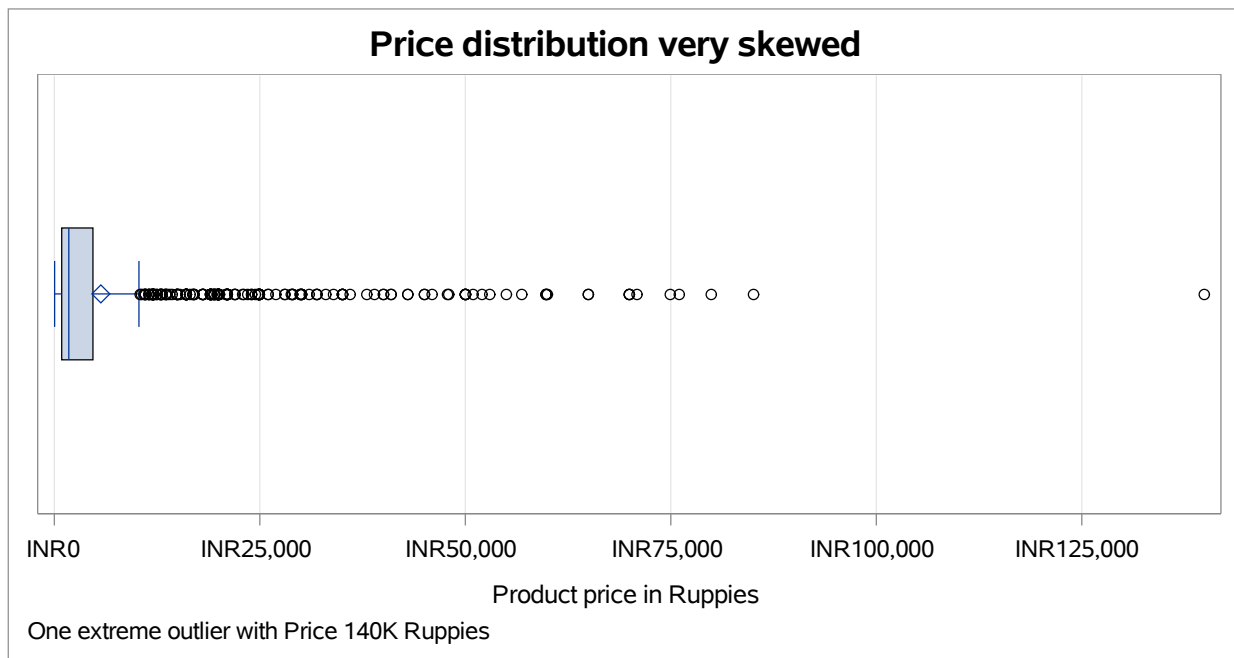


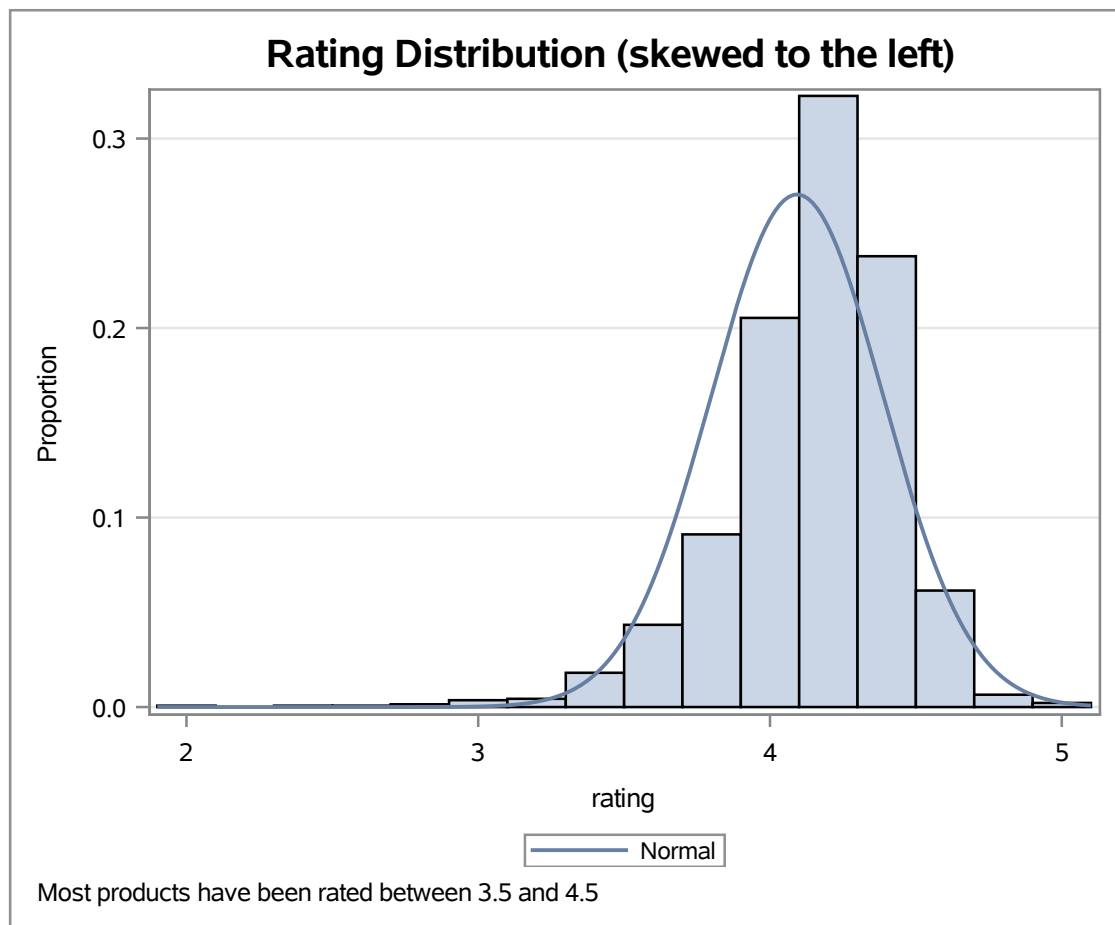
26777 Data Processing Using SAS - Autumn 2023

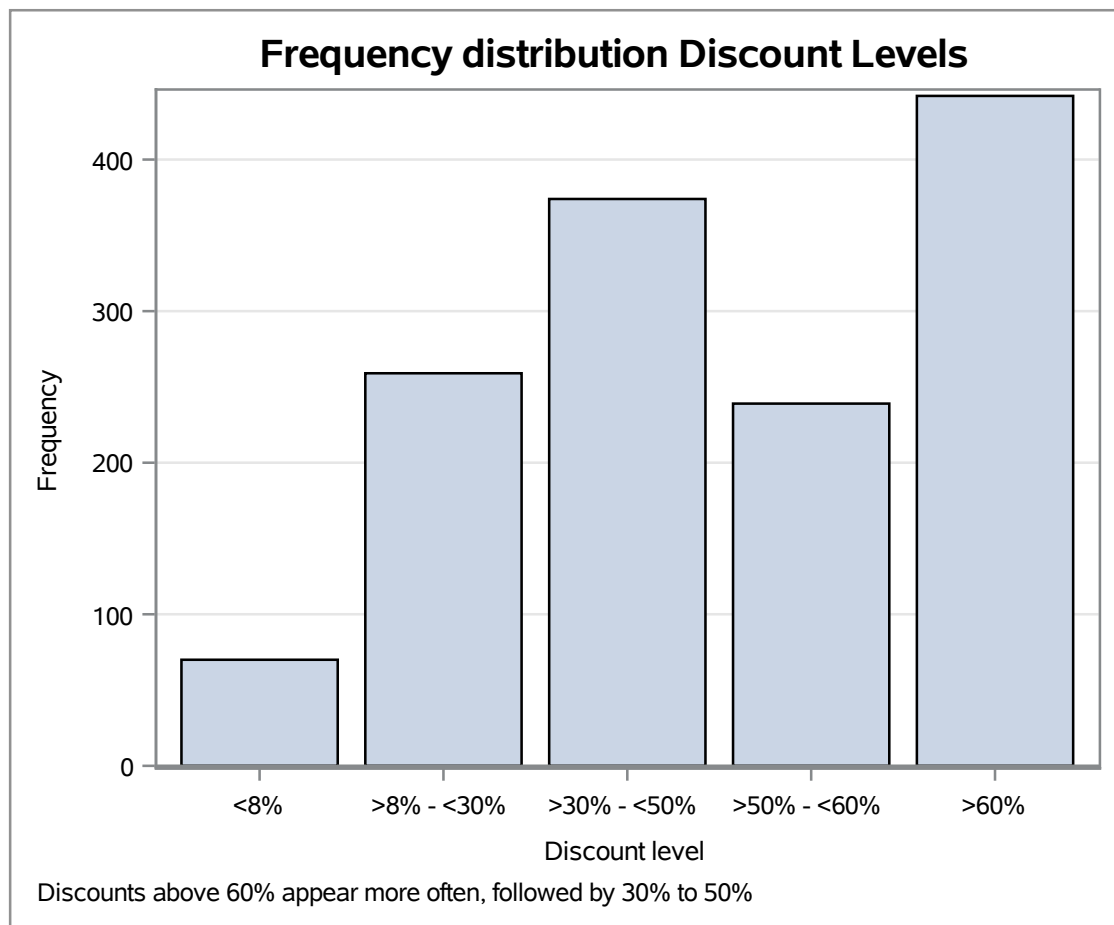
Based on the information provided in Kaggle, the file includes scrapped data from Amazon Website in India. These items are listed online and have been sold before because all of them have reviews, though no information on sales volume is available. The Dataset is mainly informative on each item scraped by the author.

Amazon Dataset analysis

Name: Laila Lima Alves
Student-ID: 14344509







Frequency table for the main product categories

5

Which categories appear more often?

Electronics, Computer and accessories, Home and Kitchen represent 97% of total!

Product main category				
main_category	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Electronics	509	36.78	509	36.78
Home&Kitchen	448	32.37	957	69.15
Computers&Accessories	389	28.11	1346	97.25
OfficeProducts	31	2.24	1377	99.49
HomeImprovement	2	0.14	1379	99.64
MusicalInstruments	2	0.14	1381	99.78
Car&Motorbike	1	0.07	1382	99.86
Health&PersonalCare	1	0.07	1383	99.93
Toys&Games	1	0.07	1384	100.00

Home Improvement + Musical

Instruments + Car&Motorbike + Health & PersonalCare + Toys & Game have 2 or less items listed

Frequency table for the main product categories versus subcategory

6

Which sub-categories appear more often?

Accessories&Peripherals | Kitchen&HomeAppliances | Mobiles&Accessories | HomeTheater,TV&Video

Percent	Table of sub_ctg1 by main_category				
	sub_ctg1(Product Subcategory 1)	main_category(Product main category)			
		Computers&Accessories	Electronics	Home&Kitchen	Total
	Accessories	0.00	0.97	0.00	0.97
	Accessories&Peripherals	23.85	0.00	0.00	23.85
	Cameras&Photography	0.00	1.19	0.00	1.19
	Components	0.37	0.00	0.00	0.37
	CraftMaterials	0.00	0.00	0.52	0.52
	ExternalDevices&DataStorage	1.34	0.00	0.00	1.34
	GeneralPurposeBatteries&BatteryChargers	0.00	1.04	0.00	1.04
	Headphones,Earbuds&Accessories	0.00	4.83	0.00	4.83
	Heating,Cooling&AirQuality	0.00	0.00	8.62	8.62
	HomeAudio	0.00	1.19	0.00	1.19
	HomeStorage&Organization	0.00	0.00	1.19	1.19
	HomeTheater,TV&Video	0.00	11.74	0.00	11.74
	Kitchen&Dining	0.00	0.00	0.07	0.07
	Kitchen&HomeAppliances	0.00	0.00	22.88	22.88
	Laptops	0.07	0.00	0.00	0.07
	Mobiles&Accessories	0.00	11.37	0.00	11.37
	Monitors	0.15	0.00	0.00	0.15
	NetworkingDevices	2.23	0.00	0.00	2.23
	PowerAccessories	0.00	0.07	0.00	0.07
	Printers,Inks&Accessories	0.82	0.00	0.00	0.82
	Tablets	0.07	0.00	0.00	0.07
	WearableTechnology	0.00	5.42	0.00	5.42
	Total	389 28.90	509 37.82	448 33.28	1346 100.00

The 3 main categories are dominated by one or two sub-categories

Frequency table for the price level

Which price level is more frequent?

Prices below 10K represent the highest proportion in the dataset : 86%

Product price level				
price_level	Frequency	Percent	Cumulative Frequency	Cumulative Percent
<2K	796	57.51	796	57.51
>2K - <10K	401	28.97	1197	86.49
>10K - <20K	91	6.58	1288	93.06
>20K - <50K	75	5.42	1363	98.48
>50K - <100K	20	1.45	1383	99.93
>100K	1	0.07	1384	100.00

One outlier with the price value above 100K

Frequency table for the discount level

8

Are discounts common usage?

Discounts are extremely frequent! 76% of the products have discounts > 30%

Discount level				
disc_perc	Frequency	Percent	Cumulative Frequency	Cumulative Percent
<8%	70	5.06	70	5.06
>8% - <30%	259	18.71	329	23.77
>30% - <50%	374	27.02	703	50.79
>50% - <60%	239	17.27	942	68.06
>60%	442	31.94	1384	100.00

Are discounts real or marketing strategy to attract customers? Comparables with other website could answer this question

Summary statistics report for Price, product category and discount level

Which additional insights are available?

The mean & median price of Electronics are much higher than other categories

Home&Kitchen have the second highest prices

The discount levels are not correlated with product prices

Analysis Variable : price Product price in Ruppies					
Product main category	Discount level	Mean	Median	Minimum	Maximum
Car&Motorbike	>30% - <50%	4000	4000	4000	4000
Computers&Accessories	<8%	827	599	39	4333
	>8% - <30%	3103	1599	99	37999
	>30% - <50%	2382	1090	249	59890
	>50% - <60%	1652	1300	299	13750
	>60%	1352	999	199	32000
Electronics	<8%	8799	1850	250	33999
	>8% - <30%	15824	13499	180	69999
	>30% - <50%	17995	3990	171	139900
	>50% - <60%	6988	2299	399	65000
	>60%	4745	1999	299	31990
Health&PersonalCare	>50% - <60%	1900	1900	1900	1900
Home&Kitchen	<8%	1294	825	79	8995
	>8% - <30%	3341	2097	100	12999
	>30% - <50%	4980	2845	299	75990
	>50% - <60%	3860	1970	199	59900
	>60%	4959	1999	499	59900
HomeImprovement	>50% - <60%	799	799	599	999
MusicalInstruments	>30% - <50%	699	699	699	699
	>60%	1995	1995	1995	1995
OfficeProducts	<8%	301	165	50	1295
	>8% - <30%	305	263	75	720
	>30% - <50%	178	178	175	180
	>50% - <60%	2999	2999	2999	2999
	>60%	800	800	800	800
Toys&Games	<8%	150	150	150	150

Details about categories with less than two products online

10

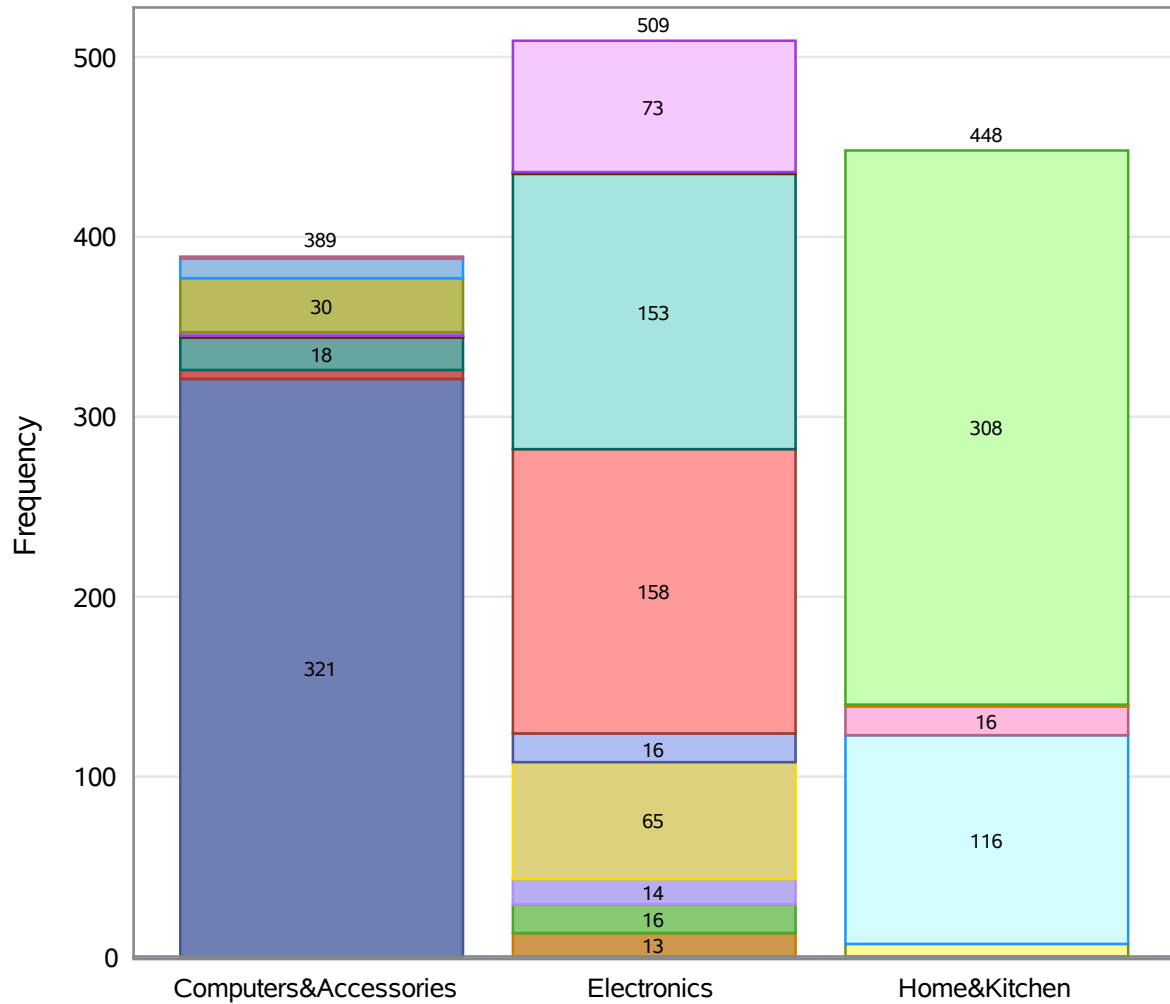
Obs	main_category	sub_ctg1	product_name	price
44	Toys&Games	Arts&Crafts	Faber-Castell Connector Pen Set - Pack of 25 (Assorted)	Rs150
262	MusicalInstruments	Microphones	Boya ByM1 Auxiliary Omnidirectional Lavalier Condenser Microphone with 20ft Audio Cable (Black)	Rs1,995
291	Health&PersonalCare	HomeMedicalSupplies&Equipment	Dr Trust Electronic Kitchen Digital Scale Weighing Machine (Blue)	Rs1,900
342	MusicalInstruments	Microphones	MAONO AU-400 Lavalier Auxiliary Omnidirectional Microphone (Black)	Rs699
410	HomeImprovement	Electrical	Gizga Essentials Cable Organiser, Cord Management System for PC, TV, Home Theater, Speaker & Cables, Reusable Cable Organizer for Desk, WFH Accessories, Organizer Tape Roll, Reusable Cable Ties Strap	Rs599
473	HomeImprovement	Electrical	ESnipe Mart Worldwide Travel Adapter with Build in Dual USB Charger Ports with 125V 6A, 250V Protected Electrical Plug for Laptops, Cameras (White)	Rs999
803	Car&Motorbike	CarAccessories	Reffair AX30 [MAX] Portable Air Purifier for Car, Home & Office Smart Ionizer Function H13 Grade True HEPA Filter [Internationally Tested] Aromabuds Fragrance Option - Black	Rs4,000

Obs	discount	rating
44	Rs150	4.3
262	Rs798	4
291	Rs899	4
342	Rs478	3.8
410	Rs249	4.5
473	Rs425	4
803	Rs2,339	3.8

These products could also have been classified in the main categories due to its characteristics

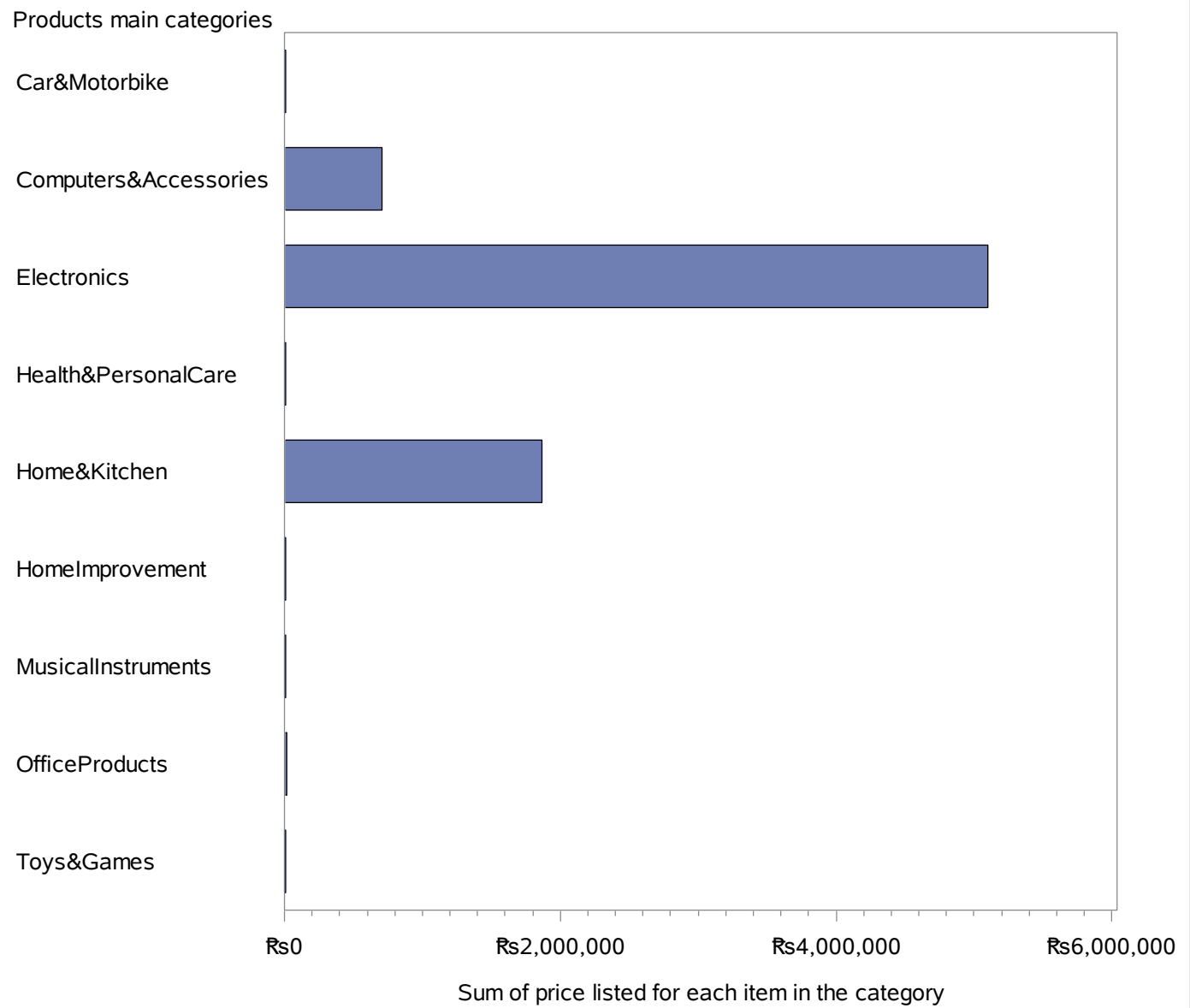
Frequency table for the 3 main product categories (Histogramm)

Graphical visualization from the table previously shown



Bar Chart Products categories and Sum of prices

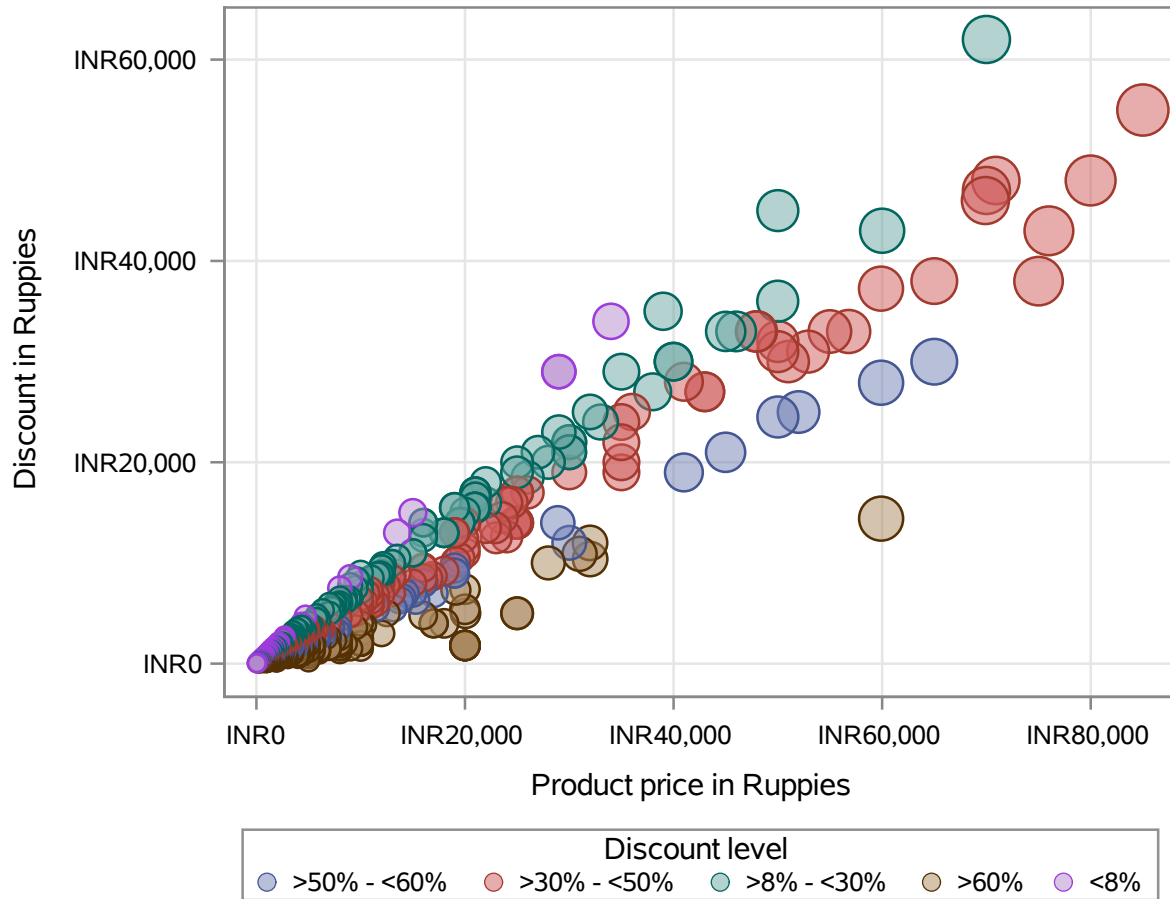
The sum of all prices confirms Electronics is by far the most expensive category



Even when outlier in electronics for 130,000 Ruppies is excluded

Strong correlation between Discount and Prices in Ruppies

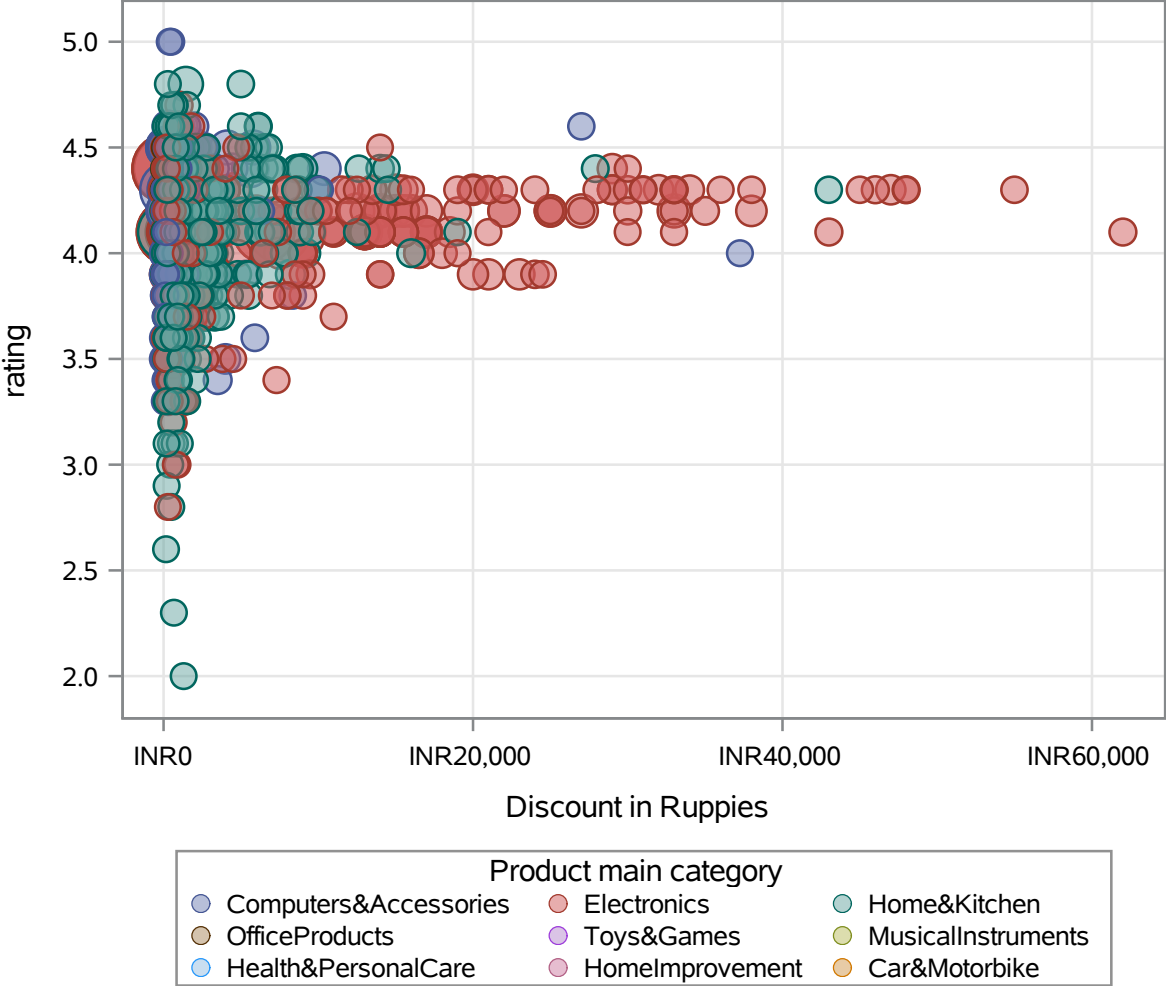
Expected as prices are high, the discounts represent higher values



Outlier excluded. The different bubble colors show the discount levels and follow an expected pattern with higher discount in the bottom

Correlation between Discount and Ratings

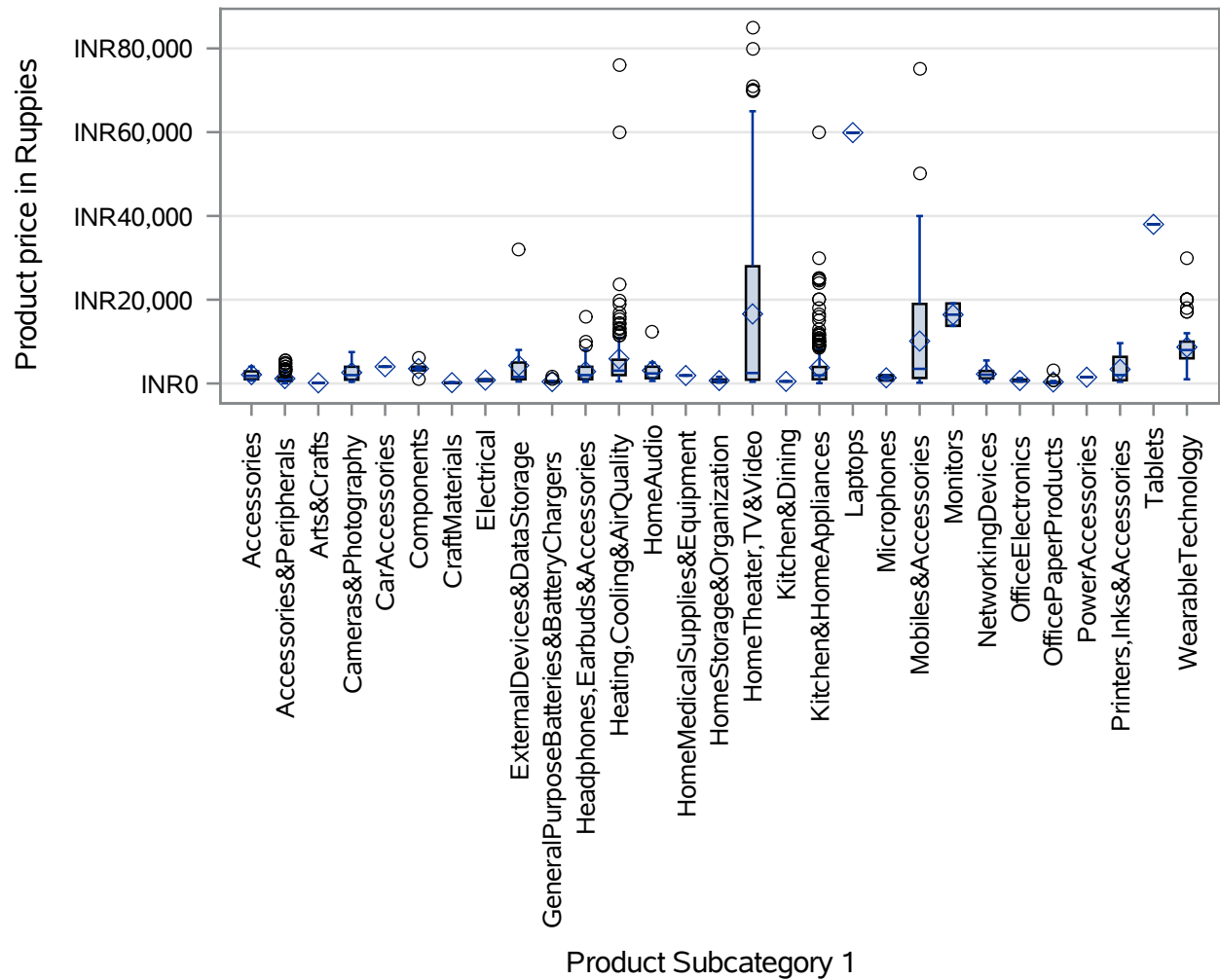
Most Ratings are between 3.5 and 4.5 showing a positive evaluation for most products



Outlier excluded. More expensive products with more discount have similar ratings than cheaper ones

Box-plot price distribution per Product sub-category

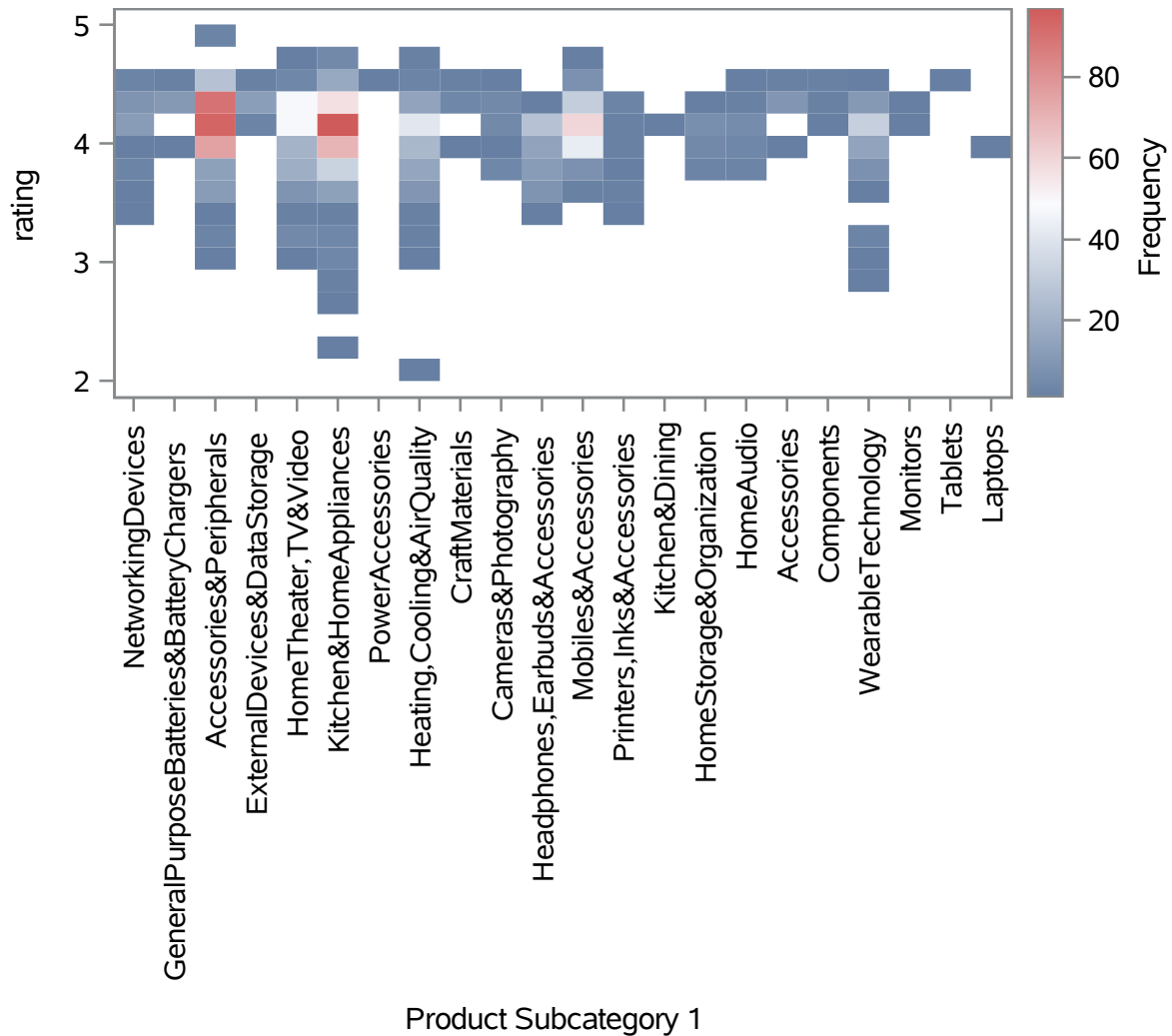
Sub-categories with highest dispersion in price: HomeTheater,TV&Video | Kitchen&HomeAppliances | Heating,Cooling&AirQuality | Mobiles&Accessories



Outlier excluded. Only 5 Sub-categories have prices above 40K

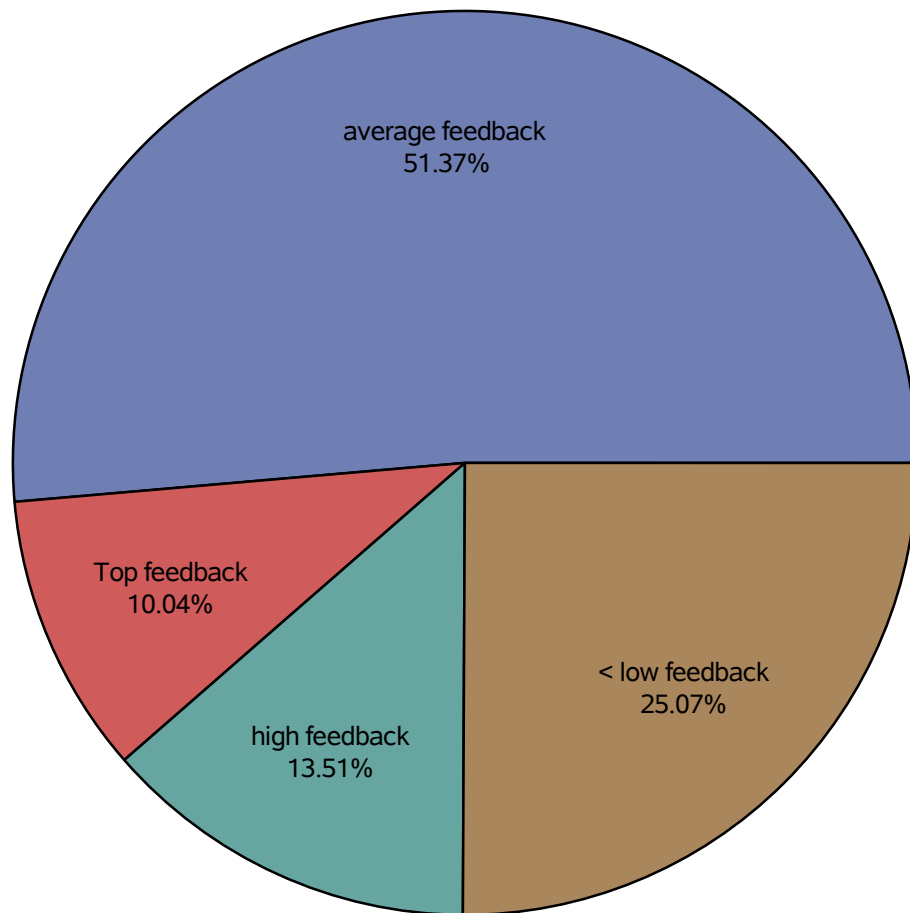
HeatMap for number of Ratings written per Product Sub-Category

Include only category Electronics, Home&Kitchen & Computers&Accessories
Kitchen&HomeAppliances has more Ratings below 3 than others



Sub-categories with more frequency have more Feedback (Accessories&Pheripherals + Kitchen&Homeappliances + Mobile&accessories)

3/4 of the products have more than 18 000 customers writting feedback

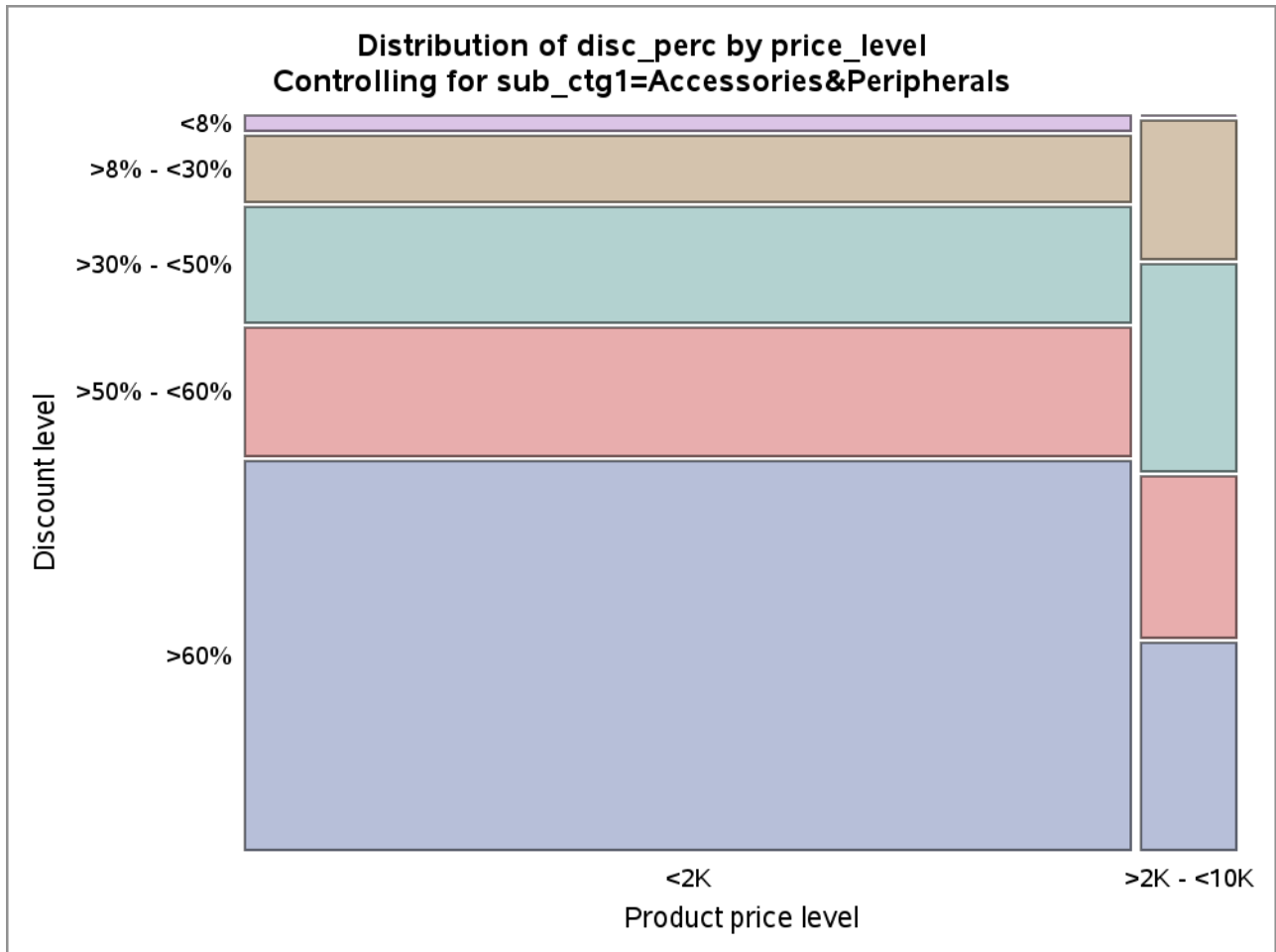


Low Feedback < 1173 | Average Feedback < 18543 | High < 40895 | Top > 40895

Mosaic Plot for Discount and Price levels for Top 4 Sub-categories

18

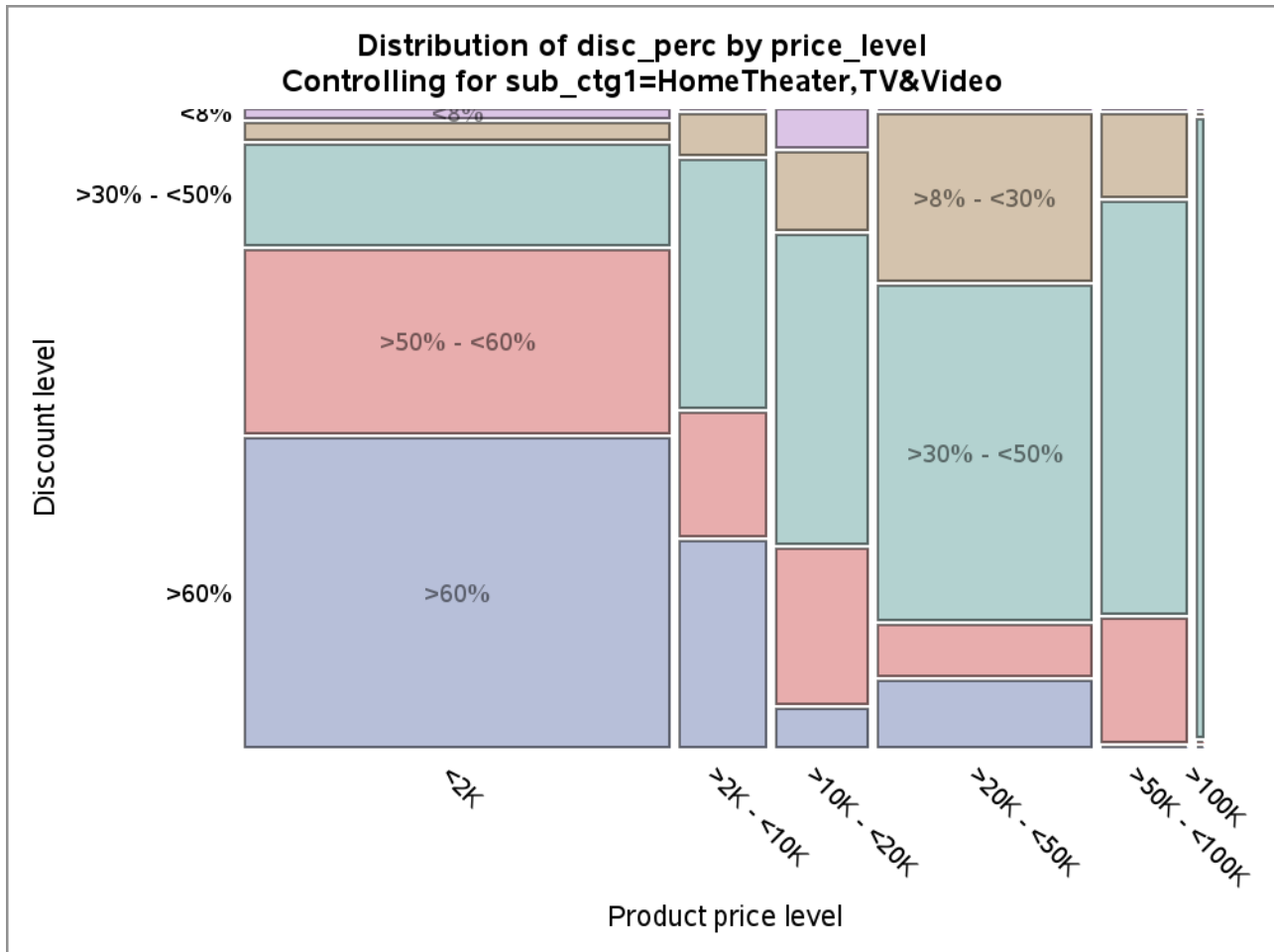
Accessories&Peripherals | Kitchen&HomeAppliances | Mobiles&Accessories | HomeTheater,TV&Video
Accessories&Peripherals have lower prices and higher discounts
Mobiles&Accessories have higher prices and less discounts



Mosaic Plot for Discount and Price levels for Top 4 Sub-categories

19

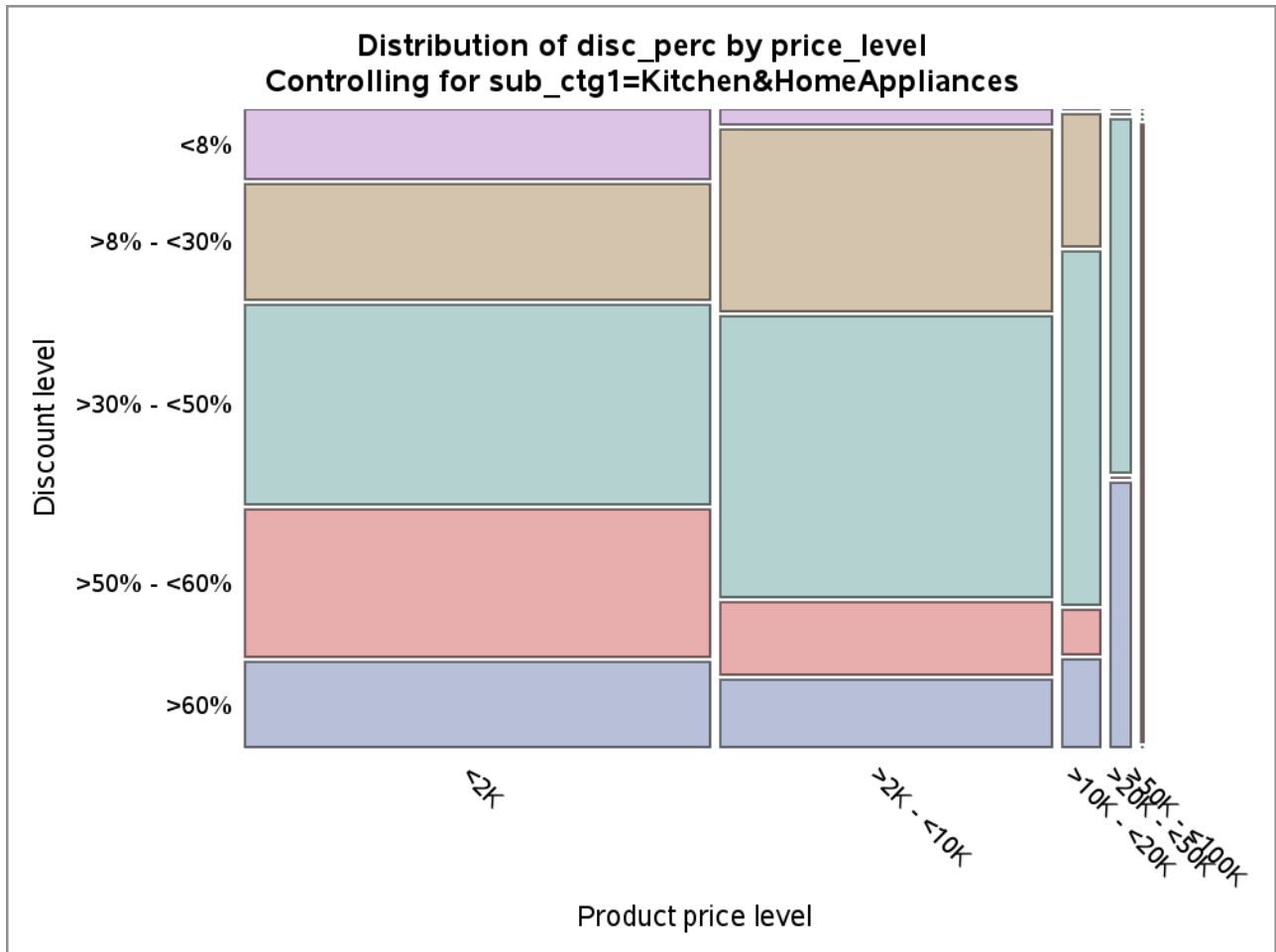
Accessories&Peripherals | Kitchen&HomeAppliances | Mobiles&Accessories | HomeTheater,TV&Video
Accessories&Peripherals have lower prices and higher discounts
Mobiles&Accessories have higher prices and less discounts



Mosaic Plot for Discount and Price levels for Top 4 Sub-categories

20

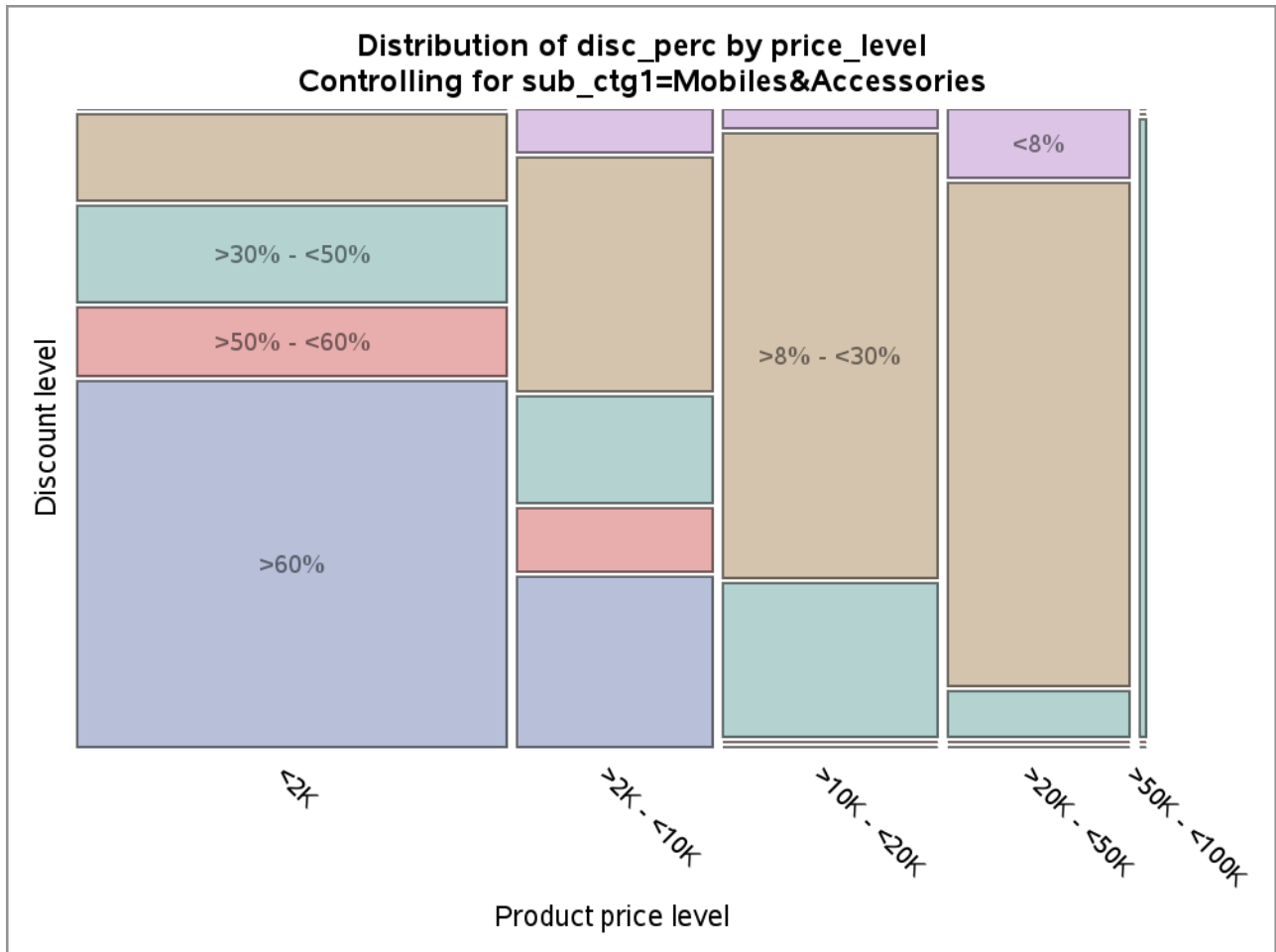
Accessories&Peripherals | Kitchen&HomeAppliances | Mobiles&Accessories | HomeTheater,TV&Video
Accessories&Peripherals have lower prices and higher discounts
Mobiles&Accessories have higher prices and less discounts



Mosaic Plot for Discount and Price levels for Top 4 Sub-categories

21

Accessories&Peripherals | Kitchen&HomeAppliances | Mobiles&Accessories | HomeTheater,TV&Video
Accessories&Peripherals have lower prices and higher discounts
Mobiles&Accessories have higher prices and less discounts



Two way frequency table for Mobiles&Accessories

22

Product main category=Electronics

Percent	Table of price_level by disc_perc						
	price_level(Product price level)	disc_perc(Discount level)					
		<8%	>8% - <30%	>30% - <50%	>50% - <60%	>60%	Total
	<2K	0.00	5.88	6.54	4.58	24.84	41.83
	>2K - <10K	1.31	7.19	3.27	1.96	5.23	18.95
	>10K - <20K	0.65	15.03	5.23	0.00	0.00	20.92
	>20K - <50K	1.96	14.38	1.31	0.00	0.00	17.65
	>50K - <100K	0.00	0.00	0.65	0.00	0.00	0.65
	Total	6 3.92	65 42.48	26 16.99	10 6.54	46 30.07	153 100.00

Confirms that prices > 10K have discount < 30%. Discounts > 60% are for prices < 2K

Two way frequency table for HomeTheater,TV&Video

23

Product main category=Electronics

Percent	Table of price_level by disc_perc						
	price_level(Product price level)	disc_perc(Discount level)					Total
		<8%	>8% - <30%	>30% - <50%	>50% - <60%	>60%	
	<2K	0.63	1.27	7.59	13.92	23.42	46.84
	>2K - <10K	0.00	0.63	3.80	1.90	3.16	9.49
	>10K - <20K	0.63	1.27	5.06	2.53	0.63	10.13
	>20K - <50K	0.00	6.33	12.66	1.90	2.53	23.42
	>50K - <100K	0.00	1.27	6.33	1.90	0.00	9.49
	>100K	0.00	0.00	0.63	0.00	0.00	0.63
	Total	2 1.27	17 10.76	57 36.08	35 22.15	47 29.75	158 100.00

Most discounts are for prices < 2K. Products with price >10K have discounts <50%

Average price for Computers&Accessories

24

Analysis Variable : price Product price in Ruppies				
N	Mean	Std Dev	Minimum	Maximum
389	1816.85	4163.41	39.00	59890.00

Frequency table of price and discounts levels for Computers&Accessories

25

Percent	Table of price_level by disc_perc						
	price_level(Product price level)	disc_perc(Discount level)					Total
		<8%	>8% - <30%	>30% - <50%	>50% - <60%	>60%	
	<2K	2.06	7.97	14.91	15.17	42.67	82.78
	>2K - <10K	0.26	3.60	5.40	3.08	3.60	15.94
	>10K - <20K	0.00	0.00	0.26	0.26	0.00	0.51
	>20K - <50K	0.00	0.26	0.00	0.00	0.26	0.51
	>50K - <100K	0.00	0.00	0.26	0.00	0.00	0.26
	Total	9 2.31	46 11.83	81 20.82	72 18.51	181 46.53	389 100.00

Discounts above 50% for goods below 2K are the most common