

Amazon Sales Analysis

Name – Abhishek Sahu

Roll. No. – 23N0458

Data Source: Kaggle

Various Steps:

- Data cleaning
- Data Transformation: like changing object data type to float data type
- Feature selection: product id, product name, discounted price, rating, rating count, category
- Splitting Columns
- Data saving

Exploratory Data Analysis and Visualization:

- Analyze the product by main category and subcategory
- Analyze the main category and subcategory by rating
- Analyze the product by rating
- Analyze the rating distribution
- Analyze the amount of main category and subcategory by rating score
- Analyze the main category and subcategory by discounted price
- Analyze the active reviewers
- Make co-relation matrices

Library Used:

- Pandas
- Matplotlib
- Seaborn

Conclusion: Now we see products like computer and accessories, electronics, home & kitchen products are more popular than other products and also We noticed that even the products which have not been sold much have good ratings which shows that customers are very satisfied with the products. But the rating of some products which are in regular demand is not good. Their quality will have to be improved to achieve good sales in future.

Conclusions which are my own

- Analyze the product by main category and subcategory
- Analyze the product by rating
- Analyze the rating distribution
- Analyze the amount of main category and subcategory by rating score (for this I get idea from Kaggle)
- Analyze the subcategory by discounted percentage
- Make co-relation matrices
- Conclusion

I did not present many things in my presentation which have In my code. I write the observation for that in google notebook

Reference:

Kaggle (some code I used but 95% of the code is mine)

