Determining Popular Phone Features in India

Adam Hall, Mohamed Hussien, Virginia Sahagun,
Nazmus Sakib Sumon, Shadman Chowdhury

Background Info & Research

- The Indian market is ultra-competitive
- Market expected to grow 10 % in 2023 to reach 175 million units[1]
- New Competitors emerging like Realme,Xiaomi etc.
- Emergence of Ecommerce site like Flipkart helps increase sale of smartphones
- New market entrants needs analysis on features that are the main drivers of popularity among the buyers

Value Proposition

?

Overview of the problem

- What are the main drivers of ratings in Indian Smartphone market?
- No predictive method to determine rating based on smartphone configuration

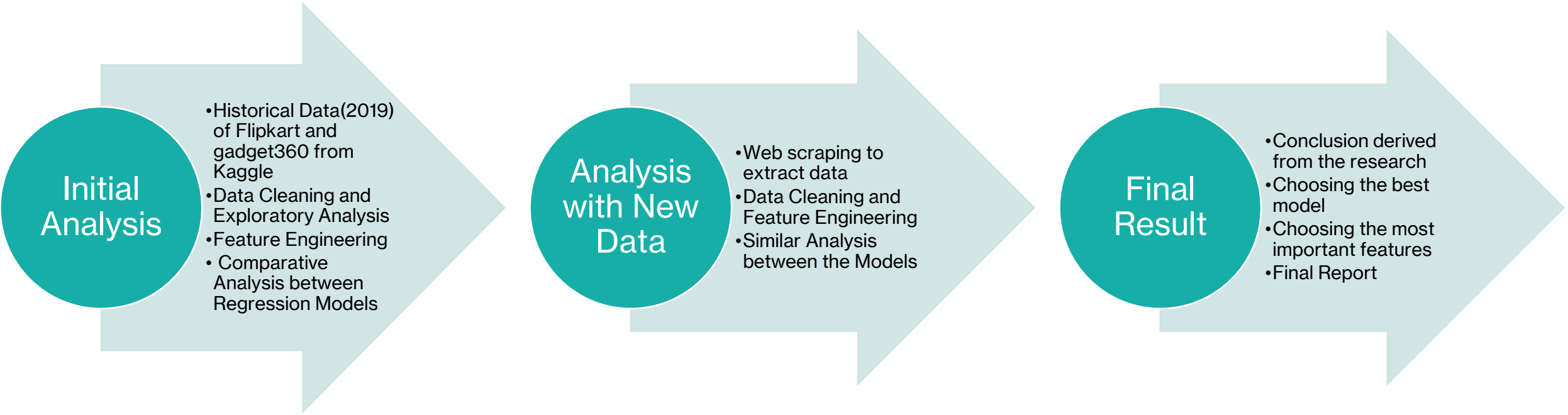
Planned Solution

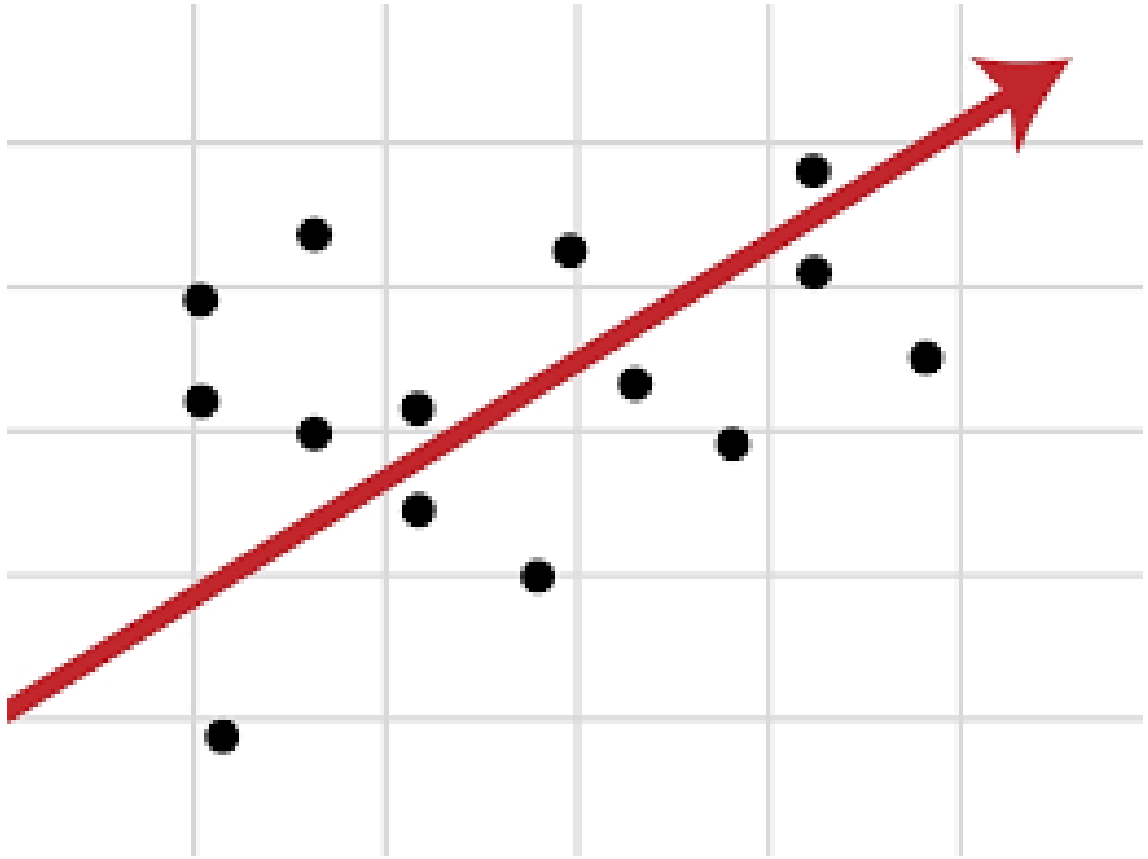
- A Regression based model to predict ratings
- Determining the features that impact the model the most(i.e- main drivers)

Initial Hypothesis

- Research show that price, brand, storage capacity, speed and battery life may be most useful predictors of popularity overall[2]
- Research done in Neighboring country Bangladesh showed that price and operating system are the main factors[3]
- The trend of choice has been similar throughout the years
- Regression based models like Xgboost regressor, Random Forrest regressor , Linear regression can be used to predict rating[4]
- The main drivers of prices are RAM and Battery Life[5]

Planned Approach





Planned ML Models

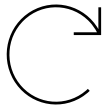
- Linear Regression
- Principal Component Analysis
- Lasso Regression
- XGBoost and Random forest regression
- Compare models with adjusted R^2, MSE, and Accuracy

Data Cleaning and Processing Process

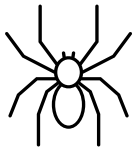
Two different datasets collected from Kaggle

Whitespace removed from the columns that will be joined together

Duplicate and null values removed

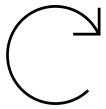


Unit conversion for RAM and Storage size

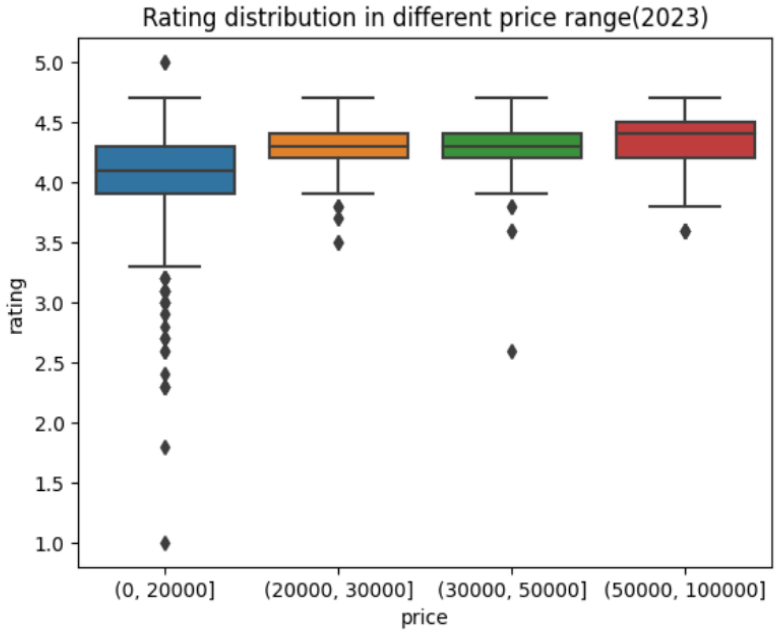


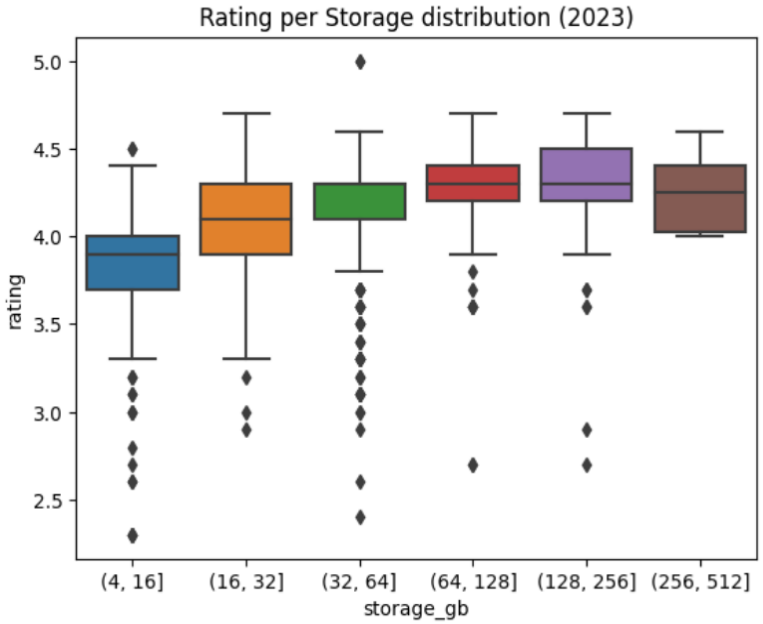
Two different datasets collected by Web scraping

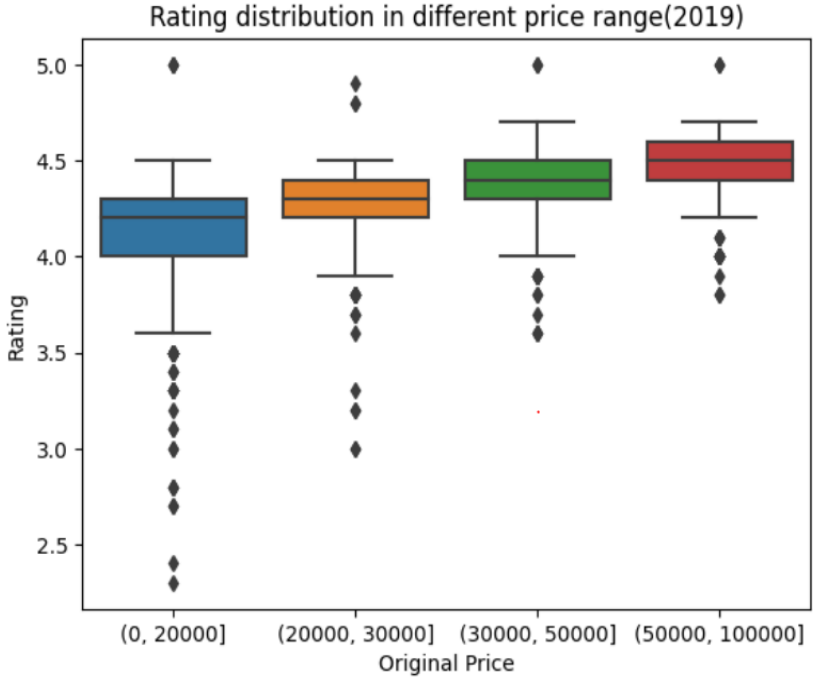
Duplicate and null values removed

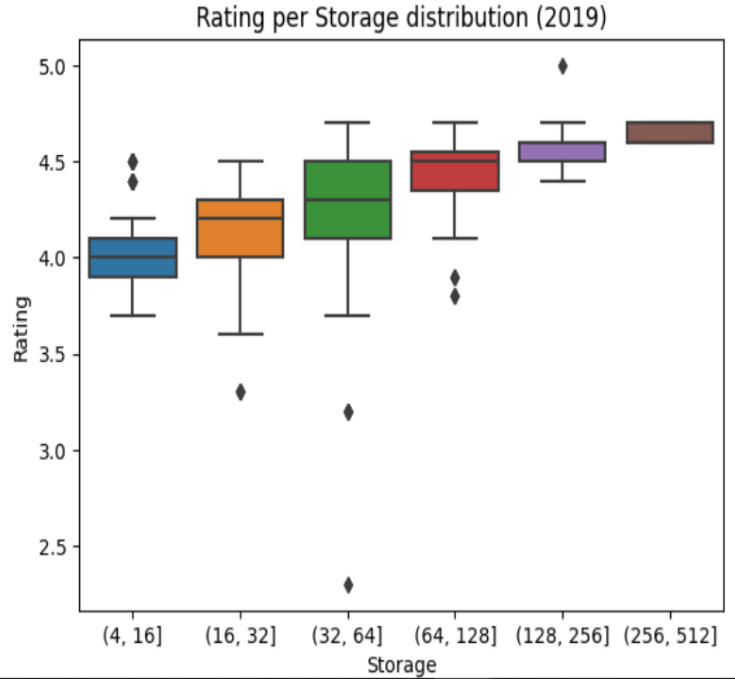


Unit conversion for RAM and Storage size









Findings from Data Visualization

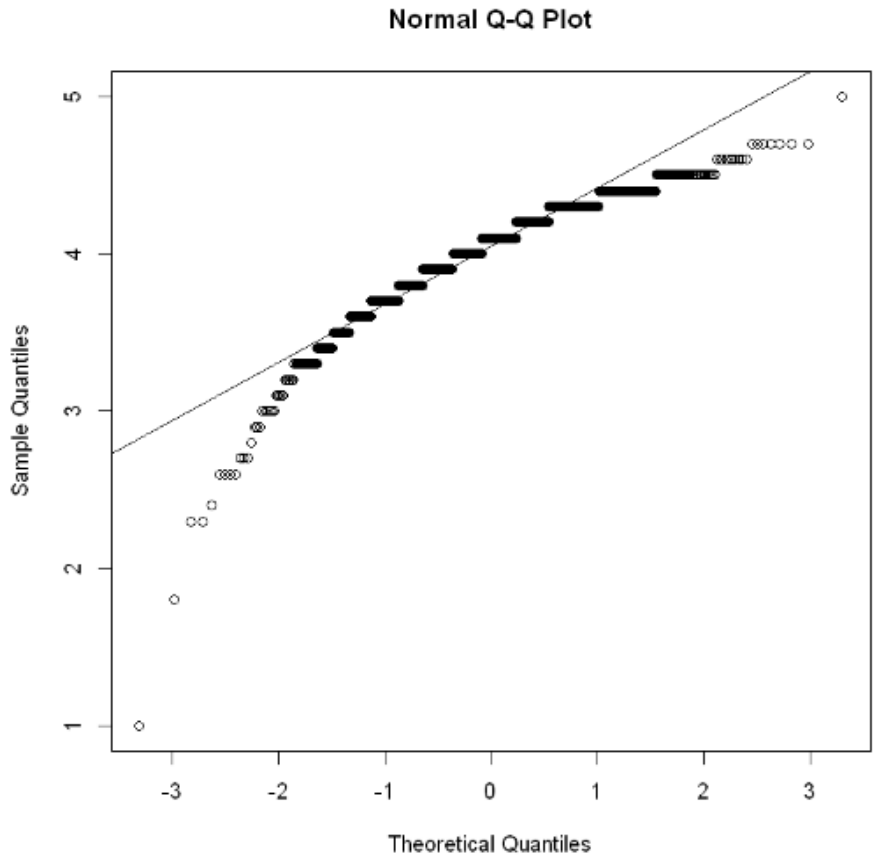
-Increased price results in higher ratings

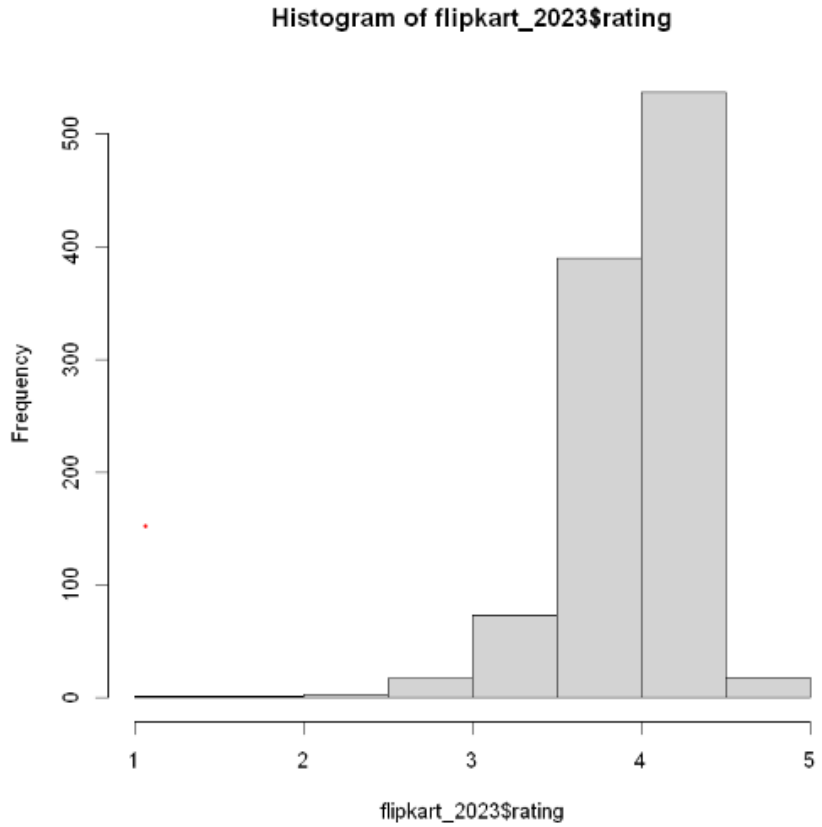
-Increased storage lead to better ratings

- Our initial hypothesis holds in the data visualization

Challenges Found through EDA

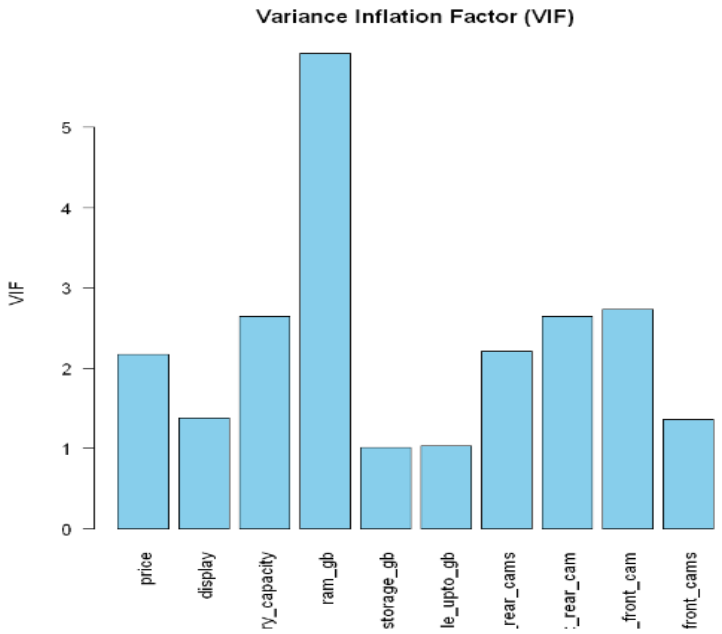
- Response variable not Normally distributed (Y-Variable transformation will be required)

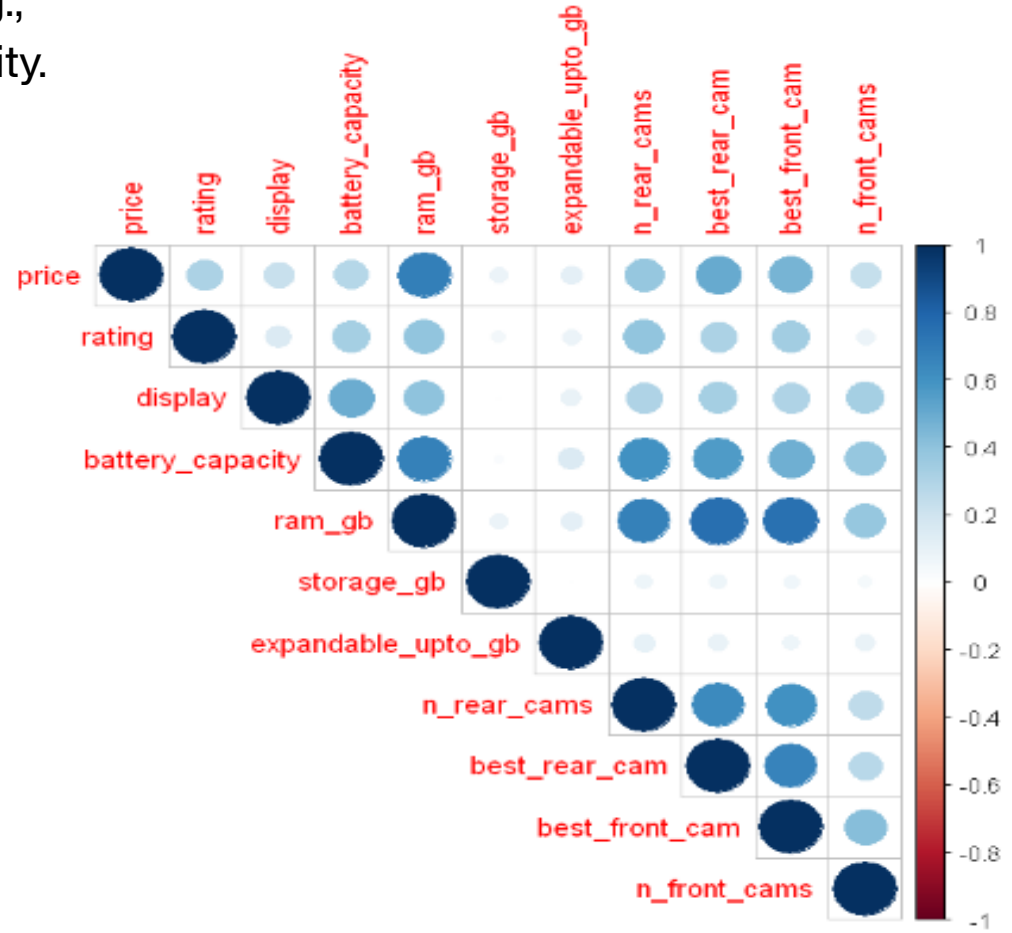




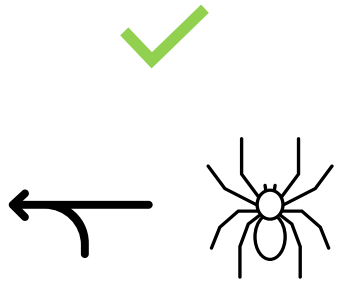
Challenges Found through EDA

- Correlation matrix shows that:
  - Price, Battery capacity, RAM and cameras have the highest correlation coefficient with rating.
  - High correlation coefficient between multiple parameters (e.g., RAM and number of cameras) indicating high risk of collinearity.
- This is further confirmed by the Variance Inflation Factor (VIF) shown below (especially for RAM).

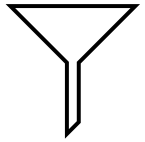




Progress Report

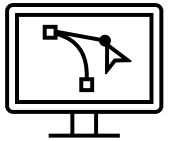


Two different versions(2019 and 2023) of two datasets(Flipkart,Gadget360) collected through Kaggle and webscrapping



Data Cleaning

Preliminary EDA



Data Visualization

Sources Cited

1. B. Standard, "Business Standard," [Online]. Available: https://www.business-standard.com/article/technology/indiasmartphone-market-to-grow-10-to-reach-175-million-units-in-2023-122122300733_1.html.
2. Pekka Kekolahti, Kalevi Kilkki, Heikki Hämmäinen, Antti Riikonen, Features as predictors of phone popularity: An analysis of trends and structural breaks, Telematics and Informatics,Volume 33, Issue 4, 2016, Pages 973-989, ISSN 0736-5853, https://doi.org/10.1016/j.tele.2016.03.001.
3. Uddin, M. R., Lopa, N. Z., & Oheduzzaman, M. (2014). FACTORS AFFECTING CUSTOMERS'BUYING DECISIONS OF MOBILE PHONE: A STUDY ON KHULNA CITY, BANGLADESH. *International Journal of Managing Value and Supply Chains*, *5*(2), 21.
4. https://medium.com/analytics-vidhya/is-it-possible-to-predict-rating-of-google-play-store-apps-based-on-the-given-information-da9a44a3ac1e
5. https://github.com/teguharia172/Phone-Price-Classification-and-Exploratory-Data-Analysis
6. Odaymat, Rommy. (2019). Factors affecting mobile phone selection. https://www.researchgate.net/publication/333461879_Factors_affecting_mobile_phone_selection