

4. Dashboarding



Overview

Discounts

Deep Dive

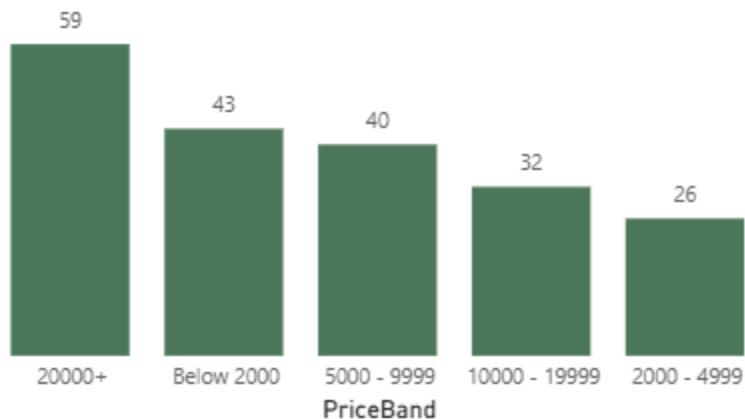
Category

- Laptops
 - Mobiles
 - Power Bank
 - Speakers
 - Tablets

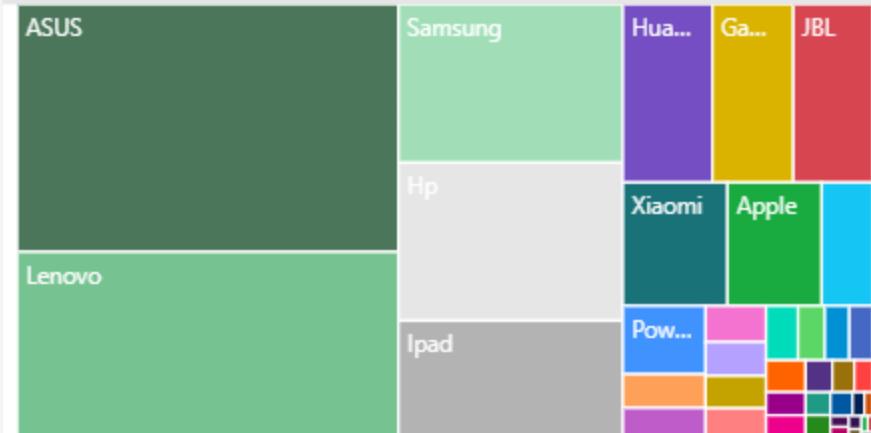
Brand

- ANKER
 - Apple
 - ASUS
 - Cell Tel
 - Choetech
 - Devia
 - Earldom
 - Gaming Lab
 - G-Tab
 - Havit

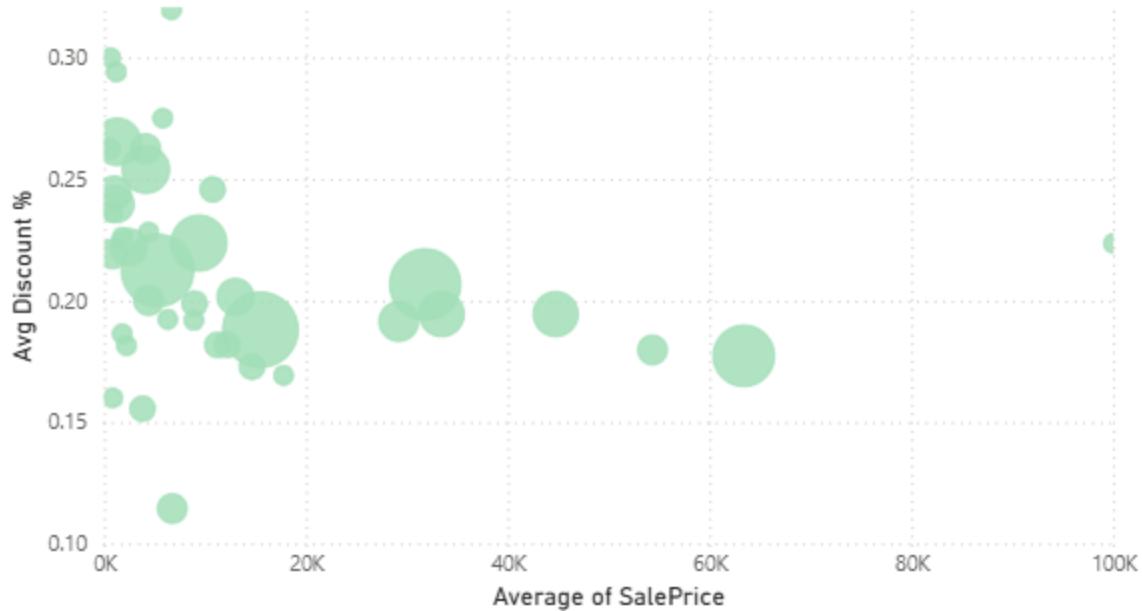
Total Products by PriceBand



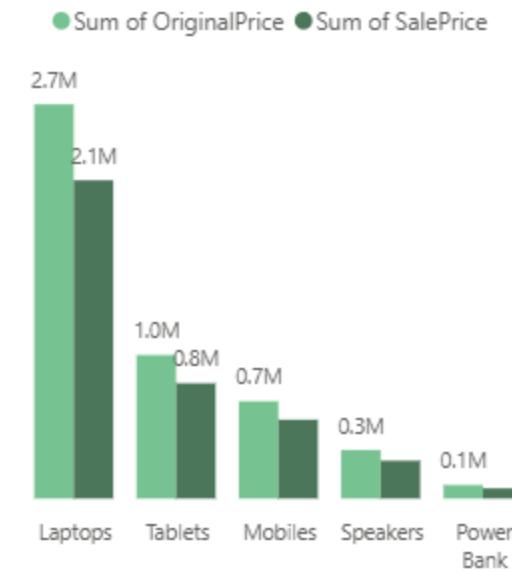
Discount Amount by Brand



Average of SalePrice vs Avg Discount % Sized with Total Products by Brand



Sum of OriginalPrice and Sum of SalePrice by Category



[Overview](#)[Discounts](#)[Deep Dive](#)

Help Q&A understand people better by adding synonyms.

[Add synonyms now](#)

Ask a question about your data



Try one of these to get started

[what is the total products by category](#)[what is the average sale price by category](#)[what is the average original price by category](#)[top categories by total products](#)[top price bands by average sale price](#)

Sum of Saved by ProductName



Total Products by Brand

